



<http://researchspace.auckland.ac.nz>

ResearchSpace@Auckland

Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage.

<http://researchspace.auckland.ac.nz/feedback>

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the [Library Thesis Consent Form](#) and [Deposit Licence](#).

Note : Masters Theses

The digital copy of a masters thesis is as submitted for examination and contains no corrections. The print copy, usually available in the University Library, may contain corrections made by hand, which have been requested by the supervisor.

Oral Corrective Feedback, Individual Differences, and L2 Acquisition of French Past Tenses

Nadja Mifka Profozic

*A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of
Philosophy in Applied Linguistics*

Department of Applied Language Studies and Linguistics, The University of Auckland, 2012.

Abstract

The research reported in this thesis aimed at investigating the differential effects of two types of oral corrective feedback in a foreign language classroom: recasts as an input-providing strategy and clarification requests as an output-prompting strategy (Ellis, 2006; Ellis & Sheen, 2011). Drawing on the Interaction hypothesis (Long, 1981a, 1983b, 1996) both strategies can be considered a type of focus-on-form. The effectiveness of oral corrective feedback was examined in relation to the acquisition of two French past tenses – the passé composé and the imparfait, which have been documented among the most difficult language structures for French L2 learners (Bardovi-Harlig, 2000; Harley, 1989, 1993; Kaplan, 1987). The effects of corrective feedback were also examined in relation to the mediating role of the learners' individual differences in language analytic ability, working memory, and anxiety.

The study was conducted with three intact classes involving 52 high school students learning French as a foreign language in New Zealand. A quasi-experimental design was employed, with a pre-test, treatment, immediate post-test and a delayed post-test. Oral corrective feedback was provided in the form of recasts in one class and clarification requests in the other class as the learners performed three picture-based, information-gap tasks designed to encourage communication and elicit the use of the two past tenses. Acquisition was measured by means of similar picture-based narrative tasks requiring relatively free oral and written production. The tasks were counterbalanced at the three testing times to ensure that there was no task-effect. The accuracy of use of the two tenses was scored by means of the target-like use analysis (Pica, 1983) which takes into account the overuse of the target forms.

Mixed design Repeated measures ANOVAs, followed by the ANCOVAs, were computed to examine the effects of the treatment. Overall, the results indicate that recasts were more effective than clarification requests for the acquisition of both target structures. Recasts resulted in a high level of uptake with repair but, by and large, this was not associated with learning, except for the long term gains in oral production of passé composé. The mediating role of individual learner differences was examined by means of Multiple regression analysis. Language analytic ability predicted the gains for both grammatical structures in oral

production, and for passé composé in written production of the Clarification Request group. In the Recast group analytic ability predicted only the long-term gains for passé composé in oral production. Working memory measured by the Speaking span test predicted only the short-term gains in oral production of passé composé in the Recast group. Language learning anxiety did not appear to have any influence on the effects of corrective feedback on learning.

The results are discussed in terms of cognitive-interactionist theories of L2 acquisition. In general they are supportive of the Interaction and Noticing hypotheses, and of the role played by input rather than of the skill-learning theory and the role of modified output. The findings suggest that teachers need not be afraid to make the corrective force of recasts clear to learners during communicative tasks.

*To the memory of my parents and grandparents:
to life at the borders,
where languages, like people, meet and interact.*

Acknowledgements

I would like to express my sincere gratitude to the following people.

I am especially grateful to Professor Rod Ellis, my main supervisor, for his expert advice and guidance over the past four years, and particularly for the time he spent reading my numerous drafts and revisions, for his prompt feedback and insightful, most valuable comments.

Dr. Helen Basturkmen, my co-supervisor, assisted at the beginning stages of my study, gave me encouragement, generous support and helpful advice.

Dr. Jenefer Philp sparked my interest in learner individual differences while I was her Masters student. Thanks to her, I embarked on this long, challenging and exciting journey.

Russell Greenwood proofread this thesis.

Staff at the Student Learning Centre provided workshops, and showed willingness to answer all my questions over the years, in particular those regarding the challenges of statistical analyses.

Thanks to my officemates and all doctoral students in the Department, for their collegial and friendly support, for sharing their commitment and hard work on weekdays and weekends.

Thanks to all staff in the Department of Applied Language studies and Linguistics, for creating an environment supportive of learning and research.

To the anonymous students, French language teachers, HODs and principals at participating schools, who generously allowed me into their classes I offer my gratitude. I am indebted for their patience and invaluable contribution to the data collection. Without their participation this research would have never been possible.

Thank you to my cheerful senior French class who were with me from the middle of 2006 for the next several years. Small in number but great in attitude, they helped me immensely to trial numerous tasks prior to conducting the Pilot study. From the bottom of my heart, thanks to Charlotte, Elizabeth, Heidi, Lisa, Suzanne, Ben, Cedric, Kennedy, Marcel and Simba.

Thanks to the NZ Ministry of Education, for providing me with the Teachers Study Award in 2008 and 2009 when I was still teaching while conducting my research.

Thanks also to the University of Auckland for providing me with a Partnership Appeal Award in 2010.

Last but not least, I acknowledge my gratitude to my closest family. To my step-mother Olga, for her incredible support and love over the years, for her example of strength and endurance. To my son and my daughter, Neven and Anja, for their unconditional love, understanding and patience. They have been my most wonderful friends, my motivation and everlasting inspiration.

Table of Contents

Abstract.....	ii
Acknowledgements.....	v
List of Figures	xiii
List of Tables	xiv
Chapter 1. Introduction.....	1
1.1. Overview	1
1.2. Theoretical significance of CF	1
1.2.1. Linguistic background	2
1.2.2. Interactionist background	2
1.2.3. Psycholinguistic background	4
1.3. Pedagogical significance of CF	4
1.4. Significance of research involving individual differences	5
1.5. Personal teaching and learning experience.....	5
1.6. Overview of the study.....	7
1.7. Outline of the thesis.....	8
Chapter 2. Theoretical and pedagogical perspectives of CF	9
2.1. Introduction	9
2.2. Theoretical perspectives	9
2.2.1. Taxonomy of oral corrective feedback	9
2.2.2. Recasts	10
2.2.3. Requests for clarification	10
2.2.4. Learner uptake, repair and self-repair.....	11
2.2.5. Universal Grammar-based approaches to CF	11
2.2.6. Cognitive approaches to CF	13
2.2.7. Interactionist approaches to CF.....	17
2.2.8. Skill learning theory	18
2.2.9. Output Hypothesis	19
2.2.10. Socio-cultural approaches to CF	19
2.3. Pedagogical perspectives	20
2.3.1. Error correction.....	20
2.4. Summary	22

Chapter 3. Research on oral corrective feedback	24
3.1. Introduction	24
3.2. Classroom descriptive studies.....	24
3.2.1. Corrective feedback and uptake	25
3.2.2. Contrasting findings	28
3.2.3. Context and instructional setting.....	28
3.2.4. Counterbalance Hypothesis	29
3.2.5. Characteristics of recasts	30
3.2.6. Learner perceptions	31
3.3. Experimental laboratory studies.....	32
3.3.1. Positive and negative evidence.....	32
3.3.2. Learner ‘readiness’	33
3.3.3. Modified output following clarification requests	34
3.3.4. Clarification requests, prompts and recasts	34
3.4. Classroom quasi-experimental studies	36
3.4.1. Interactional modification	36
3.4.2. Clarification requests	37
3.4.3. Recasts	37
3.4.4. Prompts vs. recasts	38
3.4.5. Proficiency as a factor	39
3.4.6. Explicit correction vs. recasts.....	39
3.5. Meta-analyses of oral corrective feedback.....	41
3.6. Summary and conclusion	42
Chapter 4. Individual learner differences.....	45
4.1. Introduction	45
4.2. Language aptitude	45
4.3. Language analytic ability.....	47
4.3.1. Matching Instruction to IDs and Aspects of Learning	47
4.3.2. Learning Conditions and Learning Mechanisms	48
4.3.3. Analytic Ability and Teaching Methods	49
4.3.4. Analytic Ability and Incidental Learning.....	50
4.3.5. Analytic Ability and Corrective Feedback	51
4.4. Working memory	53
4.4.1. Phonological Loop.....	53

4.4.2. Working Memory Capacity	55
4.4.3. Working Memory and Type of Instruction.....	57
4.4.4. Working Memory and Corrective Feedback	57
4.4.5. Language learning anxiety	59
4.4.6. Anxiety in the classroom.....	59
4.4.7. Anxiety and Corrective Feedback	60
4.5. Summary	62
Chapter 5. Pilot study.....	63
5.1. Introduction	63
5.2. Research questions	63
5.3. Method	64
5.3.1. Instructional setting	64
5.3.2. Design.....	64
5.3.3. Participants	65
5.3.4. Instruments.....	66
5.3.5. Procedures	69
5.3.6. Testing.....	72
5.3.7. Data analysis	75
5.4. Summary of main findings	76
5.4.1. Problems encountered in the pilot study and changes for the main study	77
5.4.2. Changes to research questions	79
Chapter 6. Methodology	81
6.1. Introduction	81
6.2. Research questions	81
6.3. Design of the study	82
6.4. Participants and research site.....	83
6.5. Target structures.....	85
6.6. Instruments for data collection	87
6.7. Procedures for data collection.....	88
6.7.1. Treatment	88
6.7.2. Coding and scoring the tests.....	95
6.8. Data analysis	100
Chapter 7. Oral production results.....	101
7.1. Introduction	101

7.2. Data Analysis	101
7.2.1. Descriptive statistics	101
7.2.2. Inferential statistics.....	102
7.3. Results for RQ 1.....	103
7.3.1. Results for passé composé: descriptive statistics	103
7.3.2. Inferential statistics.....	105
7.3.3. Results for imparfait: descriptive statistics	108
7.3.4. Inferential statistics.....	110
7.4. Summary	112
7.5. Results for RQ 2.....	114
7.5.1. Results for passé composé.....	114
7.5.2. Results for imparfait.....	116
7.6. Summary	118
7.7. Discussion.....	119
7.7.1. Input-providing vs. output-prompting CF	120
7.7.2. The target structures	122
7.7.3. Between-group comparisons.....	124
7.7.4. Counterbalance Hypothesis	124
7.7.5. Effects of starting proficiency	125
7.7.6. Concluding comments	126
Chapter 8. Written production results.....	128
8.1. Introduction	128
8.2. Data Analysis	128
8.2.1. Descriptive statistics	128
8.2.2. Inferential statistics.....	128
8.2.3. Data analysis for RQ 4.....	129
8.3. Results for RQ 3.....	129
8.3.1. Results for passé compose: descriptive statistics	129
8.3.2. Results for imparfait.....	133
8.4. Summary	137
8.5. Results for RQ 4.....	139
8.5.1. Results for passé composé.....	139
8.5.2. Results for imparfait.....	141
8.6. Summary	143

8.7. Discussion.....	144
Chapter 9. Uptake with repair and learning gains	147
9.1. Introduction	147
9.2. Data analysis	147
9.3. Results.....	148
9.3.1. Repair and gains for passé composé.....	149
9.3.2. Repair and gains for imparfait	151
9.4. Summary	156
9.5. Discussion.....	156
Chapter 10. Results for learner individual differences	162
10.1. Introduction	162
10.2. Data analysis	162
10.3. Results.....	164
10.3.1. Descriptive statistics	164
10.3.2. Multiple regression analyses	165
10.3.3. Summary	168
10.4. Discussion.....	169
Chapter 11. Conclusion	173
11.1. Introduction	173
11.2. Summary of main findings	173
11.2.1. Main findings for effects of CF on oral production.....	173
11.2.2. Main findings for effects of CF on written production	174
11.2.3. Role of uptake with repair in language learning.....	175
11.2.4. Mediating role of individual difference factors	176
11.3. Theoretical implications.....	176
11.3.1. The relative effects of recasts vs. output-prompting CF.....	177
11.3.2. The relative effects of implicit vs. explicit CF.....	179
11.3.3. The role of uptake with repair	180
11.3.4. Individual differences and corrective feedback.....	181
11.3.5. Effects of CF on the different grammatical structures	183
11.4. Pedagogical implications.....	183
11.5. Limitations of the study	185
11.6. Suggestions for future research.....	186
References	188

Appendix A	Ethics Forms	206
Appendix B	References for illustrations used in the study	211
Appendix C	Materials used for testing in oral production	212
Appendix D	Materials used for testing in written production	213
Appendix E	Language analysis test.....	215
Appendix F	Speaking span test.....	218
Appendix G	Anxiety questionnaire.....	220
Appendix H	Pilot Study : Results	222
12.1.	Research question 1.....	222
12.1.1.	Oral passé composé	222
12.1.2.	Written passé composé	222
12.1.3.	Oral imparfait.....	223
12.1.4.	Written imparfait	223
12.2.	Research question 2.....	224
12.3.	Research question 3.....	225
12.4.	Research Question 4	225
12.4.1.	Results for anxiety.....	225
Appendix I	Transcripts from the treatment tasks	227
Appendix J	Transcripts from tests	235

List of Figures

Figure 1 Design of the main study	83
Figure 2 Error treatment sequence in the present study	89
Figure 3 Mean scores for passé composé: the RE group, the CR group and the CN group.....	104
Figure 4 Repeated measures ANOVA: performance of three groups over time	107
Figure 5 Mean scores for imparfait: the RE group, the CR group and the CN group.....	109
Figure 6 Repeated measures ANOVA: performance of three groups over time	111
Figure 7 Mean scores for passé composé: the RE group, the CR group and the CN group.....	130
Figure 8 Repeated measures ANOVA: performance of three groups over time	132
Figure 9 Mean scores for imparfait in written production	134
Figure 10 Repeated measures ANOVA: performance of three groups over time	136

List of Tables

Table 1	Taxonomy of oral CF (Sheen & Ellis, 2011).....	10
Table 2	Basic meaning of past tenses, involving aspect, in French and English	86
Table 3	Descriptive statistics for passé composé.....	104
Table 4	Number of CF episodes for the PC and number of learners participating in interaction	105
Table 5	Adjusted mean scores for the RE and the CR group on post-test 1 and post-test 2.....	105
Table 6	Results of Repeated measures ANOVA: within group differences	106
Table 7	Descriptive statistics for imparfait in oral production	108
Table 8	Number of CF episodes for the imparfait and number of learners participating in interaction	109
Table 9	Adjusted mean scores for the RE and the CR group	110
Table 10	Results of Repeated measures ANOVA: within group differences	111
Table 11	Summary of results for RQ 1: Within group comparisons	113
Table 12	Summary of results for RQ 1: Between group comparisons.....	113
Table 13	Descriptive statistics for 'low' and 'high' sub-groups on each test of passé composé	114
Table 14	Adjusted mean scores for RE and CR 'low' and 'high' sub-groups on each post-test	114
Table 15	Effect sizes for each sub-group, observed between tests.....	114
Table 16	Descriptive statistics for 'low' and 'high' sub-groups on each test of imparfait	116
Table 17	Adjusted mean scores for RE and CR 'low' and 'high' sub-groups on each post-test	116
Table 18	Effect sizes for each sub-group, observed between tests.....	117
Table 19	Summary of results for RQ 2	119
Table 20	Descriptive statistics for passé composé in written production	130
Table 21	Adjusted mean scores for the RE and the CR group on post-test 1 and post-test 2	131
Table 22	Results of Repeated measures ANOVA: within group differences	131
Table 23	Descriptive statistics for imparfait	133
Table 24	Adjusted mean scores for RE group and CR group on post-test 1 and post-test 2.....	134
Table 25	Results of Repeated measures ANOVA: within group differences	135
Table 26	Summary of results for RQ 3: Within group differences.....	138
Table 27	Summary of results for RQ 3: Between group differences	138
Table 28	Descriptive statistics for 'low' and 'high' sub-groups on each test of passe compose	139
Table 29	Adjusted mean scores for 'low' and 'high' sub-groups on each post-test	139
Table 30	Effect sizes for each sub-group, observed between tests over time	139
Table 31	Descriptive statistics for 'low' and 'high' sub-groups on each test of imparfait	141
Table 32	Adjusted mean scores for 'low' and 'high' sub-groups on each post-test	141
Table 33	Effect sizes for each sub-group, observed between tests.....	142
Table 34	Summary of results for RQ 4	144
Table 35	RE group: number of CF episodes and percentage of uptake with repair	148
Table 36	CR group: number of CF episodes and percentage of uptake with repair.....	148
Table 37	RE group: learners' repair of PC and their gains in oral and written production.....	150
Table 38	RE group: learners who did not produce repair of PC and their gains.....	150
Table 39	CR group: learners' repair of PC and their gains in oral and written production.....	150

Table 40 CR group: learners who did not produce repair of PC and their gains	151
Table 41 RE group: learners' repair of IMP and their gains in oral and written production	151
Table 42 RE group: learners who did not repair IMP and their gains in oral and written production	152
Table 43 CR group: learners' repair of IMP and their gains in oral and written production	152
Table 44 CR group: learners who did not repair IMP and their gains in oral and written production	152
Table 45 Comparison of mean gain scores in oral production of PC.....	153
Table 46 Comparison of mean gain scores in written production of PC	154
Table 47 Comparison of mean gain scores in oral production of IMP.....	154
Table 48 Comparison of mean gain scores in written production of IMP	155
Table 49 Descriptive statistics for LAT scores	164
Table 50 Descriptive statistics for SST strict scores	164
Table 51 Descriptive statistics for language anxiety.....	165
Table 52 RE group: Multiple regression coefficients for the gain scores in oral production.....	166
Table 53 RE group: Multiple regression coefficients for the gain scores in written production	166
Table 54 CR group: Multiple regression coefficients for the gain scores in oral production	167
Table 55 CR group: Multiple regression coefficients for the gain scores in written production.....	167
Table 56 Descriptive statistics for oral PC tests	222
Table 57 Descriptive statistics for written PC tests.....	222
Table 58 Descriptive statistics for oral IMP tests	223
Table 59 Descriptive statistics for written IMP tests	223
Table 60 Number of recasts and uptake with repair	224
Table 61 Correlation between uptake with repair and gain scores.....	225
Table 62 Descriptive statistics for anxiety and LAA scores	225

Chapter 1. Introduction

1.1. Overview

Corrective feedback (CF) has been a much debated and researched topic in second language acquisition (SLA) over the past two decades. It has received attention from different theoretical positions including linguistic theories and approaches based in cognitive psychology. As a topic of significance to pedagogy and language teaching CF has also been a key issue for teachers and teaching methodologists. This chapter introduces the topic of CF by outlining its significance at three parallel levels: I refer to the (1) theoretical, (2) pedagogical and (3) personal reasons that have contributed to the selection of this topic for my PhD thesis.

I begin by considering the theoretical importance of CF and then examine its pedagogical significance. I also briefly explain why this thesis investigates CF in relation to individual learner differences. I also reflect on my own interest in CF as a language teacher and a language learner. The last section of this chapter presents the outline of the thesis.

1.2. Theoretical significance of CF

The theoretical significance of corrective or negative feedback draws on both linguistic and psycholinguistic sources. Although researchers disagree about numerous issues in SLA, there is one point that everyone agrees on: that is, learners need to receive *input* in order to acquire both their first (L1) and a second language (L2). Input refers to “the language to which a learner is exposed either orally or visually” (Gass, 1997, p. 28) and it is distinguished from *intake* which refers to those parts of input that are utilised by learners. Since the introduction of the Input hypothesis (Krashen, 1982, 1985) it has been widely accepted that *comprehensible input* is necessary for language acquisition. Comprehensible input is achieved by learners making use of the context to help them understand the message and also as a result of simplification and modifications in phonology, lexis and syntax (Long, 1981a, p. 262).

1.2.1. Linguistic background

However, the type of input that is available to learners and used by learners has been an issue of considerable disagreement, in both L1 and L2 acquisition. In linguistic terms, the evidence learners receive in the input can be both positive and negative. Positive evidence refers to all grammatical input, showing “what is possible in a language” (White, 1989, p. 13) whereas negative evidence refers to the information about ungrammaticality. It may be that the provision of negative evidence in the input can indicate when ungrammaticality occurs. With regard to L1 acquisition it is generally assumed, in particular from the perspective of Universal Grammar theory, that negative evidence is not relevant and is not used by children. In fact, the existence of negative evidence would “reduce the need for an innate component” (p. 14). This innate knowledge or capacity is believed to be the Universal Grammar (UG) that all humans are endowed with. Not everyone, however, shares the opinion that negative evidence is not provided and not used by children acquiring their L1 (Bohannon, MacWhinney, & Snow, 1990; Bohannon & Stanowicz, 1988; Farrar, 1990, 1992; Nelson, 1987; Saxton, 1997, 2010). Therefore Farrar (1990) noted that “the role of adult linguistic input in children’s language acquisition is a central issue for language acquisition theories” (p. 607).

In L2 acquisition the role of negative evidence may be more important, at least for adults and older children, who may no longer have access to UG. In her seminal paper on corrective feedback Schachter (1991) pointed out that “one of the most important questions in the field of second language acquisition today is: Do second language learners, adults and older children, need and/or make use of negative data?” (p. 96). The question Schachter asked twenty years ago has been answered positively in numerous experimental studies over the years, even though there are still unresolved questions regarding the type of negative evidence that best facilitates acquisition. In a broad sense negative evidence may refer to all the explicit knowledge that learners receive about an L2. More specifically, it arises in the corrective or negative feedback that learners receive while attempting to communicate in an L2.

1.2.2. Interactionist background

The Interaction hypothesis (Long, 1996) claims that L2 acquisition is facilitated by interaction and *negotiation for meaning* involving *focus-on-form*. Long (1981a, 1983b, 1996)

argues that, while learners are oriented to meaning, focus-on-form may assist their interlanguage development. Long claims that implicit CF may be of particular relevance because it allows for a focus-on-form within an overall meaning oriented activity. However, it has been shown, in particular in instructional settings (R. Ellis, Basturkmen, & Loewen, 2001; Havranek & Cesnik, 2001; Lyster & Ranta, 1997; Panova & Lyster, 2002; Sheen, 2004) that CF is used not only for the purpose of negotiating meaning. It also occurs in situations where no communication breakdown has occurred but where the learner's erroneous utterance requires correction in order to be made more target-like. Lyster (1998a) proposed the term *negotiation of form* to refer to feedback that is motivated by purely linguistic reasons.

Several types of implicit CF, both output-prompting as e.g. clarification requests, and input-providing, e.g. recasts (R. Ellis, 2006; Sheen & Ellis, 2011) have been claimed to be important for language learning. Both have been shown to be facilitative of interlanguage development, but in different ways. Clarification requests, which encourage uptake and possibly repair, have proved to be an effective strategy for assisting lower proficiency learners to clarify the input they receive and also for facilitating their own L2 production. Recasts have also been shown to be facilitative of interlanguage development, in particular in one-on-one conversations between a learner and a more competent speaker. The following two examples demonstrate the two CF strategies as they occurred in meaning oriented activities. The first is an example of a clarification request (Takashima & Ellis, 1999, p. 173), the second is an example of a recast (R. Ellis & Sheen, 2006, p. 576).

Example 1:

Learner: *Cinderella change into the beautiful girl.* (trigger)

Teacher: *Sorry?* (clarification request)

Learner: *Cinderella changed into a beautiful girl.* (uptake with repair)

Example 2:

Teacher: *When you were in school?*

Learner: *Yes. I stand in the first row.* (trigger)

Teacher: *You stood in the first row?* (recast)

Learner: *Yes, in the first row, and sit, ah, sat the first row.* (uptake)

1.2.3. Psycholinguistic background

Cognitive psychology and psycholinguistics have provided theoretical accounts as to why it is important for CF to occur during the interaction, at the time when learners are engaged with the meaning. Having introduced the construct of noticing and the Noticing hypothesis, Schmidt (1990) suggested that no learning can take place without noticing novel features and structures in the input. Moreover, during the conversation learners may notice the 'gap' or the difference between their erroneous utterance and the target-like utterance the interlocutor has provided. This phenomenon, assisted by focal attention, entails a 'cognitive comparison' (Nelson, 1987). Recasts, as an implicit type of oral CF have been claimed to be particularly beneficial in facilitating this process (Doughty, 2001; Long, 2007; Long & Robinson, 1998) .

1.3. Pedagogical significance of CF

The pedagogical significance of CF has been established in a number of studies that have demonstrated the beneficial role of CF in L2 acquisition (Ammar, 2008; Ammar & Spada, 2006; Doughty & Varela, 1998; R. Ellis, 2007; Loewen & Erlam, 2006; Lyster, 2004; Lyster & Mori, 2006; Muranoi, 2000; Sheen, 2008; Takashima & Ellis, 1999; Yang & Lyster, 2010). However, language teaching methodologists have not always agreed with the recommendations derived from theory and empirical research. While many researchers argue that CF provided during communicative tasks, more often than not, promotes learning and contributes to changes in the learners' interlanguage development, some methodologists have recommended that CF be avoided in classroom communicative activities (Edge, 1993; Harmer, 2001). It is suggested that CF should be avoided for two main reasons: first, because it interrupts the learner's flow of thoughts and development of fluency. The second reason is the potential negative affective impact that CF may have on learners by creating anxiety and de-motivation.

Although teachers trained in contemporary methods have been taught not to correct their learners' errors and mistakes when they are trying to communicate, a number of descriptive studies show that many of them actually do correct learners' errors. Descriptive classroom studies demonstrate that recasts constitute the type of feedback which is most often used in the classroom (Havranek, 2002; Loewen & Philp, 2006; Lyster, 1998a, 1998b; Lyster & Ranta, 1997). However, in terms of immediate uptake, recasts have been found not to be very effective (Lyster, 1998a, 1998b; Lyster & Ranta, 1997; Panova & Lyster, 2002; Yang &

Lyster, 2010). This has been attributed to the implicitness of recasts and their possible ambiguity, making it difficult for learners to identify them as corrective. In contrast, experimental laboratory studies (Ayoum, 2001; Braidı, 2002; Han, 2002; Iwashita, 2003; Long, Inagaki, & Ortega, 1998; Mackey & Philp, 1998; McDonough & Mackey, 2006; Philp, 2003) and also certain classroom studies (Doughty & Varela, 1998; Muranoi, 2000; Revesz & Han, 2006) have demonstrated the effectiveness of recasts for the development of the learner interlanguage.

1.4. Significance of research involving individual differences

It is probably not surprising that the results of classroom studies have not been consistent and there are sometimes conflicting findings. This may be attributed to not only the different research methods used but also to individual difference (ID) factors, which have been found to play a significant role in mediating the effect of CF. Recent research (Robinson, 2005; Sheen, 2006, 2007a, 2008; Trofimovich, Ammar, & Gatbonton, 2007) has indicated that the effects of CF need to be considered in relation to learner ID factors. Research shows that the interplay of ID factors in language learning is complex, dynamic, and context-dependent (Dörnyei, 2005; R. Ellis, 2004; Robinson, 2002c). Therefore there is need to focus not only on the effectiveness of CF but also on the mediating role of IDs in the classroom.

1.5. Personal teaching and learning experience

When I started my PhD study I already had more than ten years of teaching experience, both in secondary school and adult foreign language (FL) and second language (L2) contexts. As a teacher I was trying to follow the guidelines I received at my teacher training college. These recommended that CF be avoided. The recommendation was not to provide learners with the correct ‘answers’ in order to encourage them to think for themselves. When I trained for the Cambridge Examination of Language Teaching to Adults (CELTA), I was taught that fluency activities should not be interrupted by error correction and that learner errors should be dealt with at some point later (e.g. by the teacher providing metalinguistic explanation at the end of the lesson). I accepted these guidelines and abode by them, trying not to interfere in my students’ attempts to communicate in the target language. I considered the recommendation to be sensible, as I was aware that explicit correction might easily distract and prevent

learners from further talk. However, it was difficult to refrain from recasting, even though I tried to train myself deliberately not to recast in the classroom. Recently, however, I have changed my opinion with regard to recasts, in part as a result of my reading of the research which has shown that recasting is actually a spontaneous, natural reaction to errors. Referring to L1 acquisition, Saxton (2010) claims that recasts are “the hallmark of adult-child discourse” (p. 94) as they fall out naturally from conversation with a child. Even Krashen (Krashen & Terrell, 1983), who rejected any overt or explicit error correction, contended that “reformulations” or “restatements”, in other words recasts, could be used as “direct natural feedback” just as in real life situations (p.178).

As I immersed myself deeper in the study of CF I became more and more aware of the significance of negative feedback for language acquisition. Even if the answers were not always clear-cut, I could at least in part answer some of the questions I often asked myself as a language teacher and a language learner. I believe that as a fluent speaker of at least three different languages apart from my native Croatian I can be considered an experienced language learner. I have experienced both successes and failures in language learning over the years. As a language learner I have always felt that my *language use* benefited from the recasts I received, even though I hardly ever repeated the recast and corrected myself. However, I was aware that I had to be clear about the meaning of the utterance which was recast if the recast was to have any purpose. In my own learning experience, clarification requests have been notably less frequent. I have usually perceived clarification requests as requests to clarify my own thought or make the meaning of my utterance clearer and more comprehensible, rather than as a request to correct an error I had committed. In short, my learning experience has told me that even though both recasts and clarification requests are implicit forms of corrective feedback, they are very different and such may be their contribution to L2 learning.

I elected to investigate the acquisition of two grammatical structures, the passé composé and the imparfait. As a teacher of French I knew that both the form and use of these two verb tenses presented difficulties to my students. Research findings (Bardovi-Harlig, 2000; Harley, 1989, 1993; Lyster, 2004) have also shown that these two tenses are among the most problematic areas of French grammar for L2 learners. Another reason for the choice of these two structures was that although they are frequently heard in everyday conversations, classroom discourse does not provide many opportunities for their use, in particular the use of

the imparfait. So, focusing on the acquisition of the passé composé and the imparfait in a classroom environment was a challenge worthy of investigation. To the best of my knowledge no classroom experimental research has been conducted so far with the aim to examine the role of oral corrective feedback on the acquisition of the two French tenses. The only two similar studies were Ayoun's (2001, 2004) investigations of the effectiveness of written recasts delivered in a computer lab. However, these two studies were conducted in entirely different experimental conditions and used a very different mode of delivering CF.

1.6. Overview of the study

The current study focused on the differential effects of recasts and clarification requests on the acquisition of the two language structures in a classroom environment, which still provides the prevalent context of teaching and learning languages at schools. Three intact classes of French FL high school learners participated in the study. Their average age was 16 and they had approximately had 500 hours of French language instruction to date. It is worth noting that secondary school students have been an underrepresented group in SLA research so far (Kormos & Sáfár, 2008). This is surprising because learners of this age constitute the majority of foreign language learners world-wide.

The role of individual difference (ID) factors was also included in the research in an attempt to examine the interaction between ID factors and the effectiveness of corrective feedback. Three ID factors were closely looked at: language analytic ability and working memory as cognitive variables, and anxiety as an affective ID variable.

The study employed a pre-test – treatment – immediate post-test – delayed post-test design, with two experimental groups and a control group. While the two experimental groups received each one type of corrective feedback during the performance of three communicative tasks, the control group did not undergo any treatment and did not perform the tasks the other two groups did. Both the treatment tasks and the tests were picture-based, designed to elicit the use of the target structures. During the execution of the treatment tasks the learners received feedback only on the errors in the production of the two target structures. Other errors were ignored. Acquisition of the two target structures was measured in both oral and written production. This allowed for the observation of change after the treatment in two conditions at different levels of cognitive demand: a more demanding

condition of semi-constrained unplanned speech production, and on the other hand, in written production which allows for more flexibility and more control over the performance.

The results related to the effectiveness of CF were obtained by the use of mixed design Repeated measures ANOVAs followed by the ANCOVAs with corrective feedback episodes as a covariate. The role of the three ID factors was examined by the use of Multiple regression analyses. The results indicate that, overall, corrective recasts were significantly more effective than clarification requests for the acquisition of both target structures, in both oral and written production. Results of Multiple regression analyses suggest that analytic ability was a much stronger predictor in the Clarification Request group than in the Recast group. On the other hand, working memory was a strong predictor for short-term gains in oral production of the passé composé in the Recast group. Anxiety did not appear to interfere with the effectiveness of corrective feedback.

1.7. Outline of the thesis

The outline of the thesis is as follows: In the next chapter I present a review of literature considering both theoretical and pedagogical perspectives on oral CF. In Chapter 3 a review of research studies investigating oral CF is presented. Chapter 4 discusses the literature on three ID factors: language analytic ability, working memory and anxiety. Chapter 5 reports on the pilot study and Chapter 6 explains the methodology of the main study. Chapter 7 presents the results for the effects of CF on oral production in the study and Chapter 8 deals with the results for written production. Chapter 9 is concerned with the uptake of CF during the treatment and its relationship with learning gains in both oral and written production. Chapter 10 presents the results of the analysis that investigated the relationship between individual learner differences and the gains in oral and written production. Chapter 11 summarizes the main findings, suggests their theoretical and pedagogical implications, points out the limitations of the study and provides suggestions for further research.

Chapter 2. Theoretical and pedagogical perspectives of CF

2.1. Introduction

Corrective or negative feedback is a term used to indicate that an utterance in a learner's interlanguage is deviant and that a change or a correction is needed to make it more target-like. Corrective feedback and its role in L2 acquisition have attracted much attention over the past decades, both at theoretical and practical/pedagogical level. This thesis is concerned with oral CF. A review of literature related to oral CF in L2 acquisition is presented in this and the following chapter. I first consider the main theoretical perspectives from which CF has been approached and discussed. I start with the classification of different types of oral corrective feedback and then present the main theoretical concepts and issues related to CF. The second part of this chapter is concerned with pedagogical perspectives on oral CF.

2.2. Theoretical perspectives

In this section I present Sheen and Ellis' (2011) taxonomy of oral corrective feedback and provide definitions of the particular types of CF which are the focus of this thesis. Then I consider the theoretical issues that have contributed to the ongoing debate on CF in second language acquisition. I will consider:

- (1) Universal Grammar-based approaches,
- (2) Cognitive approaches,
- (3) Interactionist approaches, and
- (4) Socio-cultural approaches.

2.2.1. Taxonomy of oral corrective feedback

Sheen and Ellis (2011) proposed a taxonomy of oral CF based on the implicitness and explicitness of error treatment, and whether it is input-providing or output-prompting. This taxonomy clearly shows the complex nature of CF and takes into consideration the multiple functions of oral CF. In this review two implicit error correction strategies will be attended to in more detail: the recasts and the clarification requests.

Table 1 Taxonomy of oral CF (Sheen & Ellis, 2011)

	Implicit	Explicit
Input-providing	Recast	Explicit correction
Output-prompting	Repetition Clarification request	Metalinguistic explanation Elicitation Paralinguistic signal

2.2.2. Recasts

As shown in Sheen and Ellis’ taxonomy, recasts are a type of oral CF that provides input by reformulating an erroneous utterance. That is, recasts supply the correct, target-like form (phoneme, morpheme, word or a sequence of words), but they do not explicitly state that an error has been committed. Long (1996, p. 434) defined recasts as “utterances that rephrase a child’s utterance by changing one or more sentence components (subject, verb or object) while still referring to its central meaning”. A particular characteristic of recasts lies in their dual nature, i.e. at the same time they provide positive evidence and implicit negative feedback. This quality makes them unique among all the CF strategies.

2.2.3. Requests for clarification

Long (1983b, p. 137) defined clarification requests as “any expression by a native speaker designed to elicit clarification of the interlocutor’s preceding utterance”. Requests for clarification, in the same way as recasts, do not state explicitly that an error has been committed but rather indicate that there was a problem in communication. Requests for clarification do not provide the target-like form. Thus in the Sheen and Ellis’ taxonomy they are classified as output-prompting. By indicating that there is a problem, clarification requests “require that the interlocutor either furnish new information or recode information previously given” (Long, 1983b, p. 137). Hence, by encouraging learner uptake, clarification requests may result in self-repair.

2.2.4. Learner uptake, repair and self-repair

The notion of learner uptake, as Lyster and Ranta (1997) pointed out, comes from Speech Act theory (Austin & Urmson, 1962, p. 116). Lyster & Ranta define uptake as a “learner utterance that immediately follows the teacher feedback and constitutes a reaction in some way to the teacher’s intention to draw attention to some aspect of the student’s initial utterance” (1997, p. 49). Lyster and Ranta suggest that learner uptake can be realized either as ‘repair’ or as ‘needs repair’ which means that it still requires repair. This kind of feedback promotes the ‘discovery approach’ by helping students to use their own linguistic resources and in this way it may lead to a better and deeper understanding of the concepts of the target language (Hendrickson, 1978).

2.2.5. Universal Grammar-based approaches to CF

Researchers working in the paradigm of Universal Grammar have viewed corrective feedback in relation to the issue of negative evidence. Positive evidence or positive data (‘primary linguistic data’) refer to the input of well-formed utterances that are available to learners in the process of language acquisition. Long (1996, p. 413) defines positive evidence as “models of what is grammatical and acceptable (not necessarily the same) in L2”. Negative evidence, on the other hand, denotes any type of information which indicates to learners that an utterance is not target-like and needs correction. It is “direct or indirect information about what is ungrammatical” (p. 413).

The theory of UG claims that positive evidence is sufficient for the acquisition of L1. Children learn their mother tongue only from ‘impoverished’ input by means of the innately specified system of principles and parameters which limit the hypotheses they can form while learning a language (Chomsky, 1975, 1981). The principles and parameters of UG may therefore, explain the ‘logical or learnability problem’ of L1 acquisition: the fact that knowledge of L1 extends far beyond the input available to children. Research in L1 acquisition indicates that children generally do not receive correction in their L1 (Grimshaw & Pinker, 1989; Pinker, 1989). It is argued that although there are studies that provide evidence of consistent provision of negative input, in particular recasts, by caretakers to their children (Bohannon & Stanowicz, 1988), this does not mean that negative evidence is necessary or that it is always available to all children. Pinker (1989) and Grimshaw and Pinker (1989) proposed that four requirements should be met in order to establish the

relevance of negative evidence for child language acquisition: that is, one would need to prove that negative evidence (1) exists, (2) is usable, (3) is used and (4) is necessary. They claimed that no evidence that would satisfy all four conditions has been provided for L1. Long (1996) also concluded that as far as L1 is concerned, “implicit correction, and recasts in particular play an important facilitative role” but “empirical support for a role for negative evidence in L1 acquisition is far from sufficient” (p. 437). In fact, the lack of negative evidence in child language acquisition forms the basis for the claim regarding the innateness of UG because, if negative evidence was shown to be relevant and necessary for language acquisition, the grounds for claiming the existence of a ‘language acquisition faculty’ would be eliminated (Gass, 1997; White, 1989).

The issue of negative evidence in L2 acquisition is more complex and more controversial since there are at least two clear differences between L1 and L2 acquisition: the first is that adults start to learn an L2 when they have already learned their L1 system and established knowledge of the world through their L1. The second is a well-known fact that most learners fail to acquire full linguistic competence in an L2. Therefore, the provision of negative feedback may be seen as necessary to indicate what is not grammatical and not acceptable in the L2. There is a range of different positions in this respect and some researchers have also revised their positions over time. Moreover, the theory of Universal Grammar has undergone considerable changes over the years, with regard to ‘what UG is assumed to consist of’ (White, 2003). At one end are those linguists who argue that negative evidence is not relevant to L2 acquisition because UG is available in L2 acquisition as well as in L1 acquisition. For example, Beck and Eubank (1991) put forward the same set of criteria as Pinker (1989) proposed for L1 acquisition. Krashen (1982, 1985) also argued that L1 and L2 acquisition are identical and that access to UG eliminates the L2 learner’s need for negative or explicit data. So, according to this view, only positive data or ‘primary linguistic data’ have an effect on the L2 interlanguage grammar. At the opposite end are those researchers (e.g. Schachter, 1991) who argue that UG is not accessible for L2 and that the learning mechanisms guiding L1 and adult L2 are entirely different.

Between these two extremes there are those positions that claim there are both similarities and differences between L1 and L2 acquisition. For example, Schwartz (1993) posited that L2 acquisition starts with the L1 grammar but UG remains fully accessible. Schwarz contended that negative evidence may be relevant only for those elements which are not

specified by UG, for example lexis or verbal morphology. That is, negative evidence can have effect only on the type of knowledge which Schwarz termed learned linguistic knowledge (LLK) and the type of performance which she called learned linguistic behavior (LLB). White (1989) also pointed out that “UG is only one component in an acquisition theory” and it “will interact with various others, and the failure of L2 learners (where they fail) may be attributable to these other areas, and not necessarily to the non-operation of UG” (1989, p. 50). White (1991) for example, investigated whether negative evidence is needed in parameter resetting of adverb placement for Francophone learners of English (since parameters are language specific whereas principles are universal). She argued that negative evidence was significantly more effective than positive evidence alone.

The notion of negative evidence comprises a range of possible solutions as to *what*, and *how* it is provided: from simple implicit CF strategies to explicit metalinguistic explanations and more generally, to explicit instruction in general. Long (1996) reviewed the question of negative evidence in L2 acquisition related to Pinker’s, and Beck and Eubank’s four criteria and concluded that “evidence of use is proving no easier to come by than in L1 acquisition” (1996, p. 444). He also claimed that the necessity of negative feedback for L2 learning is arguable but, more relevant and more important than the necessity may be the facilitative role of negative feedback in L2 acquisition. At different levels it has been shown that positive data are not sufficient for L2 acquisition, so negative evidence or negative feedback might be a timely signal to L2 learners about the grammaticality and acceptability of their L2 utterances.

2.2.6. Cognitive approaches to CF

Cognitive approaches to L2 acquisition seek to explain language learning as a mental process similar to other higher-order mental functions (e.g. vision) and of the same kind as other types of learning or problem-solving (N. C. Ellis, 1994a, 1996). There is a range of theories and approaches in cognitive linguistics, but they all share the same or similar view about the human capacity for language learning: in short, they argue against the innateness of UG and derive their claims from usage-based approaches and human experience with the language. For cognitive theorists there is no difference between competence and performance (N. C. Ellis, 1994a; Skehan, 1998). Cognitive linguistics, for example, has been closely related to the information-processing model which seeks to explain language learning in terms of input, processing and output of linguistic data. In information-processing models, noticing performs

a “mediating role between input and the operation of memory systems” (Skehan, 1998, p. 49). Cognitive psychology has provided the theoretical basis for the Noticing hypothesis.

2.2.6.1. Noticing Hypothesis

Noticing is the term which has been used to indicate learners’ attending to new, previously unrecognised language features, or the “conscious attention paid to the input in order for input to become intake” (R. Ellis, 1994, p. 361). Schmidt and Frota (1986, cited in Schmidt, 1990) introduced and Schmidt (1990) further elaborated the Noticing hypothesis. Schmidt claimed that noticing is the necessary condition for the acquisition of all aspects of an L2 (lexis, phonology, grammatical form and pragmatics). Research and L2 acquisition theory (R. Ellis, 1995; Long, 1996; Schmidt, 1990) suggest that only those elements of the input which are attended to and noticed are likely to be stored in long-term memory and, therefore, learned. Thus, noticing requires focal attention and awareness (Schmidt, 1990). In psychology awareness is defined as subjective experience of consciousness (Bialystok & Ryan, 1985; Tomlin & Villa, 1994).

Schmidt (1990) rejected the possibility of subliminal or unconscious learning for adult learners and for the learning of new features. He distinguished three levels of awareness (1990, p. 132). According to this view, the first level is *perception* which, although involving mental representations of external events, does not have to involve consciousness. Schmidt argued that only perception can be subliminal, not language learning. The second level is *noticing* which is defined as the subjective experience of stimuli and is comparable to *apperceived input* in Gass’ (1997). Schmidt proposed that it is “the basic sense in which we commonly say that we are aware of something, but (this) does not exhaust the possibilities” (p. 132). The third level is *understanding*, which Schmidt clearly distinguishes from basic noticing described as the subjective experience of stimuli. Understanding, according to Schmidt, involves analysis and comparison. In short it involves thinking. Later Schmidt (2001) revised his position regarding awareness and consciousness, stating that implicit learning (i.e. learning without awareness) exists and has been demonstrated in experiments where more learning occurred than could be verbally reported. He also suggested that there may be a possibility of some unattended learning but he still emphasized that “attended learning is far superior, and for all practical purposes, attention is necessary for all aspects of

L2 learning” (2001, p. 3). Attention, according to Schmidt, is limited, selective, subject to voluntary control and it controls access to consciousness.

2.2.6.2. Attention and awareness

The role of attention and awareness has attracted considerable interest in discussions on L2 learning. For example, Tomlin and Villa (1994) examined studies of attention and learning and claimed that attentional processes such as detection and orientation do not have to involve awareness. They proposed a more fine-grained analysis of attention involving alertness, orientation and detection. Tomlin and Villa emphasized detection, which is the cognitive registration of some stimuli. So, only when a feature or some other stimulus has been detected, can cognitive control and awareness take place. There is substantial evidence, as Tomlin and Villa (p. 193) pointed out, that information or stimuli can be cognitively detected without awareness.

In an attempt to take into account both views, Schmidt’s and Tomlin and Villa’s, Robinson (1995) proposed a definition of noticing based on the theories of attentional allocation. He defined the concept of noticing as ‘detection plus rehearsal in short-term memory prior to encoding in long-term memory’ (p. 296). As he explained, activation in short-term memory must reach a certain level in order for awareness to occur. At the same time, activation in short-term memory requires attentional allocation from the central executive and further activation in long-term memory. Depending on the nature of the rehearsal in short-term memory, processing can be data-driven (where simple rehearsal is required) and/or conceptually-driven (involving elaborative rehearsal and activation of long-term memory). The nature of rehearsal, according to Robinson, depends on the demands of particular tasks.

Attention is selective, thus only parts of input are attended to. Robinson (2003) suggested that selection of linguistic input is “only one aspect of action control guided by the supervisory attentional system, and executive control mechanisms” (p. 635). There are different views in SLA on what guides selection of input for language learning. These can be broadly divided into two groups:

(1) The first one involves those positions that view selection as guided by innate representations of abstract phonetic and grammatical knowledge, modularly encapsulated and automatic.

(2) The second group involves constructivist accounts of SLA, which argue that selection of input is a continuous process of pattern recognition.

Different as they are, however, both views share the opinion that selection of auditory input occurs and that attention is required for this to happen, but that it does not have to involve awareness (Robinson, 2003). However, there may be differences regarding the different aspects of language. For example, N.C. Ellis (1994b) claimed that attention is necessary and sufficient for learning the perceptual aspects of novel word forms whereas learning word meanings would require both attention and explicit awareness.

2.2.6.3. Noticing and corrective feedback

The link between noticing and corrective feedback entails ‘noticing-the-gap’, or the process of ‘cognitive comparison’ or ‘fast mapping’ (Nelson, 1987) that enables learners to notice the difference between the target form and their own erroneous production. Doughty (2001) argues that cognitive comparison occurs in working memory (WM), which makes possible the connection between short-term memory (STM) and long-term memory (LTM). Therefore, the intervention provided by corrective feedback may be highly beneficial for knowledge restructuring and learning if it leads to cognitive comparison. Doughty has founded her claims on numerous psycholinguistic experiments which have shown that previously attended items stay active in STM even after the moment of initial awareness. This ‘highlighted state of activation’, she argued, is the crucial and the ideal moment for corrective feedback to take place. Such interventions interrupt the learner’s own processing and they may be seen as ‘cognitive intrusions’ (Doughty, 2001). However, these interruptions from ‘outside’ may lead to elements becoming incorporated into an utterance even though originally the speaker did not intend to make them part of his or her utterance. Doughty’s explanation is based on psycholinguistic speech error analyses, which suggest that people can attend to, internalize and use speech, environmental input or conceptual information that was not originally formulated in their own speech plan for their utterance. Doughty further claims that because of their timing which requires simultaneous processing at different levels, these ‘cognitive intrusions’ allow for mapping the form onto meaning and function. It is fundamental, as Doughty and Williams (1998) emphasized, “that meaning and use must already be evident to the learner at the time attention is drawn to the linguistic apparatus needed to get the meaning across” (1998, p. 4).

2.2.7. Interactionist approaches to CF

Much of the research and discussion with regard to noticing has been inspired by the Interaction hypothesis introduced by Long (1981a, 1981b, 1983a, 1983b) and later updated (Long, 1996). The first version of the Interaction hypothesis came as a reaction and a supplement to Krashen's Input hypothesis (Krashen, 1985). Long argued that 'comprehensible input' is necessary but not sufficient for L2 acquisition and that the process of negotiating input is important for comprehension and acquisition. The major advancement of the updated version of Interaction hypothesis, as R. Ellis (1999b) indicated, was the greater emphasis put on negative feedback and modified output, and their combined contribution to L2 development.

One of the main arguments of the Interaction hypothesis is the importance of focus on form, defined by Long (1991) as the type of the treatment which "overtly draws students' attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication" (p. 45). Thus, it is important to emphasize that focus on form is restricted to meaning-based pedagogical activities and is clearly distinguished from 'focus on forms'. Negative or corrective feedback is seen as a type of focus on form. As Long (1996) explained, noticing which is an aspect of focus on form may be facilitated through interaction and, while negotiating for meaning, learners may attend to form. The facilitative role of interaction and negotiation is possible due to their role in "connecting input, selective attention and output in productive ways" (p. 452). In other words, interaction and negotiation provide the natural environment for focus on form (including CF) and noticing.

The Interaction hypothesis, as Long and Robinson (1998) claim, proposes that L2 acquisition cannot be explained by either a purely linguistic nativist or a purely environmentalist theory. Gass (1997, 2002) also contends that the Interactionist perspective is compatible with both Universal Grammar's claims about the innateness of language mechanisms, and with cognitive models of language learning (Gass, 2002, p. 170). As Gass explained, the role of interaction in acquisition does not depend on whether there is or is not an innate mechanism that guides L2 learning.

2.2.7.1. Negotiation for Meaning and Negotiation of Form

In his Interaction hypothesis Long emphasized the importance of negotiation for meaning since it provides a natural context for focus on form. Negotiation for meaning is claimed to increase the comprehensibility of input. It also elicits negative feedback which, as Long argued (1996) may assist L2 acquisition by promoting noticing and noticing-the-gap and so enabling the learner to simultaneously connect form, meaning and use (Doughty, 2001). In their discussion of focus on form, Long and Robinson (1998) emphasized that recasts as an implicit type of negative feedback are more likely to facilitate acquisition than models which provide only positive evidence. Long and Robinson considered recasts to be an ideal type of focus on form in that they do not interrupt the “predominant focus on meaning” (1998, p. 26). However, Lyster (1998b) distinguished between “negotiation for meaning” and “negotiation of form”, explaining that negotiation of form “aims not only for comprehensibility of message but also for accuracy and precision in form, thus involving a more pedagogical and less conversational function of negotiation” (p. 53). This may assist L2 learning in two ways: one is that CF gives “clues to learners to assist them in modifying their use of non-target forms” and so learners have an opportunity to revise their hypotheses about the target language. The second is that CF may encourage student-generated repair as self-repair or peer-repair which, in Lyster’s opinion, helps in the process of proceduralising the knowledge of target language forms that students may already have partially acquired.

In short, both negotiation for meaning and negotiation of form direct learners’ attention to form, so both constitute focus on form. As already pointed out, they serve to induce noticing and noticing-the-gap, which ultimately can lead to interlanguage restructuring.

2.2.8. Skill learning theory

The functions of corrective feedback, in particular of the types which encourage students to self-repair, have also been discussed in relation to skill learning theory. Skill learning theory is based on the view that language learning involves the same processes guiding any other skill acquisition, e. g. mathematics, swimming or driving. Crucial to any skill acquisition is practice (DeKeyser, 2007a). Following the findings from cognitive psychology, DeKeyser (2007b) argues that the process of skill acquisition entails three stages: (1) declarative, (2) procedural and (3) automatic. Thus the initial step in L2 learning occurs as a result of the explicit presentation of a linguistic rule. What follows is the practice to initiate the process of

proceduralisation and finally, a learner is supposed to reach the stage of automatization where errors are reduced and the linguistic feature can be accessed easily. However, DeKeyser cautions that automatized knowledge is not the same as implicit knowledge since “absence of awareness is not a requirement for automaticity” (2007a, p. 4). Lyster, also a proponent of skill learning theory, argues that the “opportunities to engage with feedback in a productive mode via prompting” (2004, p. 427) assist learners to proceduralize their knowledge.

2.2.9. Output Hypothesis

Related to the Interaction hypothesis is the Output hypothesis, proposed by Swain (1985). Swain claimed that learners need to have not only the opportunities for negotiated interaction, but also the opportunities for output in order to learn an L2. According to Swain, one function of output is to provide opportunities for meaningful use of learners’ linguistic resources. Negotiating meaning, Swain contends, “needs to incorporate the notion of being pushed toward the delivery of a message that is not only conveyed, but that is conveyed precisely, coherently and appropriately” (p. 249). Two additional roles that can be added to output apart from the ‘contextualized’ and “pushed” language use, are that: 1) output provides opportunities to test out hypotheses learners have made of expressions (and see if they work) and 2) using the language may force the learner to move from semantic processing to syntactic processing. Later Swain (1995) identified three functions of output in L2 acquisition. These are:

- (1) noticing-the-gap,
- (2) testing hypotheses about linguistic forms and
- (3) understanding rules and knowledge about language.

The Output hypothesis is relevant to the investigation of CF since it provides the theoretical basis for the claim that learner uptake and, in particular, repair and self-repair, are important for the effectiveness of CF.

2.2.10. Socio-cultural approaches to CF

The Socio-cultural theory (SCT) of mind claims that cognition and learning emerge within dialogue. They are constructed in the process of a collaborative dialogue. The main argument of the Socio-cultural theory of mind (Lantolf & Poehner, 2008; Vygotsky, Hanfmann, & Vakar, 1962) is that cognitive functions such as attention, or problem solving, are “mediated

mental activities”. This means that their sources are social, external to the learner but they are transformed into mental ones through interaction. The process of internalisation is realised, or ‘mediated’ through semiotic tools. Language is one of the most important such tools. Swain (Swain & Lapkin, 2000) has applied the Socio-cultural theory of mind to research on L2 acquisition, claiming that known language serves to mediate further language learning. This is achieved through collaborative dialogue where learners work together to solve linguistic problems and co-construct language or knowledge about the language.

In a study of corrective feedback guided by the principles of SCT Aljaafreh and Lantolf (1994) demonstrated how feedback as regulation needs to be adjusted to individual learner’s *zone of proximal development (ZPD)*. Regulation refers to stages in the learning process, so that the first stage is object-regulation in which objects are used to mediate mental activities. The second stage is other-regulation in which parents, teachers or peers provide assistance to learners when performing the tasks they cannot do independently. The third stage is self-regulation where tasks are accomplished without any assistance and this is the goal of all instructional activities. The ZPD is the learner’s potential developmental level. The expert or a more competent speaker, together with the learner, tries to discover the learner’s ZPD in order to determine what kind of help is required. Dialogue and collaboration between the two are, therefore, of vital importance for the success of CF and in the end, for the learner’s language development. Aljaafreh and Lantolf argued that interactional moves and adjustments cannot be determined *a priori*, “rather, they must be collaboratively negotiated online with the learner” (1994, p. 466). Thus, it is necessary that mediation involves continuous assessment of ZPD and of the assistance that learners require. This was shown in Poehner and Lantolf’s (2005) and Poehner’s (2008) studies of dynamic classroom assessment.

2.3. Pedagogical perspectives

2.3.1. Error correction

Among the first studies that addressed the issue of corrective feedback in a second language classroom was Chaudron’s (1977) study, which provided a detailed description of error correction treatment. Chaudron adopted Annett’s (1969) three main functions of feedback in learning tasks and adapted them to language learning. Chaudron (p. 30) considered the following three functions of CF:

- (1) incentive i.e. stimulation of increased effort which translates into motivation,
- (2) reinforcement i. e. promoting maintenance of the learner's responses and
- (3) information which will contribute to changes in learners' responses.

Along with the detailed description of classroom error correction strategies, which I consider in the next chapter, Chaudron (1988) also pointed to the problem of ambiguity and inconsistency in teachers' provision of CF. Thus it is not surprising that the issue of CF has often resulted in more questions than definite answers. Hendrikson (1978) put forward a set of questions which probably every classroom teacher at some point asks himself or herself:

Should learner errors be corrected?

If so, when should learner errors be corrected?

Which learner errors should be corrected?

How should learner errors be corrected?

Who should correct learner errors?

These questions have been addressed in many books dealing with teacher education and training (Edge, 1993; Harmer, 2001; Hedge, 2000; Ur, 1996). Although they sometimes differ in detail, it is clear there is a difference in the positions of teacher educators and researchers. The major difference regarding the provision of CF in a classroom is summarized as follows: whereas theory so far has provided very strong reasons for focus on form and CF during communicative language tasks, teaching recommendations are generally in the opposite direction. Many teacher educators recommend that communicative, fluency oriented activities should be conducted without interruption. They argue that correction would disturb and discourage more than help (Ur, 1996). However, Ur also states that there might be situations in oral work when a 'gentle, supportive intervention can help' (1996, p. 247). Harmer (2001) for example, suggested that correction should depend not only upon the kind of mistakes being made (and the reason for them) but also on the type of activity the students are taking part in (p. 99). In his view, during communicative activities teachers should not interrupt the flow of talk, but in accuracy work correction is recommended. However, Harmer contends that there may be times, even in fluency work, when an intervention is needed but that it should be "gentle correction with tact and discretion" (2001, p. 107). It seems that such recommendations echo Krashen's (1985; Krashen & Terrell, 1983) view. He argued against direct correction in acquisition activities, emphasizing that error correction should be used only for conscious learning. However, he does not reject the use of

‘reformulations’ or ‘re-statements’ if the learner has produced an utterance which is incomprehensible, “just as in real life situations” (1983, p. 178). Krashen states that:

It is not clear that this sort of expansion is actually responsible for encouraging the students to speak more accurately and correctly. It is more likely that although some students do use this direct natural feedback for conscious inductive learning, the main contribution of the instructor’s expansion is that it provides more comprehensible input (1983, p. 178)

Clearly, pedagogical perspectives on oral CF, like the theoretical accounts, demonstrate that CF is essentially a complex and controversial matter. Therefore it is not so surprising that both the theory and pedagogy offer such different views on the role and effectiveness of CF. Furthermore, as Sheen (2011) pointed out, perhaps the most important issue for teachers is which errors should be corrected. Although selective correction is widely promoted by language teaching methodologists, “systematic engagement with selective error correction remains a challenging task for most teachers” (p. 41). There is also the issue of consistency of error correction in the classroom, which was observed in many studies but, as Sheen noted, “somewhat surprisingly, methodologists do not address this issue” (p. 42).

2.4. Summary

In this chapter I have briefly reviewed the theoretical and pedagogical accounts that have contributed to our understanding of oral CF. Drawing on UG-based approaches, cognitive-psychological and interactionist approaches, as well as on socio-cultural theory of mind, I have indicated that:

- Corrective feedback seen as negative evidence may not be necessary but its role may be facilitative for L2 learning.
- Noticing and noticing-the-gap are the key cognitive-psychological processes that make cognitive comparison possible in WM.
- Interaction and negotiation in communicative language activities play an important role in L2 acquisition: they provide the natural conditions for focus on form and for the provision of CF.
- Both recasts and clarification requests are implicit forms of CF but they differ in terms of input and output: recasts provide input whereas clarification requests ‘push’ learners to produce output.

- Corrective feedback that results in uptake with repair may promote learning by assisting learners to notice-the-gap and to test hypotheses about linguistic forms.
- Drawing on socio-cultural theory, CF should take into consideration the learners' level of development and ideally their ZPD.
- There is a difference between the views expressed by researchers and teaching practitioners: the former argue for the particular beneficial effects of CF during communication whereas the latter, in general, do not recommend error correction in fluency activities.

In the next chapter I turn to empirical research and provide a review of studies that have investigated the effect of CF, in particular recasts and clarification requests, which are the focus of this thesis.

Chapter 3. Research on oral corrective feedback

3.1. Introduction

Empirical research on oral corrective feedback (CF) has grown substantially over the past two decades. Descriptive and experimental laboratory studies were dominant in earlier years but recent years have seen an increase in classroom quasi-experimental studies. In this chapter a review of research on oral CF is presented. Classroom descriptive studies are considered first. This is followed by a brief review of selected laboratory studies relevant to the current research with regard the types of CF under investigation. Quasi-experimental studies carried out in L2 and FL classrooms are then discussed in detail. The last part of the chapter considers several meta-analyses of CF studies.

3.2. Classroom descriptive studies

The majority of early studies on corrective feedback in the classroom were descriptive. A large body of research came from the Canadian French immersion classrooms. One of the first attempts to describe a detailed model of discourse taking place when an error occurs in a classroom was provided by Chaudron (1977). He meticulously described the types of errors collected in Canadian French immersion classrooms. More than thirty features and types of teachers' corrective reactions were identified in Chaudron's description. He emphasized that most cases involved repetitions. These could be of four types:

- repetition without or with change (the latter would be 'recasts' in today's terms) and
- repetition without or with emphasis which would more overtly indicate that there was an error.

Chaudron also indicated that one type of repetition which has been mistrusted as "especially weak in helping to locate the error" (p. 41) is repetition with change alone without emphasis (i.e. a recast). Chaudron argued that such repetitions lacked saliency. He found that emphasis and reduction of repetitions were positively related to success of student 'intake' (i. e. uptake with repair) whereas there was a very low success ratio for repetitions with expansion.

3.2.1. Corrective feedback and uptake

Also in the Canadian French immersion setting, Lyster and Ranta (1997) observed six classes of content-based instruction resulting in 100 hours of audio-recorded teaching. They developed a model of coding categories for oral CF which has provided a useful tool and has been used in a number of other studies. In their classification of teacher feedback to students' erroneous utterances the following types of CF were distinguished: 1. explicit correction, 2. recasts, 3. requests for clarification, 4. metalinguistic feedback, 5. elicitation, and 6. repetition. They also examined 'uptake'. Lyster and Ranta's definition of uptake did not refer only to those utterances which incorporated the correction (utterances that may be termed 'successful uptake') but they included all kinds of student responses to the teachers' CF moves (e.g. saying 'yes', asking questions, etc). So Lyster and Ranta classified uptake as 'with repair' and 'no repair'. They found that about 62% of all student errors in the classroom receive feedback from teachers but in the end only 17% of these errors are repaired. As Lyster and Ranta recorded, recasts made up the largest proportion of all feedback types (55%). All other types of CF were used much less frequently, for example, elicitation 14%, clarification requests 11%, metalinguistic feedback 8% , explicit correction 7% and repetition only 5% of all feedback types.

The analysis of uptake indicated that 69% of all recasts did not result in any uptake and only 18% of all recasts were followed by repair whereas 13% of the recasts remained in need of repair. Clarification requests in Lyster and Ranta's analyses were followed by a much higher level of uptake with only 12 % of all clarification requests not resulting in uptake. However, repair after clarification requests occurred in only 28% of instances whilst 60% of all clarification requests followed by uptake remained in need of repair. According to Lyster and Ranta, the highest rate of repair followed elicitation moves and metalinguistic feedback: 46% and 45% respectively.

In a later study, Lyster (1998b) described four types of recasts defined as the teacher's implicit provision of a correct reformulation of all or part of a student's ill formed utterance. The types of recasts in the classification were: 1. an isolated declarative recast, 2. an isolated interrogative recast, 3. an incorporated declarative recast and 4. an incorporated interrogative recast). Lyster's definitions and examples (p. 58) are provided below:

1. an isolated declarative recast is defined as confirmation of a learner's message by correctly reformulating all or part of the utterance with falling intonation and no additional meaning, e.g:

S: *Avant que quelqu'un le prendra.*

“Before someone will take it.”

T: *Avant que quelqu'un le prenne.*

“Before someone takes it.”

2. an isolated interrogative recast seeks confirmation of the learner's message by correctly reformulating all or part of the utterance with rising intonation and no additional meaning, e.g:

S: *On pense que, qu'il est **prisonnère**, comme, um, quelque part.*

‘They think that, that he's a prisoner (fem. form) like, um, somewhere.’

T: ***Prisonnier?** ...*

‘Prisoner? ...’

3. an incorporated declarative recast provides additional information by incorporating the correct reformulation of all or part of a learner's utterance into a longer statement, e.g:

S: *Ou **une** bateau.*

‘Or a boat.’ (fem. article)

T: *Oui, c'est vrai que ca pourrait etre **un** bateau, mais là on donne des adresses.*

‘Yes, that's true that it could be a boat, but there they're giving addresses.’

4. an incorporated interrogative recast seeks additional information by incorporating the correct reformulation of all or part of a learner's utterance into a question, e.g:

S: *Elle **changer** de couleur.*

‘It change colour’

T: ***Pourquoi elle change** de couleur ?*

‘Why does it change colour?’

Lyster's data suggested that the largest category was the first group (67 %) and both the first and the second group made up 79 % of all recasts. None of the incorporated recasts led to any repair. Only 23% of isolated recasts led to repair and 13% to needs-repair, whereas 64% led to topic-continuation. Lyster concluded that recasting did not invite negotiation and that in a content-based classroom the corrective function of recasts might be less salient than their various other discourse functions. Furthermore, Lyster (1998a) separated both recasts and explicit correction from the other types of feedback which he grouped into ‘negotiation of form’, arguing that recasts and explicit correction, in contrast to the four prompting strategies, just supply correct forms that learners may or may not repeat (1998a, p. 191). However,

regarding the isolated interrogative recasts (type 2) he pointed out that they may include confirmation checks as a negotiation strategy (1998b, p.59). Some studies, for example, Oliver (1995) and Mackey and Philp (1998) provide evidence of this discourse function of recasts.

Similar results are reported by Panova and Lyster (2002) who observed 18 hours of instruction in communicative ESL classrooms in Canada. They analysed 10 hours using Lyster and Ranta's (1997) coding system and coded all errors as phonological, grammatical and lexical. In their data there were 1,716 student turns and 1,641 teacher turns, 25 % of which (412) included CF. Panova and Lyster reported that recasts and translation of learner errors were used most frequently, together accounting for 77 % of all feedback moves. About half of the CF moves were followed by some kind of uptake (192 or 47%) but only 65 of these uptake moves included learner repair. That is, only 16 % of all feedback moves and only 8% of all student errors were repaired after teacher feedback. The authors reported that clarification requests, which accounted for only 11% of all feedback moves, were followed (the same as elicitation and repetition) by the highest rate of learner uptake (100%). However, only 23% of clarification requests resulted in uptake with repair. On the other hand, recasts and explicit correction in which the learners were provided with the target form, were followed by much lower uptake, 40% and 33% respectively. The rate of repair after recasts was even lower, only 13% of all recasts resulted in repair. Similarly, Havranek's (2002) large-scale study involving 207 EFL students in Austria, found that recasts were the least effective form of CF. In contrast, elicited self-correction proved to be most effective with 65.2% accurate use of corrected forms by the students who were corrected and 69.5% success of correct use by their peers, i.e. the students who only listened and were exposed to the corrections directed at others. The participants in this study were of different ages, from 10-year old primary school pupils to university students majoring in English. Havranek provided evidence of CF effectiveness not only based on observation of immediate student reactions, but on the basis of class-specific specially designed language tests. The author identified ten different types of corrective feedback. The effectiveness of CF was evaluated separately for those learners who elicited the CF by committing an error, and for their peers who only listened. Overall, results showed that a little bit more than half (51.4%) of the learners who committed an error and were corrected were able to use the same structure in subsequent language tests. Their peers benefited even more than those who were corrected - their rate of success on the post-test was 61%. Havranek reported that the most successful way of

providing CF was to elicit self-correction but if this failed, than a recast (“repetition of the correct form”) seemed to promote learning.

3.2.2. Contrasting findings

In contrast, the results of Ellis, Basturkmen and Loewen (2001) conducted in New Zealand, differed considerably from Lyster and Ranta’s, and Panova and Lyster’s studies. Ellis, Basturkmen and Lowen analysed 12 hours of classroom interaction in adult ESL classrooms and found that in a total of 448 focus-on-form (FonF) episodes learner uptake occurred in 429 episodes. However, they took a broader perspective on uptake by examining learners’ responses to pre-emptive as well as reactive FonF. ‘Preemptive’ FonF comprised student-initiated and teacher-initiated FonF involving either a student’s question or the teacher’s explanation of a feature that may have been problematic to students. ‘Reactive’ or ‘responding’ FonF involved the teachers’ responses to a student’s utterance that either contained a linguistic error or was unclear in meaning (p. 295). Of all the episodes that contained student uptake, 235 (74.1 %) resulted in successful uptake and only 82 (25.9 %) in unsuccessful uptake. Thus the authors reported that overall, students were able to produce a target form correctly or show understanding in more than half of the total focus-on-form (FonF) episodes. Uptake was higher in student-initiated than in teacher-initiated FonF. Ellis and al.’s study also showed that there were more than two times as many episodes involving negotiation of form as negotiation of meaning, but there was significantly more uptake in episodes involving negotiation of meaning. The analysis that compared the different interactional moves showed that recasts were the most frequent type of ‘responding’ or ‘reactive’ focus on form and resulted in 71.6 % uptake, 76.3% of which was successful.

3.2.3. Context and instructional setting

Such contrasting findings in classroom descriptive studies may be less surprising if the context of the class activities and the instructional setting are taken into consideration. The role of context and type of activities in the provision of CF and the learners’ use of feedback was examined in Oliver and Mackey’s (2003) study. The research involved five experienced teachers and their young learners’ classes. Oliver and Mackey found that provision of feedback most frequently occurred in explicit language-focused and content-focused lessons but most opportunities for uptake were observed in explicit language-focused lessons

resulting in 85% of uptake, and in communicative contexts with 38 % of uptake. In contrast, zero uptake was observed in management-focused activities.

Oliver (2000) compared the different patterns of interaction depending on age and the type of lesson. In a study involving twenty teachers and their intact classes, Oliver demonstrated how the different patterns of interaction and negotiation for meaning emerged in the different contexts. This study provided evidence that in teacher-fronted lessons with adults teachers use more negotiation strategies (clarification requests, confirmation checks, repetition) than with the children, whereas in adult pair-work recasts were more often provided. Moreover, in pair-work, there were significant differences between the two age groups with the adults using more recasts than the children. In this respect Oliver's analysis corresponds to the findings reported by Doughty and Pica (1985) of North American ESL classrooms. Their study showed that more recasts occurred in pair-work than in teacher-fronted lessons. However, in a previous study Oliver (1995) also found evidence of children providing their non-native peers with recasts. She reported that children incorporated 10% of all recasts in their subsequent utterances. Oliver (1995) further observed that one third of all recasts were incorporated in subsequent speech if the learners had been given an appropriate opportunity to do so.

Different instructional settings have also been found to affect the rate of uptake. Sheen (2004) set out to analyse and compare the uptake of recasts and student repair in four different settings: Korean EFL, New Zealand ESL, Canadian immersion and Canadian ESL classrooms. This study confirmed that recasts were the most frequent type of CF in all four settings, albeit more frequent in NZ and Korean classrooms. There was a much larger percentage of learner uptake and repair following recasts in Korean EFL and NZ ESL classes compared to the Canadian immersion and adult ESL classes. Sheen's study drew attention to the fact that in different settings (foreign language/second language, immersion or content based language learning) students may behave in different ways depending on their orientation and the focus of their attention.

3.2.4. Counterbalance Hypothesis

Evidence in support of the claims about the importance of learning settings is also provided by Lyster and Mori (2006). They compared Canadian French immersion and American Japanese immersion elementary classrooms. They analysed the interaction in the two

instructional settings and found that in both settings recasts were the most frequently used CF strategy. However, of all repair moves in the French immersion repair following recasts accounted for only 38 % compared to 68% in the Japanese immersion. On the other hand, repair following prompts in the French immersion accounted for 53% and only 23% in Japanese immersion classrooms. Explicit correction was followed by low repair in both French and Japanese immersion, 9% and 19% respectively. In fact, the most striking difference, as the authors observed, was in the uptake of recasts in the two settings: in the Japanese classrooms 72% of recasts were followed by uptake whereas in the French classrooms uptake occurred after only 32% of recasts. Lyster and Mori also noted other differences between these two learning settings, for example, the distribution of time spent on different types of activities and the content focus of activities. Thus they proposed the Counterbalance hypothesis which claims that “instructional activities and interactional feedback that act as a counterbalance to a classroom’s predominant communicative orientation are likely to prove more effective than instructional activities and interactional feedback that are congruent with its predominant communicative orientation” (2006, p. 269).

3.2.5. Characteristics of recasts

Sheen (2006) further examined the recasts in the datasets from New Zealand ESL and Korean EFL classes, and found that mode, length, type of change, linguistic focus and reduction of recasts were significantly related to learner repair. Sheen defined recasts as “the teacher’s reformulation of all or part of a student’s utterance that contains at least one error within the context of a communicative activity in the classroom” (p. 365). In her coding system mode refers to the mood of the recast (declarative or interrogative). Scope refers to the extent to which a reformulation differs from the learner’s erroneous utterance, so a recast can be isolated or incorporated. Reduction refers to recasts in which the reformulation is shorter than the learner’s erroneous utterance. For example, reduced recasts were related to more repair and declarative recasts elicited more uptake than interrogative recasts. Sheen established the following features of recasts that were most likely to be effective for acquisition: recasts should be directed at features that learners are ready to acquire, they should induce noticing, should be linguistically salient and should be repeatedly directed at a single linguistic feature.

Loewen and Philp (2006) looked at the corrective functions of recasts in relation to such characteristics as stress, intonation, length, change, etc. The results of their study which

examined 17 hours of meaning-based classroom interaction involving 118 ESL learners, showed that stress, declarative intonation, one change and multiple feedback moves predicted successful uptake, while interrogative intonation, shortened length, and one change predicted post-test learning scores. The authors suggested that recasts could vary in implicitness and this may influence their effectiveness. Loewen and Philp reported that at least 50 % of recasts were beneficial as shown in post-test scores. The authors noted that in a communicative foreign language classroom there is an understanding that the main purpose is to learn the target language and that this awareness may orient the participants' attention to form. Egi (2007) also found that the length of recasts played a significant role by showing that shorter recasts were easier to notice. Egi reported a significant relationship between the length of recasts and the learners' interpretations of morpho-syntactic and lexical recasts.

3.2.6. Learner perceptions

The above classroom studies point to the fact that learners perceptions and interpretations of implicit CF, in particular recasts, may play a major role in their effectiveness. Mackey, Gass and McDonough (2000) studied the relationship between feedback and learner perceptions. They argued that learner perceptions of recasts may vary depending on the focus of recasts, and that the nature and the target of recasts affect both the learner perceptions and their modified output. Mackey et al. found that implicit negative feedback, especially recasts, were less accurately perceived when they addressed morpho-syntax than lexical and phonological errors. However, the recasts of morpho-syntactic errors were among the most frequent types of feedback provided. Carpenter, Jeon, MacGregor and Mackey (2006) conducted an experiment in which they demonstrated that the corrective function of recasts could be identified and interpreted appropriately only in the immediate discourse context. Carpenter and al. video-taped task-based interactions and showed the clips to two groups of advanced ESL students. One group was able to view the complete scenes whereas the other group was presented with the teachers' feedback without the preceding learners' utterances. The results of the study showed that the group who could not hear the utterances in context were significantly less successful in distinguishing recasts from repetitions. As Mackey (2006) suggested, the key to identifying the corrective function of recasts lies in the context since recasts are contingent on learners' errors.

3.3. Experimental laboratory studies

The advantage of laboratory studies is that they can be more tightly controlled and it is easier to exclude confounding variables that may be present in classrooms. However, classrooms have ecological validity. In this part of the chapter I focus on the experimental studies whose findings are most relevant to the types of CF investigated in the current research. Only the studies based on interaction and communication are included so the ‘focus on forms’ studies are not taken into consideration.

3.3.1. Positive and negative evidence

Following the early descriptive research in the Canadian immersion context, the majority of experimental laboratory studies have focused on the investigation of implicit corrective feedback. The question these studies addressed was whether implicit feedback promotes language development and acquisition. To answer this question it was necessary to employ a pretest – treatment – post-test design and measure the change in the learners’ interlanguage. Because descriptive studies found that much of the implicit CF in classrooms, in particular recasts, is not attended by students, the majority of the studies have focused on recasts as the type of feedback which has been shown to be the most ambiguous for learners.

There was also a need to investigate whether negative evidence plays a role in the acquisition of an L2. From this standpoint, Carroll, Swain and Roberge (1992) conducted a study with 79 intermediate and advanced learners of French who were individually trained on word formation in French. The participants in the experimental group received explicit feedback involving the correct form whereas the participants in the control group were not corrected. The results showed that the learners in the experimental group outperformed the control group on the immediate and delayed post-tests, but the correction did not help them to make morphological generalizations. Rather, corrective feedback enabled them to learn individual items. It appeared to be more useful for more advanced learners so they concluded that correction and proficiency may interact in the retention of information.

Long, Inagaki and Ortega (1998) compared the effects of models as positive evidence and recasts as implicit negative evidence and found that, at least for one structure, adverb placement in Spanish, recasts were more effective than models. There was evidence that five out of eight learners were able to produce the target structures with previously unknown

words. Leeman's (2003) study, on the other hand, aimed at examining whether the effectiveness of recasts was due to positive or to negative evidence. Her participants, learners of Spanish L2 were tested in four different conditions (1) recasts, i.e. they received negative evidence that is implicit in recasts which, at the same time, explicitly provide positive evidence, (2) negative evidence only, i. e. there was no positive evidence in the feedback, only the error repeated, (3) enhanced salience of positive evidence so that in the input the target feature was stressed to make it more salient but participants did not receive any feedback, and (4) unenhanced positive evidence (i. e. control group). Only the recast group and the group who received enhanced positive evidence in the input outperformed the other two groups. This led Leeman to argue that it was actually the positive evidence in recasts, which made them effective for language learners.

3.3.2. Learner 'readiness'

Mackey and Philp (1998) investigated the effects of intensive recasting on English question formation. The participants, 35 ESL learners, were divided into a group of 'readies' and 'unreadies' according to the level of their interlanguage. The results showed that those learners who were ready to move to a higher level demonstrated significant development. The reason why the 'unreadies' did not show much improvement may have been, as the authors hypothesized, in that the content of the recasts was not adjusted to their developmental level. This study also provided evidence that uptake or uptake with repair was not an indicator of learning. In this study uptake with repair was not frequent (only 5%) but the recasts proved to be effective, in particular for learners who were receiving recasts appropriate to their developmental level.

Philp (2003) investigated learners' noticing of recasts, using a "cued immediate recall" as a measure of noticing: on the sound of knocking the participants had to recall immediately the recast they had just heard as a response to their own utterance. The results of this study showed that the learners at higher and intermediate level of interlanguage development accurately recalled over 70% of recasts, while those at a lower level were accurate at recalling 60% of recasts. There was a significant difference in the rate of accurate recall between the high and intermediate groups on the one hand and the lower level group on the other hand. Thus Philp pointed to the developmental level or learner 'readiness' to acquire an L2 feature as a key factor in determining the effectiveness of recasts. Philp's finding is

directly linked to the observations in L1 acquisition and the claims (Farrar, 1990) that the child's 'readiness' is one of the main factors facilitating the effectiveness of recasts in the acquisition of grammatical morphemes. Related to the issue of learner 'readiness' is also Philp's (2003) finding that familiarity with a particular feature or structure may play a significant role. This involves not only the familiarity with the sound but also the possibility of recognizing the lexical items in the recasts, indicating that learners will notice more easily those features they are already acquainted with. Developmental 'readiness' was also emphasized as one of the key conditions for the success of recasts in a small scale study with ESL adult learners, reported by Han (2002). The author indicated that individual attention, consistent focus and intensity of recasts further contributed to their effectiveness.

3.3.3. Modified output following clarification requests

Since previous research had shown that clarification requests result in more modified output than recasts (Lyster & Ranta, 1997; Lyster, 1998a; 1998b; Lyster & Panova, 2002), McDonough (2005) set out to examine the effectiveness of clarification requests as an implicit form of negative feedback following ESL learner errors in question formation. She also tested Oliver's (1995, 1998, 2000) claim that uptake depends on whether the learners are given an opportunity to use the feedback or not. McDonough compared the post-test results of four groups who received either: a) clarification requests with an opportunity to modify output, b) clarification requests with enhanced opportunity to modify output by repeating the error, c) no opportunity to modify output and d) no feedback at all. Logistic regression showed that the production of modified output following negative feedback was the only predictor of ESL question development. This result corresponds to Havranek's (2002, p. 262) finding that unsuccessful elicitation of self-correction in the classroom is not conducive to learning.

3.3.4. Clarification requests, prompts and recasts

McDonough (2007) examined the differences between two implicit types of CF, recasts and clarification requests, on the emergence of the lexical activity aspect in English verbs. The participants were 74 Thai students who took part in one-on-one treatment sessions with their teachers, English native speakers, and received either recasts or clarification requests as responses to their erroneous production of past tense verbs. McDonough found both recasts and clarification requests were significantly more effective than no feedback but did not find

any significant difference between the effects of recasts and clarification requests. This research is directly comparable to the current study regarding the two types of CF. However, it differs with regard to how the CF was provided. In McDonough's study it was provided one-on-one but in the current study it took place in the whole class context. Lyster & Izquierdo (2009) also investigated two types of feedback, namely recasts and prompts, which were delivered individually to members of a group of French L2 undergraduate students. In a pre-test - treatment - post-test study Lyster & Izquierdo found that both prompts and recasts were equally effective as measured by oral production tasks and a computerized reaction-time binary-choice test.

A recent line of research has focused on examining recasts in the context of priming which has been a widely used research and treatment method in L1 acquisition (Bock, 1996; Bock & Griffin, 2000; Cleland & Pickering, 2003). Priming is thought to be an implicit cognitive process that occurs with little awareness and it suggests that language users' prior experience with language influences their subsequent language use (Trofimovich & McDonough, 2011). The phenomenon of syntactic or structural priming "refers to the tendency for speakers to produce a syntactic structure that appeared in the recent discourse" (2011, p. 8). McDonough (2006) and McDonough and Mackey (2006) conducted several experiments with the aim of exploring priming effects in the research on recasts. McDonough and Mackey (2006) found, for example, that development in ESL question formation significantly correlated with both recasts and primed production following recasts, but not with the immediate repetition of recasts. Also, the logistic regression analysis confirmed that only recasts and primed production, but not immediate uptake of the recasts, predicted the development. The authors operationalised primed production as "a learner's use of the question form provided in the recast, to ask a new question" (p. 705). Primed production occurred either immediately following a recast or within several turns of the recast. Related to this, McDonough and Mackey claimed that "productively using a form in one's own way a short time after hearing it, rather than immediately repeating or mimicking it" (p. 709), is associated with development. In another study McDonough (2006) found evidence of syntactic priming for learning prepositional datives in English but not for double object datives. This led the researcher to conclude that it would be important to further investigate if priming plays a role in L2 development.

3.4. Classroom quasi-experimental studies

Classroom studies investigating the effects of oral CF on L2 acquisition have not been as numerous as laboratory studies. However, their number has rapidly grown over the past few years. The results obtained in classroom studies have been mixed and sometimes even conflicting. This has been due mainly to the different instruments used for measuring acquisition, different instructional settings (FL vs. L2 or content-based classes), the different language structures that were examined and the different pragmatic and prosodic features of CF. The choice of studies presented in this section took into account whether the results were obtained in a pre-test – treatment – post-test research design and whether at least one of the conditions investigated included either recasts or clarification requests.

3.4.1. Interactional modification

The claims that interaction and negotiation lead to L2 development (Chaudron, 1977; Doughty & Pica, 1985, 1986; R. Ellis, 1990, 1992; R. Ellis, Tanaka, & Yamazaki, 1994; Gass & Varonis, 1994; Larsen-Freeman & Long, 1991; Long, 1981a, 1983a, 1983b, 1996; Pica, 1987, 1994; Pica, Holliday, Lewis, & Morgenthaler, 1989; Swain, 1985) have been tested not only in experimental one-on-one conditions, but in classroom studies too. For example, R. Ellis, Tanaka and Yamazaki (1994) conducted two studies of the same design, one with 79 and the other with 127 high school students, which provided evidence of the effectiveness of classroom interaction and the negotiation for meaning for vocabulary acquisition. In both studies the participants were divided into Premodified, Interactionally modified and Baseline (control) groups. Only the Interactionally modified group actually communicated with the teacher, whilst the other two groups listened to the instructions which were either baseline or premodified. The results showed that only the students who were exposed to the interactionally modified input comprehended better and learnt more new words. Interactional modifications were based on requests for clarification so that the students were trained to clarify meaning if they did not initially understand. The authors noted that there “seems to be no direct relationship between the frequency of interactional modifications and comprehension” (p. 482) since, as they explained, learning was not facilitated by the interactional modifications but rather by the modified input resulting from modifications. This study also provided evidence that not only the participants who took part in interaction improved their comprehension and learnt more words, but also those who were only exposed

to the input resulting from negotiation by others. This study therefore suggested that in a classroom environment students can learn just from listening.

3.4.2. Clarification requests

Takashima and Ellis (1999) found evidence of the superiority of clarification requests in a study carried out with 61 university students in Japan who were tested on their learning of English regular past tense. One group participated in ‘focused negotiation’, receiving clarification requests as error correction, i. e. whenever students produced an error. The other group received clarification requests only when there was a genuine breakdown in communication, so this kind of negotiation was termed ‘unfocused’. The tests were oral narratives and the target-like use analysis was applied in scoring. The results showed that the experimental, ‘focused negotiation’ group outperformed the ‘unfocused’ group in accurate use of the past tense on all the post-tests.

3.4.3. Recasts

In an attempt to examine the effectiveness of corrective recasts as a “relatively implicit focus on form” Doughty and Varela (1998) conducted a pre-test – treatment – post-test – delayed post-test study in a content-based classroom. They successfully used the repetition of a student error followed by recasting, to make recasts more salient to the learners. Without any explicit grammar instruction and with the student attention focused on conducting a science experiment, this study provided evidence that corrective recasts can be effective in a communicatively oriented lesson. The group who were treated with recasts improved their use of past tense forms in both oral and written production.

Muranoi (2000) also successfully used recasting preceded by a question for repetition. The treatment involved form-focused instruction after the task, so it was slightly different from the studies investigating only the effects of corrective feedback. Muranoi’s participants were 91 Japanese university students who received CF on English articles while conversing with the teacher in small groups in the classroom. They retained the gains evident in the immediate post-tests on the delayed post-tests administered five weeks later. A more recent study, by Revesz and Han (2006), provided evidence to support the claim that recasts are particularly effective if applied in a familiar context (Philp, 2003). Revesz and Han compared two groups who both received recasts, one of them during the work on familiar tasks and the other

working on unfamiliar tasks. The results revealed that the task and context familiarity played a significant role in the effectiveness of recasts, so that the learners working on familiar tasks improved significantly more than the other group. However, this difference was evident only in oral tasks whereas in written tasks both groups improved significantly in comparison with the control group.

3.4.4. Prompts vs. recasts

Lyster (2004) compared recasts with prompts, while carrying out a study in four French immersion classes in Canada, with 179 young learners (10 and 11-year old). Prompts, as opposed to recasts, consisted of clarification requests, elicitations and metalinguistic clues). The study lasted for five weeks and involved form-focused instruction on French grammatical gender, followed by a series of tasks and activities based on the use of the target structure. So this study did not look solely at the effect of CF but involved intensive form-focused instruction as well, which makes the comparison with other CF studies somewhat difficult. There was a pretest, immediate post-test and delayed post-tests. Apart from the two experimental groups, one group did not receive any feedback but did the communicative activities whereas the other comparison group did not do the activities either. The results on the post-tests showed significant gains in knowledge after the treatment for all three groups that received form-focused instruction, but prompts proved more effective than both recasts and no feedback, in particular in the written tests. Furthermore, the prompt group maintained their gains for a longer period of time whereas the recasts group and no feedback groups did not.

In a recent classroom study that compared the effectiveness of recasts and prompts in a Chinese university EFL context, Yang and Lyster (2010) found that prompts facilitated the acquisition of rule-based regular past tense forms while both recasts and prompts were shown to be facilitative of learning the exemplar-based irregular past. Prompts involved several prompting techniques, namely clarification requests, elicitations, metalinguistic clues and repetition. Yang and Lyster used both oral and written tests based on picture description. They explained their findings in terms of the learning benefits of modified output (Swain, 1985, 1998). They also related their findings to the dual mode system (Skehan, 1998) and speculated that “it seems that during online communication, prompts more than recasts

trigger access to the rule-based system whereas recasts and prompts alike trigger access to the exemplar-based system” (Yang & Lyster, 2010, p. 259)

3.4.5. Proficiency as a factor

Ammar and Spada (2006) also investigated the effectiveness of recasts and prompts in a study of Francophone learners of English. The target structure was the English possessive *his/her* which is difficult to acquire for French native speakers. In the Canadian bilingual situation learners usually attain high levels of comprehension and fluency in production but even after many years of L2 learning they often experience considerable problems with grammatical accuracy (Harley, 1989; Ammar & Spada, 2006). Sixty-four students in the 6th grade participated in the study which was spread over four weeks and included a pre-test, immediate post-test and a delayed post-test. This study showed that both experimental groups benefited from being exposed to corrective feedback in the classroom, but there were differences in the effects of recasts and prompts according to the learners’ proficiency level. Ammar and Spada (2006) found that recasts benefited the learners with higher proficiency whereas prompts benefited both higher and lower proficiency learners. Thus their study confirmed the results of Mackey and Philp’s (1998) laboratory study. In a later study Ammar (2008) explained that the prompt group significantly outperformed the recasts group in both written post-tests but only on the delayed oral post-test. The participants were also assigned to developmental stages established for the English 3rd person possessive and the reaction time of their oral production in the picture description task was measured. Ammar reported significantly faster reaction time on post-tests for the prompt group than for the recasts and the control group.

3.4.6. Explicit correction vs. recasts

Ellis, Loewen and Erlam (2006) reported on a classroom study which measured the effects of implicit (recasts) and explicit (metalinguistic explanation) corrective feedback on the acquisition of English regular past tense. The treatment consisted of CF in the course of two communicative tasks. The results of the study showed that the group who received explicit (metalinguistic) feedback performed significantly better than the recasts and the control groups on the delayed oral imitation post-test which was designed to measure implicit knowledge. On a grammaticality judgment test the metalinguistic group again performed significantly better than the recast group in the delayed post-test. Significant differences

between the metalinguistic group and the recasts group in the delayed post-test were found for new items and between the metalinguistic and control group for both old and new items. This indicated a beneficial role for the explicit corrective feedback for system learning. Improved accuracy in the use of past tense only in the delayed post-test was explained as due to the need for time to consolidate in learning that had taken place.

Ellis (2007) compared the effectiveness of implicit (recasts) and explicit (metalinguistic explanation) CF on learning two grammatical rules in L2 English: regular past tense considered an easy rule and comparative adjectives considered a more difficult rule. The results of the study conducted with 34 ESL students in New Zealand showed that, when both structures were taken together, there was no difference among the three groups (recasts, metalinguistic and control) except for the oral imitation test on which the metalinguistic group outperformed the control group but no differences between the two experimental groups were evident. However, the analyses comparing the effects of CF on each grammatical structure indicated that recasts had a similar effect for both structures showing an initial gain only for comparative but a small decline in scores on a delayed post-test. Metalinguistic feedback resulted in much larger initial gains for the comparative and larger subsequent gains (from immediate to delayed post-test) for the regular past tense. Moreover, only metalinguistic feedback was found to have an effect on both structures, although only on ungrammatical sentences in the oral imitation test.

Sheen (2008) also compared the effectiveness of recasts and explicit CF on the acquisition of English articles. She found that explicit correction was significantly more effective than the treatment with recasts. Language anxiety was identified as one cause of the ineffectiveness of recasts. Loewen and Nabei (2007) found that none of the three CF types they researched (recasts, clarification requests and metalinguistic feedback) had any effect on the tests of explicit knowledge while all three proved to be effective on the tests of implicit knowledge. There was no difference in the effectiveness of these three types. Also, no effect for CF was found in oral production tasks which were used to measure acquisitional stages. Such results, as the authors observed, may have been due to a very low number of participants and a very short period of treatment.

To sum up, the findings from experimental classroom studies suggest that:

- Explicit metalinguistic feedback and prompts including several prompting strategies have been shown to be more effective than recasts in classroom environments.
- Recasts have been found effective in promoting language development provided they are made salient so they are noticed. In particular, recasts can benefit more proficient learners and those who are ‘ready’ to acquire a certain language feature.
- Uptake with repair following clarification requests has been found to be associated with learning. However, studies investigating recasts have shown that learning can occur even when recasts are not followed by uptake with repair.

Even though the number of classroom studies has grown in recent years, there are still relatively few experimental classroom studies that have investigated particular feedback types and that have ensured a rigorous pre-test, immediate post-test and delayed post-test design.

3.5. Meta-analyses of oral corrective feedback

Since Norris and Ortega’s (2000) seminal meta-analysis, five meta-analyses, directly or indirectly related to CF, have been conducted. Russell and Spada’s (2006) study included both written and oral CF, whereas Keck, Iberri-Shea, Tracy-Ventura and Wa-Mbaleka’s (2006) and Mackey and Goo’s (2007) meta-analyses focused on the effectiveness of interaction. More recently, Lyster and Saito (2010) and Li (2010) also published meta-analytic studies of CF. These are more directly relevant to the current study.

Lyster and Saito focused on classroom CF and included 15 studies in their meta-analysis. Li meta-analysed 33 studies including not only published ones but also PhD theses. Lyster and Saito’s study reported that overall corrective feedback had a significant effect on learners’ acquisition of the L2 structures which were investigated, with a medium effect size for comparisons between experimental and control groups. A large effect size was found for within group contrasts. The effect of oral CF was durable, i.e. evident in delayed post-tests. A larger effect size was found for prompts than for recasts and free constructed responses yielded a larger effect size than more constrained outcome measures. The comparison of SL and FL classroom settings did not result in any significant difference. Longer treatments proved to be more effective than shorter treatments, and corrective feedback proved to be more effective with younger learners than with the adults. Li’s (2010) meta-analysis reported a medium overall effect for CF and the effect was durable. This meta-analysis included both classroom and laboratory studies, and compared them. It found stronger effects in laboratory

experiments than in classroom studies. Explicit types of CF worked better than implicit feedback in both the immediate and the short-delayed post-tests. However, Li has reported that implicit feedback resulted in a larger effect size in the long-delayed post-tests. In contrast with Lyster and Saito, Li's analysis produced stronger CF effects in FL than SL settings. Also different from the Lyster and Saito's study was the effect of treatment length – Li found that shorter treatments were significantly more effective than longer treatments. On the other hand, the finding relating to outcome measures was similar in the two meta-analyses: Li also found that free-constructed responses yielded a larger effect size than constrained-constructed responses and judgment tests.

3.6. Summary and conclusion

Based on the review of descriptive classroom studies, laboratory studies and classroom experimental studies, the following are the main findings regarding recasts and clarification requests.

Recasts:

1. Recasts are the most widely and most frequently used CF type in language classrooms.
2. Uptake with repair and the effectiveness of recasts differ significantly in different instructional settings (classroom vs. laboratory, FL classes vs. immersion and SL classes).
3. Learner 'readiness' and familiarity with the language features play a significant role in the effectiveness of recasts.
4. Immediate discourse context influences the extent to which recasts are effective for learning.
5. The characteristics of recasts (length, stress, mode, number of changes) influence their effects on learning.
6. Recasts may be ambiguous and not noticed in a classroom if they are not salient enough. Morpho-syntactic recasts are more difficult to notice than phonological and lexical recasts.

7. Uptake following recasts is an optional move. Uptake with repair after recasts is sometimes but not always associated with learning.

Clarification requests:

1. Clarification requests are less frequently used as a type of CF in language classrooms.
2. Clarification requests have been investigated most often as one type of prompting technique, but very rarely as an isolated type of CF.
3. Clarification requests have been shown to be conducive of uptake much more than recasts, but not with a high percentage of repair.
4. Clarification requests have proved to be particularly beneficial for learners at lower levels, when negotiating for meaning.
5. Clarification requests can be used to address either misunderstanding related to meaning or problems related to language accuracy (Lyster & Ranta, 1997, p. 47).

Given these points, and being aware of some of the conflicting findings resulting from the research so far, I set out to examine the effectiveness of recasts and clarification requests in a French FL classroom environment. Even though the number of classroom studies has grown in recent years, there is still a paucity of experimental classroom studies investigating particular feedback types and applying a rigorous design with a pre-test, immediate post-test and a delayed post-test. Clarification requests have so far been examined as part of a group of prompting techniques in classroom instruction but not as an isolated CF type. To the best of my knowledge only one study has looked at the difference between recasts and clarification requests in one-on-one conditions (McDonough, 2007) and only one study, of very short duration, attempted to compare the effectiveness of recasts, clarification requests and metalinguistic feedback in a classroom setting (Loewen & Nabei, 2007). There is no record of a similar study in the context of learning French as a foreign language. Considering the findings of Lyster and Ranta (1997) and Panova and Lyster (2002) studies, where recasts were found to be most frequently used but the least effective type of corrective feedback, it seemed worthwhile to undertake further investigation of recasts and clarification requests as two types of implicit negative feedback. Comparing their effectiveness in the context of French FL classroom may provide us with important information about their differential effects, in particular if the focus is on language structures (the *passé composé* and the *imparfait*) that are especially challenging for L2 learners in terms of both use and form.

Chapter 4. Individual learner differences

4.1. Introduction

Individual differences have been recognized as extremely important in the process of language learning. Research has shown that individual factors can determine not only the speed at which languages are being learnt but also the level of L2 attainment (J. B. Carroll, 1962; Ehrman & Oxford, 1995; R. Ellis, 2004; Gardner, 1985). Individual factors can be classified broadly according to the ability/aptitude, motivation and personality traits (Dörnyei, 2005) or as cognitive, motivational and affective variables (R. Ellis, 2012). The current study looked at the roles of language analytic ability (LAA) and working memory (WM) as cognitive factors and at language anxiety as an affective variable. This chapter will attempt to present a review of literature related to these individual difference (ID) factors. Each of them will be discussed separately. I start with the notion of language aptitude which has traditionally been used to ‘cover’ all or some of the individual capacities associated with easier, or faster language learning. Then I separately review the literature related to the role of analytic ability and working memory in L2 learning, including the studies investigating these IDs in interaction involving corrective feedback. However, it is not possible to avoid some overlap in the review of studies investigating analytic ability and those looking at working memory because these constructs, even though traditionally measured by different tests, are clearly related. The last part of the review is concerned with language learning anxiety.

4.2. Language aptitude

Language aptitude or a person’s special ability related to language learning is considered to be one of the most important ID factors influencing the rate and success of second language learning (J. B. Carroll, 1981; Dörnyei, 2005; Ehrman & Oxford, 1995; R. Ellis, 2004; Sawyer & Ranta, 2001; Skehan, 1989; Sternberg, 2002). Aptitude is a complex of several, mostly stable, traits. So, as R. Ellis put it, “it is best viewed as a complex construct, involving both a general cognitive dimension and a more specific linguistic dimension” (R. Ellis, 1999). The concept of aptitude which has been influential until today is based on Carroll’s extensive work and his four-component view of aptitude (J. B. Carroll, 1962, 1981) consisting of:

- (1) phonemic coding ability,
- (2) grammatical sensitivity,
- (3) inductive language learning ability, and
- (4) rote learning for foreign language materials, which was later renamed as ‘associative memory’ denoting the capacity to form links in memory.

Skehan (1989, 2002) proposed a slightly changed concept of aptitude, merging grammatical sensitivity and inductive ability into one component which he labelled ‘language analytic ability’. Thus the three modules or three phases in Skehan’s (1989) information processing model of L2 acquisition (input, processing and output) involve:

- (1) auditory processing,
- (2) language processing, and
- (3) the memory component of language learning aptitude.

In spite of evidence suggesting a close relationship between aptitude and achievement in second or foreign language learning (J. B. Carroll, 1962, 1981; Ehrman & Oxford, 1995), some linguists have strongly opposed the idea that innate ability is responsible for success in language learning. Krashen (1981, 1985) for example, claimed that aptitude may be relevant only in classrooms where language is taught in an explicit way and in ‘monitored’ test situations, and thus only relates to conscious learning. He argued that language ‘acquisition’ does not depend on aptitude. However, Skehan (1989) who also found a relationship between the speed of learning the first language and language aptitude scores measured in the same subjects a decade later, suggested that aptitude may be equally important in both formal classroom and immersion conditions. He suggested it might be even more important in informal contexts because in such conditions learners are left without assistance and they need to infer the rules for themselves. For example, Reves (1983) as cited in Skehan (1989), compared two different conditions in which a group of Arab students learned English in a formal classroom and Hebrew informally with their host families, and found that learning in both situations was closely related to their aptitude. Such findings are relevant to the present study which sought to investigate how two cognitive factors (language analytic ability and working memory) are related to the L2 learning that occurs when learners perform meaning-oriented tasks. That is, it addressed the question: What role did these cognitive abilities play in focus on form during communication?

The following sections of the chapter present a review of those studies that have looked at different aspects of language aptitude in relation to learning outcomes and to different conditions of learning. The role of aptitude related to age is also considered. The studies included in this review are relevant to the current study in that they provide information about what kind of learning conditions are susceptible to language aptitude and what conditions do not depend on innate abilities.

4.3. Language analytic ability

Language analytic ability has been identified as that component of language aptitude which correlates with general intelligence (Sasaki, 1996), in particular with those aspects of intelligence which are known as fluid intelligence (Ehrman & Oxford, 1995; Sternberg, 2002). Analytic ability is essential for making inferences and generalizations, and thus contributes to all aspects of abstract, formal, logical or mathematical thinking. Sternberg contends that “all languages might require some measure of logical-mathematical intelligence” (2002, p. 21). In the most complex and widely used aptitude battery until today, the Modern Languages Aptitude Test – MLAT (Carroll & Sapon, 1959), language analytic ability is related to the sub-test measuring grammatical sensitivity (Words and Sentences part of the MLAT). It should be noted, however, that inductive language learning ability, which Carroll (1962) also considers a component of aptitude, is actually not measured by the MLAT. On the other hand, a test specially designed for young learners by Pimsleur (1966), contains a measure of inductive ability. For Skehan (1998), as already pointed out, inductive learning ability and grammatical sensitivity together constitute ‘language analytic ability’ which is involved in the processing phase of language learning.

4.3.1. Matching Instruction to IDs and Aspects of Learning

The fact that aptitude is complex and that there are notable differences among L2 learners led some early researchers to examine whether particular learning conditions could be matched to particular learners depending on their different aptitude scores. For example, Wesche (Wesche, Edwards, & Wells, 1982) assigned the students to two different learning conditions, so that those with higher analytic ability but with less phonemic decoding ability were assigned to more analytic instructional conditions, while those with high overall aptitude scores were taught using more memory oriented teaching methodologies. Wesche found no difference in the final achievement of the two groups. She also reported that both groups of

students were equally satisfied. Nagata, Aline and Ellis (1999) examined how different components of aptitude influenced different aspects of L2 learning. Their study, for example, showed that analytic ability as measured by the MLAT grammatical sensitivity sub-test was closely related to comprehension but not to vocabulary learning. On the other hand, the scores on the MLAT memory sub-test correlated significantly with both comprehension and vocabulary learning. Furthermore, the study showed that sound-symbol associations of the PLAB sub-test were related to short-term vocabulary learning but not to long-term learning.

4.3.2. Learning Conditions and Learning Mechanisms

Harley and Hart (1997, 2002) conducted two studies with the aim of investigating whether different components of aptitude, namely analytic ability and associative memory, were more related to the learning conditions or to the age at which the learners began learning the L2. In the first study (1997) they compared two groups of grade 11 French L2 immersion students in Canada, those who started early in grade 1 and those who started later, in grade 7. They administered the ‘associative memory’ sub-test of MLAT and the PLAB analytic ability sub-test for young learners. They found different patterns of relationship between these tests and proficiency measures in the two groups of learners. It appeared that those who started early showed stronger correlations with the memory measures whereas for the late starters there was a significant relationship with language analytic ability. Harley & Hart concluded that this may have been due to the different methods of teaching since the teaching was more explicit with the late starters.

The second study by Harley and Hart (2002) was carried out with a group of adolescent students who started L2 French in a ‘core’ programme in grade 3 or 4 and took part in an immersion three-month exchange with French families in Quebec. The same aptitude tests were administered as in the previous study. This time significant correlations between analytic ability measures and scores on an oral imitation and a cloze test were found. The memory sub-tests did not appear to be significantly related to proficiency after the three-month immersion. This study suggested that the age of starting to learn an L2 rather than the learning setting may be more important where analytic ability is concerned.

Strong support for the claim that the role of analytic ability is clearly related to the age of L2 onset has been provided by DeKeyser (2000) who tested a group of 57 Hungarian immigrants to the US. He was able to show that analytic ability as measured by the MLAT grammatical

sensitivity sub-test correlated significantly with the scores on a grammaticality judgment test for those L2 speakers who arrived in the US after the age of 17. The GJT scores of ‘young arrivals’ did not correlate with their analytic ability scores.

DeKeyser administered the same test as that used by Johnson and Newport (1989) to examine the Critical period hypothesis. He also provided evidence that those ‘late comers’ who achieved the highest, near-native level in L2 English had all very high analytic ability. This was explained by the argument put forward by Lenneberg (1967; 1975), who claimed that only children before the age of puberty can use their implicit learning mechanisms for learning a language whereas adults rely on their general cognitive problem-solving abilities when faced with the challenge of learning a second or foreign language.

4.3.3. Analytic Ability and Teaching Methods

There is presently substantial evidence indicating that adult L2 learning is related to both analytic ability and memory - the individual capacities which are captured by the MLAT sub-tests. There is a large number of studies, including Ehrman and Oxford’s (1995) study, that strongly support the claims put forward by Carroll and Sapon (1959) suggesting that “MLAT may be addressing learning abilities that are independent of methodology” (Ehrman & Oxford, 1995, p. 77). However, there are still questions which call for more research, in particular regarding teaching methodologies in different learning environments and L2 development. One of the studies which has often been cited as providing evidence that analytic ability plays a role in communicative as well as in formal classroom settings is Ranta’s (2002) research. This was conducted in Quebec with young L1 French speakers in an intensive ESL programme. It is worth noting, however, that the task used to measure the analytic ability of French L1 children in this study tested grammatical knowledge in written French rather than analytic ability, and thus really only showed a relationship between L2 proficiency and formal grammar knowledge in the L1. Ranta argued that the part of MLAT measuring grammatical sensitivity is actually a set of metalinguistic tasks (2002, p.162). However, even though there may be some overlap, the link is not straightforward. Ranta’s metalinguistic task did not tap learners ability to analyse novel sentences or to generalize, both of which are essential components of language analytic ability. So, it is not really possible to claim that analytic ability plays a significant role in communicative language

teaching in general. It is, however, clearly important to investigate the interaction between the instructional approach and language analytic ability.

Such an investigation can be found in Erlam's (2005) study of learners of L2 French at secondary school level. The target structures were direct object pronouns in French. Having administered the tests of grammatical sensitivity (MLAT) and phonemic coding ability (PLAB), she found that the deductive instruction based on the explicit presentation of the rule followed by practice in oral production was superior to the inductive (rule-search) instruction and the structured input instruction in which explicit grammar explanation was followed by input-based activities. Furthermore, deductive instruction was the only condition which benefitted all learners irrespective of their abilities since only in the 'deductive group' there was no evidence of a significant correlation between the learners' IDs and their gain scores. It is worth noting that the deductive instruction group not only received explicit explanation of the rule but also a chart that classified the use of pronouns – which may have assisted the students' learning. DeGraaff (1997) used an artificial language and the MLAT sub-tests to investigate the differences in learning between a group who received an explanation of the grammatical rules and a group who did not receive any explanation. On all post-test proficiency measures the group who received explicit explanation performed better. Language aptitude (i.e. grammatical sensitivity and memory sub-tests) was found to be relevant to learning outcomes in both groups. The only exception was the judgment task without pressure where the scores did not correlate with language aptitude in either of the two groups.

4.3.4. Analytic Ability and Incidental Learning

An important line of research, which has systematically investigated the relationships between learning conditions, L2 learning outcomes and cognitive IDs, is Robinson's theory of aptitude-treatment interactions. Robinson (1997a) examined four types of instructional conditions for the learning of both difficult and easy grammatical rules, by assigning Japanese learners of English to four groups: explicit (with explicit presentation of the grammar rule), implicit (as memorization of the text with the in-built rule), rule-searching and incidental (with an orientation to meaning). In all the first three conditions gains in acquisition correlated strongly with the aptitude scores, whereas incidental learning was not related to aptitude measured by the MLAT sub-tests. In this condition the participants were

given a meaning-related task with the rule built into the material. Such results suggest that aptitude might not be as important when the focus is on meaning and negative feedback is provided implicitly. This study was replicated (Robinson, 2002a) using a new language, Samoan, and the same four conditions. Again the results indicated no significant correlation with aptitude in the group who were assigned to the incidental learning condition. Robinson emphasized that the task demands determine the role that IDs play in each learning condition. He further argues that in real classroom conditions it is more likely that aptitude will determine the learning success in any condition. He claims that “when an individual’s abilities match the demands of the task, aptitude often leads to awareness and to superior levels of learning” (Robinson, 1997b).

4.3.5. Analytic Ability and Corrective Feedback

Until recently there were only few studies which looked at the relationship between corrective feedback (CF) and IDs. Even though the first such studies did not investigate IDs in relation to specific CF strategies and did not target specific language features, they afford some interesting insights regarding possible interactions between IDs and CF. Commenting on the lack of studies that have examined the effectiveness of error correction in a classroom context, DeKeyser (1993) noted that this may be because the provision of error correction depends on a number of individual students’ characteristics. Thus he conducted a study in which he looked at corrective feedback in relation to extrinsic motivation, previous achievement, grammatical sensitivity and anxiety, hypothesizing that these factors were the most likely to predict success in L2 learning. The study involved two classes of Dutch students learning French as an L2 in Belgium. It extended over a full school year and it included corrective feedback on all errors ‘concerning morpho-syntax’ during different communicative activities in one class, while the other class completed similar activities but without any CF. DeKeyser found no significant differences in language proficiency scores between the two classes at the end of the year. Although grammatical sensitivity was a significant predictor of the grammar test scores, the hypothesis that students with strong grammatical sensitivity would benefit most from error correction was not confirmed. The second study was carried out by Havranek (Havranek & Cesnik, 2001). She analysed the data from 207 subjects of different ages and proficiency levels in a range of different settings, from primary school to university. The data were collected during class time and included ten different types of CF. This study looked at the relationship between the effectiveness of CF

and proficiency (measured by the term mark), and various individual learner variables, including verbal and non-verbal intelligence. The study showed that those who were ‘better language learners’ and who achieved higher scores on proficiency tests benefited more from the error correction which was provided.

A recent study by Sheen (2007b, 2008, 2011) looked specifically at the effectiveness of recasts and explicit metalinguistic CF mediated by language analytic ability. The targeted structures were English articles. The mediating effect of analytic ability was examined in relation to direct and direct metalinguistic CF provided both orally and in written form.

The test used to measure analytic ability was the same as the one used in the present study. The results of Sheen’s research revealed that analytic ability scores correlated significantly with the gains of those participants who received explicit metalinguistic correction (written and oral) as well as written direct CF. In contrast, the gains of the recast group (which were lower than those in the metalinguistic group) were not related to the participants’ analytic ability.

In another study, using a computerized task, Trofimovich, Ammar and Gatbonton (2007) investigated the relationship between several IDs including analytic ability (the grammatical sensitivity sub-test of MLAT) and production accuracy scores in a picture description task. The participants were provided with recasts while performing the task. The results showed a significant correlation between analytic ability and production accuracy scores after about 12 minutes. However, no significant correlation was found with the learners’ self-reported noticing of recasts. It should be noted that these results refer to the production accuracy scores very shortly after the recasts had been heard (12 minutes).

Finally, it is interesting to mention an unpublished study by Robinson and Yamaguchi (1999), reported in Robinson (2005). This study revealed that at the end of a five-week period in which students received recasts during task-based instruction, learning gains were not associated with the scores on the grammar sensitivity sub-test of MLAT. However, they correlated significantly with the scores of phonetic sensitivity and (rote) memory ability. These results, along with those obtained in Robinson’s other studies (Robinson, 1996, 1997a, 2002c) now led to the conclusion that “across contexts, laboratory studies of incidental learning and classroom studies of focus on form during task-based learning” indicate that

there is “non-influence of IDs in grammatical sensitivity on incidental learning processing for meaning” (Robinson, 2005, p. 56).

4.4. Working memory

The concept of working memory has attracted much attention over the past decades, following the seminal work of Baddeley and Hitch (1974), who introduced their model of working memory. As they explained, working memory is “the short-term memory system, which is involved in the temporary processing and storage of information” (Gathercole & Baddeley, 1993). Working memory is a limited-resource system and its capacity may change over a person’s life-span. It has been found that people may differ considerably in their WM capacity.

The role of working memory is evident in a range of complex everyday cognitive activities, for example in comprehension, reasoning, problem-solving or long-term learning. So, the relationship between working memory and language and its importance for both L1 and L2 learning has figured frequently in empirical research on working memory. Miyake and Friedman (1998) suggested that “working memory for language may be one (if not the) central component of language aptitude” (p. 339).

There have been two main strands of research into the complex functioning of working memory, involving (1) the British group and (2) the Canadian group of researchers. The first one has generated an impressive number of empirical studies by looking into the role of working memory and in particular the role of the ‘phonological loop’ in both children and adults. The non-word repetition task or a serial digit span tasks have been used most often (A. D. Baddeley, Gathercole, & Papagno, 1998; A. D. Baddeley & Logie, 1999; Gathercole & Baddeley, 1993; Miyake & Shah, 1999). The second line of research has focused on the human capacity to process and store information simultaneously (Daneman, 1991; Daneman & Carpenter, 1983; Daneman & Green, 1986; Daneman & Merikle, 1996; Just & Carpenter, 1992). For this purpose the reading span, listening span and speaking span test have been designed.

4.4.1. Phonological Loop

Baddeley and Hitch (1974) first proposed a three-part model of WM consisting of (1) the phonological loop, (2) the central executive and (3) the visuo-spatial sketchpad. The fourth component, the episodic buffer was added later (Baddeley & Logie, 1999). The episodic buffer is assumed to be a temporary storage system or “a temporary representation accessible to conscious awareness” (Baddeley, 2007, p. 139) . The central executive, which is “the most important but least understood component of working memory” (p.117) , is thought to be responsible for attentional control. It does not involve storage itself (Gathercole & Baddeley, 1993). The phonological or articulatory loop has been the most extensively researched component of WM. It is well established that the phonological loop increases significantly in children between the age of four and eight, so older children and adults have a much larger WM capacity than young children. It has been found that the phonological short-term memory plays an important role in the acquisition of new vocabulary between the age of four and five but that after the age of five the role of STM in vocabulary acquisition diminishes (Gathercole & Baddeley, 1993).

The role of the phonological loop is also thought to be very important for L2 vocabulary learning and the acquisition of grammar (Baddeley, Gathercole & Papagno, 1998). The phonological loop has two components: the phonological short-term store and the sub-vocal rehearsal mechanism. It is thought that the short-term store enables the sub-vocal rehearsal process to take place during the acquisition of new phonological forms. Baddeley et al. (1998) suggest that the phonological loop may have a critical role not only in learning the phonological forms of new words but also in syntactic learning. For example, in a developmental study (French, 2006) of 54 French L1 children (age 11 and 12) attending an intensive 5-month ESL course in Canada, it was found that phonological memory was closely associated with vocabulary acquisition both initially and after five months. However, after five months the correlations were significantly greater for the low proficiency group than for the high proficiency group suggesting that the role of phonological WM is more significant at the initial stages of L2 learning, and that the existing L2 knowledge increasingly mediates L2 development at later stages. It was also found that phonological memory predicted the scores of L2 grammar ability at the beginning, but when vocabulary knowledge had been accounted for, PSTM was no longer a significant predictor. Significant correlations between phonological WM (PSTM) and the adult L2 grammar learning were also found in

experimental studies, using both artificial languages (N. C. Ellis & Schmidt, 1997) and natural languages, for example Welsh (N. C. Ellis & Sinclair, 1996) and Italian (Williams, 1999). However, one possibility advanced by Bates and Goodman (1997) is that grammatical development depends upon the size of a person's lexicon, and that phonological memory ability influences grammar development only indirectly, by impacting on the acquisition of vocabulary. Williams (1999) showed that the memory for input may be more important at the beginning stages of L2 learning. The same is suggested by the results of the experiments reported in Williams and Lowatt (2003), where there was an initial strong association between the participants' learning scores and memory but this decreased with the amount of training.

4.4.2. Working Memory Capacity

The second line of research has been more focused on the examination of the individuals' WM capacity to process and store information simultaneously. This is what usually happens in reading for comprehension, problem solving, parsing syntactically complex structures and other complex cognitive tasks that people perform in everyday life. Daneman and Carpenter (1983) developed a reading span test and in a meta-analysis of 77 studies (Daneman & Merikle, 1996) the average weighted correlation between reading span and global measures of L1 reading comprehension was .41, much higher than the correlation of reading measures with the tests measuring the phonological loop. N. C. Ellis (2005) contends that when the span test is used (a reading span, a listening span or a speaking span test, for example) then individual differences in working memory relate to language learning aptitude (i.e. analytic ability). Ellis explains that "phonological short-term memory should contribute to the memory of form ... and to the ability to hold phonological feedback in the form of recasts" whereas 'the supervisory attentional system as tapped more in the Canadian group should be more associated with explicit learning' (p. 339).

The differences between the two constructs (PSTM and WM capacity) measured by different types of tests can be clearly seen in Harrington and Sawyer (1992) and Kormos and Sáfár (2008) studies. Harrington and Sawyer's study provided evidence that WM capacity is an important factor in L2 reading comprehension by imposing heavy demands on WM for processing. While there were significant correlations involving reading measures and the reading span test, such correlations were not found for the simple digit span or word span

tests. Similarly, Kormos and Sáfár (2008) reported the results of a study with 121 Hungarian young learners of English who took the Cambridge First Certificate Exam (FCE) and were tested for PSTM using the non-word span test and for WM capacity using the backward digit span. All measures of the FCE were only weakly related to PSTM scores but were highly significantly associated with WM capacity scores, suggesting that the two types of WM tests were really measuring two different constructs.

It is evident from these and other studies (Just & Carpenter, 1992; Miyake, Carpenter, & Just, 1994; Miyake & Friedman, 1998; Miyake & Shah, 1999) that the role of WM capacity is crucial in L2 comprehension and production tasks. However, there are still many unresolved questions. Miyake and Friedman suggest, for example, that WM limitations can constrain specific language processes, such as complex sentence syntactic parsing or ambiguity resolution (1998, p. 348). On the other hand, Juffs (2004, 2005) found only weak evidence for the role of WM in an online processing task that involved ambiguity resolution, and claimed that differences in WM (measured by the reading span test) “do not seem to play a role where online processing is concerned” (Juffs, 2004, p. 220). He contended that variability in L2 parsing and comprehension in online tasks is linked primarily to the influence of the learners’ first language (Juffs, 2005, p. 213).

The type of task is clearly an important factor in determining the scope of WM involvement. Thus Daneman and Green (1986) argue that processing differences in WM are “highly task-specific” (p. 15). They developed the speaking span test with the aim of confirming their hypothesis that “vocabulary production fluency is related to the individual’s ability to coordinate the processing and storage functions of working memory” (p. 10). According to Daneman and Green, WM capacity determines how well individuals can use context to both comprehend and produce words. This means that, when speaking, a person has to retrieve and use semantically most appropriate words for the intended idea, for a particular context, and use them appropriately in relation to syntax. In this process both the storage and the processing functions of working memory, and their coordination, seem to be crucial. The more demanding the task is, the more activation of working memory is needed. So, for individual differences in WM capacity to emerge, the task performed has to be complex (Just & Carpenter, 1992). Using a speaking span test based on Daneman and Green (1986) and Daneman (1991), Guar Tavares (2008) found significant correlations between WM scores and accuracy in oral (narrative) tasks without planning, whereas for the group who had time

for planning there were significant correlations of WM capacity only with the fluency and complexity measures. It has been found that the speaking span test predicts fluency in the speech of L1 and L2 speakers (Daneman, 1991; Daneman & Green, 1986) and may predict accuracy in the speech of L2 speakers when performing the task online.

4.4.3. Working Memory and Type of Instruction

There are only few studies that have explored possible interactions between L2 learners' working memory and the type of instruction. For example, in a study published only in Japanese, Ando, Fukunaga, Kurahachi, Suto, Nakano, & Kage (1992) investigated, as cited in R. Ellis (2012), the mediating effects of WM on Japanese young students' learning of L2 English. They used both reading and listening span tests and found that in the traditional grammar-oriented approach WM correlated highly with the learning scores on a delayed test. In contrast, in the group who were taught by the more communicative approach those students with lower WM spans benefited more in short term. Erlam (2005) investigated the role of working memory in three different teaching methodologies (already mentioned in the section related to analytic ability). She used a test designed to assess the processing of information and found significant positive correlations between learning gains and WM only in the group assigned to the structured input condition. For the other two groups there were no significant correlations between learning scores and WM scores. Erlam pointed out that the WM test used in this study measured only the processing of information and this may explain why she only found a relationship between WM and learning gains in the group that had to process input. In the case of the study reported in this thesis, WM was investigated in relation to gains after relatively implicit corrective feedback in oral interaction.

4.4.4. Working Memory and Corrective Feedback

Robinson (1995, 1997b, 2002c, 2003, 2005b) developed a complex framework of aptitude-treatment interaction research involving profiling learners' cognitive IDs and matching these profiles to instructional treatments including CF. In such a model, for example, learners with stronger 'memory for contingent speech' and noticing-the-gap would benefit from implicit CF such as recasts, whereas learners with lower phonological memory capacity and speed (both involved in memory for contingent speech) would have difficulty in responding to recasts (Robinson, 2002c, p. 118). A study that provided evidence for such a conclusion was already mentioned Robinson and Yamaguchi (1999) in which learners received recasts during

the task-based instruction. Robinson compared the results of this study with similar results obtained by the Mackey, Philp, Egi, Fujii and Tatsumi study (2002) who measured phonological working memory in laboratory conditions. Mackey, Philp and al. (2002) looked at the possible interaction of recasts and L2 development by investigating the role of PSTM and WM. For that purpose a non-word recall test and a listening span test were used. The results suggested that there was a relationship between WM capacity measured by a composite score and reports of noticing the recasts: learners with larger WM reported more noticing whereas those with a lower WM composite score noticed less. Also, those with low WM scores showed the most initial interlanguage change whereas learners with higher WM scores showed more development in the delayed post-tests. As R. Ellis (2012) pointed out, these studies suggest there are interactions between working memory, type of instruction, and whether the effects of the instruction are immediate or delayed.

In a more recent study Mackey, Adams, Stafford and Winke (2010) examined the role of WM in the learners' production of modified output during interaction with more proficient interlocutors. WM capacity was measured by the listening span test. Learners of Spanish L2 were provided with corrective feedback which was not described in detail but most often involved clarification requests. Mackey et al. found that WM capacity predicted the amount of modified output (i.e. the learners with larger WM capacity produced more modified, target-like output). In the study mentioned earlier, which used a computerized task with L1 French speakers learning English, Trofimovich, Ammar and Gatbonton (2007) looked not only at analytic ability, but also at the role of PSTM measured by a non-word recognition task and the processing and executive aspects of WM. No significant correlations were found for any of these, neither with self-reported noticing nor with the production accuracy measures. However, the authors reported that, when taken together, the combined contribution of attention control, analytic ability and PSTM could explain 40% of the variance in production accuracy scores. Sagarra (2007) also used tasks in which recasts were delivered via a computer and investigated their effects on gender agreement in L2 Spanish. This study found beneficial effects for recasts in both written post-tests and delayed face-to-face post-tests. Having administered the reading span WM test, Sagarra reported a "significant main effect for WM" (p. 242) which was used as a covariate in the analysis. However, it is worth noting that all the research on working memory over many years has had to rely on "an inherently correlational paradigm" because, as Daneman and Green (1986) noted, "working memory (like many other individual differences parameters) cannot be

manipulated readily in the laboratory” (p. 17). Hence either correlation or regression analyses are preferred in research on individual learner differences.

4.4.5. Language learning anxiety

Language learning anxiety has usually been classified as an affective learner variable. However, Dörnyei (2005) points to the ambiguity of the concept of anxiety by stating that it is not clear whether it is a motivational component, or a personality trait, or emotion. Despite this theoretical uncertainty, empirical research has provided substantial evidence that anxiety can seriously impact on L2 learning. Horwitz, Horwitz and Cope (1986) have conceptualized the construct of foreign language anxiety as they named it, as a type of situation-specific anxiety arising from the problems some learners may experience in L2 learning. They developed a 33-item instrument, the Foreign Language Classroom Anxiety Scale (FLCAS) which has been widely used in language learning environments and has provided valuable data on the kind of problems L2 learners may go through. Such quantitative data have been complemented with learners’ qualitative accounts in journals and dairies.

Different from trait anxiety which is a personality characteristic and state anxiety, language learning anxiety has most often been associated with the notion of a situation-specific anxiety, which can be defined as a tendency to experience a state anxiety reaction during language learning or communication. MacIntyre (1999) explained that language learning anxiety can develop into state anxiety, which, “is the transient emotional state of feeling nervous that can fluctuate over time and vary in intensity” (p. 28). State anxiety has an effect on emotions, cognition and behaviour. When people experience state anxiety they “ruminate over real and imagined failures, and often try to plan ways to escape from the situation” (p. 29). As MacIntyre explained, a person with a high level of language anxiety may frequently experience state anxiety. It is also possible, as Ganshaw and Sparks (1993) suggested, that anxiety is “an early indicator of basic language problems” (p. 199). Hence, it is likely to be a consequence rather than the cause of language difficulties. Sparks & Ganshaw argue that FL difficulties may have their source in a native language phonological difficulty which is a language-related cognitive deficit.

4.4.6. Anxiety in the classroom

Different sources of classroom anxiety have been identified. These include speaking activities in the classroom and oral presentations as the most intensive causes of anxiety (Woodrow,

2006). Young (1990) has indicated that unrealistic learner beliefs, such as beliefs about how quickly a language can be learnt, are a major contributor to apprehension and language anxiety (p. 428). Students with high levels of language anxiety tend to underestimate their ability to speak, comprehend, and write using the second language. One effect of language anxiety may be “overstudying” (Horwitz et al., 1986). It has been reported that the students who experience anxiety feel the need to compensate for its negative effects by making increased efforts to learn. Interestingly, it has been observed as well, that anxiety may sometimes induce positive outcomes (i.e. it does not always have debilitating effects on learning).

MacIntyre and Gardner (1994) have developed a means of measuring language anxiety which takes into consideration the process of language learning, including three phases:

- (1) the initial stage of L2 learning,
- (2) the processing stage, and
- (3) the output stage of language learning.

Based on this model, they developed three scales for measuring anxiety: the input anxiety scale (IAS), the processing anxiety scale (PAS) and the output anxiety scale (OAS). Having conducted a series of factor analyses to examine the psychometric properties of the MacIntyre & Gardner’s three scales, Onwuegbuzie, Bailey and Daley (2000) validated the model with 258 university students and proposed a similar three-stage model designed to measure the dynamic nature of language learning anxiety. The questionnaire they produced consists of eight items referring to the input, processing and output stages of learning. This questionnaire was used in the current study to measure students’ anxiety in all three stages of learning.

4.4.7. Anxiety and Corrective Feedback

As already pointed out, only a few studies have looked at the relationship between anxiety and CF, although anxiety has been one of the main arguments against the use of CF in the classroom. For example, Krashen (1983) argued that:

overt error correction of speech even in the best of circumstances is likely to have a negative effect on the students’ willingness to try to express themselves. Thus the question is whether or not the possible positive effects of the correction of speech errors

outweigh the intrinsic negative effects with regard to raising the affective filter (1983, p. 177).

DeKeyser's (1993) study was the first study which looked at the effects of CF in interaction with the learners' individual characteristics. The study was carried out with two classes of senior high school students learning French as FL. One class was receiving corrective feedback over the whole school year whilst the other class served as a control group and did not receive CF. The two classes had the same programme over the year and the criterion variables measured in post-tests were: written grammar, oral fluency and oral accuracy. DeKeyser administered the anxiety questionnaire based on the test developed by Gardner (1985). The results of this study suggested that individual learner differences such as aptitude, anxiety and extrinsic motivation interacted with the students' achieved proficiency "in complex ways". DeKeyser found that the students with the lowest levels of anxiety benefitted from correction only in written grammar. He also pointed out that the strong interaction between error correction and anxiety may have obscured any potential interaction with the aptitude. Havranek & Cesnik's (2001) study investigated the effects of CF on a large group of 207 participants who were asked to fill in a questionnaire measuring achievement motivation, anxiety and self-concept. Havranek's findings were somewhat "unexpected and difficult to interpret" (p. 114). She found, for example, "that high inhibiting anxiety and high promotive anxiety are both more conducive to learning from CF than low anxiety of either type" (p. 115). The reason for this, as the author speculated, may have been the fact that the standardized questionnaires used were not developed specifically for language learning situations.

More recently, Sheen's research (2008, 2011) revealed a very interesting relationship between anxiety and the effect of implicit and explicit, oral and written corrective feedback on the acquisition of English articles. As already mentioned, she conducted a study which looked at the effect of CF mediated by the impact of individual learner variables. In the case of anxiety, the most interesting finding concerned the learners' who received recasts during an oral production task. Post-tests showed that the recast group was outperformed by the explicit metalinguistic CF group. However, when the students in the recast group were divided into those with lower and higher levels of anxiety (as measured by a self-report questionnaire), it was revealed that the 'low anxiety group' outperformed the 'high anxiety' group and the control group. This group of learners performed similarly on the post-tests to those groups that received metalinguistic CF. Furthermore, the 'low anxiety group' also

produced more modified output or uptake with repair. So, oral recasts proved to be facilitative of learning but only for those learners with low language anxiety. This study is important because it shows that CF in the classroom is closely related to both learners' performance during the task and to their learning.

4.5. Summary

This chapter focused on the role of analytic ability, working memory and anxiety in L2 learning. The main findings were as follows:

- Most of the studies to date report a strong relationship between language analytic ability and the success of adult L2 learning, in both formal classroom environments and informal settings. However, there are studies which provide evidence that analytic ability is not related to L2 grammar acquisition in incidental learning that arises from the meaning-oriented tasks.
- Working memory has been found to play a major role in both L1 and L2 acquisition. There have been two main directions in research on WM related to language: one has been concerned with short-term memory and the role of the phonological loop in the acquisition of new words and sequences of words. Measures of WM in tests measuring the phonological loop have been found to be related to the acquisition of new vocabulary and syntax – but more in the initial stages of L2 learning. The second line of research has been concerned with the simultaneous processing and storage functions of WM. This capacity is considered to be critical in explaining individual differences in L2 comprehension of complex sentences as well as in reading and speaking skills. It has also been established that the task demands determine the level of the involvement of WM.
- Language anxiety has been shown to be a debilitating factor in language learning, in particular with regard to speaking in a classroom. However, as yet it is unclear whether language anxiety is the primary source of difficulties some learners experience in L2 learning, or whether it is a side-product of cognitive deficits specific to FL learning.

Based on these findings I set out to investigate whether language analytic ability, working memory and language anxiety play a role in the acquisition of L2 grammar resulting from the corrective feedback that learners receive when they perform communicative tasks.

Chapter 5. Pilot study

5.1. Introduction

This chapter reports on the methodology of the pilot study that was conducted five months prior to the main study. This small scale study enabled the piloting of the instruments for the main study, and served to indicate possible problems in the methodology. I start with the research questions and then proceed with the description of the methodology, the instructional setting, the design of the study, the participants and the instruments involving treatment and testing tasks. This is followed by an account of treatment and testing procedures including coding and scoring, and the methods of data analysis. Results for the oral and written production tests are presented in Appendix H. The chapter concludes with a brief discussion identifying the problems encountered in the pilot study, the changes for the main study decided on and the formulation of the research questions for the main study.

5.2. Research questions

1. What effect does oral CF (recasts) have on L2 learners' acquisition of French linguistic structures: a) the passé composé and b) the imparfait past tense?
2. What effect does student noticing of the target structures during the CF have on learners' acquisition of these structures?
3. To what extent do individual learner difference factors (anxiety and language analytic ability) influence learners' noticing of the target linguistic structures during the oral corrective feedback?
 - a) Does anxiety influence learners' noticing of the linguistic structures?
 - b) Does language analytic ability influence learners' noticing of the linguistic structures?
4. To what extent do individual learner difference factors mediate learners' acquisition of the two target structures?
 - a) Does anxiety mediate learners' acquisition of the target structures?
 - b) Does language analytic ability mediate learners' acquisition of the target structures?

5.3. Method

In this section the instructional setting, design of the study, participants, instruments, procedures and the analysis of data are presented. A section about the target structures is included in the chapter on methodology of the main study.

5.3.1. Instructional setting

The pilot study was conducted in two Auckland secondary schools which offer French as a foreign language at senior level. The two schools were chosen after their teachers responded to my request put on the listserv of the NZ French Teachers Association. At the time when the pilot study started the participants had had on average 590 hours of French instruction. At both schools students sit the National Certificate of Educational Achievement (NCEA) exams. French curriculum is driven by the NCEA assessments and is based on a functional syllabus. That is, the teaching content at each level of instruction is designed to enable the students to use language in order to perform different functions. The teachers at both schools claimed that their teaching practices favoured the communicative approach, but that there was also an emphasis on grammar and accuracy.

The pilot study was carried out over a period of five weeks with a two-week school holiday between week 1 and week 2 of the study. It started in the last week of Term 3 (22nd September) and had to be completed by 5th November when senior students in Auckland secondary schools went on study leave for the external exams. Prior to the beginning of the pilot study Ethics approval was obtained from the University of Auckland Human Participants Ethics Committee (Ref. 2008/360) and the consent forms were signed by the participating students, their French language teachers, principals and the Boards of Trustees of the two schools. The Ethics forms are presented in Appendix A.

5.3.2. Design

The study employed a quasi-experimental, pretest-treatment-posttest-delayed posttest design. The overall design of the pilot study involved a treatment group and a control group. The study included:

a) The execution of two types of focused tasks (one written and one oral) which were used as a pre-test. They were used in order to test the participants' level of control over the target

structures before the treatment. At both schools the target structures (passé composé and imparfait) had been introduced two years earlier so the participants already had some explicit knowledge of the target structures and presumably had developed some procedural knowledge. So, the aim of the study was to investigate what effect the instruction had on the accuracy with which the learners used two structures that they had already partially acquired.

- The performance and recording of two oral treatment tasks which were also designed as focused tasks and required interaction between the participants and the teacher/researcher. The tasks were designed to elicit the target structures. During the performance of the tasks oral corrective feedback (CF) in the form of recasts was provided by the teacher. Participants in the control group did not receive any feedback and did not complete the treatment tasks.
- The application of the stimulated recall protocol after the performance of the two treatment tasks.
- The administration of a written and oral focused task to test the participants' level of acquisition of the target structures following the treatment a day after the second treatment task.
- The administration of a written and oral focused task which were used as a delayed post-test to test the participants' level of acquisition two weeks after the treatment.
- The administration of a questionnaire to measure the participants' level of anxiety.
- The administration of a test to measure the participants' language analytic ability.

5.3.3. Participants

Participants in the study were 18 students of French Year12: 13 participants in the treatment group and 5 in the control group. The participants' mean age was 17 years (17.0 for the treatment group, 17.1 for the control group). In the treatment group all participants were females. Out of 13 students, 10 were English native speakers, 1 was a Chinese Mandarin native speaker, 1 was a Cantonese native speakers and 1 was a German native speaker. In the control group there were 3 females and 2 males. Out of 5 participants, 2 were English native speakers, 2 Chinese Mandarin and 1 was a German native speaker. All English L2 speakers had native or near native level of English, having completed all or almost all their education in New Zealand.

The imbalanced distribution of participants was due to the fact that at one school, in a class of 18 students, only seven agreed to take part (one of them turned up only once so her data were not taken into consideration) and in another class of 12 students, only five participants signed the consent form. Due to the time constraint it was not possible to find more students for the control group because that would have involved the inclusion of one more class, probably at another school, which was not manageable.

5.3.4. Instruments

The study used tasks, stimulated recall and tests as instruments.

5.3.4.1. Treatment tasks

Two information-gap focused narrative tasks were used for the purpose of the recasts treatment. The treatment tasks were picture-based and designed to elicit the use of the two past tenses in French. Each story was introduced by a sentence indicating that it happened either 'last week' or 'yesterday'. Prior to the beginning of the pilot study I trialed a large number of this type of tasks with my own students of French Year 12 and 13. Students' oral and written performance was observed and analysed. Moreover, two independent French native speakers were asked to retell the stories, to establish that the tasks would elicit the use of the passé composé and the imparfait past tense.

These tasks were two-way focused information gap tasks (R. Ellis, 2003). In order to complete them, i.e. to retell the story shown in a series of six pictures, the participants first had to find out information that was missing from their set of pictures. The sheets they were given contained three pictures showing what happened in the story, presented in the form of a jig-saw. The teacher provided information about the other three pictures. Once the students obtained all information needed to tell the story, they had to retell the whole story on their own, taking turns on a voluntary basis. New vocabulary related to each of the two picture-based tasks was introduced at the beginning and put on the whiteboard for reference. The students were also allowed to ask for the meaning of any French word they did not know. Each treatment task lasted for 20-25 minutes. During the performance of the tasks participants received feedback in the form of recasts. All errors other than the two target structures were ignored. The tasks were audio recorded by a digital voice recorder.

5.3.4.2. Stimulated recall protocol

Stimulated recall protocol was used in order to measure the learners' noticing during the treatment tasks. Listening to the audio-recording of interactions immediately after the task served to prompt the participants to recall the thoughts they had while performing the tasks. Stimulated recall is an introspective method in which learners are asked to explore and articulate their thoughts referring to the task after the task has been completed. Gass and Mackey (2000) suggested that stimulated recall be carried out with some support, for example showing a videotape or playing an audio-recording, so that the learners can watch or listen to themselves and vocalize what their thought processes were at the time of the original activity. In this study the stimulated recall followed the performance of the tasks. Replaying the audio-recording of the oral interaction served to prompt the participants to recall any thoughts they had while performing the tasks.

5.3.4.3. Tests

Tests for target structures

For the purpose of pre-testing, post-testing and delayed post-testing picture-based focused narrative tasks were used. Each test consisted of six pictures, requiring learners to retell/narrate the story either orally or in written form. From pictures it was possible to make up a story involving the plot and foreground, and background information. Three different story-based tests were designed for the purpose of oral testing and another three for the purpose of written testing. In order to account for possible differences in the difficulty of the pre-test, post-test and a delayed post-test, the tests were counter-balanced using a three-split block design. Thus on each occasion participants were split into three groups and each group used a different story so that those who had story A for the pre-test had story B for the post-test and story C for the delayed post-test while those who had B for pre-test, later had C for the post-test and A for the delayed post-test, etc. Obligatory occasions were then determined for each structure taking into consideration the context and the adverbs or connectives between two or more verbs or phrases. The following are some examples:

Example 1 : *Il a descendu (passé composé) l'escalier très vite et il est tombé. (passé c.)*

[he went down the stairs very quickly and he fell]

but:

Il est descendu (passé composé) très vite et il est tombé. (passé composé)

[he went down very quickly and he fell]

but :

Il descendait (imparfait) l'escalier quand il est tombé. (passé c)

[he was going down the stairs when he fell]

Example 2: *Les garçons ont vu une dame qui se faisait bronzer (imparfait)*

[the boys saw a lady who was sunbathing]

but:

Elle s'est fait bronzer hier après-midi (passé composé)

[she was sunbathing/ she got sunbathed yesterday afternoon]

The materials used for testing can be found in Appendices C and D.

Tests for individual differences

a) Language analytic ability

In order to measure the participants' language analytic ability a test designed by Ottó (2002) and used by Schmitt, Dörnyei, Adolphs and Durow (2004) was administered. The test, consisting of 14 multiple choice items, is based on an artificial language and requires the participants to analyse the grammatical markers supplied in a glossary, deduce the rule and select the correct sentences. The language analytic ability test was scored on a discrete item basis so that one point was given for each correct answer. The total score was 14 points and the final scores were calculated as percentages. The language analysis test (LAT) is presented in Appendix E.

b) Language learning anxiety

In order to measure language learning anxiety, the questionnaire developed by MacIntyre and Gardner (1994) and validated on 258 participants by Onwuegbuzie, Bailey and Daley (2000), was used. The questionnaire, consisting of 8 items, contains two items from MacIntyre and Gardner's input anxiety scale, three items from their processing anxiety scale and three items from the output anxiety scale. Scores on the questionnaire were coded on a 6-point Likert scale so that the lowest score for each question was 1 and the highest was 6. As there were eight items of the questionnaire, the scores for each question were added up and divided by 8 in order to obtain a mean score for each participant. The anxiety questionnaire is presented in Appendix G.

5.3.5. Procedures

5.3.5.1. Treatment

Each treatment task lasted for 20-25 minutes. Each picture-based story gave opportunities for elicitation of obligatory contexts for the use of, approximately, 15 different verbs in perfect form (*passé composé*) and 10 verbs in the imperfect form (*imparfait*). During the execution of the two tasks (one each week) the students received treatment in the form of recasts. The participants in the control group did not receive any treatment and they did not complete the information-gap tasks. They continued with their regular lessons.

Noticing was operationalised as students' accurate perception / recognition of corrective feedback in the interaction and the immediate recall of the target form. Therefore, the instances of uptake with repair were identified as evidence of noticing: that is, those instances in which students showed that they had noticed the target form in the recast and were able to recall it immediately. Uptake with repair episodes were scored and calculated as a percentage of the total recast episodes.

The following are some examples of recasts and the students' responses showing either uptake with repair or absence of uptake with repair or absence of any uptake:

a) recasts followed by uptake with repair:

Example 3:

S9: ...*parce que la voiture est renversée* (reflexive pronoun missing –pronominal verb)

[...because the car is overturned]

T: *s'est renversée?*

[overturned ?]

S9 : *s'est renversée*

[overturned]

Example 4:

S6: *et cette année ils décidé/er d'aller...* (auxiliary missing)

[and this year they decide to go....]

T: *ils ont décidé ?*

[they decided ?]

S6: *ils ont décidé d'aller à un lac et...*

[they decided to go to a lake...]

Example 5:

S6: *et ils prepare/er ils sont préparé* (auxiliary first missing, then the incorrect auxiliary supplied)

[and they prepare they are prepare]

T: *ils ont préparé ?*

[they prepared ?]

S6: *oui ils ont préparé ah ah toutes les choses oui...*

[yes they prepared ah ah all the things yes....]

Example 6 :

S10 : *parce que ... je ne sais pas...parce que Philippe ne donnait/donné/er pas la bouteille à Paul* (auxiliary missing)

[because... I don't know...because Philippe not give]

T : *n'a pas donné ?*

[did not give]

S10 : *n'a pas donné la bouteille*

[did not give the bottle]

Example 7:

S13: *il faisait beau et il y a beaucoup d'arbres* (verb 'avoir' used in present)

[the weather was nice and there are many trees]

T: *il y avait beaucoup d'arbres?*

[there were many trees ?]

S13: *ummm oui*

[ummm yes]

T: *il y avait?*

[there were ?]

S13: *il y avait beaucoup d'arbres...*

[there were many trees...]

b) recast followed by absence of uptake/uptake with repair:

Example 8:

S4: *et il fait du soleil* (present tense used)

[and it is sunny]

T: *il faisait du soleil?*

[it twas sunny ?]

S4: *oui d'accord et la voiture...*

[yes all right and the car...]

Example 9:

S1: *il y a un petit forêt un petit* (present tense used)

[there is a little forest a little]

T: *il y avait un forêt?*

[there was a little forest?]

S1: *oui et aussi une tente*

[yes and also a tent]

Example 10:

S4: *Paul perdu l'équilibre* (auxiliary missing)

[Paul lose his balance]

T: *Paul a perdu léquilibre?*

[Paul lost his balance ?]

S4: *eh oui et er et tombé/er sur la rue*

[eh yes and er er fall on the street]

5.3.5.2. Stimulated recall

Following the treatment tasks, the stimulated recall was administered to provide an additional measure of noticing. The aim was also to see if the recasts had induced anxiety in the participating students. While listening to the recording of the tasks I stopped the recorder at episodes involving recasts and asked the student four questions in the following order (if the answer was given to question 1 then there was no need for question 2, and sometimes there was no need for question 3):

- 1) What did you think here?/ or: Can you remember what you thought here?
- 2) Did you notice anything?
- 3) Why do you think I interrupted you/her? or: Why do you think I said that?
- 4) How did you feel about it/or: when I interrupted you /her?

However, during the administration of the stimulated recall there were some logistical problems. First, the conditions for stimulated recall were not adequate: I had to administer it in a classroom and there was often too much noise. The participating students often did not concentrate enough and so failed to make the effort needed to listen and focus on the recording (although they were generally cooperative). Also, because of the time constraint it was not possible for the stimulated recall to be administered as planned, with each student listening to the whole task. Because of the problems with the administration of the stimulated recall the data that were collected were not analysed.

5.3.6. Testing

All participants were tested on their use of the two target structures (the imparfait and the passé composé) separately. They were asked to retell the story presented on a set of six pictures, both in writing and orally. Three different stories were used for the written tests and another three stories for the oral tests, following the procedure as described in the design of the study. The teacher/researcher administered all the tasks. The oral tests were audio recorded by a digital voice recorder and transcribed by the researcher.

For all the written tests the participants were given the following instruction: ‘You saw this event yesterday/two days ago/ and you need to write a report about it for your school magazine’. For the oral tests they were given the following instruction: ‘You need to report on this event for your school broadcasting programme’. Additionally, there was a sentence introducing the story and setting it in the past. The participants were asked to use at least two sentences for each picture and to include as much descriptive information as possible (e. g. description of the weather, surroundings, etc). They were given a list of words that may have been new to them, including verbs to help them tell the story. The participants were allowed to ask for the French translation of any word they wanted to use, including verbs, which were presented to them in the infinitive form.

For each student there were from 5 to 14 obligatory contexts for the passé composé and from 4 to 14 for the imparfait, with an average of 9 obligatory occasions established for the passé composé and 8 obligatory occasions for the imparfait per one 6-picture story.

5.3.6.1. Coding and scoring

Written and oral tests were scored separately for each tense so there were four groups of data obtained from the tests:

- a) data on oral passé composé
- b) data on oral imparfait
- c) data on written passé composé
- d) data on written imparfait

Scores on the tests were calculated using the target-like use analysis (Pica, 1983). This takes into consideration the overuse of the target form. For each target-like use in an obligatory context including the correct form participants were given 2 points (1 point for correct usage, 1 point for correct form). First, the number of obligatory contexts was established for each tense separately (the imparfait and the passé composé) for each narrative. This number then became the denominator, which was multiplied by 2. If there was an oversuppliance of a tense used, then this number, multiplied by 2, was added to the denominator. The number of scores obtained for each correct suppliance in obligatory contexts, as described above, was used as a numerator and multiplied by 100 in order to obtain the percentage of accurate obligatory use.

The formula used to obtain percentage scores is presented below:

$$\frac{\mathbf{n} \text{ correct suppliance in contexts (including correct form)}}{\mathbf{n} \text{ obligatory contexts} \times 2 + \text{oversuppliance in non-obligatory contexts} \times 2} \times 100 = \% \text{ accuracy}$$

Coding and scoring of written data

Using the above formula, written data were coded so that each tense structure was taken as a composite of two parts (passé composé being composed of an auxiliary and the past participle, and imparfait being a composite of a stem and the ending).

The following is an example of the coding for the passé composé:

Il est sorti (passé composé in appropriate context, correct form) = 2 points

Il a sorti (correct use, partly inappropriate form-inaccurate auxiliary) = 1.5 points

Il est sorté (correct use, partly inappropriate form-inaccurate past participle) = 1.5 points

Il a sorté (correct use, inappropriate form in both auxiliary use and past part.) = 1 point

The following is an example of coding the imparfait of the verb *peindre* (to paint):

Il peignait (imparfait in appropriate context, correct form) = 2 points

Il peignais (partly inappropriate form: inaccurate ending) = 1.5 points

Il peindrait (partly inappropriate form: inaccurate stem) = 1.5 points

Il peindra (inappropriate form: both stem and ending inaccurate) = 1 point

Coding and scoring of spoken data

The same method as in writing was used for coding the passé composé in oral tests. However, when coding the imparfait in oral tests it was not possible to distinguish between the endings *-ais* and *-ait* as they sound the same. The following is an example of coding the imparfait in oral tests:

Il peignait /s (imparfait in obligatory context, correct form) = 2 points

Il peindrait/s (correct use, partly inappropriate form: inaccurate stem) = 1.5 points

Transcription

Oral data were transcribed using a broad transcription so that the time spent on producing an utterance or the length of pauses was not measured. Also, the difference, for example, between an ‘open e’ and a ‘closed e’ in French was ignored - because it is very difficult if not impossible for learners to achieve this difference in pronunciation. This is important since in the learners’ utterances containing imparfait of the verbs ending in *-er* it is not possible to distinguish the infinitive (*parler*), imperfect (*parlait*) and past participle (*parlé*). Therefore, when coding the data I compared the spoken data with the written data of each participant in order to see if the participant had used any imperfect verb tense in writing (other than the imperfect forms of *avoir* and *être*). If the participant had not used any imparfait form in the written tasks it was an indication that they had used the infinitive form of the verb given to

them in the oral tasks. Therefore such examples were not coded as imparfait and were not given any points.

Repairs

When coding a distinction was made between a delayed and an immediate repair. If it was a delayed repair (preceded by hesitation) no points were given.

For example:

... *et les infirmiers viennent er er vien err sont venus* (0 points)

[...and nurses come er er com are come]

or: ...*le taxi attend ...uhmm le taxi ...uh était devant la maison* (0 points)

[... the taxi wait...uhmm the taxi ...uh was in front of the house]

However, if it was an immediate repair, and the utterance spoken rapidly, I coded it for correct target-like use, for example:

il vu il a vu (2 points) – the learner immediately self-corrected using the auxiliary,
without hesitation (PC of verb *voir*- to see)

il a tom il est tombé (2 points) – self-correction of the auxiliary before articulation of
the verb is completed (PC of verb *tomber* – to fall)

il a il est venu (2 points) – self-correction of the auxiliary before the articulation of
the main verb (PC of verb *venir* – to come)

In such examples it was clear that the student self-corrected her/his erroneous output before it was fully articulated.

5.3.7. Data analysis

The SPSS package was used to perform the statistical analyses. The following analyses were used to answer the research questions: first, descriptive statistics for the four tests of the target structures (oral passé composé, oral imparfait, written passé composé, written imparfait) were computed, followed by the descriptive statistics for language analytic ability test and language learning anxiety.

An Independent samples t-test was used to examine if there were any significant statistical differences between the two groups (Recasts and Control) on the pre-tests. In order to answer Research question 1) the mixed design factorial ANOVA with repeated measures was computed for each target structure and for oral and written tests separately. These analyses were performed in order to investigate if there was an effect of the treatment involving recasts, by comparing the two groups on the immediate and on a delayed post-test. Alpha level for p-value was set at $p < .05$.

In order to answer Research question 2) Pearson correlation coefficients were calculated to see if the scores obtained by measuring student noticing of errors as evidenced in uptake with repair (separately for each tense) correlated with the scores obtained on the post-tests and delayed post-tests. In order to answer Research question 3) Pearson correlation was computed for the treatment (recast) group, to establish if there was a relationship between the ID factors and student noticing of the target structures as evidenced in uptake with repair. In order to answer Research question 4) Pearson correlation coefficients were computed to examine if there was a relationship between the individual difference factors and the scores obtained on all posttests.

5.4. Summary of main findings

Detailed results are not presented here as the purpose of the pilot study was to establish whether the design and the materials used in the study were satisfactory. Readers interested in the results are referred to Appendix H. However, a brief summary of the main findings is provided below.

Research question 1:

The results of mixed design factorial ANOVA with Repeated measures did not reveal any significant differences between the treatment group and the control group for any of the structures under investigation (oral PC, oral IMP, written PC, written IMP). On the pre-test for oral IMP there was a significant difference between the two groups, so an ANCOVA using pre-test scores as a covariate was employed, and t-tests used to examine group differences on the immediate and the delayed post-test separately.

Research question 2:

A strong, significant correlation was found between uptake with repair and both short-term and long-term gain scores (on the immediate and the delayed post-test) for oral PC and the short-term gain scores for written PC. No other correlations between uptake and gain scores were apparent.

Research question 3:

No significant relationship was found between uptake with repair and ID factors.

Research question 4:

A statistically significant negative correlation ($r = -.800$, $p < .05$) was found between anxiety and long-term gain scores for oral IMP in the treatment group. No such correlation was evident in the control group.

5.4.1. Problems encountered in the pilot study and changes for the main study

The pilot study confirmed that the type of tasks and the materials used in treatment and tests were appropriate for the current developmental level of participating students, and that the chosen tasks could elicit a large number of the IMP and the PC structures. The tests used to measure anxiety and analytic ability were also found to be appropriate and easy to manage. However, a number of problems emerged and are considered below.

Problem and solution 1: Stimulated recall

There was no suitable location to administer the stimulated recall and generally it is very difficult to find one in a secondary school. Also, the age of the students (teenagers) made the administration of the stimulated recall difficult and quite possibly unreliable. For these reasons, the data collected by means of the stimulated recall may not have been reliable. Reluctantly, I will need to omit the stimulated recall from the main study.

Problem and solution 2: Insufficient sample size

The obvious problem of this study was the small and imbalanced number of participants (particularly, the control group with only five participants). I was aware of the problem at the

time when the pilot study was carried out but there was not enough time to work with one more group of students to make the control group larger and more balanced. As stated earlier, there was the issue of time constraint at the very end of the school year, as well as an unexpectedly small number of students who agreed to take part in the study. Efforts will be made to ensure that the sample size for the main study is larger and sufficient to carry out statistical analyses.

Problem and solution 3: Difference in proficiency

The groups of students participating in the pilot study were not of the same proficiency level. This became evident in the significant difference on the pre-test for oral IMP. This problem could be resolved in the main study by choosing participants carefully and balancing the groups at the beginning. Every effort will be made to ensure the participating groups of students are equivalent in terms of their knowledge of the target structures at the beginning of the study.

Problem and solution 4: Scheduling the study

Inconvenient timing at the end of the school year, after the school exams and before the external exams, was an issue which greatly impacted on the results of the pilot study. Although the control group did not receive any treatment, they studied intensively while preparing for the exams with their teacher. The revision of past tenses was one of the most important tasks in their agenda for the exam preparation. To address this problem I decided the main study should be carried out soon after the beginning of the next school year so that there is enough time to complete all the tasks and tests.

Problem and solution 5: Scoring the passé composé and the imparfait

The different procedure for scoring the written and oral tests is probably unnecessary. Moreover, the distinction between the written endings of IMP tense (*-ais* and *-ait*) is actually not relevant to the research questions of the current study as it does not show the effect of the oral corrective treatment. The oral production of French IMP forms in learners' interlanguage is, in fact, very difficult, almost impossible to measure with 100 per cent accuracy, at least for the verbs ending in *-er*. How do we know whether a learner has used an imparfait, an infinitive, or a past participle (omitting the auxiliary in the use of the PC) when they all sound the same? In order to address these problems, I decided to revise the coding

and scoring procedure for the main study. Also, related to the issue of the IMP in oral production, I decided to limit the analysis to only those verbs that could be clearly distinguished in oral production. Such a decision is supported by research evidence suggesting that at the beginning and intermediate stages of learning L2 French a very limited number of verbs (only of state and activity lexical aspect) are used in the imparfait.

I also decided that for the main study one more treatment group should be introduced – a group that would receive treatment in the form of requests for clarification. Additionally, working memory as an important individual difference factor would be included in the design of the study.

5.4.2. Changes to research questions

Based on the findings of the pilot study and the problems encountered, I decided to revise the research questions for the main study.

The question related to the effectiveness of oral CF, will not undergo changes. The main study will examine the effectiveness of two implicit CF strategies: recasts and clarification requests. Their effectiveness will be investigated for oral production and written production separately.

Since previous research has shown that the effectiveness of CF may be mediated by the learners' proficiency level I decided to introduce a question that will look at the mediating role of proficiency on the effectiveness of CF.

A question relating to the extent of repair following CF will be introduced in the main study. Even though the frequency of repair following CF was recorded and analysed in the pilot study it was not included as a separate research question. A question relating to the relationship between uptake with repair and acquisition gains will be introduced in the main study. Since stimulated recall has been abandoned, *noticing* of corrective feedback will be operationalized in terms of *uptake with repair*.

The question related to the role of individual differences will investigate the predictive power of analytic ability, working memory and anxiety for the acquisition of the target structures in the two treatment groups.

The research questions for the main study are presented at the beginning of the next chapter which describes the methodology of the main study.

Chapter 6. Methodology

6.1. Introduction

This chapter presents the research design of the main study and the methods used to collect the data. The study adopted a quantitative approach, using a pre-test – treatment – post-test – delayed post-test design. As a classroom study it can be classified as quasi-experimental research since it aimed at “comparing group behaviour in probabilistic terms under controlled conditions using intact groups” (Brown & Rodgers, 2002, p. 12). Brown and Rodgers also emphasize that many researchers think that conducting an experiment is the most appropriate way to investigate a question about language learning (p. 195). That is, only when there is a pre-test and a post-test following the treatment it is possible to show what the learners have learnt, i. e. to see if any change has occurred between the pre-test and the post-test.

I begin with the research questions for the main study and proceed with the description of the overall design. Detailed information about the participants and the research site is provided. This is followed by a brief presentation and a discussion on the two target structures in the focus of this study. Further to this, the instruments used in the study are described in detail. The chapter concludes with a detailed description of the procedures used for data collection.

6.2. Research questions

1. Does oral corrective feedback (recasts and requests for clarification) have an effect on L2 learners’ acquisition of French linguistic structures – the passé composé and the imparfait – as manifested in oral production?

a) Do recasts have an effect on the L2 acquisition of passé composé and imparfait?

b) Do requests for clarification have an effect on the L2 acquisition of passé composé and imparfait?

2. Does the learners’ starting proficiency mediate the effect of the oral corrective feedback on the learners’ acquisition of the passé composé and the imparfait, as manifested in oral production?

3. Does oral corrective feedback (recasts and requests for clarification) have an effect on L2 learners' acquisition of French linguistic structures – the passé composé and the imparfait – as manifested in written production?

a) Do recasts have an effect on the L2 acquisition of passé composé and imparfait?

b) Do requests for clarification have an effect on the L2 acquisition of passé composé and imparfait?

4. Does the learners' starting proficiency mediate the effect of the oral corrective feedback on the learners' acquisition of the passé composé and the imparfait, as manifested in written production?

5. a) To what extent do learners repair their errors, following correcting feedback?

b) Do learners who produce uptake with repair demonstrate greater gains in accuracy than learners who were just exposed to uptake with repair?

6. Do individual difference factors predict the learners' acquisition of the target structures in the two experimental groups?

a) Does language analytic ability predict the acquisition of the target structures in (1) the Recast group and (2) the Clarification request group?

b) Does working memory predict the acquisition of the target structures in (1) the Recast group and (2) the Clarification request group?

c) Does language anxiety predict the acquisition of the target structures in (1) the Recast group and (2) the Clarification request group?

6.3. Design of the study

The design of this study, as already established in the pilot study, is best described as quasi experimental with a pre-test, treatment, post-test and a delayed post-test. The overall design of the main study involved two treatment groups and a control group. One treatment group received recasts and the other group received clarification requests as corrective feedback during the performance of the treatment tasks. The control group did not receive any treatment and did not perform the tasks the other two groups did. The data collection started at the beginning of March 2009 (week 5 of Term one) and the entire process of data collection lasted for 14 weeks, including a two week school holiday between Term one and

Term two of the school year. Only the delayed post-test and the administration of the working memory test were administered in Term two.

The figure below presents the design of the main study:

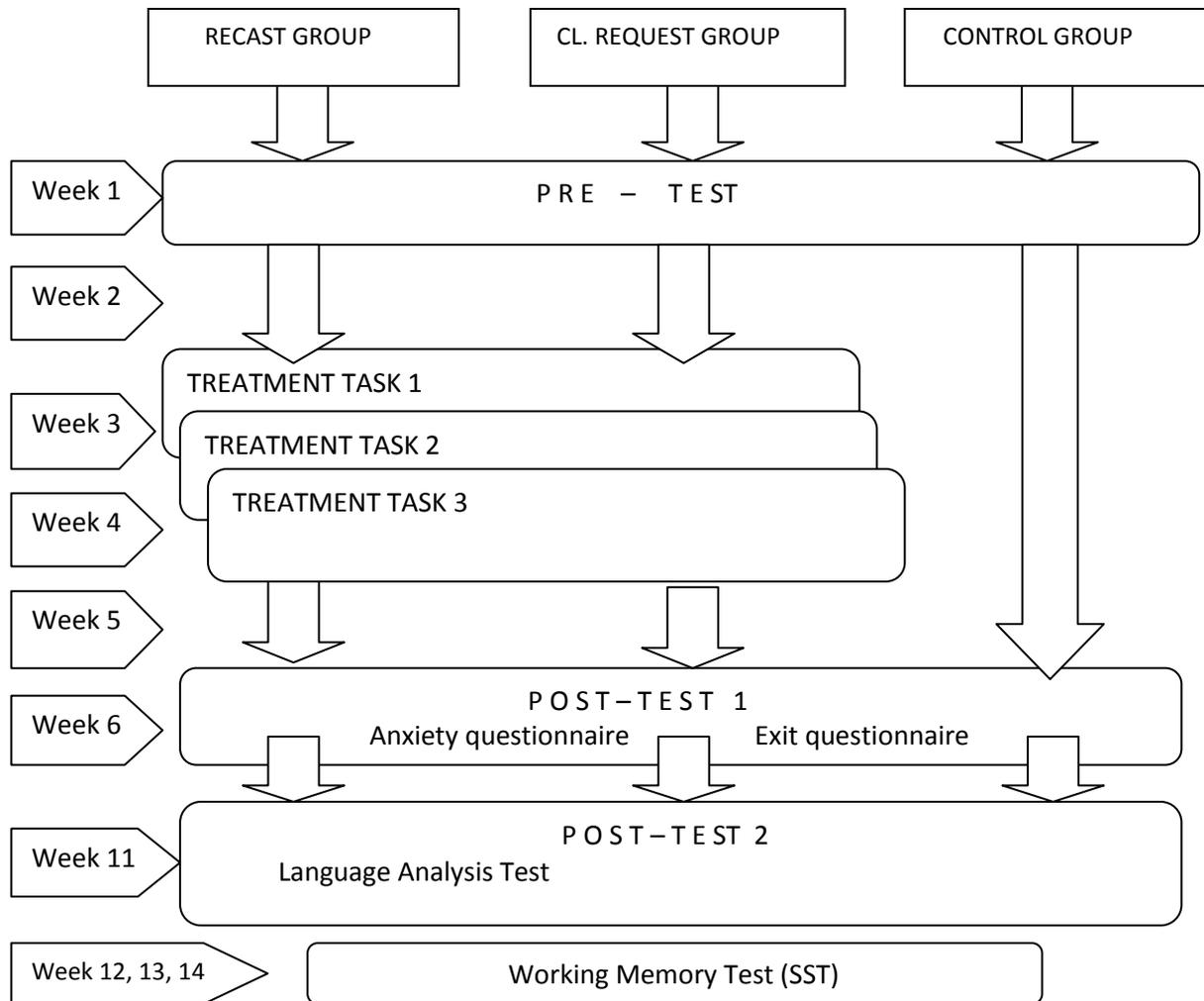


Figure 1 Design of the main study

6.4. Participants and research site

The participants in the study were recruited from the secondary school student population, i. e. students who chose to learn French as part of their high school curriculum. Each of the three groups of students in the study involved intact French language classes. Both treatment groups were classes from the same school and the control group students were from another school. Both treatment groups were taught by the same teacher and the control group students

were taught by a different teacher. In one treatment group the students were from Year 11 (an ‘advanced’ Year 11, according to their teacher) and the other group involved students from both Year 11 and 12 (a ‘low’ year 12 according to their teacher). The control group was also a mixed class of students from Years 12 and 11.

Altogether 52 students were involved: 18 in each of the treatment groups and 16 in the control group. Their average age was 16 (15.76). In the Recast (RE) group there were 9 males and 9 females, while in the Clarification Request (CR) group there were 5 males and 13 females. In the Control (CN) group there were 5 males and 11 females. All three groups involved speakers whose native language was not English. However, they all reported English native or near-native competency since all their schooling had been in an English speaking environment. In the RE group 11 students were English native speakers, one was bilingual native English/German, two were German native speakers, two Korean, one Chinese and one a Filipino Tagalog language native speaker. In the CR group there were ten English native speakers, four Chinese, two Korean, one German and one Romanian native speaker. In the CN group 14 participants were English native speakers, one was German and one was a native speaker of Serbian.

On average they had had around 500 hours of French learning to date. The majority of the participants started learning French in Year 9 but some of them started in Year 7 or 8 (at Intermediate school). At both schools students sit the National Certificate of Educational Achievement (NCEA) exams. French is an optional subject and the French curriculum is driven by the NCEA assessments. It is based on a functional syllabus, as explained in the chapter reporting the pilot study. The teachers at both schools claimed that their teaching practices favoured the communicative approach, but that there was also an emphasis on grammar and accuracy. The target structures (the passé composé and the imparfait of some verbs) had been introduced about a year or a year and a half earlier in both of the schools, so the participants already had some explicit knowledge of the target structures and presumably had also developed some procedural knowledge.

The school from which the treatment groups were recruited is a large multicultural, decile 10 school with a roll of around 3100 students. This school runs a 6-day cycle and each period lasts for 45 minutes. Students at this school start learning French as a foreign language in Year 9, receiving instruction every day for one 45 minute period. At Year 10 they continue with seven 45 minute periods over six days. The same frequency of foreign language

instruction is followed at senior level (Years 11, 12 and 13). Moreover, senior level students at this school have either individual or small group (2-3) regular meetings once a week (for about 20 minutes) with a French native language assistant in order to develop their fluency. The school from which the control group was recruited is also a large multicultural, decile 7 school, with a roll of around 2500 students. At this school French and German are compulsory languages at Year 9 (each language for half a year). Starting from year 10 French can be chosen as an optional subject, with three one-hour periods per week. At senior level there are four one-hour (60 minute) periods of French per week. Language classes at this school are relatively small at senior level, so usually there are combined classes including students of different levels. Generally, at both participating schools flexibility in taking classes at different levels (depending on students' abilities and preferences) is one of the characteristics of their curriculum.

6.5. Target structures

The two target structures that are the focus of this study have been chosen partly because they are considered some of the most problematic grammar features for French L2 learners (Bardovi-Harlig, 2000; Harley, 1989; Harley & King, 1989). However, they are used very frequently in everyday conversation related to past events and past activities. The difficulties in the acquisition of these two past tenses in French arise from both their formal complexity and their functional complexity, i.e. they have multiple semantic meanings (discourse functions). The formal complexity of the French *passé composé* and the *imparfait* lies in both the nature of the rules governing their form and also because they involve numerous items that have to be learnt. The *passé composé* employs two different auxiliaries (*avoir* and *être*) in a structure consisting of the auxiliary and the past participle which requires agreement with the subject according to relatively complex grammar rules. However, the *passé composé* has one basic meaning: perfective or completed action. On the other hand, *imparfait* has three semantic meanings:

- a) imperfective (ongoing action in the past with out-of-focus endpoints)
- b) iterative (habitual events and repeated events in the past, and
- c) durative (states of being in the past)

The following table, adapted from Katz and Blyth (2007) presents the basic meanings of the French *passé composé* and the *imparfait*, compared to their English meanings.

Table 2 Basic meaning of past tenses, involving aspect, in French and English

Tense	Meaning	Example
passé composé	perfective (completed)	J'ai lu... <i>I read, I did read, I have read</i>
Imparfait	imperfective (ongoing)	Je lisais... <i>I was reading</i>
	iterative (habitual)	Le samedi, je lisais <i>On Saturdays, I would/used to read</i>
	durative (state)	J'étais fatigué(e) <i>I was tired</i>

From the above table it is clear that the meanings of past narratives in French, involving both tense and aspect, do not correspond entirely to English past narratives. While the tense distinctions usually do not present particular problems for English speakers learning French, the aspectual distinction between English and French is much less transparent since the French and the English aspectual systems display some similarities but also some crucial differences (Ayoun & Salaberry, 2005). Whereas the tense is concerned with 'placing events on a time line', the aspect refers to the distinction among the kinds of action that are described by verbs. The French language has two aspects, perfective and imperfective, whereas English distinguishes simple and progressive aspect.

As far as aspect is concerned, briefly, linguists typically identify four different classes of verbs according to their lexical aspect, following the classification proposed by Vendler (1967) and Comrie (1976). The four lexical aspects are: a) states, b) activities, c) accomplishments and d) achievements. It has been found that there is a strong correlation between lexical aspect and grammatical aspect, so that the verbs encoding states and activities typically involve the use of the imperfective aspect (imparfait) whereas verbs encoding accomplishments and achievements typically involve the use of the perfective aspect (passé composé). Furthermore, these distinctions are reflected in the use of the two tenses when narrating past events: the passé composé is usually employed when referring to the foreground (le premier plan) or expressing the plot (line of events) whereas the imparfait is used with reference to the background (l'arrière-plan) in a narrative.

A particular difficulty for English learners of French is the fact that the strategy of matching an English form to a French form may sometimes result in a correct target structure use but can also lead to an overgeneralisation and erroneous equation of the English simple past tense with the French passé composé. Another difficulty is that the difference between these two tenses is sometimes difficult to perceive (for example: j'ai regardé vs. je regardais). Moreover, in a foreign language classroom past narrative discourse does not typically occur often so the students usually have little opportunity to contrast the passé composé and the imparfait.

The research so far has found that, for example, immersion students in Canada, even when achieving communicative competence, were struggling with the appropriate use of particular language features such as narrative past tenses which employ the notion of the aspectual distinction: perfect vs. imperfect /completeness vs. incompleteness. Bardovi-Harlig (2000, p. 112) proposed four principles based on the longitudinal research into the acquisition of tense-aspect morphology. These are: 1. the development of temporal expression is slow and gradual, 2. the acquisition of form often precedes the function, 3. irregular morphology precedes regular morphology and 4. learners tend to avoid discontinuous marking such as aux. + V-tense inflection and instead use only suffixed inflections. Harley's (1989) study suggested that learners first start to use the passé composé, in certain cases not yet entirely productive and then imparfait but only with *être*, *avoir* and modals. Kaplan (1987) conducted a study with 16 English speakers learning French, based on a semi-structured interview, and found that first year students had more difficulty with the accuracy of passé composé forms than with the imparfait forms (in a limited number), while second year students manifested errors in distribution of the imparfait – the use of the imparfait made 82% of all errors.

6.6. Instruments for data collection

The treatment tasks and the tests are described in detail in the pilot study chapter. One more treatment task was introduced in the main study, so while there were two treatment tasks in the pilot study, the main study involved three tasks.

The tests administered in the main study were the same as the tests in the pilot study. The materials used for testing can be found in Appendix C (oral tests) and Appendix D (written tests).

The tests for individual differences are presented in Appendices E, F and G. Three ID factors were involved in the main study: language analytic ability (LAA), working memory (WM) and anxiety. The tests for language analytic ability and anxiety are described in the pilot study chapter. The WM test was used only in the main study. This was the Speaking span test, based on Daneman (1991) and Daneman and Green (1986), and adapted from Mizera (2006).

This test consists of a set of 65 common English words divided into five sets, each comprising five sub-sets (starting from two up to five words). Participants work in front of a computer and are tested individually, one by one. They view sets of words, one word at a time. Each word is seen for 1 second. After the last word of each set, participants must produce original sentences for each word in the set, saying these aloud. For example, if on the computer screen the words ‘fingers’ and ‘noticed’ appear, when they disappear the student has to produce two unrelated sentences each containing one of these two words (for example, “She has long fingers” and “I noticed him in the crowd”) . If three words appear on the screen, for example ‘machine’, ‘results’, ‘compass’, then the participant has to produce three sentences, each containing one of these three words. And so forth, up to the set of five words which require the participant to make up five sentences each containing one of the words listed. The test allows for calculation of a *strict score* and a *lenient score*. The strict score is calculated if all sentences are produced in the same order as the words appear on the screen and in the same form as the words on the screen (e.g. the nouns in plural, or singular). In the current study the strict score was used for the analyses because it is considered to be a more stringent measure of working memory capacity.

6.7. Procedures for data collection

This section is concerned with the procedures for data collection and the coding and scoring of the data. Since the coding and scoring systems were revised for the main study I hereby present a detailed description and explanation of the methods used.

6.7.1. Treatment

6.7.1.1. Coding of CF episodes

In the current study Lyster and Ranta’s (1997) definition of an error treatment sequence was modified. As Lyster and Ranta explained, an error treatment sequence consists of a trigger

utterance (an utterance containing an error), the teacher’s corrective feedback move (recast or request for clarification in this case) and either the learner’s immediate response to the feedback (uptake) or the topic continuation either by the teacher or by the student. In the current study, if a CF episode was not followed by uptake with repair it was repeated. Only then, after the second opportunity given to students for noticing and repair, was the topic continued either by the student or by the teacher.

The following chart presents the error treatment sequence in the present study:

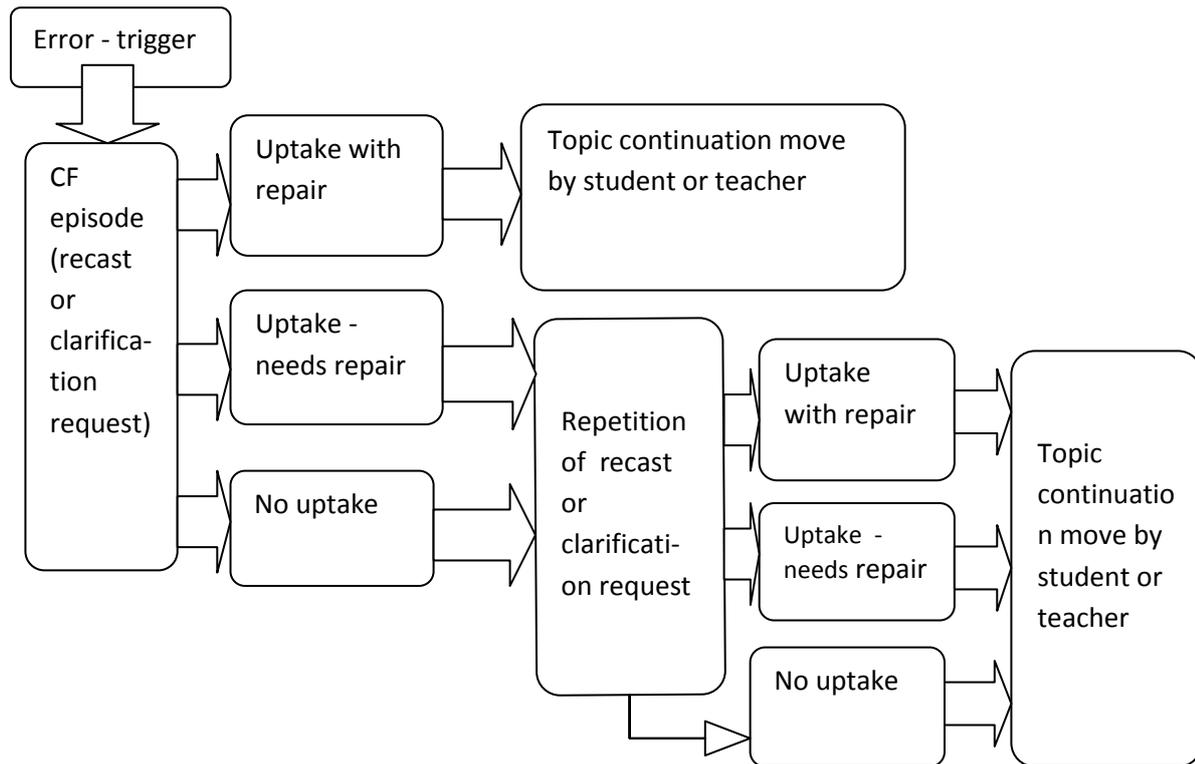


Figure 2 Error treatment sequence in the present study

Recasts were operationalised, following Long’s definition of corrective recasts (2007, p. 77), as the teacher’s reformulation of all or part of a learner’s utterance in which one or more nontarget-like grammatical items is/are replaced by the corresponding target language form and where, throughout the exchange the focus of the interlocutors is on meaning. The recasts were delivered in a rising tone.

The definition of *clarification requests* has been adopted from Lyster and Ranta (1997, p. 47) who used the term introduced by Long (1983b, p. 137), indicating feedback which shows students “that their utterance has been misunderstood by the teacher or that the utterance is ill-formed in some way that a repetition or a reformulation is required”. According to Lyster

and Ranta, this is a type of corrective feedback that can refer to problems in either comprehensibility or accuracy or both.

Uptake was operationalised as “a student’s utterance that immediately follows the teacher’s feedback and that constitutes a reaction in some way to the teacher’s intention to draw attention to some aspect of the student’s initial utterance” (Lyster & Ranta, 1997, p. 49). *Uptake with repair* was operationalised as a learner’s utterance that immediately follows the teacher’s feedback and contains repair of the error which initially triggered the teacher’s feedback.

Examples of recasts

Recasts followed by uptake with repair:

Example 1:

S2: *ummm les deux garçons un garçon acheté et* (auxiliary missing)

[ummm the two boys a boy buy and]

T: *ils ont acheté ?* (recast)

[they bought ?]

S2: *ils ont acheté une coca*

[they bought a coke]

T: *bien bien*

[good good]

S2: *et et ils ont acheté seulement un coca*

[and and they bought only one coke]

Example 2

S6: *..... boisson.... et les deux garçons s’est disputé* (inaccurate auxiliary, sing)

[...drink... and the two boys argues]

T: *ils se sont disputés ?* (recast)

[they argued ?]

S6: *oui ils se sont disputés ummmm parce que le garçon n’est umm n’est donné*

[yes they argued ummm because the boy is not umm is not given]

T : n'a pas donné ? (recast)

[did not give ?]

S6 : n'a pas donné la bouteille à à le autre garçon uhmmm ils s'est battre

[did not give the bottle to to the other boy uhmmm they to fight]

T : ils se sont battus ?

[they fought ?]

S6 : ils se sont battus ummm et... je sais pas ooh uh

[they fought ummm and ... I don't know ooh uh]

Uptake with repair after the repetition of a recast:

Example 3:

T : et la voiture de police umm qu'est-ce qu'ils ont fait ? Est-ce qu'ils ont continué ou ... ?

[and the police car umm what did they do? Did they continue or...?]

S6 : arrêté

[stop]

T : il se sont arrêtés ? (recast 1st time)

[they stopped ?]

S6 : oui

[yes]

T : ils se sont arrêtés ? (recast 2nd time)

[they stopped?]

S : ils se sont arrêtés et err er il appelé une ambulance (uptake with repair followed by a new error with the verb 'appeler')

[they stopped and err er he call an ambulance]

Recasts followed by no uptake or uptake with no repair:

Example 4:

S12: et le père a relaxé

[and dad has relaxed]

T : se relaxait ? (recast 1st time)

[was relaxing?]

S12: *oui*

[yes]

T: *il se relaxait ?* (recast 2nd time)

[was relaxing?]

S12: *oui*

[yes]

Example 5 :

T: *alors seulement un garçon avait de l'argent ? je ne sais pas*

[then only one boy had some money? I don't know]

S2: *er um un garçon devient devient jalous*

[er um one boy becomes becomes jealous]

T: *oh il est devenu jalous?* (recast 1st time)

[oh he became jealous?]

S2: *um um mm*

T: *il est devenu ?* (recast 2nd time)

[he became?]

S2: *jalous d'autre garçon et l'autre garçon dit um um*

[jealous of the other boy says um um]

Examples of requests for clarification

Request for clarification followed by uptake with repair:

Example 6 :

S19: *il porte une casque*

[he wears a helmet]

T: *comment ... ?* (clarification request)

[how ... ?]

S19: *uuuh, il portait une casque*

[uuuh he was wearing a helmet]

Example 7 :

S34 : *il y a deux frères ... Philippe et Paul um il um a joué le foot*

[there are two brothers ... Philippe and Paul um he um played soccer]

T : *hein ? qu'est-ce qu'ils faisaient ?* (request for clarification)

[uh ? what were they doing?]

S34: *um? sorry I don't understand... jouaient/ joué au foot ?*

[were playing/play soccer ?]

T: *um bien ... ils jouaient au foot*

[um well ... they were playing soccer]

S34 : *et um il faisait beau avec beaucoup de soleil... il faisait très chaud donc ils ont très soif*

[and um the weather was nice with a lot of sun ... it was hot so they are very thirsty]

T : *hein? comment ?* (request for clarification)

[uh ? how ?]

S34 : *ils étaient ? er er er ... il avait/aient*

[they were ? er er er ... he had

Uptake with repair after the repetition of a clarification request:

Example 8:

S34 : *très soif... et ils a acheté...*

[very thirsty ... and they has bought]

T : *umm tu peux répéter ?* (CR 1st time)

[umm can you repeat ?]

S34 : *ils ssssssssss..sont ils sont*

[they aaaaaaaaa...are they are]

T : *hein ?* (CR 2nd time)

[uh ?]

S34 : *ils ont acheté*

[they bought ?]

T : *bien*

[good]

Uptake with no repair or no uptake after a clarification request:

Example 9:

T: et quel temps faisait –il?

[and what was the weather like?]

S35: c'est l'après –midi

[it's afternoon]

T: ummm je n'ai pas entendu ? (CR 1st time)

[ummm I haven't heard ?]

S35: c'est l'après-midi

[it's afternoon]

*T: mais.. est-ce qu'il y avait du soleil ou... ?...le temps pas l'heure (CR and a hint regarding
the meaning of the question – what was the weather like ?)*

[but ...was it sunny or...? ... the weather not the time]

*S35: il avait du soleil et il fait beau (repetition of the verb 'avait' but failing to correct the
original error)*

[it had some sun and it is nice]

Example 10 :

S27: ils sont très contents

[they are very pleased]

T: uhmm pardon ? (CR 1st time)

[uhmm pardon me?]

S27: ils sont très contents

[they are very pleased]

T: hein ... ? (CR 2nd time)

[uh...?]

S27: ? –

Example 11 :

T: qu'est-ce qu'il s'est passé quand ils sont sortis du supermarché ?

[what happened when they left the supermarket ?]

S32 : um what they did after ? ummm il y avait battre ?

[ummm there was to fight?]

T: hein ? quoi ? (CR 1st time)

[uh? what?]

S32 : there was a fight ? okay Paul a poussé? (use of English, accurate use of the verb 'a poussé', but no repair of the verb 'se battre')

[okay Paul pushed?]

T: ummm? (CR 2nd time)

S32: et Philippe est tombé (correct use of the verb, 'est tombé', but still no uptake of the verb 'se battre')

[and Philippe fell]

6.7.1.2. Scoring of uptake

As already stated, in the current study corrective feedback was repeated if it did not result in successful uptake after the teacher's first attempt at correction. To take account of this successful uptake was scored in two ways: (a) only uptake moves following the first CF move were included and (b) all uptake moves in a CF episode were included. Thus, (a) provides an indication of whether a learner could self-correct immediately following CF while (b) provides an overall measure of the extent to which uptake following CF was successful. Uptake episodes were coded and scored separately for the passé composé and for the imparfait.

6.7.2. Coding and scoring the tests

All tests were coded and scored separately for each of the two target structures. Oral production tasks were performed individually, each participant was recorded by a digital voice recorder and the recordings were transcribed by the researcher. The number of obligatory use contexts was established for each tense (the passé composé and the imparfait)

separately. In addition, all instances of overuse of the two structures were identified. The tests used in oral production (Appendix C) elicited a range of 5 to 15 obligatory occasions for the use of passé composé with an average of 10 per test per student. The same tests elicited a range of 3 to 13 obligatory occasions for imparfait, with an average of 6 per student per test.

The only difference between scoring the oral and the written tests was with regard to the imparfait: whereas in oral production only a limited number of verbs were used for analysis (as explained in the next section), in written production all imparfait structures were taken into consideration because written form allows for the different verb endings in French to be clearly distinguished. The tests used in written production (Appendix D) elicited a range of 3 to 15 obligatory occasions for passé composé with an average of 8 per test per student. The tests used in written production elicited a range of 3 to 11 obligatory occasions for imparfait, with an average of 6 per student per test.

In order to ensure the reliability of coding and scoring, 15 % of each set of tests were coded and scored for obligatory use of the passé composé and imparfait by a French native speaker, a teacher at Alliance Française in Auckland. The following are the correlation coefficients for each set of tests:

Oral production		Written production	
PC Pre-test	.91	PC Pre-test	.98
PC Post-test 1	.93	PC Post-test 1	.87
PC Post-test 2	.89	PC Post-test 2	.90
IM pre-test	.95	IM Pre-test	.98
IM Post-test 1	.97	IM Post-test 1	.93
IM Post-test 2	.92	IM Post-test 2	.95

I first explain the analysis used to obtain the scores and then describe coding and scoring for each target structure in detail.

6.7.2.1. Target-like use analysis

Scores were calculated using Target-like use analysis (Pica, 1983) which takes into consideration the overuse of the target form. For each correct suppliance in obligatory context including the correct target-like form participants were given 2 points: one point was given

for the correct use of the verb tense and one point was given for the correct form. For example, (a) if a particular context required the tense which was actually used and it was formed correctly participants were given 2 points, (b) if a particular context required the tense which was used but did not contain the fully correct form participants were given 1 point and (c) if a particular context did not require the tense which was actually used the score was 0.

Each obligatory occasion was scored as follows:

One point – the learner used the correct tense but not the target-like form

e.g. passé composé: *Il a entré dans la chambre* (passé composé using incorrect auxiliary)

e.g. imparfait : *Il pleuverait* (the inflection indicates the use of imparfait but the target-like form should read “pleuvait” /infinitive: pleuvoir)

Two points – the learner used the correct tense with the target-like form

e.g passé composé: *Le professeur s’est réveillé à sept heures*

Il a pris le petit-déjeuner

e.g imparfait : *Il faisait très froid*

Il ne pouvais pas marcher

To calculate the target-like use for each structure the following formula (Pica, 1983) was employed:

$$\frac{\mathbf{n} \text{ correct suppliance in contexts (including correct form)} \times 100}{(\mathbf{n} \text{ obligatory contexts} \times 2) + (\mathbf{n} \text{ suppliance in non-obligatory contexts} \times 2)} = \text{per cent (\%)} \text{ accuracy}$$

The number of obligatory occasions was doubled to allow for the fact that two points were available for each obligatory occasion. Also, the number of overuses of the two tenses was doubled so that these were weighted similarly to the number of obligatory occasions. The formula provides a percentage of target-like use for each structure.

6.7.2.2. Coding and scoring of passé composé

Since for the passé composé the target structure consists of an auxiliary and a past participle, 2 points were given for the use of the correct auxiliary and a verb form (past participle),

whilst 1 point was given for the use of an incorrect auxiliary and a verb form (past participle). If no auxiliary was used the structure was not identified as a passé composé. Possible incorrect forms of the past participle were ignored.

Examples:

Elle a ouvert la porte = 2 points (correct auxiliary + verb form: past participle)

Elle est ouvert la porte = 1 point (incorrect auxiliary + verb form: past participle)

Il est sorti = 2 points (correct auxiliary + verb form: past participle)

Il a sorti = 1 point (incorrect auxiliary + verb form: past participle)

Elle a ouvré/ or ouvrié/ or ouvert la porte = 2 points (correct auxiliary + verb form)

Il est sorti/ or sortié = 2 points (correct auxiliary + verb form)

Il a sorti / sortié = 1 point (incorrect auxiliary + verb form)

6.7.2.3. Coding and scoring of imparfait in oral and written production

Oral production data were transcribed using broad transcription so that the difference, for example, between an ‘open e’ (e.g. in French imparfait inflected endings –ait, -ais) and a ‘closed e’ (e.g. in French infinitive and past participle endings –er and -é) was ignored. Therefore all verbs ending in –er in the infinitive (e.g. parler, regarder, manger, etc) were excluded from coding and scoring for imparfait in oral production since in learner language it is impossible to distinguish whether an imparfait, or an infinitive, or a past participle form was used. Only the verbs which could be identified unambiguously in their imparfait forms were coded and scored for imparfait (e.g. avoir, être, faire, pouvoir, savoir, pleuvoir, vouloir, dormir, peindre, attendre, etc) and given 2 points. If there were inaccuracies in the form, but the inflection undoubtedly suggested the use of the imparfait, 1 point was given (e.g. ‘peindrait’ instead of ‘peignait’, or ‘attendrait’ instead of ‘attendait’).

Examples:

Il y avait du vent = 2 points

Il pleuvait = 2 points

Elle était contente = 2 points

Un homme peignait les fenêtres = 2 points

Un homme peindrait les fenêtres = 1 point

The written production tests, as already noted, allow for the distinction between the forms of all verbs. Therefore, in contrast to oral production, all verbs used in the imparfait were taken into consideration for analysis. The characteristic imparfait inflections (for example *-ais*, *-ait*, *-aient*) were taken as indication of the use of imparfait in all verbs. Possible spelling mistakes (e. g. *-ais* instead of *-ait* or *-aient*) were ignored.

6.7.2.4. Hesitations, false starts, repetitions and self-repairs

Hesitations and false starts in oral production were excluded from coding and scoring. For example:

‘... *parce que la femme est la uhh est pein uhhh* ‘ (no obligatory occasion established)

[...because the woman is the uhh is pain uh]

‘... *il se il a il est uhh no...*’ (no obligatory occasion established)

[... he himself he has he is uhh no..]

Repeated attempts at the target verb forms were excluded from coding and scoring. The first utterance only was coded and scored:

‘... *elle a décidé elle a décidé de se promener avec son chien* (1 obligatory occasion for
passé composé , correct use of passé composé : 2 points)

[...she decided she decided to take her dog for a walk]

Self-repair was coded as two obligatory occasions and the first utterance given 0, the second utterance given 2 points:

‘...*il se veille oh il s’est veillé à huit heures...*’ (2 obligatory occasions, 1 target-like
form : 2 points)

[...he wakes up oh he woke up at eight o’clock...]

If self-repair took place before the utterance was fully articulated, it was coded for one obligatory occasion and scored according to the repair:

‘... *il a il est tombé...* (1 obligatory occasion) - 2 points

‘...*il est il a oublié ...* (1 obligatory occasion) - 2 points

6.8. Data analysis

The SPSS package was used to perform the statistical analyses. The following chapters present the results of the analyses obtained for each research question. The procedures and the types of analyses used to answer each research question are explained at the beginning of each chapter in relation to each particular question. In the next chapter I present the results for oral production. This will be followed by the chapter presenting the results for written production.

Chapter 7. Oral production results

7.1. Introduction

This chapter considers research questions 1 and 2 using the oral production data:

1. Does oral corrective feedback (recasts and clarification requests) have an effect of L2 learners' acquisition of French linguistic structures: the passé composé and the imparfait past tenses?
 - a) Do recasts have an effect on the L2 acquisition of the passé composé and the imparfait past tenses?
 - b) Do requests for clarification have an effect on the L2 acquisition of passé composé and the imparfait past tenses?
2. Does the learners' starting proficiency mediate the effect of the oral corrective feedback on the learners' acquisition of:
 - a) the passé composé past tense?
 - b) the imparfait past tense?

7.2. Data Analysis

The SPSS package was used to perform the statistical analyses. The following analyses were used to answer Research question 1 and Research question 2:

7.2.1. Descriptive statistics

- a) Descriptive statistics for each structure (imparfait and passé composé) were calculated.
- b) The frequencies of corrective feedback episodes for each target structure in each treatment group were calculated.
- c) Since the calculation of frequencies of corrective feedback episodes for each target structure revealed a considerable difference in CF episodes between the two treatment groups, an ANCOVA was used with the covariate CF episodes to take account of this, at each level of post-testing, for each target structure separately. This made possible to adjust the mean scores at each post-test, taking into account the number of corrective feedback episodes.

7.2.2. Inferential statistics

a) In order to answer Research question 1, a mixed design ANOVA with repeated measures was calculated, using the Bonferroni adjustments for multiple comparisons to examine the change that occurred in each treatment group and the control group after the pre-test (within group comparisons). The Bonferroni post-hoc tests were used because they are considered the most conservative (strict), which means that if the Bonferroni adjustments show a significant difference it must have really occurred. Separate repeated measures ANOVAs were calculated for the *passé composé* and for the *imparfait*.

b) Since it was assumed that the different number of CF episodes may have affected the performance of the treatment groups, the between-group differences were looked at separately, first the differences between the Control (CN) group and each of the treatment groups: using the Independent samples t-test, the Recast (RE) group was compared with the CN group at each level of post-testing, and the Clarification Request (CR) was compared with the CN group at each level of post-testing separately.

c) For comparison between the two experimental groups, which had to take into account the differences between the number of corrections each group received, a series of ANCOVAs were used (with the covariate CF episodes), for each target structure separately, at each level of post-testing separately.

7.2.2.1. Analyses for Research Question 2

In order to answer Research question 2, the participants in each group were divided into 'high' and 'low' based on their pre-test scores. For the *passé composé* the median score for the whole sample was 36.6. In each group, all learners above the median were placed in the 'high' proficiency sub-group and all those below 36.6 in the 'low' proficiency sub-group. For the *imparfait* the median score for the whole sample was zero. All learners above zero were deemed 'high' proficiency' and all those at zero the 'low' proficiency.

Once the 'low' and 'high' sub-groups were formed, an ANCOVA was performed at each level of post-testing for the four treatment sub-groups (RE-High, CR-High, RE-Low, CR-Low) using the covariate CF episodes.

Separately, a series of ANOVAs with the post-hoc tests, were used to examine the differences between each treatment group's 'low' and 'high' and the CN group's low' and 'high' proficiency participants, at each level of post-testing.

Effect sizes were estimated for within-group and between-group observed differences. In most situations it was possible to calculate Cohen's d , so that the difference between the two means was divided by the pooled standard deviation of the two groups (Field, 2009; Norris & Ortega, 2000). The following formulas were used:

$$d = \frac{\text{Mean 1} - \text{Mean 2}}{S(\text{pooled})} \qquad S(\text{pooled}) = \frac{(N_1 - 1) S_1^2 + (N_2 - 1) S_2^2}{(N_1 - 1) + (N_2 - 1)}$$

(S = standard deviation, N = sample size)

In order to estimate the magnitude of contrasts in ANCOVAs, the correlation coefficient r was used as a measure of an effect size (Field, 2009) since it was possible to derive r from the output of ANCOVAs. The following formula was used:

$$r_{\text{contrast}} = \sqrt{\frac{t^2}{t^2 + df}}$$

It is to be noted that t-statistics in an ANCOVA have N-2 degree of freedom (df) where N is the total sample size.

Effect sizes were interpreted according to Norris and Ortega (2000) and following Cohen's recommendation, so that effects lower than ES= .5 were considered low effect sizes, effects lower than ES = .8 were considered medium and effects above ES = .8 were considered large effect sizes.

7.3. Results for RQ 1

7.3.1. Results for passé composé: descriptive statistics

Table 3 and Figure 3 show the descriptive data for oral production of the passé composé past tense. On the pre-test the RE group had a mean percentage score of 29.35 (SD=26.43), the CR group had a mean percentage score of 37.44 (SD=27.19) whilst the CN group had a mean

of 36.62 (SD=17.33). An analysis of variance (ANOVA) showed no statistically significant differences among the three groups at the time of pre-testing:

$$F(2, 51) = .492, p = .614 (p > .05)$$

As indicated in Table 3 and Figure 3, on the immediate post-test (post-test 1) the RE group achieved a mean percentage score of 50.82 (SD=30.10), the CR group obtained a mean percentage score of 38.13 (SD=22.92) and the CN group increased slightly achieving a mean of 37.59 (SD=15.12). On the delayed post-test (post-test 2) the RE group accuracy score slightly decreased compared to the immediate post-test, with a mean percentage score of 48.49 (SD=30.86), the CR group slightly increased with a mean of 39.63 (SD=23.00) and the CN group slightly decreased obtaining a mean score of 34.14 (SD=17.28).

Table 3 Descriptive statistics for passé composé

GROUP	Pre-test		Post-test 1		Post-test 2	
	Mean	SD	Mean	SD	Mean	SD
RECAST (N=18)	29.35	26.44	50.82	30.1	48.49	30.86
CL.REQUEST (N=16)	37.44	27.19	38.13	22.92	39.63	23.01
CONTROL (N=16)	36.62	17.33	37.60	15.12	34.14	17.28

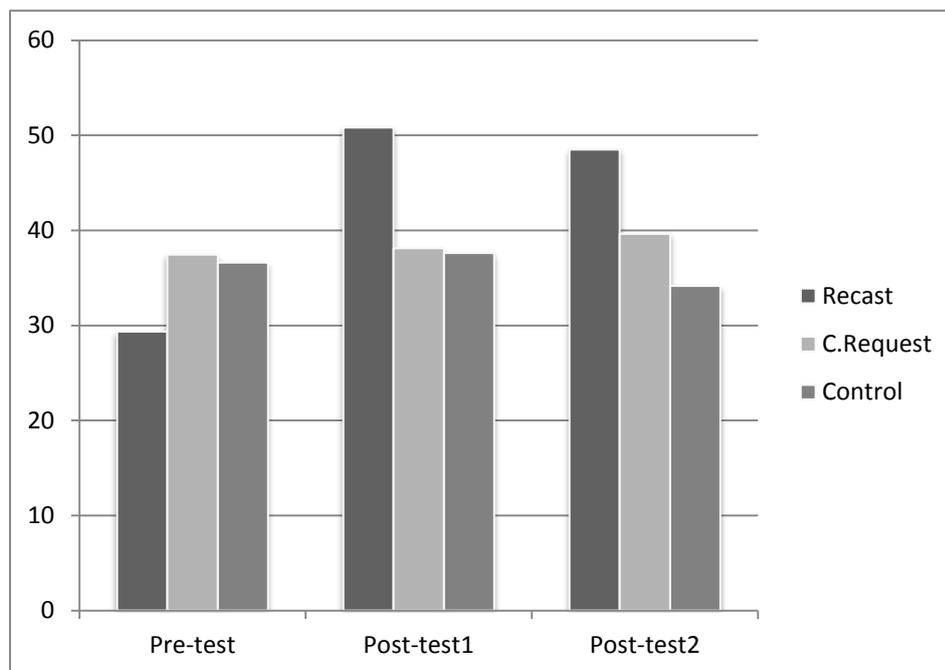


Figure 3 Mean scores for passé composé: the RE group, the CR group and the CN group

b) Table 4 presents the data related to the frequency of corrective feedback for passé composé in each of the two treatment groups, the RE and the CR group. These results indicate that the RE group received considerably more corrective feedback than the CR group, due to the fact that the RE group produced more errors in the target structure than the CR group. As is evident from Table 4, there were 33 CF episodes for the passé composé recorded in the RE group and 17 CF episodes for the passé composé in the CR group.

Table 4 Number of CF episodes for the PC and number of learners participating in interaction

GROUP	No of learners	CF episodes
RECAST	12	33
CLAR. REQUEST	11	17

c) Taking into consideration the different number of corrective feedback episodes each treatment group received, an ANCOVA with the covariate CF episodes was calculated at each level of post-testing. This produced the adjusted mean scores for each treatment group on post-test 1 (immediate) and post-test 2 (delayed). Table 5 presents the adjusted mean scores of the oral passé compose for two treatment groups on post-test 1 and post-test 2.

Table 5 Adjusted mean scores for the RE and the CR group on post-test 1 and post-test 2

GROUP	Post-test 1		Post-test 2	
	Mean	St.Err	Mean	St.Err
RECAST (N=18)	49.23	6.32	46.66	6.38
CL.REQUEST (N=16)	39.91	6.72	41.68	6.77

7.3.2. Inferential statistics

a) Table 6 and Figure 4 show the results of a Repeated measures ANOVA which was used in order to examine the change over time in all three groups (within group comparison). This analysis is based on the original scores of each group, without taking into consideration the number of CF episodes. The analysis revealed the following results using the Bonferroni adjustments for multiple comparisons:

For the RE group there is a statistically significant difference between the pre-test and the immediate post-test ($p = .000$) with a close to large effect ($ES = .76$), and between the pre-test and the delayed post-test ($p = .002$) with a medium effect size ($ES = .67$). There is no significant difference between the two post-tests (post-test 1 and post-test 2) and the effect is negative ($ES = -.08$).

In the CR group there is no evidence of statistically significant difference between any of the post-tests compared with the pre-test, neither between the post-tests themselves. The effect sizes are very small (pre-test to post-test1 $ES = .03$, pre-test to post-test2 $ES = .09$, post-test1 to post-test2 $ES = .07$). The same results, showing no statistically significant difference between any of the tests, and very small effect sizes are evident for the CN group.

Table 6 Results of Repeated measures ANOVA: within group differences

	RE group		CR group		CN group	
	p	ES	p	ES	p	ES
Pre-test - Post-test 1	.000	.76	>.05	.03	>.05	.06
Pre-test - Post-test 2	.002	.67	>.05	.09	>.05	-.01
Post-test 1 - Post-test2	>.05	-.08	>.05	.07	>.05	-.02
	df=17		df=15		df=15	

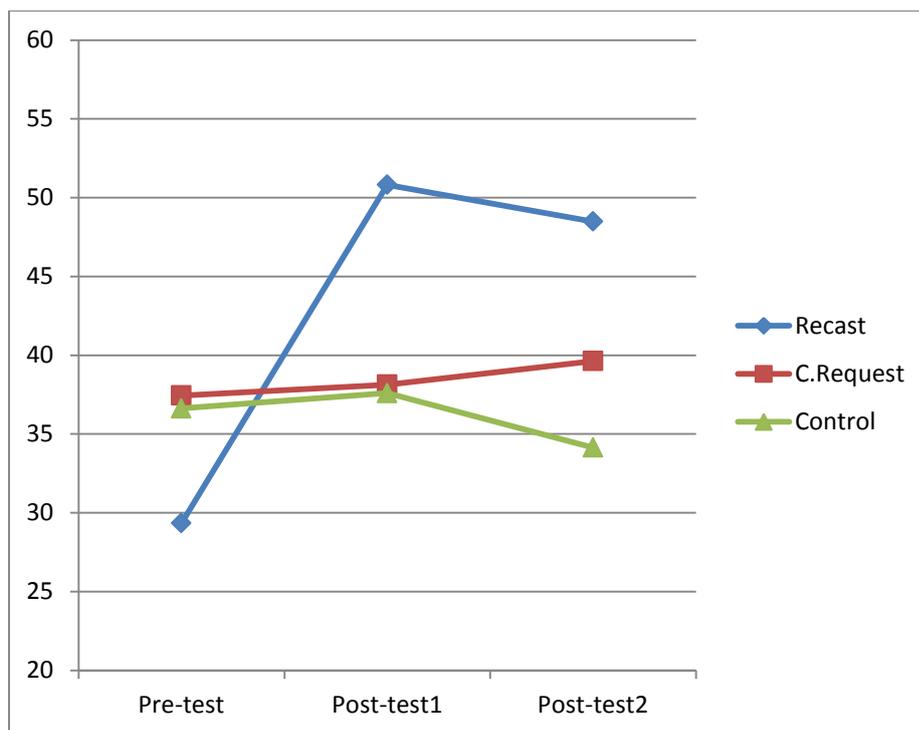


Figure 4 Repeated measures ANOVA: performance of three groups over time

b) Since it was assumed that the number of CF episodes may have affected the performance on post-tests, separate analyses were carried out in order to compare each treatment group with the CN group at each level of post-testing,

The results of a t-test between the RE and the CN group show the following:

t-test on post-test 1: $t(32) = 1.59$, ns, $ES = .27$

t-test on post-test 2: $t(32) = 1.69$, ns, $ES = .29$

A comparison between the CR and the CN group, using the

t-test, shows the following results:

t-test on post-test 1: $t(31) = -.02$, ns, $ES = .04$

t-test on post-test 2: $t(30) = .51$, ns, $ES = .09$

These t-tests show no significant differences between each of the two treatment groups compared with the control group on post-test 1 and post-test 2.

c) In order to examine the differences between the RE group and the CR group, an ANCOVA was computed at each level of post-testing using the covariate CF episodes for the passé composé. With the covariate CF episodes evaluated at $CF = 1.41$, the tests of Between

Subjects Effects show no significant differences between the two groups: $F(1, 31) = 1.13$, ns and the effect is small, $ES = .18$

On post-test 2 with the covariate CF episodes evaluated at $CF = 1.41$, the Tests of Between Subjects Effects also show no significant differences between the two groups, the effect is small:

$F(1, 31) = .279$, ns, $ES = .12$

7.3.3. Results for imparfait: descriptive statistics

Table 7 and Figure 5 show the descriptive data for oral production of the imparfait past tense. On the pre-test the RE group had a mean percentage score of 15.79 ($SD=19.53$), the CR group had a mean percentage score of 16.52 ($SD=23.17$) whilst the CN group had a mean of 18.48 ($SD=21.85$). An analysis of variance (ANOVA) showed no statistically significant differences among the three groups at the time of pre-testing:

$F(2, 51) = .492$, $p = .614$ ($p > .05$)

As indicated in Table 7 and Figure 5, on the immediate post-test (post-test 1) the RE group achieved a mean percentage score of 35.22 ($SD=31.65$), the CR group obtained a percentage mean of 21.99 ($SD=26.15$) and the CN group decreased with a mean of 16.10 ($SD=23.31$). On the delayed post-test (post-test 2) the RE group accuracy score increased again achieving a mean percentage score of 47.72 ($SD=29.17$), the CR group slightly increased with a mean of 23.81 ($SD=29.64$) and the CN group increased obtaining a mean score of 19.72 ($SD=20.26$).

Table 7 Descriptive statistics for imparfait in oral production

GROUP	Pre-test		Post-test 1		Post-test 2	
	Mean	SD	Mean	SD	Mean	SD
RECAST (N=18)	15.79	19.53	35.22	31.65	47.72	29.17
CL.REQUEST(N=16)	16.52	23.17	21.99	26.15	23.81	29.64
CONTROL (N=15)	18.48	21.85	16.10.	23.31	19.72	20.26

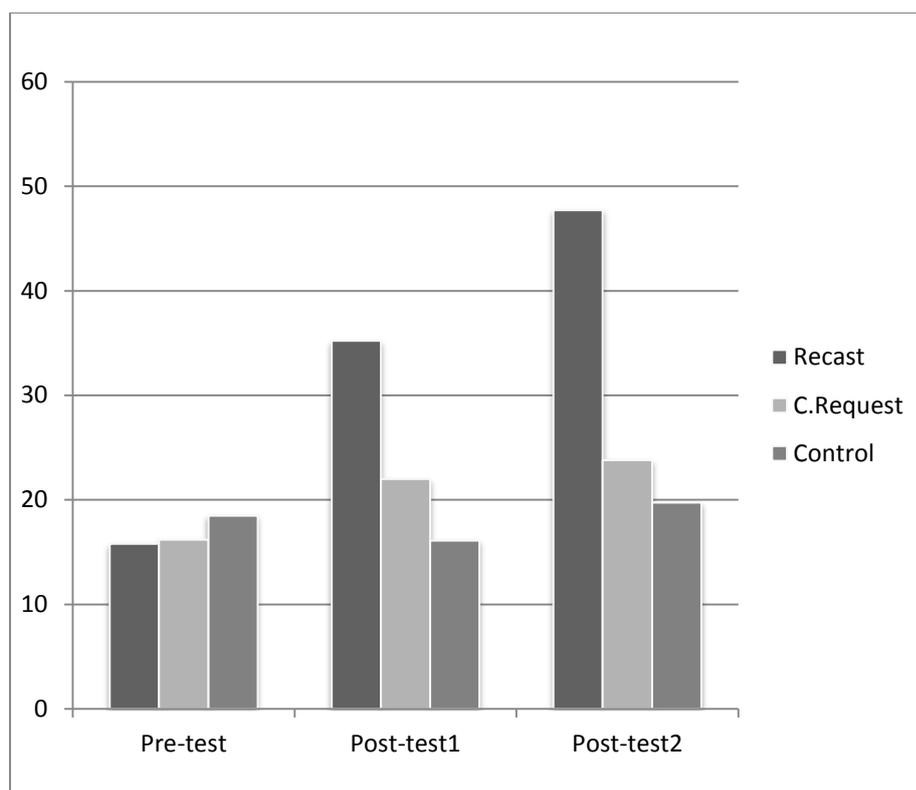


Figure 5 Mean scores for imparfait: the RE group, the CR group and the CN group

Table 8 presents the data related to the frequency of corrective feedback for the imparfait in each of the two treatment groups, the RE and the CR group. These results indicate that the RE group received considerably more corrective feedback than the CR group, due to the fact that the RE group produced more errors of the target structure than the CR group. As evident from Table 8, there were 35 CF episodes for the imparfait recorded in the RE group and 23 CF episodes for the imparfait in the CR group.

Table 8 Number of CF episodes for the imparfait and number of learners participating in interaction

GROUP	No of learners	CF episodes
RECAST	13	35
CLAR. REQUEST	12	23

Taking into consideration the different number of corrective feedback episodes each treatment group received, an ANCOVA with the covariate CF episodes was calculated at each level of post-testing for the oral imparfait. This produced the adjusted mean scores for

each treatment group on post-test 1 (immediate) and post-test 2 (delayed). Table 9 presents the adjusted mean scores of the oral imparfait for two treatment groups on post-test 1 and post-test 2.

Table 9 Adjusted mean scores for the RE and the CR group

GROUP	Post-test 1		Post-test 2	
	Mean	St.Err.	Mean	St.Err.
RECAST (N=18)	34.36	7.00.	46.82	7.04
CL.REQUEST(N=16)	22.96	7.44	24.82	7.48

7.3.4. Inferential statistics

Table 10 and Figure 6 show the results of a Repeated measures ANOVA which was used in order to examine the change over time in each of the three groups (within group comparison). This analysis is based on original scores, without taking into account the number of corrective feedback episodes each group received. The analysis revealed the following results using the Bonferroni adjustments for multiple comparisons:

For the RE group there is a statistically significant difference between the pre-test and the immediate post-test ($p = .002$) with a large effect size ($ES = .76$), and between the pre-test and the delayed post-test ($p = .000$), again with a large effect ($ES = 1.05$). There is also a significant difference between the two post-tests, post-test 1 and post-test 2 ($p = .044$), with a medium effect size ($ES = .41$).

In the CR group there is no evidence of statistically significant difference between any of the post-tests compared with the pre-test, neither between the post-tests themselves. Effect sizes are small: from pre-test to post-test1 $ES = .22$, from pre-test to post-test2 $ES = .28$, and between post-test1 and post-test2 $ES = .07$. The same results, showing no statistically significant difference between any of the tests, and small effect sizes are evident for the CN group.

Table 10 Results of Repeated measures ANOVA: within group differences

	RE group		CR group		CN group	
	p	ES	p	ES	p	ES
Pre-test - Post-test 1	.002	.76	>.05	.22	>.05	-.11
Pre-test - Post-test 2	.000	1.05	>.05	.28	>.05	.06
Post-test1 - Post-test2	.044	.41	>.05	.07	>.05	.17
	df=17		df=15		df=14	

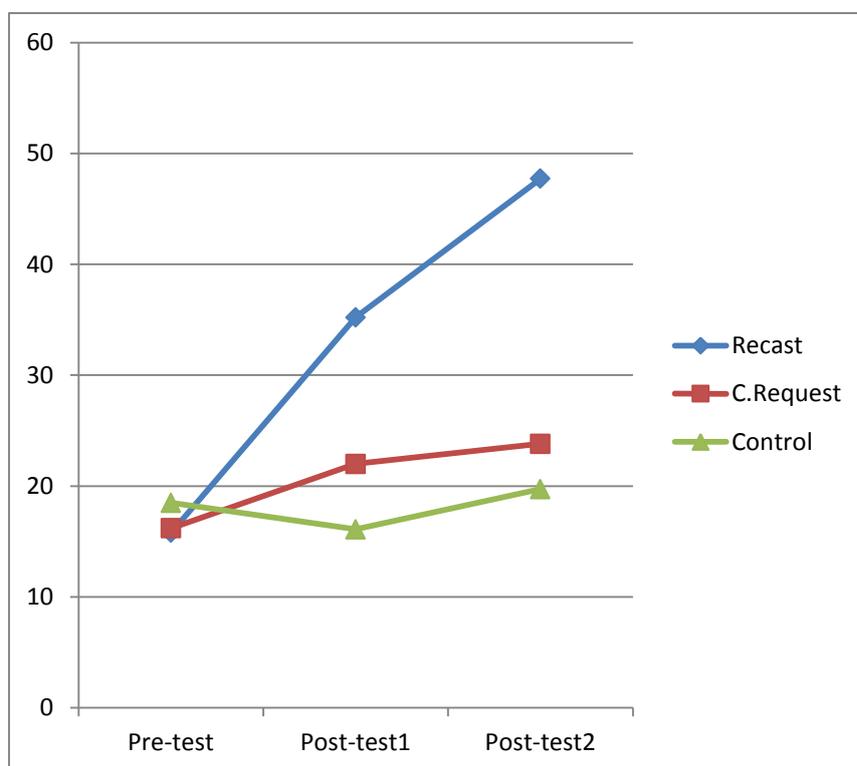


Figure 6 Repeated measures ANOVA: performance of three groups over time

Since it was assumed that the number of CF episodes may have affected the performance of the two treatment groups, separate analyses were computed in order to compare each treatment group with the CN group at each level of post-testing.

The results of a t-test for comparison between the RE and the CN group show the following:

t-test on post-test 1: $t(31) = 1.10$, $p = .055$, $ES = .19$

t-test on post-test 2 : $t(31) = 3.24$. $p = .003$, $ES = .50$

For comparison between the CR and the CN group the results of the t-test show the following:

t-test on post-test 1 : $t(30) = .76$, ns, ES= .14

t-test on post-test 2: $t(29) = .45$, ns, ES= .08

These results indicate that for the imparfait target structure there is a close to significant difference between the RE and the CN group on post-test 1 and a significant difference between these two groups on post-test 2. Between the CR and the CN group there are no significant differences either on post-test 1 or on post-test 2.

In order to examine the differences between the RE group and the CR group taking into consideration the number of CF episodes, an ANCOVA was computed at each level of post-testing using the covariate CF episodes for the imparfait. With the covariate CF episodes evaluated at $CF = 1.58$, the tests of Between Subjects Effects at the immediate post-test show no significant differences between the two groups: $F(1, 31) = 1.12$, ns, with a small effect size (ES = .18).

On the delayed post-test, with the covariate CF episodes evaluated at $CF = 1.58$, the tests of Between Subjects Effects show a statistically significant difference between the two groups: $F(1, 31) = 4.47$, $p = .043$, ES= .36

7.4. Summary

A summary of the oral production results for Research question 1, as displayed in Table 11 and Table 12, indicates that regarding the within group (between tests) results for the passé composé target structure, only the participants in the RE group have significantly improved from pre-test to post-test 1 and from pre-test to post-test 2. For the imparfait target structure again, only the RE group have significantly improved from pre-test to post-test 1, from pre-test to post-test 2 and from post-test 1 to post-test 2.

When the different number of CF episodes has been accounted for, there are no significant differences between the two treatment groups (RE and CR) neither on post-test 1 nor on post-test 2 for the passé composé. There are also no statistically significant differences for the passé composé on either of the post-tests between each of the treatment groups when compared with the CN group. For the imparfait target structure, when the number of CF

episodes is accounted for, there are no statistically significant differences between the RE and the CR group on post-test 1 but there is a statistically significant difference between the two treatment groups, evident on post-test 2. Comparison of each treatment group separately with the CN group shows that there is a close to significant difference between the RE and the CN group on post-test 1 and a statistically significant difference between these two groups on post-test 2 for the imparfait. There are no significant differences between the CR and the CN group on any of the post-tests for the imparfait.

Table 11 Summary of results for RQ 1: Within group comparisons

GROUP	Passé composé	Imparfait
RECAST	Post-test 1 > Pre-test (p = .000) Post-test 2 > Pre-test (p = .002) Post-test 2 = Post-test 1 (p > .05)	Post-test 1 > Pre-test (p = .002) Post-test 2 > Pre-test (p = .000) Post-test 2 > Post-test 1 (p = .044)
CL. REQUEST	Post-test 1 = Pre-test (p > .05) Post-test 2 = Pre-test (p > .05) Post-test 2 = Post-test 1 (p > .05)	Post-test 1 = Pre-test (p > .05) Post-test 2 = Pre-test (p > .05) Post-test 2 = Post-test 1 (p > .05)
CONTROL	Post-test 1 = Pre-test (p > .05) Post-test 2 = Pre-test (p > .05) Post-test 2 = Post-test 1 (p > .05)	Post-test 1 = Pre-test (p > .05) Post-test 2 = Pre-test (p > .05) Post-test 2 = Post-test 1 (p > .05)

Table 12 Summary of results for RQ 1: Between group comparisons

	Post-test 1	Post-test 2
Passé composé	Recast=Cl.Request (p > .05) Recast = Control (p > .05) Cl. Request = Control (p > .05)	Recast=Cl.Request (p > .05) Recast = Control (p > .05) Cl. Request = Control (p > .05)
Imparfait	Recast = Cl. Request (p > .05) Recast = Control (p = .055) Cl. Request = Control (p > .05)	Recast > Cl.Request (p=.043) Recast > Control (p = .003) Cl. Request = Control (p > .05)

7.5. Results for RQ 2

7.5.1. Results for passé composé

Table 13 displays the descriptive data for oral production of the passé composé ‘low’ and ‘high’ proficiency sub-groups and Table 14 shows the adjusted mean scores on post-tests for the treatment sub-groups once the covariate CF episodes has been accounted for. Table 15 presents the effect sizes for each ‘low’ and ‘high’ sub-group, indicating the magnitude of the observed effect between tests.

Table 13 Descriptive statistics for 'low' and 'high' sub-groups on each test of passé composé

GROUP	Pre-test		Post-test 1		Post-test 2	
	Mean	SD	Mean	SD	Mean	SD
RE-Low(N=9)	4.52	4.65	26.16	17.66	21.31	12.22
RE-High (N=9)	54.18	8.74	75.50	15.65	75.66	14.61
CR-Low (N=8)	14.88	11.83	31.32	21.18	36.13	23.55
CR-High(N=8)	60.00	16.75	44.93	23.91	43.13	23.48
CN-Low (N=7)	20.47	9.71	25.89	7.44	25.51	4.19
CN-High (N=9)	49.18	9.31	46.70	13.19	40.86	20.75

Table 14 Adjusted mean scores for RE and CR 'low' and 'high' sub-groups on each post-test

GROUP	Post-test 1		Post-test 2	
	Mean	St.Err	Mean	St.Err
RE-Low (N=9)	26.16	6.4	21.20	6.2
RE-High (N=9)	73.80	6.6	73.60	6.3
CR-Low(N=8)	32.50	6.6	38.10	6.7
CR-High (N=8)	45.20	6.8	43.60	6.5

Table 15 Effect sizes for each sub-group, observed between tests

	Effect size					
	RE- Low	RE-High	CR-Low	CR-High	CN-Low	CN-High
Pre-test - Post-test 1	1.94	1.75	1.0	-.74	.62	-.22
Pre-test - Post-test 2	1.99	1.85	1.2	-.85	.72	-.55
Post-test1 - Post-test 2	-.32	.01	.22	-.08	-.07	-.34

7.5.1.1. Tests comparing each treatment sub-group with the control sub-groups

An ANOVA on post-test 1, looking at the differences between the RE-Low, CN-Low, RE-High and the CN-High sub-groups reveals a significant difference among the four sub-groups: $F(3, 30) = 22.87, p = .000$

The Bonferroni post-hoc tests performed for these four sub-groups show that there is a significant difference between the RE-High and the CN-High ($p = .001$) with a large effect size ($ES = 2.0$) but there are no significant differences between the RE-Low and the CN-Low sub-group ($p > .05$).

On post-test 2 for the same four sub-groups an ANOVA indicates a significant difference again:

$F(3, 30) = 24.74, p = .000$

The post-hoc tests reveal a significant difference between the RE-High and the CN-High ($p = .000$), with a large effect ($ES = 1.97$) whereas no significant differences are evident between the RE-Low and the CN-Low sub-groups ($p > .05$).

The same analyses performed among the four sub-groups of the CR and the CN group (CR-Low, CN-Low, CR-High, CN-High) show the following results:

An ANOVA among the four sub-groups shows a close to significant result:

$F(3, 29) = 2.77, p = .059$

The post-hoc tests show no significant differences between the CR-Low and the CN-Low ($p > .05$), neither between the CR-High and the CN-High sub-groups ($p > .05$). There are only significant differences between the 'low' and the 'high' sub-groups.

On post-test 2, for the same sub-groups an ANOVA shows no significant result:

$F(3, 28) = 1.13, ns (p > .05)$

7.5.1.2. Tests comparing the treatment sub-groups

The result of the tests Between Subjects of an ANCOVA performed on post-test 1 among four treatment sub-groups (RE-Low, CR-Low, RE-High, CR-High), taking into account the covariate CF episodes, shows a significant difference among the four groups: $F(3, 30) = 10.15, p = .000$

The follow-up pair-wise comparisons show a significant difference between the RE-High and the CR-High sub-group ($p = .005$), with a large effect size ($ES = 1.57$). The difference between the RE-Low and the CR-Low sub-groups is not significant ($p > .05$).

The result of the tests Between Subjects Effects of an ANCOVA on post-test 2 (RE-Low, CR-Low, RE-High, CR-High), taking into account the covariate CF episodes, again show a significant difference among the four sub-groups:

$$F(3, 29) = 11.96, p = .000$$

The follow-up pair-wise comparison indicates a significant difference between the RE-High and the CR-High sub-groups ($p = .003$), with a large effect ($ES = 1.73$). The difference between the RE-Low and the CR-Low sub-groups is not significant ($p > .05$).

7.5.2. Results for imparfait

Table 16 displays the descriptive data for oral production of the imparfait ‘low’ and ‘high’ proficiency sub-groups, whereas Table 17 shows the adjusted mean scores for the same sub-groups when the covariate CF episodes has been accounted for. Table 18 shows the effect sizes for each ‘low’ and ‘high’ sub-group as measured between tests over time.

Table 16 Descriptive statistics for 'low' and 'high' sub-groups on each test of imparfait

GROUP	Pre-test		Post-test 1		Post-test 2	
	Mean	SD	Mean	SD	Mean	SD
RE-Low (N=10)	0.00	0.00	19.08	24.80	32.36	24.86
RE-High (N=8)	35.54	11.15	55.39	28.39	66.91	22.66
CR-Low(N=9)	0.00	0.00	8.20	12.50	12.40	24.80
CR-High(N=7)	37.04	19.06	39.75	29.10	38.47	32.71
CN-Low (N=7)	0.00	0.00	5.94	10.43	3.57	9.45
CN-High (N=8)	34.64	17.73	25.00	28.28	33.85	15.97

Table 17 Adjusted mean scores for RE and CR 'low' and 'high' sub-groups on each post-test

GROUP	Post-test 1		Post-test 2	
	Mean	St. Err	Mean	St. Err
RE-Low (N=10)	18.53	7.7	31.70	8.3
RE-High (N=8)	54.90	8.5	66.50	9.2
CR-Low(N=9)	9.20	8.3	13.80	8.9
CR-High(N=7)	39.90	8.5	39.10	9.8

Table 18 Effect sizes for each sub-group, observed between tests

	Effect size					
	RE-Low	RE-High	CR-Low	CR-High	CN-Low	CN-High
Pre-test - Post-test 1	1.54	1.00	1.31	.11	1.14	-.42
Pre-test - Post-test 2	2.60	1.86	1.00	.06	.76	-.04
Post-test1 - Post-test 2	0.50	0.45	0.23	-.04	-.24	.40

7.5.2.1. Tests comparing each treatment sub-group with the control sub-groups

An ANOVA on post-test 1, looking at the differences between the RE-Low, CN-Low, RE-High and the CN-High sub-groups reveals a significant difference among the four sub-groups: $F(3, 29) = 5.66$, $p = .004$

The Bonferroni post-hoc tests performed for these four sub-groups show that there are no significant differences between the RE-High and the CN-High ($p = .085$) but the effect size is large ($ES = 1.07$). There are no significant differences between the RE-Low and the CN-Low sub-group ($p > .70$) either.

On post-test 2 for the same four sub-groups an ANOVA indicates that a significant difference:

$$F(3, 30) = 12.75, p = .000$$

The post-hoc tests reveal a significant difference between the RE-High and the CN-High ($p = .015$), with a large effect size ($ES = 1.71$), and a significant difference between the RE-Low and the CN-Low sub-groups ($p > .039$), again with a large effect ($ES = 1.54$).

The same analyses performed among the four sub-groups of the CR and the CN group (CR-Low, CN-Low, CR-High, CN-High) show the following: an ANOVA among the four sub-groups indicates a significant result: $F(3, 28) = 4.22$, $p = .014$. However, there are no significant differences between the two ‘low’ and the two ‘high’ sub-groups.

On post-test 2 an ANOVA again indicates a significant result $F(3, 27) = 4.37$, $p = .012$ but the post-hoc tests show no significant differences between the two ‘low’ and the two ‘high’ sub-groups.

7.5.2.2. Tests comparing the treatment sub-groups

The result of the tests of Between Subjects Effects of an ANCOVA performed on post-test 1 among four treatment sub-groups (RE-Low, CR-Low, RE-High, CR-High), taking into account the covariate CF episodes, indicates a significant difference among the four sub-groups: $F(3, 30) = 6.01, p = .002$

The follow-up pair-wise comparisons show no significant differences between the RE-High and the CR-High sub-groups ($p > .05$), nor between the RE-Low and the CR-Low sub-groups ($p > .05$). There are only significant differences within each group.

The result of the tests of Between Subjects Effects of an ANCOVA on post-test 2 (RE-Low, CR-Low, RE-High, CR-High), taking into account the covariate CF episodes, show a significant difference among the four groups:

$F(3, 29) = 5.7, p = .003$

The pair-wise comparisons indicate a significant difference between the RE-High and the CR-High sub-group ($p = .043$) with a large effect size ($ES = 1.04$) but no significant differences between the RE-Low and the CR-Low sub-groups ($p > .05$)

7.6. Summary

In summary, when each group is divided into 'low' and 'high' according to the score on the pre-test, and when the analyses are performed in the same way as for the three main groups – so that for the differences between the treatment groups an ANCOVA was used accounting for the number of CF episodes, and for the differences between each treatment group and the control group an ANOVA was used at each level of post-testing. The results show that the 'high' RE group outperformed both the 'high' CR and the 'high' CN group on both post-tests for the passé composé, but the 'low' sub-groups did not differ.

For the imparfait, on post-test 1 there were no significant differences between any of the sub-groups, neither 'high' nor 'low', but on the delayed post-test (post-test 2) the 'high' RE group outperformed both the 'high' CR and the 'high' CN group. Also, on delayed post-test for the imparfait the 'low' RE group outperformed the 'low' CN group. The summary of results for Research question 2 is presented in Table 19.

Table 19 Summary of results for RQ 2

	Post -test 1	Post -test 2
PASSE COMP.	RE-High > CR-High (p=.005)	RE-High > CR-High (p=.003)
	RE-Low = CR-Low (ns)	RE-Low = CR-Low (ns)
	RE-High > CN-High (p=.001)	RE-High > CN-High (p=.000)
	RE-Low = CN-Low (ns)	RE-Low = CN-Low (ns)
	CR-High =CN-High (ns)	CR- High = CN-High (ns)
	CR-Low= CN-Low (ns)	CR-Low = CN-Low (ns)
IMPARFAIT	RE-High = CR-High (ns)	RE-High >CR-High (p= .043)
	RE-Low = CR-Low (ns)	RE-Low = CR-Low (ns)
	RE-High = CN-High (ns)	RE-High > CN-High (p=. 015)
	RE-Low = CN-Low (ns)	RE-Low > CN-Low (p = .039)
	CR-High = CN-High (ns)	CR-High = CN-High (ns)
	CR- Low = CN-Low (ns)	CR-Low = CN-Low (ns)

7.7. Discussion

The analysis of the results for the two treatment groups (RE and CR) was complicated by the fact that the RE group received notably more corrections than the CR group for both the passé composé and the imparfait. This makes it difficult to compare the two types of CF as any difference between the two CF groups may reflect the frequency of the corrections they received rather than the type of correction. The explanation for the differences in the number of corrections each group received is that the learners in the CRs group made far fewer errors in both past tense structures than the learners in the RE group. It is not clear why this happened as the pre-tests indicated that there were no statistically significant differences in the two groups on either structure prior to the treatment.

I will consider the results for Research question 1 first. For the whole sample, there is evidence that the corrective feedback benefited the RE group more than the CR group. There were significant gains for the RE group over time (i.e. from pre-test to post-test 1 and from pre-test to post-test 2) for both structures but there were none for either the CR or the CN groups. Clearly, the recasts benefited the learners, enabling them to improve in the accuracy with which they used both structures. Given that the CN group received no corrections, no improvement was to be expected. However, the CR group did receive corrections and yet did not improve. Why then did the RE group improve over time while the CR group did not?

There are two possible explanations. One is that for CF of any type to be effective there must be an adequate number of corrections. In the case of the RE group the 33 corrections this group received for passé composé and the 35 they received for imparfait were clearly sufficient to induce improvement. It is possible that the 17 corrections the CR group received for passé composé and the 23 for imparfait were not sufficient. However, this explanation is not entirely convincing as both groups received sufficiently focused CF for an effect to become apparent. Here are some examples from other studies: In Yang and Lyster's (2010) study which looked at English regular and irregular past tenses, the Prompt group received 27 corrections in total and the RE group 23 corrections in total (in the current study it would be 40 corrections in total for clarification requests and 68 in total for the recasts). In Ellis (2007) the RE group received 42 corrections for past tense and 22 for comparative (in total 66) whilst the Metalinguistic group received 25 corrections for past tense and 19 for comparative (in total 44). In Ellis et al.'s (2006) study the RE group received 52 and the Metalinguistic group 40 corrections, whereas in Loewen and Nabei (2007) the RE and CR groups received each 18 corrections and the Metalinguistic group less than 6 corrections on average. In other words, the number of corrections received by the CR group is comparable to several other studies that found an effect for CF. The second explanation is that it was the type of correction that was responsible for the differential effects of the two types of CF over time. That is, input-providing CF enabled the learners to improve over time whereas the output-prompting CF did not. If this second explanation is correct, an explanation for why the input-providing CF led to improvement and output-prompting CF did not need to be found.

7.7.1. Input-providing vs. output-prompting CF

It has been claimed that pushing learners to produce output assists acquisition (cf. Output hypothesis, Swain, 1985). Where corrective feedback is concerned, self-repair and self-discovery are considered effective in helping L2 learners to develop automaticity. There is clear evidence that requests for clarification and other prompting strategies can facilitate acquisition (McDonough, 2005; Nobuyoshi & Ellis, 1993; Pica, 1991; Pica et al., 1989; Takashima & Ellis, 1999; Yang & Lyster, 2010). On the other hand, it has been argued that recasts as an input-providing strategy may assist in noticing-the-gap (Schmidt, 1990) or may induce a 'cognitive comparison' (Nelson, 1987). According to such a view, recasts not only supply a correction but also help learners to realize the 'link' between the form used and its meaning. Doughty (2001) has named this "cognitive mapping among forms, meaning and

use” (p. 211). This kind of ‘clicking’ between a linguistic form and its linguistic meaning (which can refer to both the semantic meaning of a feature and to its use in context) has been the main argument in favour of recasts, in accordance with Long’s (1996) Interaction hypothesis. Thus, in the present study recasts may have assisted the learners by focusing their attention on the ‘gaps’ in their use of the target language and in this way stimulating the ‘cognitive comparison’.

Another argument in favour of the input-providing strategy in this particular case is the fact that recasts reduce the processing load (Skehan, 1998) leaving enough capacity for students to process not only the meaning but also the structures at which their attention is briefly directed. As Skehan contended, recasts can be facilitative of managing the processing load of formally complex structures. In comparison, clarification requests usually help learners to retrieve and consolidate existing knowledge but that may use up the learners’ limited cognitive resources while they are primarily engaged with the meaning.

The tasks in this study were purely communicative and meaning oriented. That is, the students’ attention was directed at meaning, not at form. Although the recasts negotiated form rather than meaning, it was vital that they occurred while the learners were engaged with meaning. This constitutes the basic requirement for recasts to be effective: that meaning and use must already be evident to the learner at the time attention is drawn to form (Doughty, 2001). Research in L1 acquisition (Camarata, Nelson, Gillum, & Camarata, 2009; Farrar, 1990, 1992; Saxton, 1997, 2010) has shown that recasts in conversation, while children’s attention is engaged with the meaning, result in significantly more accurate use of grammatical morphemes and complex sentence structures than imitation strategies conducted out of a conversational context. These researchers argue that ‘corrective recasts’ provide negative evidence or negative input as a response to errors and thus facilitate children’s language ‘growth’ by making transparent the comparison between an error and its correction when both are embedded in a meaningful conversation.

Input-providing CF also resulted in frequent exposure to the target forms, i.e. the target structures were heard frequently enough to be incorporated into the learners’ existing knowledge. The Mann-Whitney independent samples test showed that there was a significant difference between the RE and the CR group in the number of target structures occurring in the input (Asymp.sig =.008 for the passé composé, and .007 for the imparfait). It is important to emphasize that it was not the number of corrections received by each group but the number

of target structures in the input which was significant. It follows that the input-providing CF leads to a higher frequency of exposure to target structures than the output-prompting CF.

7.7.2. The target structures

It may be useful to compare the features in focus of the current study with the features in Lyster's (2004), Ammar and Spada's (2006) and Ammar's (2008) studies. In these studies the participants had to make a binary choice between the two gender marked (M/F) articles in French and the two possessive determiners in English (his/her). So, if they used one of these two incorrectly, the prompts pushed them to choose the other one and they obviously helped learners to do this. However, the features under investigation in the present study were different and involved more complex decisions on the part of the learners. They would first have to decide which tense to use (the *passé composé* or the *imparfait*) and then, for example, if the *passé composé* was chosen it had to be used following the rule 'auxiliary + past participle'. The third step would be to choose the correct auxiliary, again following the rule depending on semantic aspects of the main verb. So, the whole 'process' for an L2 learner would involve at least three 'steps' (there are also gender and number agreement rules) – all of which, in spoken communication, happens online and there is no time available for the application of the pedagogic rules, even if known. The demands of unplanned spoken communication do not allow for the application of the pedagogic rules.

As already noted, the complex system of French past tenses in spoken language including aspectual distinctions could not be captured entirely in the current study, partly because of the level of interlanguage of the participating students and partly because of the features of French spoken language (in which the forms of the infinitive, past participle and *imparfait* often sound the same). Longitudinal research to date has provided evidence that (1) the development of temporal expression is slow and gradual, (2) the acquisition of form often precedes the function and (3) the irregular morphology precedes the regular morphology (Bardovi-Harlig, 2000; Kaplan, 1987). The use of the *imparfait* in the current study was limited to the verbs with inherent lexical aspects denoting states and only a few activities (Vendler, 1967). This is consistent with the research which has shown that some of the state verbs are usually learnt early as lexically bound chunks (Harley, 1989; Harley & King, 1989). Hence the results for the *imparfait* in oral production of the current study clearly suggest that

recasts facilitate improvement of the target-like use of the imparfait for only a limited, lexically bound set of verbs, i.e. they address ‘exemplar-based’ learning.

On the other hand, the passé composé which is perceived as pragmatically easier because of its single aspectual meaning, involves morpho-syntactic complexity that makes it more challenging in terms of structural formation. The results of the present study suggest that the learners improved their accuracy in both (1) the use of the passé composé in context and (2) the choice of the auxiliary associated with the passé composé. The scoring system rewarded knowledge of the distinction between the two auxiliaries (être and avoir) in French passé composé as one point was given for the correct auxiliary and no point was given if the auxiliary was incorrect. Based on such a scoring system, the results in the current study clearly indicate that the participants owed their higher scores to both the target-like use of the tense (depending on the context) and the accurate use of the auxiliary. There is evidence of improved accuracy in students’ utterances where on the pre-test a student said “il allé/-er, il fait” (inaccurate in the context of past activities) while on the immediate post-test the same student said “il est allé, elle s’est habillée, il a fait” (target-like use of the passé composé in the context of past activities). There is also evidence of the target-like choice of auxiliary after the treatment, where a learner who used the inappropriate auxiliary on the pre-test, e.g. “il a arrivé” used the target-like auxiliary on the post-test, e. g. “il est arrivé”.

Although the improved accuracy in the use and the form of the passé composé may suggest that the learners have acquired an abstract rule, in fact it is not possible to make such claims. Additional analyses would be needed to provide evidence for this, but they were beyond the scope of this thesis. Previous research has shown that at the beginner and early intermediate stages of L2 French there are more passé composé errors of form than of incorrect use and overall, greater accuracy has been observed with the use of the passé composé than with the use of the imparfait (Bardovi-Harlig, 2000; Harley, 1989, 1993; Kaplan, 1987). The passé composé is a morphologically complex structure which presents considerable challenges to L2 learners. Therefore, the main difference between the passé composé and the imparfait for the students in this study can be explained by the structural complexity of the two verb tenses, i.e. passé composé involves a morphologically complex structure whereas imparfait is essentially lexical in nature. Both, however, seem to be learnt as ‘chunks’, at least at the beginner and pre-intermediate stages. Kaplan (p. 56) noted that chunking or memorized holistic phrases may explain learners’ accurate use of passé composé in early production.

Classroom studies such as Harley (1989), Harley and King (1989), Myles, Hooper and Mitchell (1998) and Myles, Mitchell and Hooper (1999) have shown that those learners who had best memory for 'chunk' learning were also the ones who first succeeded in analysing and in creative use of the structures.

7.7.3. Between-group comparisons

Although there are strong arguments, as previously explained, for claiming that the input-providing quality of recasts is why the students in the RE group significantly improved over time and those in the CR group did not, the analyses of covariance (ANCOVA) with the number of corrections as a covariate, showed no group differences at each level of post-testing. The only exception was the imparfait delayed post-test in favour of recasts. Such results are similar to those of the Loewen and Nabei (2007) and McDonough (2007) studies, who also did not find any differences between the effects of recasts and clarification requests. These studies suggest the same or similar beneficial effects of both CF strategies.

Some other studies (e.g. Yang & Lyster, 2010; Lyster & Izquierdo, 2009; Ammar, 2008; Lyster & Mori, 2006; Ammar & Spada, 2006) have also looked at the effects of recasts but in comparison with prompts in general which were found to be superior to recasts because they overtly push learners to self-repair and also because they may be more explicit and less ambiguous than recasts. Yang and Lyster (2010) did find that recasts were effective but only for learning exemplar based English past tenses. They argued that prompts facilitate the learning of the rule-based regular past tense. Both recasts and prompts were found to be effective in Lyster and Mori's (2006) study which introduced the Counterbalance hypothesis. This states that the type of feedback and the type of instruction are counterbalanced depending on how much emphasis there is on accuracy and controlled production in a communicative classroom. Lyster and Izquierdo's (2009) study also found the same effect for recasts and prompts but in their study the feedback was provided individually, out of class. Finally, Ammar and Spada (2006) reported an effect for recasts but only for those learners who were of higher proficiency while prompts benefited all.

7.7.4. Counterbalance Hypothesis

The Counterbalance hypothesis may also be considered in order to explain why recasts were more effective in the present study. In Lyster and Mori's (2006) study the comparison

between two different settings, French immersion and Japanese immersion, indicated that there was a relationship between the type of instruction and the type of feedback. The hypothesis claims that the effect of the feedback depends on the counterbalance between the instruction and the feedback. For example, as Lyster and Mori argue, in the Japanese classroom where there was more emphasis on controlled production and more attention paid to accuracy, recasts as an implicit and naturalistic CF strategy were effective, whereas in Canadian French immersion classrooms which are more naturalistic and experiential, with no emphasis on accuracy and controlled production, prompts as the more explicit strategy were found to be more effective. With regard to the present study it should be noted that the participants in the study were in foreign language classrooms where attention is paid to accuracy, sometimes through controlled practice which may take different forms - although not of choral repetition or speaking in isolation as in the Lyster and Mori Japanese immersion classroom. So, even though the overall orientation of the classes in the present study was communicative, there were expectations that the language had to be used accurately. Thus, as in Lyster and Mori's Japanese immersion classroom, recasts may have been more effective for this reason.

7.7.5. Effects of starting proficiency

The answer to RQ 2 which looked at the effect of the starting proficiency while controlling for the different number of CF episodes by using an ANCOVA at each post-test, suggests that the more proficient learners who received recasts were able to outperform their peers in the group that received clarification requests. The analyses of covariance with the CF episodes as covariates revealed that the 'high' proficiency RE group outperformed the 'high' CR and the 'high' CN group on both post-tests for the passé composé and on post-test 2 for the imparfait. This supports the claim that it is the type rather than the frequency of corrections that was more important in bringing about the effectiveness of the corrective feedback.

Actually, the results suggesting beneficial effects of recasts for higher proficiency learners are not surprising when considering research on oral CF as a whole. The role of recasts in assisting those learners who are developmentally ready to acquire a specific target feature has been established in a number of laboratory studies (Han, 2002; Iwashita, 2003; Leeman, 2003; Mackey & Philp, 1998; Philp, 2003). There is also classroom research reporting the effectiveness of recasts for higher proficiency learners, for example, Ammar and Spada

(2006). In L1 acquisition, as Farrar (1990) contends, the child's readiness may have significant influence and determine whether the recasts will be effective. In other words, if a learner is at the stage of development when he or she is able to 'assign' the meaning to a particular grammatical morpheme then he/she will benefit from the recast. While higher proficiency learners in the present study benefited from recasts as they were probably developmentally ready to do so, higher proficiency learners in the CR group did not benefit from the feedback they received despite the fact they were presumably at the same level of development – because the feedback did not provide them with any positive evidence. The results were slightly different for the imparfait. The higher proficiency learners in the RE group outperformed the higher proficiency learners in both the CN and the CR group only on post-test 2, but the lower proficiency learners in the RE group also outperformed the lower CN group.

Thus the main difference in the effects of proficiency on the two language structures is that whereas it was only the 'high' proficiency learners in the RE group who were superior to the other two groups in the case of passé composé, it was both the 'low' and the 'high' proficiency learners in the RE group in the case of the imparfait. A possible explanation for this lies in the difference between the two target structures. As mentioned previously, the passé composé presents a formally complex structure. On the other hand, the imparfait has a simpler form and was limited to only a set of stative and activity/durative verbs. It may be that the lower proficiency learners in the RE group did not outperform their peers (in the other two groups) for passé composé because of its formal complexity but succeeded in acquiring the limited set of imparfait verbs because of their simpler form and lexically bound nature.

7.7.6. Concluding comments

The current study suggests that it is the quality, rather than the quantity of corrections that makes CF effective in the classroom. It shows the superiority of recasts as a type of corrective feedback which simultaneously provides implicit negative feedback and explicit positive evidence. The result of such a brief dialectical process is implicit negative evidence that may assist learners in the acquisition of both lexically bound 'chunks' and morphologically complex structures. In this respect the current findings lend support to Long's (2007) claims regarding the exceptional value of recasts in language learning. The use

of corrective recasts by parents addressing their children's errors, as reported and discussed in Nelson (1987), Farrar (1990, 1992) and Saxton (1997, 2010) suggests that children do attend to recasts and subsequently repair their utterances. If children, when acquiring their first language, attend to recasts in a meaningful conversation, then in a similar way corrective recasts may work for second language acquisition even though the processes in L1 and L2 acquisition are not entirely the same. As Long (1996) argued, recasts may induce those processes which are crucial for L2 learning, such as attention and noticing (with or without awareness). The present study shows that such processes may happen in foreign language classrooms. Furthermore, because of their nature containing positive evidence, recasts contribute to the frequency of the target input and this makes them very different from other types of implicit negative feedback. At the level of participants in the current study and for the acquisition of the target language structures, clarification requests were not sufficient enough to lead to successful modified output and learning gains which could be observed in unplanned semi-constrained oral production.

Chapter 8. Written production results

8.1. Introduction

This chapter considers research questions 3 and 4 using the written production data:

3. Does oral corrective feedback (recasts and requests for clarification) have an effect on L2 learners' acquisition of French linguistic structures (the passé composé and the imparfait past tenses) as measured in their written production?
 - a) Do recasts have an effect on the L2 acquisition of passé composé and imperfect as measured in written production?
 - b) Do requests for clarification have an effect on the L2 acquisition of passé composé and imparfait as measured in written production?
4. Does the learners' starting proficiency mediate the effect of the oral corrective feedback on the learners' acquisition of: (a) the passé composé and b) the imparfait, as measured in written production?

8.2. Data Analysis

In order to analyse the written data the same procedures were followed as for the analysis of oral production data.

8.2.1. Descriptive statistics

Descriptive statistics for each structure (passé composé and imparfait) were calculated.

8.2.2. Inferential statistics

a) In order to answer Research question 3, a mixed design Repeated measures ANOVA was calculated using the Bonferroni adjustments for multiple comparisons to examine the change that occurred in each group after the pre-test (within group comparisons). Separate Repeated measures ANOVAs were calculated for the passé composé and for the imparfait.

b) The between-group differences were looked at separately. First the differences between the CN group and each of the treatment groups were examined using the Independent samples t-test.

c) For comparison of the two experimental groups, which had to take into account the differences between the number of corrections each group received, a series of ANCOVAs were used, with the CF episodes as a covariate.

8.2.3. Data analysis for RQ 4

In order to answer Research question 4, the participants in each group were divided into 'high' and 'low' based on their pre-test scores. For passé composé the median score for the whole sample was 45.77. For each group, all learners above the median were placed in the 'high' proficiency sub-group and all those below in the 'low' proficiency sub-group. For imparfait the median score for the whole sample was 20.00. Two participants whose mean score on the imparfait pre-test was 20.00 were both assigned to the lower group since the mean for the whole sample was 22.89. Therefore it seemed logical that those whose score was below the mean would be considered of lower proficiency.

Once the 'low' and 'high' sub-groups were formed, an ANCOVA was performed at each level of post-testing for the four treatment sub-groups (RE-High, CR-High, RE-Low, CR-Low) using CF episodes as a covariate. A series of separate ANOVAs with post-hoc tests were used to examine the differences between each treatment group's 'low' and 'high' and the CN group's 'low' and 'high' proficiency participants, at each level of post-testing.

Effect sizes were calculated in the same way as for oral production.

8.3. Results for RQ 3

8.3.1. Results for passé composé: descriptive statistics

Table 20 and Figure 7 show the descriptive data for written production of the passé composé past tense. On the pre-test the RE group had a mean percentage score of 35.02 (SD=30.05), the CR group had a mean percentage score of 45.51 (SD=28.08) whilst the CN group had a mean of 36.55 (SD=26.33). An analysis of variance (ANOVA) showed no statistically significant differences among the three groups at the time of pre-testing:

$F(2, 47) = .681, p = .511 (p > .05)$

As indicated in Table 20 and Figure 7, on the immediate post-test (post-test 1) the RE group achieved a mean percentage score of 50.19 (SD=35.67), the CR group obtained a mean percentage score of 47.02 (SD=33.74) and the CN group's mean decreased to 34.19 (SD=24.65). On the delayed post-test (post-test 2) the RE group accuracy score increased again achieving a mean percentage score of 61.88 (SD=30.31), the CR group increased with a mean of 49.80 (SD=32.12) and the CN group just slightly decreased to 34.13 (SD=27.98). The standard deviations for all three groups in all the tests were high, indicating considerable within group variation.

Table 20 Descriptive statistics for passé composé in written production

GROUP	Pre-test		Post-test 1		Post-test 2	
	Mean	SD	Mean	SD	Mean	SD
RECAST (N=17)	35.02	30.05	50.18	35.67	61.88	30.31
CL.REQUEST(N=17)	45.51	28.08	47.02	33.74	49.80	32.12
CONTROL (N= 16)	36.55	26.33	34.19	24.65	34.13	27.98

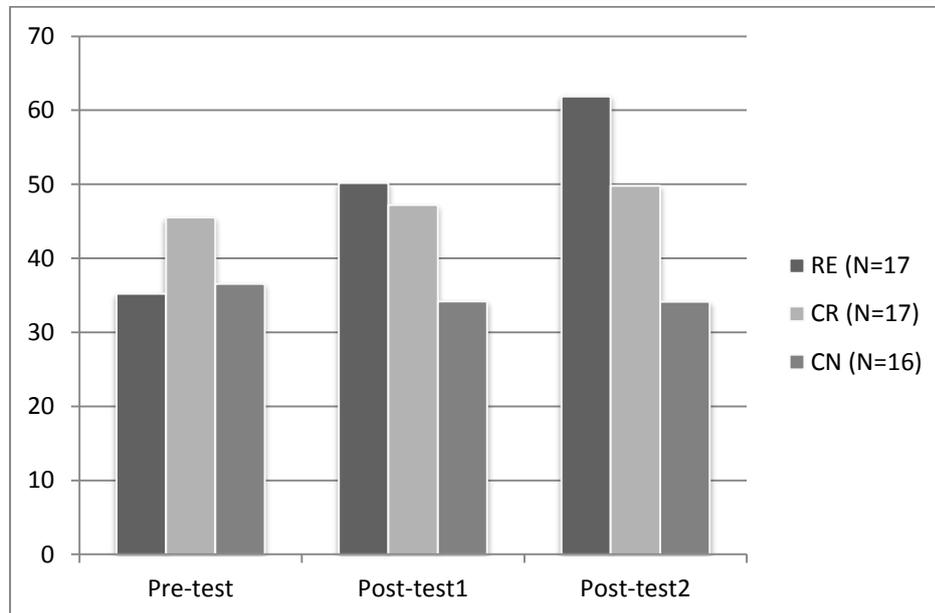


Figure 7 Mean scores for passé composé: the RE group, the CR group and the CN group

Taking into consideration the different number of corrective feedback episodes each treatment group received (see Table 4 in previous chapter) an ANCOVA with the covariate CF episodes (evaluated at CF episodes for passé composé = 1.4) was calculated at each level of post-testing. This produced adjusted mean scores for each treatment group on post-test 1

(immediate) and post-test 2 (delayed). Table 21 presents the adjusted mean scores of the oral passé composé for two treatment groups on post-test 1 and post-test 2.

Table 21 Adjusted mean scores for the RE and the CR group on post-test 1 and post-test 2

GROUP	Post-test 1		Post-test 2	
	Mean	St.Err.	Mean	St.Err.
RECAST (N=17)	47.91	7.78	60.63	7.47
CL.REQUEST(N=17)	49.29	7.78	51.05	7.47

Table 22 and Figure 8 show the results of a Repeated measures ANOVA which was used in order to examine the change over time in all three groups (within group comparison). This analysis is based on the original scores of each group, without taking into consideration the number of CF episodes [see Note 1]. The analysis revealed the following results using the Bonferroni adjustments for multiple comparisons:

For the RE group there is a statistically significant difference between the pre-test and the immediate post-test ($p = .038$) with a medium effect size ($ES = .46$) and between the pre-test and the delayed post-test ($p = .000$), with a large effect ($ES = .89$). The difference between the two post-tests (post-test 1 and post-test 2) approaches a significant level with a small effect size ($ES = .35$). In the CR group there is no evidence of statistically significant differences between any of the post-tests compared with the pre-test, nor between the post-tests themselves. Effect sizes are small (from pre-test to post-test1 $ES = .05$, from pre-test to post-test2 $ES = .14$, from post-test1 to post-test2 $ES = .08$). The same results, showing no statistically significant difference between any of the tests, are evident for the CN group.

Table 22 Results of Repeated measures ANOVA: within group differences

	RE group		CR group		CN group	
	p	ES	p	ES	p	ES
Pre-test - Post-test 1	.038	.46	>.05	.05	>.05	-.09
Pre-test - Post-test 2	.000	.89	>.05	.14	>.05	-.09
Post-test 1 - Post-test2	0.61	.35	>.05	.08	>.05	.00
	df=16		df=16		df=15	

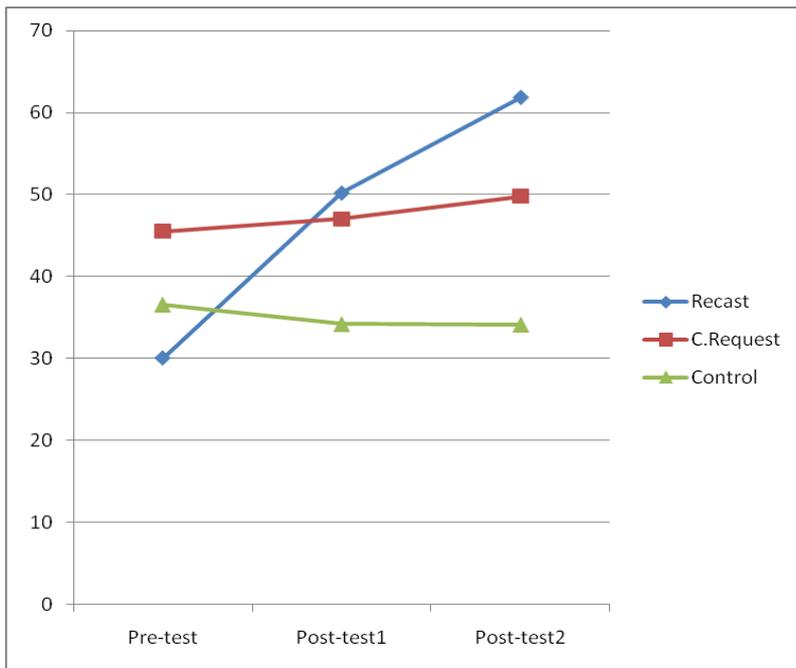


Figure 8 Repeated measures ANOVA: performance of three groups over time

b) Since it was assumed that the number of CF episodes may have affected the performance on post-tests, separate analyses were carried out in order to compare each treatment group with the CN group at each level of post-testing,

The results of a t-test between the RE and the CN group show the following:

t-test on post-test 1: $t(31) = 1.25$, ns, ES = .22

t-test on post-test 2: $t(31) = 2.78$, $p = .010$, ES = .45

A comparison between the CR and the CN group shows:

t-test on post-test 1: $t(31) = 1.24$, ns, ES = .22

t-test on post-test 2: $t(31) = 1.49$, ns, ES = .26

These t-tests show no significant difference between the RE and the CN group on post-test 1 but there is a significant difference between these two groups on post-test 2.

Between the CR and the CN group there are no significant differences on the two post-tests.

c) In order to examine the differences between the RE group and the CR group, an ANCOVA was computed at each level of post-testing using the covariate CF episodes for the passé composé. With the covariate CF episodes evaluated at $CF = 1.4$ the tests of Between Subjects Effects show no significant differences between the two groups: $F(1, 31) = .016$, $p > .05$, ES = .02

On post-test 2 with the covariate CF episodes evaluated at $CF = 1.4$ the tests of Between Subjects Effects again show no significant differences between the two treatment groups:

$F(1, 31) = .812, p > .05, ES = .15$

8.3.2. Results for imparfait

Table 23 and Figure 9 present the descriptive data for written production of the imparfait past tense. On the pre-test the RE group had a mean percentage score of 25.52 (SD=26.17), the CR group had a mean percentage score of 16.01 (SD=20.72) whilst the CN group's mean was 27.41 (SD=25.77).

An analysis of variance ANOVA, performed for the three pre-tests, showed no statistically significant differences among the three groups at the time of pre-testing:

$F(2, 47) = 1.056, p = .356 (p > .05)$

As indicated in Table 23 and Figure 9, on the immediate post-test (post-test 1) the RE group achieved a mean percentage score of 54.29 (SD=35.67), the CR group obtained a percentage mean of 29.20 (SD=23.44) and the CN group decreased to a mean of 22.71 (SD=27.57). On the delayed post-test (post-test 2) the RE group increased its accuracy with a mean score of 64.99 (SD=26.95), the CR group increased to 37.84 (SD=26.02) and the CN group decreased again to a mean score of 20.65 (SD=25.10). The standard deviations for all three groups in all the tests were high, indicating considerable within group variation.

Table 23 Descriptive statistics for imparfait

GROUP	Pre-test		Post-test 1		Post-test 2	
	Mean	SD	Mean	SD	Mean	SD
RECAST (N=17)	25.52	26.17	54.28	35.67	64.99	26.94
CL.REQUEST(N=17)	16.01	20.72	29.20	23.44	37.84	26.02
CONTROL (N=16)	27.41	25.77	22.71	27.57	20.65	25.10

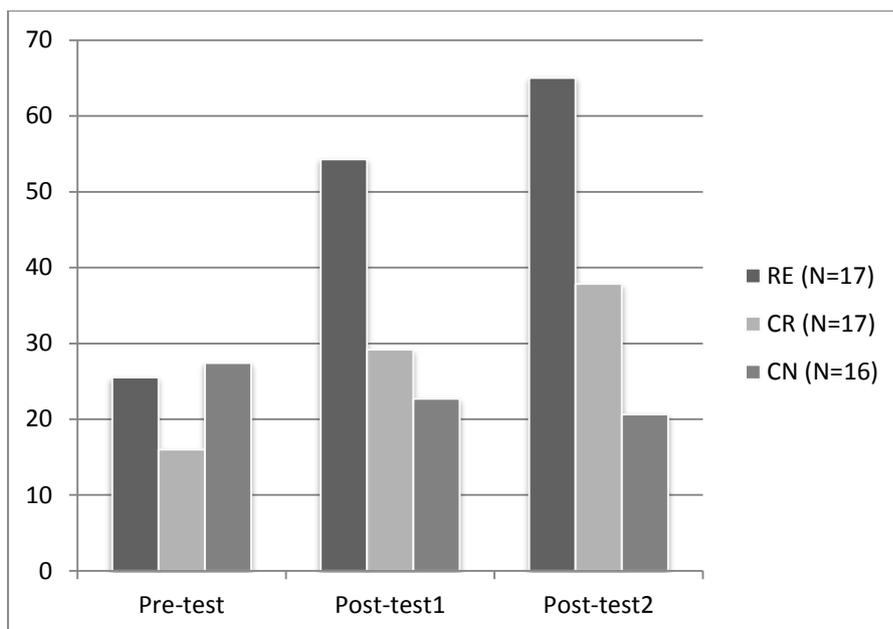


Figure 9 Mean scores for imparfait in written production

Taking into consideration the different number of corrective feedback episodes each treatment group received (see Table 8 in previous chapter), an ANCOVA with the covariate CF episodes was calculated at each level of post-testing for the written imparfait. This produced the adjusted mean scores for each treatment group on post-test 1 (immediate) and post-test 2 (delayed). Table 24 presents the adjusted mean scores of the written imparfait for two treatment groups on post-test 1 and post-test 2.

Table 24 Adjusted mean scores for RE group and CR group on post-test 1 and post-test 2

GROUP	Post-test 1		Post-test 2	
	Mean	St.Err.	Mean	St.Err.
RECAST (N=17)	53.91	7.5	63.51	6.3
CL.REQUEST(N=17)	29.57	7.5	39.32	6.3

Table 25 and Figure 10 show the results of a Repeated measures ANOVA which was used in order to examine the change over time in each of the three groups (within group comparison). This analysis is based on original scores, without taking into account the number of corrective feedback episodes each group received. The analysis revealed the following results using the Bonferroni adjustments for multiple comparisons:

For the RE group there is a statistically significant difference between the pre-test and the immediate post-test ($p = .000$) with a large effect ($ES = .93$), and between the pre-test and the delayed post-test ($p = .000$) with a large effect ($ES = 1.49$). There is no significant difference between the two post-tests, the immediate and the delayed ($p > .05$) and the effect size is small ($ES = .34$).

In the CR group there is no evidence of a statistically significant difference between the pretest and the immediate post-test ($p > .05$) but the effect size is medium ($ES = .59$). Moreover, there is a significant improvement between the pre-test and the delayed post-test ($p = .003$), with a large effect size ($ES = .93$). The difference between the two post-tests (post-test 1 and post-test 2) is not significant ($p > .05$) and the effect size is small ($ES = .37$). The CN group did not show any significant differences between any of the post-tests and the pre-test nor between the two post-tests, with very small, even negative effects.

Table 25 Results of Repeated measures ANOVA: within group differences

	RE group		CR group		CN group	
	p	ES	p	ES	p	ES
Pre-test - Post-test 1	.000	.93	>.05	.59	>.05	-.18
Pre-test - Post-test 2	.000	1.49	.03	.93	>.05	-.27
Post-test1 - Post-test2	>.05	.34	>.05	.37	>.05	-.08
	df=16		df=16		df=15	

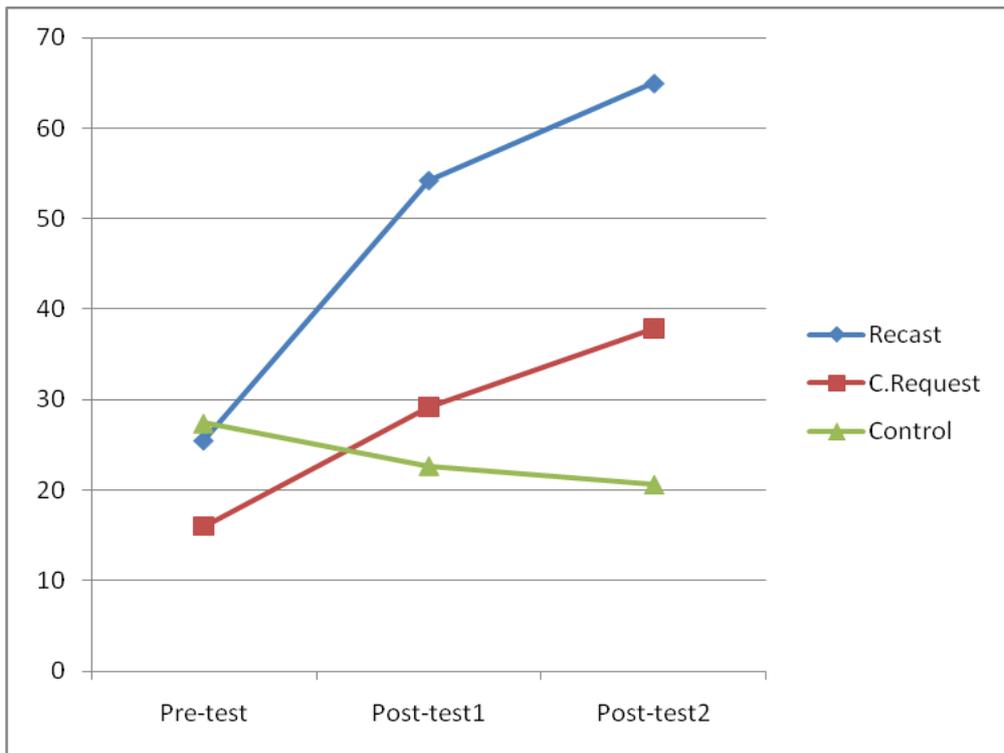


Figure 10 Repeated measures ANOVA: performance of three groups over time

Since it was assumed that the number of CF episodes may have affected the performance of the two treatment groups, separate analyses were computed in order to compare each treatment group with the CN group at each level of post-testing.

The results of a t-test for comparison between the RE and the CN group show:

t-test on post-test 1: $t(31) = 2.85$, $p = .008$, $ES = .46$

t-test on post-test 2: $t(31) = 4.88$, $p = .000$, $ES = .66$

A comparison between the CR and the CN group shows:

t-test on post-test 1: $t(31) = .73$, ns, $ES = .13$

t-test on post-test 2: $t(31) = 1.93$, $p = .063$, $ES = .33$

These results indicate that for the imparfait in written production there is a significant difference between the RE and the CN group on post-test 1, with a medium effect size, and a significant difference between these two groups on post-test 2, also with a medium effect. Between the CR and the CN group there is no significant difference on post-test 1 but the difference on post-test 2 approaches significance ($p = .063$).

In order to examine the differences for written imparfait between the RE group and the CR group taking into consideration the number of CF episodes, an ANCOVA was computed at each level of post-testing using CF episodes for the imparfait as a covariate. With the covariate CF episodes evaluated at CF = 1.7, the result on the immediate post-test shows a significant effect for the RE group: $F(1, 31) = 5.15, p = .030, ES = .38$. On the delayed post-test, with the covariate CF episodes evaluated at CF = 1.7, the test again shows a statistically significant effect for the RE group: $F(1, 31) = 7.14, p = .012, ES = .43$.

8.4. Summary

A Summary of written production results for Research question 3, as displayed in Table 26 and Table 27, indicates that regarding the within group (between tests) results for the passé composé target structure, only the participants in the RE group have significantly improved from pre-test to post-test 1 and from pre-test to post-test 2. For the imparfait target structure the RE group have significantly improved from pre-test to post-test 1, and from pre-test to post-test 2, but the participants in the CR group have also significantly improved from pre-test to post-test 2.

When the different number of CF episodes has been accounted for, there are no significant differences between the two treatment groups (RE and CR) on either of the two post-tests for the passé composé (immediate and delayed). There are also no statistically significant differences for the passé composé between the RE and the CN group on post-test 1 but on post-test 2 the RE group outperformed the CN group. There are no significant differences between the CR and the CN group on either of the post-tests.

For the imparfait target structure, when the number of CF episodes is accounted for, there are statistically significant differences between the RE and the CR group on both post-tests. Comparison of each treatment group separately with the Control group shows a significant difference between the RE and the CN group on both post-test 1 and post-test 2. There are no significant differences between the CR and the CN group on the immediate post-test for written imparfait but on the delayed post-test there is a close to significant difference between these two groups.

Table 26 Summary of results for RQ 3: Within group differences

GROUP	Passé composé	Imparfait
RECAST	Post-test 1 > Pre-test (p = .038)	Post-test 1 > Pre-test (p = .000)
	Post-test 2 > Pre-test (p = .000)	Post-test 2 > Pre-test (p = .000)
	Post-test 2 = Post-test 1 (p > .05)	Post-test 2 = Post-test 1 (p > .05)
CL. REQUEST	Post-test 1 = Pre-test (p > .05)	Post-test 1 = Pre-test (p > .05)
	Post-test 2 = Pre-test (p > .05)	Post-test 2 > Pre-test (p = .003)
	Post-test 2 = Post-test 1 (p > .05)	Post-test 2 = Post-test 1 (p > .05)
CONTROL	Post-test 1 = Pre-test (p > .05)	Post-test 1 = Pre-test (p > .05)
	Post-test 2 = Pre-test (p > .05)	Post-test 2 = Pre-test (p > .05)
	Post-test 2 = Post-test 1 (p > .05)	Post-test 2 = Post-test 1 (p > .05)

*The mean difference is significant at the .05 level

Table 27 Summary of results for RQ 3: Between group differences

	Post-test 1	Post-test 2
Passé composé	Recast=Cl.Request (p > .05)	Recast=Cl.Request (p > .05)
	Recast = Control (p > .05)	Recast > Control (p = . 010)
	Cl. Request = Control (p > .05)	Cl. Request = Control (p > .05)
Imparfait	Recast >Cl.Request (p=.030)	Recast >Cl.Request (p =.012)
	Recast > Control (p = .008)	Recast > Control (p = .000)
	Cl. Request = Control (p > .05)	Cl. Request=Control (p = .063)

*The mean difference is significant at the .05 level

8.5. Results for RQ 4

8.5.1. Results for passé composé

Table 28 displays the descriptive data for written production of the passé composé ‘low’ and ‘high’ proficiency sub-groups and Table 29 shows the adjusted mean scores on post-tests for the treatment sub-groups once the covariate CF episodes have been accounted for. Table 30 presents the effect sizes for each ‘low’ and ‘high’ sub-group, as observed between tests over time.

Table 28 Descriptive statistics for 'low' and 'high' sub-groups on each test of passe compose

SUB-GROUP	Pre-test		Post-test 1		Post-test 2	
	Mean	SD	Mean	SD	Mean	SD
RE Low (N=9)	10.00	13.64	23.47	23.76	42.78	29.08
RE High (N=8)	63.18	11.83	80.24	17.71	83.36	11.70
CR Low (N=7)	19.20	23.98	36.59	37.07	34.49	37.70
CR High (N=10)	63.93	10.15	54.32	31.06	57.71	26.79
CN Low (N=10)	18.84	12.50	22.59	18.38	24.34	18.83
CN High (N=6)	66.07	11.27	53.53	22.28	50.45	34.62

Table 29 Adjusted mean scores for 'low' and 'high' sub-groups on each post-test

GROUP	Post-test 1		Post-test 2	
	Mean	St.Err	Mean	St.Err
RE - Low (N=9)	25.48	9.2	43.43	9.4
RE - High (N=8)	75.19	10.1	81.74	10.3
CR - Low (N=7)	40.62	10.6	39.79	10.8
CR - High (N=10)	53.72	8.6	57.52	8.8

Table 30 Effect sizes for each sub-group, observed between tests over time

	Effect size					
	RE- Low	RE-High	CR-Low	CR-High	CN-Low	CN-High
Pre-test - Post-test 1	0.72	1.15	.57	-.46	.24	-.75
Pre-test - Post-test 2	1.54	1.76	.49	-.33	.35	-.68
Post-test1 - Post-test 2	.73	.22	-.07	.11	.09	-.11

8.5.1.1. Tests comparing each treatment sub-groups with the Control sub-groups

An ANOVA on post-test 1 for passé composé in written production, looking at the differences between the RE-Low, CN- Low, RE-High and the CN-High sub-groups reveals a significant difference among the four sub-groups: $F(3, 29) = 15.29, p = .000$

The post-hoc tests performed for these four sub-groups show that there is a significant difference between the RE-High and the CN-High ($p = .023$) with a large effect ($ES = 1.36$) but there are no significant differences between the RE-Low and the CN-Low sub-group ($p > .05$)

On post-test 2 for the same four sub-groups an ANOVA indicates a significant difference again: $F(3, 29) = 9.08, p = .000$

The post-hoc tests reveal a significant difference between the RE-High and the CN-High ($p = .017$) with a large effect ($ES = 1.55$), whereas there are no significant differences between the RE-Low and the CN-Low sub-groups ($p > .05$).

The same analyses performed among the four sub-groups of the CR and the CN group (CR-Low, CN-Low, CR-High, CN-High) show the following:

On post-test 1 an ANOVA among the four sub-groups shows no significant result:

$F(3, 29) = 2.68, p = .065, ns$

On post-test 2 for the same sub-groups an ANOVA again shows no significant result:

$F(3, 29) = 2.44, p = .084, ns$

8.5.1.2. Tests comparing the treatment sub-groups

The result of the Tests Between Subjects of an ANCOVA performed on post-test 1 among the four treatment sub-groups (RE-Low, CR-Low, RE-High, CR-High), taking into account CF episodes as a covariate, shows a significant difference among the four groups: $F(3, 29) = 4.43, p = .011$

The follow-up pair-wise comparisons show no significant differences between the RE-High and the CR-High sub-group ($p > .05$), and no difference between the RE-Low and the CR-Low sub-groups ($p > .05$).

The result of the Tests Between Subjects Effects of an ANCOVA on post-test 2 (RE-Low, CR-Low, RE-High, CR-High), taking into consideration CF episodes as a covariate, show a significant difference among the four sub-groups again:

$$F(3, 29) = 3.15, p = .040$$

However, the follow-up pair-wise comparisons indicate no significant difference between the RE-High and the CR-High sub-groups ($p > .05$) and no difference between the RE-Low and CR-Low sub-groups ($p > .05$)

8.5.2. Results for imparfait

Table 31 displays the descriptive data for written production of the imparfait ‘low’ and ‘high’ proficiency sub-groups, whereas Table 32 shows the adjusted mean scores for the same sub-groups when the covariate CF episodes have been accounted for. Table 33 shows the effect sizes for each ‘low’ and ‘high’ sub-group, indicating the change that occurred between tests (from pre-test to post-test 1 and to post-test 2).

Table 31 Descriptive statistics for 'low' and 'high' sub-groups on each test of imparfait

SUB-GROUP	Pre-test		Post-test 1		Post-test 2	
	Mean	SD	Mean	SD	Mean	SD
Recast Low (N=9)	4.31	8.55	36.24	29.77	52.82	30.79
Recast High (N=8)	49.39	15.86	74.58	31.71	78.68	13.10
Cl.Request Low (N=10)	1.25	3.95	23.88	23.66	32.17	28.24
Cl. Request High (N=7)	37.11	15.45	36.80	22.60	45.94	21.88
Control Low (N=7)	4.90	8.52	9.18	18.77	5.95	10.45
Control High (N=9)	44.92	20.07	33.23	29.60	32.08	27.63

Table 32 Adjusted mean scores for 'low' and 'high' sub-groups on each post-test

GROUP	Post-test 1		Post-test 2	
	Mean	St.Err	Mean	St.Err
Recast-Low (N=9)	35.22	9.3	49.96	8.2
Recast-High (N=8)	74.51	9.7	78.48	8.5
Cl.Request-Low (N=10)	25.03	8.9	35.39	7.8
Cl.Request-High(N=7)	36.55	10.4	45.25	9.1

Table 33 Effect sizes for each sub-group, observed between tests

	Effect size					
	RE-Low	RE-High	CR-Low	CR-High	CN-Low	CN-High
Pre-test - Post-test 1	1.76	1.06	.88	-.07	.31	-.47
Pre-test - Post-test 2	2.46	2.02	1.92	.47	.11	-.54
Post-test1 - Post-test 2	.55	.18	.32	.41	-.22	-.04

8.5.2.1. Tests comparing each treatment sub-group with the control sub-groups

An ANOVA on post-test 1 for written imparfait, looking at the differences between the RE-Low, CN-Low, RE-High and the CN-High sub-groups reveals a significant difference among the four sub-groups: $F(3, 29) = 6.95, p = .001$

The post-hoc tests performed for these four sub-groups show that there is a significant difference between the RE-High and the CN-High ($p = .033$), with a large effect size ($ES = 1.35$). There is no significant difference between the RE-Low and the CN-Low sub-group ($p > .05$)

On post-test 2 for the same four sub-groups an ANOVA again indicates that there is a significant difference: $F(3, 29) = 13.51, p = .000$

The post-hoc tests reveal a significant difference between the RE-High and the CN-High ($p = .002$) with a large effect ($ES = 2.24$) and a significant difference between the RE-Low and the CN-Low sub-groups ($p > .002$) with a large effect ($ES = 2.12$).

An ANOVA on post-test 1, performed for the four sub-groups of the CR and the CN group (CR-Low, CN-Low, CR-High, CN-High) shows no significant differences, but the same analysis on post-test 2 of the written imparfait indicates a significant result: $F(3, 29) = 3.39, p = .031$. Although at the level of chance, the effect of the CR-Low over CN-Low is large ($ES = 1.24$). There are no significant differences between the two ‘high’ sub-groups.

8.5.2.2. Tests comparing the treatment sub-groups

The result of the tests Between Subjects Effects of an ANCOVA performed on post-test 1 among the four treatment sub-groups (RE-Low, CR-Low, RE-High, CR-High), taking into

account CF episodes as a covariate, indicates a significant difference among the four sub-groups: $F(3, 29) = 5.22, p = .005$.

The follow-up pair-wise comparisons show a significant result between the RE-High and the CR-High sub-groups ($p = .012$) with a large effect ($ES = 1.37$) but there are no differences between the RE-Low and the CR-Low sub-groups.

The result of the tests Between Subjects Effects of an ANCOVA on post-test 2 for written imparfait (RE-Low, CR-Low, RE-High, CR-High), taking into account CF episodes as a covariate, again shows a significant difference among the four groups:

$F(3, 29) = 4.93, p = .007$

The pair-wise comparisons indicate a significant difference between the RE-High and the CR-High sub-groups ($p = .013$) with a large effect size ($ES = 1.91$), but no significant difference between the RE-Low and the CR-Low sub-groups ($p > .05$).

8.6. Summary

In summary, when each group is divided into ‘low’ and ‘high’ according to the score on the pre-test, the results show that in written production for passé composé the ‘high’ RE sub-group outperformed the ‘high’ CN sub-group but there were no significant differences between the ‘low’ sub-groups of the RE and CN. Also, there were no significant differences between the two ‘high’ and the two ‘low’ of the RE and the CR group. The analyses were performed in the same way as for the three main groups – so that for the differences between the treatment groups an ANCOVA was computed accounting for the number of CF episodes, and for the differences between each treatment group and the CN group an ANOVA was used at each level of post-testing,

For the imparfait in written production, on post-test 1 the ‘high’ RE group outperformed both the ‘high’ CN and the ‘high’ CR group but no differences were evident between any of the ‘low’ sub-groups. On post-test 2, the results are similar, i.e. the ‘high’ RE sub-group outperformed both the ‘high’ CN and the ‘high’ CR sub-groups. Also, on post-test 2 the ‘low’ RE sub-group outperformed the ‘low’ CN sub-group. The summary of results for Research question 4 referring to written production is presented in Table 30.

Table 34 Summary of results for RQ 4

	Post -test 1	Post -test 2
PASSE COMPOSE	RE High= CR High (ns)	RE High=CR High (ns)
	RE Low=CR Low (ns)	RE Low= CR Low (ns)
	RE High > CN High(p=.023)	RE High > CN High (p= .017)
	RE Low= CN Low (ns)	RE Low = CN Low (ns)
	CR High=CN High (ns)	CR High = CN High (ns)
IMPARFAIT	CR Low= CN Low (ns)	CR Low = CN Low (ns)
	RE High>CR High (p=.012)	RE High> CR High (p=.013)
	RE Low = CR Low (ns)	RE Low = CR Low (ns)
	RE High >CN High (=0.033)	RE High > CN High (p= .002)
	RE Low = CN Low (ns)	RE Low > CN Low (p = .002)
	CR High = CN High (ns)	CR High = CN High (ns)
	CR Low = ControlLow (ns)	CR Low = CN Low (ns)

*The mean difference is significant at the .05 level

8.7. Discussion

The third and the fourth research question were concerned with the effects of recasts and requests for clarification on the acquisition of the two French structures, the passé composé and the imparfait when measured by written narrative tasks. Written production differs from oral production in that the learners have more time at their disposal for both the formulation of messages and the editing of them using explicit knowledge. Thus higher scores can be expected in the written tests. This is evident when the mean scores in the various oral and written tests are compared. For example, the mean score of the whole sample on the oral pre-test for passé composé was 33.85 %, while for passé composé in written production the mean was 39.08 %. For the imparfait in oral production the mean of the whole sample on pre-test was 17.43 %, whilst for the imparfait measured in written tasks the mean was 23.00 %. A comparison of the oral and written scores on the two post-tests reveals similar differences.

Statistically significant results for the oral and written data show the following: Five out of six of the time comparisons conducted on the oral data for the whole RE group were statistically significant. Four were significant for the written data. For the CR group, none of the time comparisons were significant in the oral data, whereas one (between the pre-test and post-test 2 for the imparfait) was significant for the written data. These results indicate that although similar, the performance of the RE group was slightly better in oral production whereas the CR group showed evidence of significant gains only in written production. The

whole group comparisons with the CN group also reveal some differences between the oral and written data but only for recasts. For the oral data, there was evidence that the RE group outperformed the CN group on post-test 2 for imparfait, whereas for the written data the RE group outperformed the CN group on both post-test 1 and post-test 2 for imparfait and also on post-test 2 for passé composé.

Overall, the results for written production again point to the superiority of recasts in assisting the learners to achieve greater accuracy in the use of the two tenses. Recasting learners' oral errors clearly had a greater effect than requesting clarification. The reasons for the superiority of recasts were discussed in the previous chapter. There I argued that the superiority cannot be explained by the fact that the RE group received more corrections as even when the effect of the number of corrections was accounted for by the use of an ANCOVA recasts were shown to have a greater effect than clarification requests. All in all, the results for written production provide further evidence that the input-providing nature of recasts assists acquisition to a greater extent than the output-prompting nature of clarification requests. In this respect there is further evidence that the instructional context under investigation (i.e. a high school foreign language classroom) in comparison with the immersion context investigated by Lyster (2004) favours the use of recasts. I argued that this might be because for learners of this kind it is the noticing and noticing-the-gap that recasts facilitate (R. Ellis, 1992; Nicholas, Lightbown, & Spada, 2001; Schmidt, 1990) that is most important for learning.

The results for the comparisons involving 'high' and 'low' sub-groups of the RE, CR and the CN group are similar for the oral and written data. However, there is evidence of the improved performance of the CR group on written tests, and the 'high' RE sub-group did not outperform the 'high' CR sub-group on either of the two passé composé post-tests.

It is also worth commenting briefly on the difference between the results for the two verb tenses. The two main differences between oral and written results are (1) for the imparfait the RE group outperformed both the CR and the CN groups on post-test 1 as well as on post-test 2 in written production but only in post-test 2 in the oral production and (2) for the passé composé the RE group outperformed the CN group in written post-test 2 but not in oral post-test 2. I have suggested that the two structures involve different levels of complexity – morphologically complex structure in the case of the passé composé and morphologically simpler, lexically bound structure in the case of the imparfait. Recasts proved effective for

learning of both structures at the level of participants' interlanguage. Recasts may be particularly beneficial for learning morphologically complex structures because of their input-providing nature which can also be facilitative of managing the processing load (Skehan, 1998) of such complex structures. Requests for clarification, as the results show, are beneficial for the acquisition of a morphologically simpler structure. Furthermore, recasts proved their effectiveness in both oral and written production whereas clarification requests were clearly more effective in written production which gives greater opportunity for the application of explicit knowledge and allows greater control over performance.

A general conclusion is that although a number of differences can be seen in the oral and written results, the overall picture is clear: irrespective of how learning is measured, foreign language learners of French in a high school classroom benefit to a greater extent from corrective feedback involving recasts than from clarification requests.

Note: In Repeated measures ANOVA, which was used to examine the within group (between tests) effects over time, the covariate CF episodes was not used because (a) the Control group did not receive any treatment and (b) the within group effect over time is measured independently for each group.

Chapter 9. Uptake with repair and learning gains

9.1. Introduction

This chapter considers Research question 5:

1. To what extent do learners repair their errors, following correcting feedback?
2. Do learners who produce uptake with repair demonstrate greater gains in accuracy than learners who were just exposed to uptake with repair?

Uptake with repair was operationalised as a learner's utterance that immediately follows the teacher's feedback and contains repair of the error which initially triggered the teacher's feedback.

9.2. Data analysis

a) In order to answer Research question 5 it was first necessary to examine the frequency of uptake with repair during the provision of corrective feedback in the two treatment groups. Uptake with repair was coded and scored separately for the passé composé (PC) and for the imparfait (IMP). As already stated, in the current study corrective feedback was repeated if it did not result in repair after the teacher's first attempt at correction. To take account of this, repair was scored in two ways: (a) only uptake with repair moves following the first CF move were included and (b) all uptake moves in a CF episode were included. Thus, (a) provides an indication of whether a learner could self-correct immediately following CF while (b) provides a measure of the extent to which uptake following CF was successful overall.

b) Part 2 of Research question 5 was investigated by comparing the gain scores of those students who produced repair with those who did not, in each of the two groups: the RE group and the CR group. This analysis was chosen as in both groups about the same number of learners participated in the interaction during the execution of the tasks and a similar number of students in each group produced uptake with repair following the teacher's corrective feedback. Such an analysis made possible the comparison between the acquisition (gain) scores of those students who produced repair and those who did not. It is to be noted that those who did not take part in the interaction during the task were also exposed to input consisting of the uptake produced by other students.

Since the gain scores of the two groups of learners (those who repaired and those who did not) in both the RE and the CR group were normally distributed, an Independent samples t-test was used to compare their mean gain scores.

9.3. Results

Table 35 and Table 36 present the results related to the uptake of CF in each of the two treatment groups, the RE and the CR group.

Table 35 RE group: number of CF episodes and percentage of uptake with repair

Structure	Corrective feedback episodes	Uptake with repair after the 1 st CF move		Total uptake with repair	
		Count	%	Count	%
Passé Composé	33	26	78.8%	28	84.8%
Imparfait	35	25	71.4%	25	71.4%
Total	68	51	75%	53	77.9%

Table 36 CR group: number of CF episodes and percentage of uptake with repair

Structure	Corrective feedback episodes	Uptake with repair after the 1 st CF move		Total uptake with repair	
		Count	%	Count	%
Passé Composé	17	7	41.2%	11	64.7%
Imparfait	23	8	34.8%	10	43.5%
Total	40	15	38.5%	21	52.5 %

Overall the RE group received more corrective feedback than the CR group (for both target structures) as the RE group produced more non target-like forms which required correction. There were 33 CF episodes for the passé composé and 35 CF episodes for the imparfait

recorded in the RE group, whereas in the CR group there were 17 CF episodes for the passé composé and 23 CF episodes for the imparfait.

The analysis which considered only the uptake moves following the first CF move produced the following results: in the RE group there were 26 (78.8%) moves consisting of uptake with repair for the passé composé and 25 (71.4%) for the imparfait, while in the CR group there were 7 (41.2%) moves consisting of uptake with repair for the passé composé and 8 (34.8%) for the imparfait.

The analysis which provided an overall measure of the extent to which uptake was successful (taking into account uptake with repair after the first and second CF moves) produced the following results: in the RE group there were 28 (84.8%) episodes involving repair for the passé composé and 25 (71.4%) for the imparfait, whereas in the CR group there were 11 (64.5%) episodes involving repair for the passé composé and 10 (43.5%) for the imparfait.

9.3.1. Repair and gains for passé composé

Tables 37, 38, 39 and 40 display the data related to the uptake with repair and the gain scores for passé composé: the gains of those learners who produced repair of the passé composé are presented alongside the gain scores of those learners who did not produce repair. Table 37 shows the gains of the learners in the RE group who produced repair of passé composé whereas table 38 shows the gains of those learners who did not produce repair of passé composé. These learners either failed to produce repair or did not attempt to uptake at all. Table 39 presents the repair and the gain scores of those learners in the CR group who produced repair and Table 40 shows the gains of the learners who did not produce repair of passé composé in the CR group. In the RE group there were eleven learners who repaired their passé composé errors and in the CR group there were eight learners who repaired their passé composé errors.

Table 37 RE group: learners' repair of PC and their gains in oral and written production

LEARNER	REPAIR (number of episodes)	ORAL PRODUCTION		WRITTEN PRODUCTION	
		Gain 1	Gain 2	Gain 1	Gain 2
L 1	4	8.02	13.85	27.14	-17.50
L 2	3	33.34	27.92	27.78	27.78
L 3	1	13.18	17.73	0.00	0.00
L 4	1	36.60	33.75	13.12	15.97
L 6	6	39.58	20.67	-	56.25
L 7	5	23.08	22.22	21.43	33.33
L 8	2	8.33	15.00	11.28	10.00
L 11	2	23.62	19.45	21.43	10.32
L 14	2	0.00	14.36	0.00	18.75
L 16	1	32.91	27.78	7.14	14.28
L 17	1	-2.78	28.89	0.00	37.50

Table 38 RE group: learners who did not produce repair of PC and their gains

LEARNER	ORAL PRODUCTION		WRITTEN PRODUCTION	
	Gain 1	Gain 2	Gain 1	Gain 2
L 5	25.00	33.33	40.91	10.91
L 9	45.83	27.27	47.14	80.00
L 10	21.67	1.67	- 7.57	33.33
L 12	0.84	0.00	- 6.57	38.89
L 13	1.11	1.11	42.86	70.00
L 15	46.15	9.09	- 8.33	10.42
L 18	30.00	8.33	20.00	62.50

Table 39 CR group: learners' repair of PC and their gains in oral and written production

LEARNER	REPAIR (number of episodes)	ORAL PRODUCTION		WRITTEN PRODUCTION	
		Gain 1	Gain 2	Gain 1	Gain 2
L 19	1	-6.25	-25.00	-12.73	2.27
L 20	3	10.72	14.29	34.61	33.33
L 21	1	-10.41	-13.72	-13.18	11.23
L 27	1	-30.52	-21.43	32.14	7.14
L 28	1	-9.72	-13.13	-50.00	-27.27
L 34	1	31.25	24.05	27.78	38.39
L 35	1	-37.35	19.64	-13.89	-15.91
L 36	2	-	-	-0.96	6.25

Table 40 CR group: learners who did not produce repair of PC and their gains

LEARNER	ORAL PRODUCTION		WRITTEN PRODUCTION	
	Gain 1	Gain 2	Gain 1	Gain 2
L 22	-24.11	-37.92	31.95	17.17
L 23	47.50	69.64	5.83	17.14
L 24	44.44	46.40	-23.85	-13.75
L 25	2.50	-33.34	-7.14	-21.43
L 26	-0.71	-13.49	0.00	0.00
L 29	0.00	26.39	0.00	0.00
L 30	18.75	20.14	-	-
L 31	-25.00	-37.50	75.00	-60.00
L 32	0.00	10.00	0.00	77.78
L 33	-3.08	-	0.00	0.00

9.3.2. Repair and gains for imparfait

Tables 41, 42, 43 and 44 refer to the uptake with repair and the gain scores for imparfait. The errors in the use of imparfait were repaired by eleven students in the RE group and by seven students in the CR group. Table 41 displays the data for the RE group related to the uptake with repair and the gain scores of those learners who produced repair of the imparfait whereas Table 42 displays the gain scores of those learners who did not produce repair of the imparfait in the RE group. These learners either attempted uptake but did not repair or did not attempt uptake at all. Tables 43 and 44 show the data for the CR group: Table 43 displays the uptake with repair episodes and the gain scores of those students who produced repair of the imparfait whereas table 44 shows the gain scores of those learners who did not produce repair of the imparfait in the CR group.

Table 41 RE group: learners' repair of IMP and their gains in oral and written production

LEARNER	REPAIR (number of episodes)	ORAL PRODUCTION		WRITTEN PRODUCTION	
		Gain 1	Gain 2	Gain 1	Gain 2
L 1	1	10.00	66.66	16.67	100.00
L 2	1	20.00	27.78	33.33	25.00
L 3	6	12.50	14.28	42.86	62.50
L 7	4	23.34	58.34	8.33	16.66
L 8	1	41.70	54.17	44.45	22.23
L 10	2	0.00	33.33	46.67	30.00
L 11	3	38.10	46.67	80.00	91.67
L 12	2	8.33	7.14	-31.25	56.25
L 14	3	10.00	33.33	20.00	31.25
L 17	1	50.00	33.33	66.67	56.25
L 18	1	25.00	33.33	0.00	40.00

Table 42 RE group: learners who did not repair IMP and their gains in oral and written production

LEARNER	ORAL PRODUCTION		WRITTEN PRODUCTION	
	Gain 1	Gain 2	Gain 1	Gain 2
L 4	50.00	26.67	16.67	16.67
L 5	75.00	80.00	40.00	0.00
L 6	0.00	0.00	0.00	0.00
L 9	5.63	30.63	55.00	75.00
L 13	-40.00	-23.33	33.33	25.00
L 15	0.00	22.22	-18.75	31.25
L 16	20.00	30.00	35.00	22.50

Table 43 CR group: learners' repair of IMP and their gains in oral and written production

LEARNER	REPAIR (number of episodes)	ORAL PRODUCTION		WRITTEN PRODUCTION	
		Gain 1	Gain 2	Gain 1	Gain 2
L 19	1	25.72	48.22	-3.33	9.53
L 20	1	25.00	31.82	36.67	40.00
L 21	2	22.22	0.00	-7.70	17.50
L 29	1	0.00	0.00	0.00	0.00
L 32	1	-20.00	-20.00	40.00	71.43
L 34	3	25.00	41.66	66.66	78.57
L 35	1	-8.96	-41.66	-16.89	-16.89

Table 44 CR group: learners who did not repair IMP and their gains in oral and written production

LEARNER	ORAL PRODUCTION		WRITTEN PRODUCTION	
	Gain 1	Gain 2	Gain 1	Gain 2
L 22	33.33	20.00	3.03	-8.33
L 23	6.25	66.66	-33.33	4.17
L 24	25.00	0.00	11.11	33.33
L 25	9.09	0.00	33.33	33.33
L 26	0.00	0.00	42.86	50.00
L 27	0.00	0.00	25.00	-8.33
L 28	0.00	0.00	0.00	0.00
L 30	0.00	25.00	-	-
L 31	-50.00	-50.00	0.00	0.00
L 33	0.00	-	40.00	25.00
L 36	-	-	-13.33	41.67

Tables 45, 46, 47 and 48 present the results for the comparison of mean gain scores of those learners who produced uptake with repair and those who did not produce repair but were exposed to it. Table 45 displays the data for oral production of the passé composé and Table 46 presents the data for written production of the passé composé. Tables 47 and 48 display

the data for oral and written production of the imparfait. Each table shows the number of students (N) who did or did not produce uptake with repair and their respective mean gain score in oral and written production. Each table includes the results for the RE group and the CR group.

Table 45 Comparison of mean gain scores in oral production of PC

GROUP	Learners producing repair			Learners exposed to repair		
	N	Gain 1 mean	Gain 2 mean	N	Gain 1 mean	Gain 2 mean
RE	11	17.63	24.3	7	24.37	11.54
CR	8	-7.5	-2.2	10	5.6	3.8

The results of the Independent samples t-tests comparing the gains of learners producing repair and the learners exposed to repair in oral production of passé composé in the RE group show no significant difference for gain 1 , but there is a significant difference for gain 2 scores. This indicates that the learners who produced repair demonstrate significantly greater gains than the learners who did not produce repair:

Gain 1 : $t(16) = -.59$, ns, ES= .15

Gain 2 : $t(16) = 2.54$, $p = .022$, ES= .54

In the CR group there were no significant differences in gains between those who produced repair and those who did not:

Gain 1 : $t(15) = -1.14$, ns, ES= .28

Gain 2 : $t(14) = -.35$, ns, ES= .09

Table 46 Comparison of mean gain scores in written production of PC

GROUP	Learners producing repair			Learners exposed to repair		
	N	Gain 1 mean	Gain 2 mean	N	Gain 1 mean	Gain 2 mean
RE	11	12.93	15.04	7	18.35	43.72
CR	8	.47	6.9	10	2.42	1.88

The results of t-tests for written passé composé in the RE group, comparing the gains of learners producing repair and those exposed to repair show no significant difference for gain 1 scores but there is a significant difference for gain 2 scores, indicating that those learners who were exposed to repair actually achieved significantly higher gains than those who produced repair:

Gain 1: $t(15) = -.64$, ns, ES= .16

Gain 2: $t(15) = -2.23$, $p = .041$, ES= .50

In the CR group there were no significant differences between the two groups' gain scores:

Gain 1: $t(15) = -.12$, ns, ES= .03

Gain 2: $t(15) = -.34$, ns, ES= .09

Table 47 Comparison of mean gain scores in oral production of IMP

GROUP	Learners producing repair			Learners exposed to repair		
	N	Gain 1 mean	Gain 2 mean	N	Gain 1 mean	Gain 2 Mean
RE	11	21.72	37.12	7	15.8	23.74
CR	7	9.85	8.58	11	2.37	6.85

Results of t-tests for oral imparfait, comparing the gains of the students producing repair and the students exposed to repair, do not show any significant differences in either the RE group or in the CR group:

RE group:

Gain 1 : $t(16) = .47$, ns, ES= .12

Gain 2 : $t(16) = 1.15$, ns, ES= .28

CR group:

Gain 1 : $t(15) = .73$, ns, ES= .19

Gain 2 : $t(14) = .12$, ns, ES= .03

Table 48 Comparison of mean gain scores in written production of IMP

GROUP	Learners producing repair			Learners exposed to repair		
	N	Gain 1 mean	Gain 2 mean	N	Gain 1 mean	Gain 2 Mean
RE	11	29.79	45.50	7	23.03	24.34
CR	7	16.49	28.59	11	10.87	17.08

The results of t-tests for written imparfait also do not show any significant difference between the gains of the two groups of students (those who repaired their errors and those who did not) in each of the two treatment groups:

RE group:

Gain 1: $t(16) = .48$, ns, ES= .12

Gain 2: $t(16) = .15$, ns, ES= .04

CR group:

Gain 1: $t(15) = .42$, ns, ES= .14

Gain 2: $t(15) = .82$, ns, ES= .20

9.4. Summary

The results for Research question 5 indicate that:

There was a higher percentage of uptake with repair in the RE group than in the CR group for both target structures. In both groups there was a higher percentage of uptake with repair for the passé composé than for the imparfait past tense. In both groups about the same number of learners attempted uptake and there was no significant difference between the two groups regarding the number of learners who produced uptake with repair.

Results comparing the mean gain scores of those learners who produced repair with the learners who did not produce repair but were exposed to it, show that only in the RE group and only for the passé composé were there significant differences between the gains of those learners who repaired and those who did not repair. However, significant differences were only evident in the long-term gain scores (gain 2) but in a different way for oral and written production: in oral production the learners who produced uptake with repair had a significantly higher mean gain 2 score for passé composé than the learners who did not produce repair. However, in written production the learners who were only exposed to repair demonstrated a significantly greater mean gain 2 score for passé composé than those learners who actually produced uptake with repair.

9.5. Discussion

Research question 5 asked (a) to what extent learners repair their errors, and (b) whether learners who produced uptake with repair demonstrate greater gains in accuracy than learners who were just exposed to uptake with repair. In the following discussion I will first address the question related to (a) the level of uptake with repair in the treatment groups. This will be followed by the discussion (b) of the relationship between repair and gains in acquisition measured as improved accuracy.

Aiming at answering the first part of Research question 5, I will discuss the results of the current study with reference to the following three points: (1) the relevance and influence of the instructional setting, (2) the enhanced saliency of recasts and (3) possible ambiguity of clarification requests.

As the results show, the current study found a higher level of uptake with repair following recasts than following requests for clarification. This result contrasts with those of a number of studies that have investigated the effects of oral corrective feedback, which found that prompts (including clarification requests) in the classroom were followed by much higher rates of uptake and self-repair than recasts (Chaudron, 1988; Lyster, 1998b, 2004; Lyster & Ranta, 1997; Yang & Lyster, 2010). These studies suggest that prompting techniques are more conducive to self-correction and thereby to learning, than recasts which are claimed to be too implicit, not easy to notice and ambiguous for learners. This appears to be especially the case in the immersion classroom settings. However, the situation may be different in foreign language classrooms where learners are more focused on language itself and probably expect to be corrected, even when engaged in communicative tasks and focused on meaning. Sheen (2004) found not only more recasts being provided by teachers but also a considerably higher rate of repair after recasts in English FL and ESL classes in Korea and New Zealand than in Canadian French immersion and Canadian ESL classes. Ellis et al. (2001) also reported a high percentage of successful uptake following recasts in New Zealand communicative ESL classes. Lyster and Mori (2006) showed how different instructional settings and the culture of instruction (in Canadian French immersion and Japanese immersion classes) affect the rate of uptake and repair in the classroom. Clearly, as these studies indicate, the context in which learning takes place can considerably influence the overall success of uptake following recasts. This is reflected in the current study as well.

A higher rate of repair following recasts also suggests that recasts may have been more salient, or salient enough to be recognized as corrective. Although implicit in nature they must have been noticed by those learners who produced uptake with repair, which often occurred immediately after the first CF move. Discourse context and discourse cues such as focus on a particular target, interrogative mode (rising tone), short recasts (Sheen, 2008; Loewen & Philp, 2006; Han, 2002) as well as provision of sufficient time to allow for the opportunity to notice and notice-the-gap (Braid, 2002; Long, 2007; Oliver, 1995, 2000) may have contributed to the level of repair in the RE group. However, the difference between the percentage of repair of the *passé composé* and of the *imparfait* indicates that the saliency of recasts also depends on the linguistic feature in focus (Carpenter, Jeon, MacGregor & Mackey 2006). Lower saliency of the *imparfait* forms has already been discussed in literature (Bardovi-Harlig, 2000; Kaplan, 1987) and as such is related to the difficulties which French L2 learners experience with the *imparfait*.

A lower rate of uptake and repair following clarification requests in the current study may be seen as a surprising result considering the studies, not only in immersion classrooms, which found prompting techniques lead to a high level of uptake and repair (Lyster, 1998a, 1998b; Lyster, 2004; Lyster & Panova, 2002; Lyster & Ranta, 1997; Yang & Lyster, 2010). However, clarification requests are only one form among several others (elicitations, repetitions and metalinguistic clues) which together constitute the group of prompts. In the Lyster and Ranta's (1997) study, for example, clarification requests accounted for only 11% of all CF and they were followed by a very high percentage of uptake 88%, but only 28% of all clarification requests led to repair, leaving a rather high 60% of clarification requests in need of repair. Furthermore, the reason for a lower level of both uptake and repair in the current study may be found in the explanation provided by conversation analysis (Seedhouse, 2004) where clarification requests are classified as 'an open repair initiator'. Seedhouse argues that:

...open repair initiators do not even indicate the presence of a linguistic error: They are frequently used by listeners to initiate repair when the speaker has clearly not made a linguistic error... The use of open repair initiators by the listener may therefore actually imply to the speaker that some form of trouble other than a linguistic error has occurred (2004, p. 162)

Therefore, the use of clarification requests indicating that a breakdown has occurred in the understanding of meaning and thus negotiation of meaning is needed (Long, 1981a, 1983a, 1983b; Pica, 1994), may be ambiguous for learners if the focus of CR is on form. That is, the clarification of form, or negotiation of form (Lyster, 1998a, 1998b) is required. There is evidence of such ambiguity in the following excerpt from the task transcript where the meaning of the utterance could be easily comprehended but the non-target like linguistic form required correction. The learner switched to L1 (English) to indicate the problem:

T: *Qu'est-ce qu'ils ont fait quand ils sont sortis du supermarché ?*

L20: *Paul et son frère s'est battu* (error - trigger)

T: *Hein? Tu peux répéter?* (request for clarification)

L20: I don't know how to say ? Ils ont... ?

Thus, a possibility that the students may have been puzzled and did not know where to locate the problem, has to be taken into consideration. It can explain the lower level of repair following clarification requests in the current study.

The second part of Research question 5 was concerned with the association of repair and the subsequent short-term and long-term gains in accuracy for each target structure separately.

As presented in the Results section, the current study found a significant relationship between repair and learning gains only in the RE group and only for the passé composé long-term gains (Gain 2). Therefore the following discussion will focus on the three main findings in the RE group: (1) the relationship between the gains and repair of those learners who produced repair and those who were only exposed to it, (2) the relationship between repair and gains which only became evident in the long-term and (3) the differential effects of repair on the gains for passé composé and for imparfait.

1. The results for passé composé show that in the RE group there was a significant difference between the long-term gains of those learners who produced repair and those who did not produce it. In oral production those learners who produced uptake with repair demonstrated greater Gain 2 scores, whereas in written production the result was reversed: interestingly, the learners who were only exposed to repair actually demonstrated significantly greater long-term gains for passé composé than those who produced uptake with repair. This result is somewhat perplexing and not easy to explain, but it may be that the learners who showed significantly greater long-term gains in oral production related to their uptake with repair, demonstrated an effect of ‘training’. They may have already mastered the structure of passé composé to the level that they were able to notice-the-gap and repair (during the treatment) and thus later showed gains in the same type of production.

On the other hand, those learners who were only exposed to repair and thereby to more input, may have benefited from it even though they did not participate in interaction. Uptake is an optional move, as Mackey and Philp (1998) have shown, thus no uptake does not mean that noticing did not occur. It may be that the learners who only observed needed more time to consolidate their knowledge which they were able to show later in written production where more control over performance is available. However, it is interesting that in oral production the mean short-term gain score of the ‘observers’ group was higher than the mean gain of those who repaired. Such a result, that those who only listen and observe benefit from being exposed, corresponds to the research findings (R. Ellis et al., 1994; Havranek, 2002; Mackey, 2006) which suggest that the production of uptake and repair in the classroom may not be necessary for acquisition.

2. It is interesting that the relationship between repair and gain scores in the RE group for passé composé only became apparent at a later point, in long-term gains. It may be that the time was needed for the ‘training’ effect to be consolidated and become evident in oral

production of those learners who participated in interaction and repaired, and in written production of those learners who were only exposed to repair. Also, it may be that those learners who were only exposed to repair, and thereby to more input, showed greater gains in oral production immediately after the treatment (which is indicated by their gain 1 score) but did not retain that gain in the long-term (a comparison of individual gains shows that the scores of five students out of seven in this group decreased in the long-term). On the other hand, those who produced repair retained their gains in the long-term. However, in written production six out of seven students in the ‘observers’ group were able to show significantly greater long-term gains (gain 2) for *passé composé* – which is an indication that they needed time to consolidate their knowledge and benefited when completing the delayed written task.

3. The differential associations between repair and subsequent learning gains in the RE group for the *passé composé* and for the *imparfait* does not come as a surprise taking into consideration already known problems related to the use and acquisition of *imparfait* in French L2. As less salient, linguistic features of *imparfait* have already been documented in literature as more difficult to notice. The current study provides further evidence in this regard, indicating that the level of uptake with repair in both treatment groups was higher for the *passé composé* than for the *imparfait*. As Mackey (1999, 2006) has demonstrated, repair depends on the target structure and some forms are more salient and easier to notice than some other forms. Secondly, in the current study, the *imparfait* forms included in the analysis of oral production were limited to a set of verbs so any firm conclusion regarding the relationship between repair and gains of *imparfait* in oral production is not warranted.

Clearly, there is no evidence of any relationship between repair and learning gains for the *imparfait* in written production either, in any of the two treatment groups. This may be because of the later acquisition of *imparfait* in French L2 compared with the *passé composé* (Andersen & Shirai, 1994; Bardovi-Harlig, 2000; Harley, 1989; Salaberry & Shirai, 2002). Research so far has provided strong evidence that, even though for L2 learners the formal properties of the *imparfait* are easier to acquire than the morphology of the *passé composé*, the pragmatic functions of the *imparfait*, involving its lexical aspects, are usually acquired very slowly and much later than the *passé composé*. Therefore the gains in the *imparfait* did not have any significant relationship with the uptake with the repair produced during the treatment tasks. Also, no relationship was found between repair and the gains for the *passé composé* in the CR group. As already indicated, a significantly lower frequency of target-structures in the class input resulting from the nature of the clarification requests was

augmented by the occasional ambiguity in the student interpretation of clarification requests. As a result, there is no relationship between the repair during the treatment tasks and learning gains in either the passé composé or the imparfait.

Chapter 10. Results for learner individual differences

10.1. Introduction

This chapter considers Research question 6:

Do individual difference factors predict the learners' acquisition of the target structures in the two experimental groups, one that received CF in the form of recasts and the other that received CF in the form of clarification requests?

- Does language analytic ability predict the acquisition of the target structures in (1) the RE group and (2) the CR group?
- Does working memory predict the acquisition of the target structures in (1) the RE group and (2) the CR group?
- Does language anxiety predict the acquisition of the target structures in (1) the RE group and (2) the CR group?

10.2. Data analysis

1. To measure the participants' analytic ability the Language analysis test (LAT) was administered. The test was developed by Ottó (2002) and published in Schmitt, Dörnyei, Adolphs and Durow (2004). It consists of 14 multiple choice items, based on an artificial language. The test requires the participants to analyse the grammatical markers supplied in the glossary, deduce the rule and select the correct sentence as a translation from English. The LAT was scored by awarding one point for each correct answer (see Appendix E). The total score was 14 points and the final scores were calculated in percentages. The descriptive statistics of the LAT for each group were calculated and an Independent samples t-test was used to establish if the two groups were comparable.
2. In order to measure the construct of working memory the Speaking span test (SST) was used. The test was used based on Daneman and Green (1986) and Daneman (1991) SST, adapted from Mizera (2006). The participants were tested individually, one by one. The test consisted of 65 items divided into five sets, each comprising five

sub-sets (starting from two up to five words) so that the students, having read the single words in a sub-set, had to produce sentences using these words. For example, if the words 'fingers' and 'noticed' appeared on the computer screen, when they disappeared the student had to produce two unrelated sentences, each containing one of these two words (for example, "She has long fingers" and "I noticed him in the crowd"). If three words appeared on the screen, for example 'machine', 'results', 'compass', then the participant had to produce three sentences, each containing one of these three words. And so forth, up to the set of five words which required the participant to make up five sentences, each containing one of the words listed. (See Appendix F). Although I calculated two types of scores, a *strict score* and a *lenient score*, only the strict score was used for the analyses because it is considered to be a more stringent measure of working memory capacity. Also, the correlation of the strict and the lenient scores for the whole sample resulted in a highly significant coefficient: $r = .92^{**}$, $p = .000$ ($N=34$). The strict score was generated when all the sentences produced contained the target word in the exact form and order of presentation. Descriptive statistics for SST strict score were calculated for both groups of participants and an Independent samples t-test was computed to examine if the two treatment groups were different.

3. In order to measure the level of language anxiety, the questionnaire was used developed by MacIntyre and Gardner (1994) and validated on 258 participants by Onwuegbuzie, Bailey and Daley (2000). Scores on the questionnaire were coded on a 6-point Likert scale so that the lowest score for each question was 1 and the highest was 6 (see Appendix G). The mean score for each participant and the descriptive statistics for language anxiety were first calculated. An Independent samples t-test was used to see if there was any significant difference in the level of anxiety of the two treatment groups.

Cronbach Alpha was used to measure the reliability of scores on each of the three tests used for learner individual differences.

Multiple regression stepwise analysis using an enter selection method was computed to examine which of the three individual difference factors predicted the acquisition of the target structures in the two experimental groups. Three constructs: language analytic ability, working memory capacity and language anxiety were used as independent predictor variables

and the gain scores for each linguistic structure (short-term Gain 1 and long-term Gain 2) were used as dependent, outcome variables. Multiple regression analysis was chosen for it was felt that it would give a clearer picture of which ID factors were significantly associated with the gain scores. However, it is acknowledged that the sample size in the current study is rather low. Although it is recommended that Multiple regression analysis should be performed on a larger sample size depending on the number of predictor variables, some authors think that the minimum should be 10 cases for each predictor variable (Field, 2009, p. 222). This requirement has been met in the current study. Because of the low numbers I report not only the R2 value but also the adjusted R2. This indicates what the R2 value would be if the model were derived from the general population rather than the sample in the study (p. 235).

10.3. Results

10.3.1. Descriptive statistics

Table 49 displays the descriptive statistics for the Language analysis test (LAT), Table 50 presents the descriptive statistics for the Speaking span test (SST) strict scores and Table 51 presents the descriptive statistics for anxiety scores:

Table 49 Descriptive statistics for LAT scores

GROUP	Mean	SD	Md	Min	Max
Recast (n=16)	71.43	23.18	82.14	28.57	100.00
Cl.Request (N=18)	64.28	21.64	71.43	14.28	92.86

As shown in Table 49, the Recast group had a mean of 71.43 and the Clarification Request group had a mean of 64.28 (%). An Independent samples t-test showed no significant difference between the two groups: $t(32) = .92$, ns. Cronbach Alpha indicated the reliability coefficient for the test was .80.

Table 50 Descriptive statistics for SST strict scores

GROUP	Mean	SD	Md	Min	Max
Recast (N=16)	39.44	8.40	40.00	21.00	51.00
Cl.Request (N=18)	37.94	8.95	38.00	24.00	54.00

As shown in Table 50 the Recast group had a mean of 39.44 strict score on the SST whereas the Clarification Request group had a mean of 37.94. An Independent samples t-test showed no significant difference: $t(32) = .500$, ns. The test reliability coefficient measured by Cronbach Alpha was .77.

Table 51 Descriptive statistics for language anxiety

GROUP	Mean	SD	Md	Min	Max
Recast (N=17)	3.57	0.85	3.75	1.63	4.75
Cl.Request (N=18)	3.27	0.85	3.06	1.87	5.37

As demonstrated in Table 51, the mean scores for anxiety were very similar in the two groups. The RE group had a mean of 3.57 and the CR group 3.27. An Independent samples t-test showed no significant differences between the two groups' anxiety scores:

$t(33) = 1.04$, ns. Cronbach Alpha indicated that the reliability of the test was high, with a coefficient .82.

The correlation analysis between the LAT scores and SST strict scores for the whole sample found a significant Pearson correlation coefficient, $r = .417^{**}$ ($p = .009$) which indicates that there is a relationship between the two constructs. The Pearson correlation coefficient between the LAT scores and anxiety indicated a significant negative relationship: $r = -.406^*$ ($p = .019$). The correlation coefficient between the SST strict scores and anxiety was also negative and significant: $r = -.483^{**}$ ($p = .002$).

10.3.2. Multiple regression analyses

Table 52 and Table 53 present the results of the Multiple regression stepwise analyses for oral and written production of the RE group.

Table 52 RE group: Multiple regression coefficients for the gain scores in oral production

IDs	ORAL PC				ORAL IMP			
	GAIN 1		GAIN 2		GAIN 1		GAIN 2	
	Beta	sig	Beta	sig	Beta	sig	Beta	Sig
LAA			.56	.047				
WM	.731	.005						
Anxiety								
	R=.731		R=.560					
	R2=.535		R2=.313					
	AdjustedR2=.492		AdjustedR2=.251					
	Excluded: LAA		Excluded: WM		Excluded: LAA		Excluded: LAA	
			Anxiety		WM		WM	
					Anxiety		Anxiety	

Note. LAA = language analytic ability
WM = working memory

R = the multiple correlation coefficient
R2= R squared

Table 53 RE group: Multiple regression coefficients for the gain scores in written production

IDs	WRITTEN PC				WRITTEN IMP			
	GAIN 1		GAIN 2		GAIN 1		GAIN 2	
	Beta	Sig	Beta	sig	Beta	sig	Beta	Sig
LAA							-.656	.015
WM	-.554	.051						
Anxiety								
	R=.554						R=.656	
	R2=.308						R2=.431	
	AdjustedR2=.165						AdjustedR2=.379	
	Excluded: LAA		Excluded: LAA		Excluded: LAA			
	Excluded: LAA		WM		WM		Excluded: LAA	WM
			Anxiety		Anxiety			Anxiety

Note. LAA = language analytic ability
WM = working memory

R = the multiple correlation coefficient
R2= R squared

As shown in Table 52 and Table 53, in the RE group, WM capacity was a significant predictor of short-term gain scores (Gain 1) for oral PC (WM Beta=.731**, p=.005), whereas language analytic ability (LAA) predicted the long-term gain scores (Gain 2) for oral PC (Beta= .56*, p=.047). Analytic ability was also a predictor for long-term gains in written IMP but with a negative regression coefficient (Beta= - .656*, p = .015). That is, the learners with

lower levels of analytic ability demonstrated greater gains. The regression coefficient for WM predicting PC short-term gains (Gain 1) in written production was also negative and approaching significance (Beta = - .554, p = .051). Anxiety did not emerge as a significant predictor variable in any of the regression analyses.

Table 54 and Table 55 display the results of the analyses for oral and written production of the CR group.

Table 54 CR group: Multiple regression coefficients for the gain scores in oral production

IDs	ORAL PC				ORAL IMP			
	GAIN 1		GAIN 2		GAIN 1		GAIN 2	
	Beta	Sig	Beta	sig	Beta	sig	Beta	Sig
LAA	.725	.008	.578	.062	.575	.016	.707	.002
WM	-.566	.030						
Anxiety								
	R=.655				R=.575		R=.707	
	R2=.429				R2=.331		R2=.500	
	AdjustedR2=.348				AdjustedR2=.286		AdjustedR2=.464	
	Excluded:				Excluded:		Excluded:	
	Anxiety		Excluded: WM		Excluded: WM		Excluded: WM	
			Anxiety		Anxiety		Anxiety	

Note. LAA = language analytic ability
WM = working memory
R = the multiple correlation coefficient
R2= R squared

Table 55 CR group: Multiple regression coefficients for the gain scores in written production

IDs	WRITTEN PC				WRITTEN IMP			
	GAIN 1		GAIN 2		GAIN 1		GAIN 2	
	Beta	Sig	Beta	sig	Beta	sig	Beta	sig
LAA	.674	.003	.767	.000				
WM								
Anxiety								
	R=.674		R=.767					
	R2=.455		R2=.588					
	AdjustedR2=.418		AdjustedR2=.561					
	Excluded:				Excluded: LAA		Excluded: LAA	
	WM		Excluded: WM		WM		WM	
	Anxiety		Anxiety		Anxiety		Anxiety	

Note. LAA = language analytic ability
WM = working memory
R = the multiple correlation coefficient
R2= R squared

Table 54 and Table 55 show that in the CR group analytic ability was a significant predictor for short-term gains in oral PC (Beta=.725**, p=.008), and it approached significance for long-term gains in oral production of PC (Beta= .578, p=.062). Analytic ability also predicted both short-term and long-term gain scores for oral IMP (Gain 1: Beta= .575*, p= .016, Gain 2: Beta =.707**, p = .002). Furthermore, the short-term and long-term gain scores for PC in written production were predicted by LAA (Gain 1: Beta=.674**, p=.003; Gain 2: Beta=.767**, p=.000). Working memory was a predictor only for short-term gains (Gain 1) in oral PC, but with a negative regression coefficient (Beta= -.566*, p=.030). The gain scores of IMP in written production were not predicted by any of the ID factors. Language anxiety was not found to be a significant predictor for any of the gain scores in either oral or written production of the CR group.

10.3.3. Summary

Multiple regression stepwise analyses with ID factors as predictor variables reveal the following:

- In the RE group, WM capacity predicted the short-term gain scores for PC in oral production whereas analytic ability predicted the long-term gains for PC. Analytic ability was a predictor for long-term gains in written IMP but with a negative regression coefficient. WM capacity with a negative regression coefficient approached significance in predicting the short-term gains for written PC.
- Analytic ability was a much stronger predictor variable in the CR group: both short-term and long-term gain scores for oral IMP and for written PC were predicted by the LAT scores. Also, the short-term gain scores for oral PC were predicted by the LAT, and the regression coefficient predicting long-term gains in oral PC approached the level of significance. Working memory predicted only the short-term gains for PC in oral production, but had a negative regression coefficient. The gain scores for written imparfait (Gain 1 and Gain 2) in the CR group were not predicted by any of the ID factors.
- Anxiety did not appear to be a significant predictor for any of the gain scores either in the RE or in the CR group.

10.4. Discussion

Research question 6 asked if language analytic ability and working memory as cognitive ID factors, and language anxiety as an affective ID factor predicted the acquisition of the target structures in the two experimental groups. The results of the Multiple regression analyses show that language anxiety was not a significant predictor to learning in either of the two groups. We may assume that the learners in these two groups were comfortable with the type of tasks they participated in. Most of the students were probably motivated to learn. The students who did not want to participate in interaction were not forced to do so. This may be the reason why language anxiety did not play a role in L2 learning. So the present study does not support the findings of Horwitz, Horwitz and Cope (1986) and Sheen (2008). However, a significant negative correlation between anxiety and both analytic ability ($r = -.406^*$, $p = .019$) and working memory ($r = -.483^{**}$, $p = .002$) for the whole sample suggests that the students with higher levels of anxiety were in fact those who had lower levels of cognitive abilities. This is in line with Ganshaw and Sparks' (1993) claims that anxiety may be an indicator of 'basic language problems' rather than an indicator of situation-specific and situation-driven feelings of uneasiness.

Language analytic ability and working memory as cognitive ID factors appear to be much stronger predictors of the acquisition gains. A significant Pearson correlation coefficient between these two factors for the whole sample suggests that the Language analysis test (LAT) and the Speaking span test (SST) of WM capacity are tapping on similar cognitive abilities. However, the Multiple regression analysis indicates that they can be distinguished. There is evidence that analytic ability and WM capacity predicted gains in oral and written production of the two groups in a different way. Thus I will discuss the results of the Multiple regression analyses in terms of the outcome variables, i. e. in terms of the effects that the two ID factors had on the gains in oral and written production of the two experimental groups.

In the RE group WM appears to be a strong predictor of short-term gains (Gain 1) in oral PC. In the CR group WM is also a predictor of short-term gains of oral PC but with a negative coefficient. The results for the RE group differ from those in Ando, Fukunaga, Kurahachi,

Suto, Nakano and Kage' s (1992) study, as reported in R. Ellis (2012), which found that lower WM scores were associated with the immediate gains in a communicatively taught class. Mackey et al. (2002) also found evidence of an advantage for learners with lower WM in an immediate post-test. The results of these two studies support the 'less is best' hypothesis (Miyake & Friedman, 1998), which posits that a limited working memory constrains the amount of the input that is 'intaken' and thus enhances processing while larger working memory allows more input to be attended to and overloads the processing mechanisms. One explanation for the difference of the results in this study and those in the other two studies is that the studies used different tests to measure WM and also learning gains. On the face of it, the 'less is best' hypothesis is applicable to the CR group. One explanation for this might be that those learners with lower WM were better able to hold the repaired utterances of their fellow students following a clarification request in their short term memories. However, as only two learners in this group showed positive immediate gains in PC care needs to be taken not to over-interpret the result for the CR group. Furthermore, it should also be noted that Mackey et al. (2002) did report a significant positive correlation between WM and learners' noticing of the target features in recasts. This might also have happened in the present study but, in this case, with the noticing translating into gains in accuracy in the RE group. In Robinson's framework of Aptitude complex hypothesis (2002c, p. 119) in which cognitive abilities are related to particular learning conditions, learners with stronger 'memory for contingent speech' and noticing-the-gap benefit from implicit CF such as recasts.

Research into WM has clearly established that WM capacity is determined by the task being performed. In the RE group the predictive power of WM was more evident in oral production than in written production, and only for the PC. This can be explained by the fact that oral production imposes high demands on L2 learners. As Bialystok and Ryan (1985) pointed out, online oral production increases the burden on the control dimension (p. 215). The need to focus on meaning interferes with the control of other aspects of language (e.g. formal accuracy). In the current study, the oral production task which required the learners to narrate a story based on a set of pictures was a cognitively demanding task for L2 learners. It imposed high demands on their working memory and their need to 'plan, temporarily store and execute' online (Daneman & Green, 1986). Research has established that, for individual differences in WM capacity to emerge, the task performed has to be complex (Conway et al., 2005; Daneman & Green, 1986; Just & Carpenter, 1992). The fact that PC is a complex

structure also has to be considered. Producing the PC is more difficult than producing the IM at this level of development (Bardovi-Harlig, 2002; Kaplan, 1987). Therefore WM emerged as a predictor for the gains in oral production of PC but not for the gains in oral IM.

In the CR group analytic ability was a much stronger predictor. In oral production the CR group's analytic ability strongly predicted not only the short-term gains of PC but also the short-term and long-term gains of IM (the regression coefficient for long-term gains in oral PC approached significance). Analytic ability also strongly predicted both short-term and long-term gains of written PC in the CR group. Such a strong link between analytic ability and acquisition gains in the CR group may be explained by the fact that this group received only implicit negative feedback: the learners were required to self-correct and work out on their own how to repair their errors. This is cognitively a very demanding task and apparently only those learners with high levels of analytic ability were able to do that, and subsequently apply their knowledge in the post-tests. As already noted, at their level of interlanguage development, the formal complexity of the PC presents more difficulty than the IM. Thus analytic ability predicted gains in both oral and written PC, and in oral IM but not the gains in written IM. The reason for this may be the following: on the one hand, writing tasks allow for more flexibility and poses less pressure on the learner, and on the other hand, the form of IM did not present as much difficulty as the form of PC. It is worth noting that in the CR group only the gains in written IM reached significance in the within-group comparisons from pre-test to post-test 2.

In contrast to the CR group, analytic ability was not a strong predictor variable for the RE group. The scores on the LAT predicted only the long-term gains of PC in oral production. This might be explained by the role of analytic ability in retaining the learning gains in long-term memory (Baddeley, 2007). Language analytic ability in the RE group also predicted the long-term gain scores in written IMP, but with a negative regression coefficient. That is, the learners with lower levels of analytic ability demonstrated greater gains. As already noted, written production is less pressured so learners can use their knowledge more easily. Since the learners in the RE group were provided with positive evidence simultaneously with the negative feedback, it is possible that even the less able learners attended to recasts and noticed the 'gap' (even if they did not participate in interaction). Later in the written tasks they were able to apply their knowledge. A similar process may have occurred in the written production of the PC in the RE group, for whom working memory was a negative predictor for short-term gains.

In conclusion, the results for the CR group corroborate the findings of those studies that established a strong link between language analytic ability and L2 learning (de Graaff, 1997; Ehrman, Leaver, & Oxford, 2003; Ehrman & Oxford, 1995; Ranta, 2002). However, the results for the RE group are more in line with Robinson's finding (1996, 1997a, 2002b, 2002c) that in incidental learning analytic ability is not a significant factor. Such results in the RE group also correspond to those of the Sheen's (2007a, 2011) research which found a positive relationship between learning gains and analytic ability but only for metalinguistic feedback, not for recasts. The current results thus support the previous findings in that analytic ability may not be a significant factor in the incidental learning of grammatical structures.

Chapter 11. Conclusion

11.1. Introduction

The previous chapters examined the effectiveness of two oral corrective feedback strategies – the recasts and the clarification requests, and the mediating role of learner individual difference factors, namely language analytic ability, working memory and anxiety. The target features investigated in this study were two French language structures: the passé composé and the imparfait, both considered difficult for L2 learners, the first one because of its complex morphology, the second one due to the functional complexity arising from subtle aspectual differences. In a quasi experimental classroom study the research design involved a pre-test, treatment, immediate post-test and a delayed post-test. The effectiveness of the two CF strategies was examined by comparing the effects of the treatment with no treatment. The CF was provided in the context of the learners performing three communicative tasks.

In this chapter I first present a summary of the main findings and then discuss the theoretical and pedagogical implications of the results. Following this, I consider the limitations of the study and then I propose some directions for further research.

11.2. Summary of main findings

11.2.1. Main findings for effects of CF on oral production

Research question 1 asked if oral CF (recasts and clarification requests) had an effect on L2 learners' acquisition of French linguistic structures - the passé composé and the imparfait – as observed in oral production.

- The Recasts (RE) group improved significantly from pre-test to post-test 1 and also from pre-test to post-test 2 in accurate oral use of both the passé composé and the imparfait. The RE group also improved from post-test 1 to post-test 2 in the use of the imparfait but not in the use of the passé composé.
- The Clarification request (CR) group showed no improvement over time for either structure.

- The Control (CN) group also did not improve over time.
- There were no group differences between any of the three groups for passé composé on either post-test 1 or post-test 2.
- The RE group outperformed the other two groups for imparfait on post-test 2 and there was a near significant difference ($p = .055$) between the RE and the CN group on post-test 1.
- Overall, then, the recasts proved more effective in promoting acquisition of both verb tenses.

Research question 2 asked if the learners' starting proficiency mediated the effect of oral CF on the learners' acquisition of passé composé and imparfait as manifested in oral production. The result was that the 'high' proficiency learners who received recasts outperformed both the learners who received CR and the learners in the CN group. However, there were no group differences for passé composé among the 'low' proficiency learners. The 'high' proficiency learners in the RE group also outperformed the 'high' CR and the 'high' CN group for the imparfait on post-test 2. Furthermore, the 'low' proficiency learners in the RE group outperformed the 'low' CN group on post-test 2 for the imparfait.

11.2.2. Main findings for effects of CF on written production

Research question 3 asked if oral CF (recasts and clarification requests) had an effect on L2 learners' acquisition of the passé composé and the imparfait, as observed in written production. The results show that:

- The RE group improved significantly from pre-test to post-test 1 and also from pre-test to post-test 2 in accurate use of both the passé composé and the imparfait.
- The CR group showed no improvement over time for the passé composé, but improved significantly from pre-test to post-test 2 for imparfait in written production.
- The CN group did not improve over time for either of the two structures.
- There were no group differences on post-test 1 for the passé composé. However, on post-test 2 the RE group outperformed the CR group. There were no significant

differences between either the RE and the CR group or the CR and CN group on post-test 2.

- The RE group outperformed both the CR group and the CN group for imparfait on both post-test 1 and post-test 2.

Overall, then, the results for written production were similar to those of oral production, with an exception regarding the clarification requests. Although there was a clear advantage for recasts, the CR group showed a significant improvement over time in the acquisition of the imparfait.

Research question 4 asked if the learners' starting proficiency mediated the effect of oral CF on the learners' acquisition of passé composé and imparfait in written production. Similarly to the results in oral production, the effect of the recasts treatment was more evident in the case of 'high' proficiency learners. This proved to be true for both the passé composé and the imparfait in the RE vs. CN group comparisons, but only for the imparfait in the RE vs. CR group comparison. The effectiveness of the recasts treatment was also evident in the superiority of the RE 'low' proficiency learners compared to the CN 'low' learners in the case of the imparfait.

11.2.3. Role of uptake with repair in language learning

Research question 5 asked (a) to what extent learners repair their errors following corrective feedback and (b) if learners who produce uptake with repair demonstrate greater gains in accuracy than learners who are just exposed to uptake with repair. Uptake with repair was found to be greater in the RE group than in the CR group.

In oral production of the passé composé (but not the imparfait) the learners in the RE group who produced uptake with repair outperformed those who just listened as shown by their long-term gain scores. However, there was no difference between the repairers and auditors in the RE group in the case of the imparfait. Also no difference in gains for either structure was observed between repairers and auditors in the CR group.

The results for written production differed from those for oral production. Auditors in the RE group outperformed the repairers in the case of the passé composé (but not the imparfait) over the long term. There was no evidence that repair was beneficial for the acquisition of either structure in the CR group.

11.2.4. Mediating role of individual difference factors

Research question 6 asked if ID factors (language analytic ability, working memory and language anxiety) predicted the acquisition of the target structures in the RE group and the CR group. The results show that LAA predicted the gain scores in the CR group and that this was true for both verb tenses in oral production and for passé composé in written production. In the RE group, LAA predicted only the long-term gain scores for passé composé in oral production. In contrast, WM was not a strong predictor for gain scores in the CR group. However, in the RE group WM predicted short-term gains for the passé composé in oral production and it negatively predicted short-term gains of the passé composé in written production. Language learning anxiety was not related to gain scores in either of the two groups.

11.3. Theoretical implications

The current study was undertaken within an interactionist-cognitive framework. In such a framework corrective feedback is viewed as potentially facilitative of L2 acquisition. The results of this study lend support to this general theoretical framework as CF was found effective in promoting increasing levels of accuracy in the passé composé and the imparfait. This study reported a clear effect for CF in semi-constrained oral production and written production, suggesting that CF had significantly contributed to the participants' improved linguistic performance.

Interactionist-cognitive theories vary in which type of CF they view as most likely to contribute to L2 acquisition. They vary with regard to:

1. The relative effects of recasts vs. output-prompting CF
2. The relative effects of implicit vs. explicit CF
3. The role of uptake with repair

11.3.1. The relative effects of recasts vs. output-prompting CF

This study found that recasts significantly improved the learners' levels of accuracy in use and form of both the *passé composé* and the *imparfait*. It provides support for Doughty's (2001) claim that attention to form and, in particular, 'cognitive comparison' where own erroneous output is contrasted to the target input, promote acquisition by assisting learners to map the form onto meaning and use. Doughty argues that an essential condition for acquisition to take place is the 'noticing' (Schmidt, 2001) or a 'highlighted state of activation' (N. C. Ellis, 2005) that arises when learners are attempting to communicate meaningfully. In this respect the findings of the current study support the Interaction hypothesis (Long, 1996) and focus-on-form as explained by cognitive accounts of L2 acquisition. Recasts do not interrupt the flow of communication and they are not obtrusive although they do interrupt the learners' cognitive processes while they are engaged with the meaning (Doughty, 1991, 2001; Doughty & Varela, 1998; Long 1981a, 1983b, 1996, 2007). This is consistent with the research in L1 acquisition which has found that language development is assisted by the provision of recasts in the context of conversational interaction (Farrar, 1990, 1992; Leonard, 2011; Nelson, 1987; Saxton, 1997, 2010; Speidel & Nelson, 1989). Farrar (1990) argued that the beneficial effect of recasts is due to the fact that they constitute "a semantically contingent response" (p. 609) to an error which occurred in a conversation.

The beneficial effects of recasts were evident in both the oral and the written production post-tests that required the use of the target structures in context. In particular, the fact that the results were obtained in spontaneous oral production constrained by the given context, suggests that the provision of recasts during communicative tasks resulted in restructuring of the learners' interlanguage and contributed to the development of the learners' implicit knowledge. These learners were able to demonstrate progress in their accurate (spontaneous) use of the target language in context, relative to their level of development. As N. Ellis (2005, p. 308) pointed out, "in fluency, both language processing and language tallying are typically unconscious; our implicit system automatically processes the input, allowing our conscious selves to concentrate on the meaning rather than form". In this respect the findings of the current study are significant and rather unique: they provide information about the change resulting from instruction and measured in a way which is different from the majority of 'type – of – instruction' studies that are "based much more extensively on the application of

explicit declarative knowledge under controlled conditions, without much requirement for fluent, spontaneous use of contextualised language” (Norris & Ortega, 2000, p. 486). Actually, as Norris and Ortega report, over 90 % of the outcome variables in their seminal meta-analysis required the application of L2 pedagogic rules in “highly focused and discrete ways”. Therefore it is not surprising that the results of the meta-analysis confirmed the preference for explicit instructional methods.

Although the present findings corroborate the results of laboratory studies that have clearly shown a positive effect of recasts on language acquisition (e.g. Braid, 2002; Egi, 2007; Han, 2002; Iwashita, 2003; Leeman, 2003; Mackey, 2006; Mackey & Philp, 1998; McDonough & Mackey, 2006; Philp, 2003), they differ from a number of classroom studies, especially those undertaken in immersion settings, which have shown that recasts may be less effective than output-prompting CF. Indeed, the current study did not find clarification requests as effective as recasts. Thus, it failed to support the claims of those cognitive linguists who contend that prompting techniques are more effective than recasts in assisting L2 acquisition (e.g. Chaudron, 1977; Havranek, 2002; Havranek & Cesnik, 2001; Lyster and Ranta, 1997; Lyster, 1998, 2004; Schachter, 1991). However, previous studies differ from the current study in that they compared the effects of recasting with those of multiple prompting strategies. It should be noted that the present study used only one specific type of prompting - clarification requests - and therefore its findings cannot be extended to all types of prompting. The main reason why clarification requests did not result in much progress may have been in the fact that they constitute a strategy used to negotiate meaning rather than to point to a problem with the form. Clarification requests can be a helpful tool at the beginning stages of language learning when they encourage learners to make their utterances more comprehensible (e. g. Ellis, Tanaka & Yamazaki; Nobuyoshi & Ellis, 1993; Oliver, 1995, 1998, 2000; Pica, 1992, 1994; Pica, Holliday, Lewis & Morgenthaler, 1989; Takashima & Ellis, 1999). However, at later stages, when learners are already able to produce utterances in a comprehensible manner but lack the target-like accuracy, clarification requests may be less effective, in particular if they do not result in repair. Lyster’s studies have shown that clarification requests have a high percentage of uptake but not repair. A possible explanation for this is that learners treat clarification requests as requiring them to clarify the meaning of their utterances rather than to correct the linguistic form. Lyster and Ranta (1997), for example, noted that clarification requests can refer to problems in either comprehensibility or accuracy, or both (p. 47) and thus may be ambiguous.

11.3.2. The relative effects of implicit vs. explicit CF

The extent to which recasts are to be viewed as an implicit form of CF has recently been debated. A number of researchers (e.g. Lyster & Ranta, 1997; Lyster, 1998, 2004; Panova & Lyster, 2002; Yang & Lyster, 2010) argued that this type of feedback is too implicit and functionally ambiguous. Ellis (2007), Ellis et al. (2006) and Sheen (2008) compared the effects of recasts as a type of implicit CF with the metalinguistic CF as a type of explicit CF, and they all reported that metalinguistic feedback was more effective than the recasts. However, the current study demonstrates that recasts can be effective in a classroom context if they are perceived by learners as corrective. This concurs with the claims put forward by Nicholas et al. (2001) who argued that recasts are most effective when learners have a clear perception of recasts as a reaction to the accuracy of form. Therefore, some researchers have suggested that the implicitness of recasts is likely to vary according to the context: in fact, there is an emerging consensus (e.g. Ellis, 2010; Lyster & Saito, 2010; Sheen, 2010) that oral CF is best seen as a continuum from implicit to explicit and that there are no fixed boundaries for the different CF techniques, including recasts. If this view is adopted, it can be argued that recasts in the current study were of the more explicit kind because the learners apparently perceived them as corrective. Moreover, the context of the study (FL classroom) may have influenced the participants' perceptions so that more attention was paid to the language itself. In this regard the results of the present study support the Counterbalance hypothesis (Lyster & Mori, 2006). The results provide evidence of the effectiveness of recasts in a classroom environment where learners are expected to pay attention to accuracy. Although the treatment tasks were communicative and meaning oriented, and the learners were primarily engaged with meaning, they appeared to have recognized the corrective force of recasts to a much greater extent than the young students in immersion settings.

The recasts in the current study were short, focused on specific forms, involved only one change, were in interrogative mode (rising intonation) and were repeated if needed. In short, the recasts were highly salient to the learners. "The perceived strength of stimuli" as salience has been referred to (N. Ellis & Collins, 2009, p. 331) was clearly assisted by prosodic cues such as the mode, brevity, one change and focus on specific linguistic forms. It is also important to emphasize, however, that the notion of salience can be subjective, depending on the learner's capacity to notice the corrective function of a recast. A learner who understands the message will notice more easily the contrast between his/her own production and the corrective recast, whereas a learner who does not understand the interlocutor's utterance will

not be able to interpret a recast even if it was ‘highlighted’ by the tone, pitch or stress. In the current study repetition was used in order to increase the saliency but it did not significantly contribute to the effectiveness of CF. The participants’ reactions were similar to the student behaviour in Doughty and Varela’s (1998) study – that is, the students learnt very quickly what they were expected to do, i.e. when they noticed that the teacher was repeating the recast if they had not repaired, they began to produce uptake with repair immediately after the first recast.

11.3.3. The role of uptake with repair

The association between learning outcomes and uptake with repair in the current study is evident only in the RE group and only for long-term gains in oral production of the *passé composé*. In written production of the *passé composé* in the RE group, however, those learners who were only exposed to repair showed greater gains. In the CR group, no link between uptake with repair and gains was found. This may indicate that producing uptake with repair may not be of crucial importance for learning to take place (Ellis, Tanaka & Yamazaki, 1994; Havranek, 2002; Mackey, 2006; Mackey, Oliver & Leeman, 2003; Mackey & Philp 1998; Muranoi, 2000; Pica, 1992). In particular, where recasts are involved, it has already been acknowledged that uptake is an optional move depending solely on the learner’s willingness to repeat the interlocutor’s recast.

Nevertheless, the current findings suggest that uptake and repair may contribute to the durability of gains in oral production. Numerous studies (Loewen, 2005; Lyster, 2004; Nabei & Swain, 2002; Swain & Lapkin, 1998; Tocalli-Beller & Swain, 2005) have found a clear relationship between successful uptake and learning. These support the Output hypothesis (Swain, 1985) which claims that comprehensible output plays an important role in helping learners achieve advanced levels of grammatical accuracy. Swain (1985) argued that one function of output is to provide opportunities for meaningful use of learners’ linguistic resources. Two additional roles that can be added to output apart from the ‘contextualized’ and “pushed” language use, are that : 1) output provides opportunities to test out hypotheses learners have made of expressions (and see if they work) and 2) using the language may force the learner to move from semantic processing to syntactic processing. The question arises then as to why in the present study the pushed output in uptake with repair was not related to gains in accuracy except for long term gains in the RE group. Overall, the results of the study

indicate that it is the provision of the target forms in the input and their frequency that have significantly contributed to the change.

11.3.4. Individual differences and corrective feedback

Interactionist theories also address the role that individual learner difference factors play in mediating the effects of different kinds of CF. This study investigated four individual difference factors: the learners' L2 proficiency, language analytic ability, working memory and language anxiety.

The findings of the present study support previous research which has shown that for recasts to be effective they need to satisfy certain conditions, in particular the learner developmental readiness. If learners are at the stage when they are ready to acquire the target structure, and if the meaning of the context is clear to them, they will most probably perceive the recast as a correction and may benefit from it. The readiness of the learner and familiarity with the context have been addressed in both L1 research (Farrar, 1990, 1992; Nelson, 1987; Speidel & Nelson, 1989) and L2 research (Doughty, 2001; Han, 2002; Iwashita, 2003; Mackey, 2006; Mackey & Philp, 1998; Philp, 2003). Farrar contended that developmental readiness was crucial for a child's use of corrective recasts. Classroom studies in L2 found that recasts were effective for learners of higher proficiency (Ammar, 2008; Ammar & Spada, 2006) whereas clarification requests facilitated L2 acquisition in both lower and higher proficiency learners. The current study confirmed the superiority of recasts for higher proficiency learners when their post-test results were compared with their peers' who received clarification requests. However, when learning is measured within groups, i.e. between the pre-test and the post-tests then both the lower and the higher proficiency learners in the RE group showed significant improvement. In the CR group, on the other hand, there is evidence of significant between-test improvement and larger effect sizes only for learners of lower proficiency. Such findings contrast the results of the previous similar studies, but could be explained by the different target structures. In the current study, the learning of the passé composé complex morphological form, for example, probably requires more frequent input of the target structures and more scaffolding which may be available through the provision of recasts rather than through the provision of clarification requests.

This study in part confirms the findings concerning the mediating role of cognitive ID factors analytic ability and working memory. Language analytic ability was found to be a strong

predictor in the CR group (for all gains except for the written imparfait). This finding is consistent with the body of research in L2 acquisition that has provided strong evidence of the important role of language aptitude, specifically language analytic ability, in L2 learning (de Graaf, 1997; DeKeyser, 2000; Ehrman & Oxford, 1995; Ranta, 2002; Sheen, 2007, 2008). On the other hand, the results for the RE group concur with the Robinson's findings (1996, 1997a, 2002c) demonstrating that aptitude, i. e. analytic ability may not play a major role in incidental language learning. This is an indication of the particular value of recasts: that is, when provided in a conversational context, they can facilitate the learning of all learners, irrespective of their analytic ability. An exception to this may be the long-term gains for passé composé in oral production, suggesting that aptitude may play a role when formally complex structures have to be retained and produced in a spontaneous, unplanned speech.

Working memory predicted the short-term gains for passé composé in oral production in the RE group but not in the CR group. In contrast, the gains in written production of the passé composé in the RE group were related to lower levels of WM. These results clearly show that working memory is determined by the task being performed (Conway et al., 2005; Daneman, 1991; Daneman & Green, 1986; Just & Carpenter, 1992). As researchers working with the reading and speaking span tests have often emphasized, individual differences in working memory become evident only in tasks that are cognitively more demanding. Clearly, semi-constrained and unplanned speech production is more cognitively demanding than the same kind of task in written production and therefore the higher levels of accuracy in speech production may be expected from learners with larger working memory. Robinson (2005) also found that incidental learning, which was not related to analytic ability, significantly correlated with the scores on Daneman and Carpenter's (1980) reading span test. Indeed, the current findings lend support to Robinson's (2002b) theoretical framework of Aptitude complex hypothesis in which cognitive abilities are closely related to particular learning conditions (p.119). In Robinson's framework learners with stronger 'memory for contingent speech' and noticing-the-gap benefit from implicit CF such as recasts. According to Robinson (2005a), the phonological WM capacity and the processing speed jointly contribute to the 'memory for contingent speech' and enable the recast and the trigger (erroneous) utterance to be maintained in working memory long enough for 'cognitive comparison' to take place.

Finally, the present study failed to confirm the role of language anxiety in mediating the effects of oral CF. There may be two reasons why this finding differs from Sheen's (2008).

The first one is that the participants in the current study took part in interaction on a voluntary basis so they were not forced to speak. The second reason is that those who participated in interaction were probably comfortable when communicating in L2. Therefore, language anxiety did not affect their performance.

11.3.5. Effects of CF on the different grammatical structures

The current study showed that recasts were, more than clarification requests, facilitative of the acquisition of both structures: the passé composé as a morphologically complex structure and the imparfait as a morphologically simpler form. The term *complexity* is used here to refer to the complexity of form (DeKeyser, 2005) which is one aspect of the structure complexity defined from a linguistic perspective - the other two are *complexity of meaning* and *complexity of form-meaning mapping*.

As pointed out in the discussion related to oral production (Chapter Seven), it is not possible to claim that the learners in the present study have fully acquired the abstract rules of the passé composé and the imparfait even though they have clearly improved their level of accuracy in the use of the two structures. It is possible that both the passé composé and the imparfait were learnt as exemplars, or chunks. As N. C. Ellis argued, “that morphology can be described as rule-like behavior does not imply that morphology is rule-governed” (1998, p. 56). There is clear evidence in the current study that, due to the input-providing nature of recasts, there was a significant difference between the frequencies of the target structures in the two treatment groups. This suggests a usage-based explanation for the acquisition of the two structures. It is possible, of course, that in language acquisition both chunking procedures and creative abstract rule-construction are at work “in a dynamic tension that drives forward the overall development of the L2 system” (Myles et al., 1999, p. 76).

11.4. Pedagogical implications

The current findings clearly show that recasts as a CF strategy can be effective in the foreign language classroom. The findings also suggest that recasts can contribute to the development of the kind of implicit knowledge needed for free oral production. As already noted, previous studies have shown recasts to be effective in laboratory conditions but less effective in classrooms contexts. This study, however, has demonstrated that recasts can be effective in a

class of about 20 students. In terms of pedagogy and teaching methodology, the current findings suggest that:

- corrective feedback, specifically recasts, can be successfully incorporated into communicative tasks and teacher-student classroom interaction,
- focused tasks present a viable pedagogical tool for providing a rich communicative environment in which interaction can take place, and
- focus on form can be provided in a natural manner.

The focused tasks that were used in this study presented an appropriate context for L2 learning and provided ample opportunities for: a) conversational interaction involving focus on form, b) comprehensible input and c) comprehensible output. Focused communicative tasks, as Ellis (2003) emphasized, are valuable to teachers because they make possible the teaching of specific linguistic features in a communicative way. Furthermore, they are valuable to researchers because they provide evidence of what learners can do spontaneously, i. e. they show what the learners have acquired. Ellis also indicated that the challenge for both teachers and researchers when designing focused tasks is in making the use of the target structure “natural” or “essential” (2003, p. 153). In designing the tasks used in the current study it was important to choose a picture story that would allow for a narrative involving both foregrounding and backgrounding so that the use of the *passé composé* and the *imparfait* would be natural. The task instructions clearly indicated that a report should be prepared either for the school magazine or the school broadcasting programme, and the first sentence was given with reference to the past event. The tasks designed in this way were clearly successful in eliciting obligatory contexts for the two target structures in sufficient numbers.

Ellis (2003) suggested that “if the task is to assist acquisition it should be directed at a structure that learners are in the process of acquiring” (p. 155). It is also important to underscore that the learners were provided with all vocabulary needed to complete the task as the context had to be made clear to them. However, it is not always easy to find a ‘matching’ task for a particular structure and the learners’ interlanguage level, and certainly there are language structures for which focused tasks cannot be designed. Nevertheless, it is worth investigating and trialing different tasks because the results can be rewarding.

The current study confirms that oral corrective feedback can be successfully incorporated in communicative tasks in the classroom. In particular, this is true for recasts which are

facilitative of learning, yet also unobtrusive. As already pointed out, some teaching methodologists do not recommend the use of corrective feedback in fluency activities. However, there is now sufficient research evidence to indicate that CF, in particular recasts, are beneficial when provided during the performance of communicative tasks and that they need not interfere unduly with the primary focus on meaning.

A further implication of this study is that CF is more effective if it is directed at a grammatical structure that learners are already in the process of acquiring. CF directed at target features that are beyond the learner's developmental level may be less or not effective.

11.5. Limitations of the study

There are several limitations to this study.

- First, it may be argued that the sample size was not large, involving 52 participants in total. It was balanced, with the same number of subjects (18) in each treatment group and 16 subjects in the control group. However, this is the average or even larger than the average FL class in NZ secondary schools, so it can be claimed that the study had ecological validity. Second, given the type of testing used, with oral narratives performed individually and WM tests taken individually, any larger sample size would not have been possible logistically. With these class numbers it was very difficult to complete all the tests and tasks required by the research design.
- The fact that the researcher took the role of the teacher and participated in the treatment tasks with both experimental groups, delivering recasts in one class, and clarification requests in the other class, may be regarded as a limitation to the study. Although all care was taken in order to deliver both types of corrective feedback in a highly professional manner strictly guided by the research design, it is acknowledged that the dual role of the teacher/researcher might have impacted on the outcomes of the study. However, taking into consideration the findings from conversation analysis in classroom studies (Foster & Ohta, 2005; Seedhouse, 2004), it does not seem that the results would be much different if a different teacher/researcher conducted the study.
- At first sight, the difference in the number of corrections provided to each treatment group may constitute a limitation of the study. However, this problem was resolved

and the difference accounted for by the use of ANCOVA, with the number of corrections as a covariate. As eventually became clear, it was not the quantity of CF episodes but the nature of each type of corrective feedback, that proved to be essential for causing the differential effects in the two experimental groups.

- The choice of Speaking span test as a test of working memory may be seen as a limitation, too, because the test used was designed to measure WM capacity in relation to fluency. The current study, however, was only interested in the learners' accuracy in the production of the two target structures. In retrospect, a test of phonological short-term memory would have been an appropriate instrument to be administered along with the SST, since PSTM has been claimed to be predictive of learners' acquisition of grammatical structures (N. C. Ellis & Sinclair, 1996).
- A further limitation may have been the use of Multiple regression analysis with the current sample size. As already indicated in Chapter Ten, Multiple regression analysis requires a larger sample size. However, some authors think that 10 subjects per predictor variable should be the smallest number – this condition was satisfied in the current study.
- The type of school at which the participants in the experimental groups were recruited may be considered one more limitation. It was a Decile 10 school - that is, one of the top schools in New Zealand, the highest decile indicating a high socio-economic status of the students' families. So, the sample probably does not represent the larger population of foreign language learners in New Zealand. However, this was the only school in Auckland that responded to my request on the mail list of New Zealand French Teachers' Association to allow me to conduct a study with the students of French. The participants in the control group were my own students, at a Decile 7 school.

11.6. Suggestions for future research

Based on the findings of the current study some directions for future research can be proposed. This study has provided interesting data which would need to be further corroborated by future research. The current findings are consistent with the majority of laboratory experimental studies investigating the effects of recasts on L2 acquisition. They are also in line with the research in L1 acquisition, in particular with the findings obtained in

the area of child L1 specific language impairment. However, the current findings are in contrast with a number of classroom studies, mostly in immersion contexts, that have found recasts to be ambiguous or ineffective. Therefore, more research in foreign language classrooms is clearly needed. It would be important to further examine the effectiveness of recasts in classroom settings, by manipulating the extent to which recasts are made more or less salient depending on the target structure and the learners' proficiency level. It would be interesting, as well, to examine the effect of time or opportunity given to learners to respond to recasts as proposed by Long (2007) and Oliver (1995, 1998).

The specific type of prompts investigated in the present study were clarification requests which, as I noted above, may not be seen as 'corrective'. They might be ambiguous and even face threatening (Foster & Ohta, 2005) if they are focused exclusively on language form. It should not be concluded, however, that prompts are less effective in promoting acquisition. Studies are needed that examine the effectiveness of different kinds of prompts in promoting different types of learning.

A desired feature of future classroom experimental studies should be the inclusion of communicative tasks as instruments, not only in treatment but in testing procedures as well. Focused tasks, as the current study shows, can constitute an appropriate instrument to elicit particular target structures and to measure the acquisition rather than the explicit knowledge. Although such a research design may be more challenging to manage, there is a greater chance that it will provide more reliable data that can show a clearer picture of classroom language learning.

The inclusion of learner individual differences in an experimental research design is also important since ID factors may reveal how different IDs interact with different types of CF. A desirable future project would aim at investigating and matching the types of feedback with the types of learners based on their ID characteristics, so that the best CF strategies can be chosen for particular learner profiles.

References

- Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the Zone of proximal development. *The Modern Language Journal*, 78(4), 465-483.
- Allwright, D. (1988). *Observation in the language classroom*. London ; New York: Longman.
- Ammar, A. (2008). Prompts and recasts: Differential effects on second language morphosyntax. *Language Teaching Research*, 12(2), 183-210.
- Ammar, A., & Spada, N. (2006). One size fits all? Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition*, 28(4), 543-574.
- Andersen, R. W., & Shirai, Y. (1994). Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition*, 16(02), 133-156.
- Annett, J. (1969). *Feedback and human behaviour: the effects of knowledge of results, incentives and reinforcement on learning and performance*. Harmondsworth,: Penguin.
- Austin, J. L., & Urmson, J. O. (1962). *How to do things with words : the William James Lectures delivered at Harvard University in 1955*. Oxford: Clarendon Press.
- Ayoun, D. (2001). The role of negative and positive feedback in the second language acquisition of the passe compose and the imparfait. *The Modern Language Journal*, 85(2), 226-243.
- Ayoun, D. (2004). The effectiveness of written recasts in the second language acquisition of aspectual distinctions in french: A follow-up study. *The Modern Language Journal*, 88(1), 31-55.
- Ayoun, D. (2008). *Studies in French applied linguistics*. Amsterdam ; Philadelphia: John Benjamins Pub. Company.
- Ayoun, D., & Salaberry, M. R. (2005). *Tense and aspect in Romance languages : theoretical and applied perspectives*. Amsterdam ; Philadelphia, PA: J. Benjamins.
- Baddeley, A. D. (2007). *Working memory, thought, and action*. Oxford ; New York: Oxford University Press.
- Baddeley, A. D., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review January*, 105(1), 158-173.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47-89). New York: Academic Press.

- Baddeley, A. D., & Logie, R. H. (1999). Working memory: the multiple component model. In A. Miyake, & Shah, P. (Ed.), *Models of working memory : mechanisms of active maintenance and executive control* (pp. 28-59). Cambridge; New York: Cambridge University Press.
- Bardovi-Harlig, K. (2000). *Tense and aspect in second language acquisition : form, meaning and use*. Oxford: Blackwell.
- Bates, E., & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes*, 12(5-6), 507-584.
- Beck, M.-L., & Eubank, L. (1991). Acquisition theory and experimental design. *Studies in Second Language Acquisition*, 13(01), 73-76.
- Bialystok, E., & Ryan, E. B. (1985). A metacognitive framework for the development of first and second language skills. In D.-L. Forrest-Pressley, G. E. MacKinnon & T. Gary Waller (Eds.), *Metacognition, Cognition, and Human Performance* (pp. 207-252). Orlando, Florida: Academic Press, Inc.
- Bock, K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18(3), 355-387.
- Bock, K. (1996). Language production: Methods and methodologies. *Psychonomic Bulletin & Review*, 3(4), 395-421. doi: 10.3758/bf03214545
- Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General*, 129(2), 177-192.
- Bohannon, J. N., III, MacWhinney, B., & Snow, C. (1990). No negative evidence revisited: Beyond learnability or who has to prove what to whom. *Developmental Psychology*, 26(2), 221-226.
- Bohannon, J. N., III, & Stanowicz, L. (1988). The issue of negative evidence: Adult responses to children's language errors. *Developmental Psychology*, 24(5), 684-689.
- Braidi, S. M. (2002). Reexamining the role of recasts in native-speaker/nonnative-speaker interactions. *Language Learning*, 52(1), 1-42. doi: 10.1111/1467-9922.00176
- Brown, J. D., & Rodgers, T. S. (2002). *Doing second language research*. Oxford: Oxford University Press.
- Camarata, S., Nelson, K. E., Gillum, H., & Camarata, M. (2009). Incidental receptive language growth associated with expressive grammar intervention in SLI. *First Language*, 29(1), 51-63.
- Carpenter, H., Jeon, K. S., MacGregor, D., & Mackey, A. (2006). Learners' interpretations of recasts. *Studies in Second Language Acquisition*, 28(2), 209-236.

- Carroll, J. B. (1962). The prediction of success in foreign language training. In R. Glaser (Ed.), *Glaser, R. (1962). Training research and education*. Pittsburgh: University of Pittsburgh Press.
- Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude*. Rowley, Mass.: Newbury House.
- Carroll, S., Swain, M., & Roberge, Y. (1992). The role of feedback in adult second language acquisition: Error correction and morphological generalizations. *Applied Psycholinguistics*, 13(02), 173-198.
- Chaudron, C. (1977). A descriptive model of discourse in the corrective treatment of learners' errors. *Language Learning*, 27(1), 29-46. doi: 10.1111/j.1467-1770.1977.tb00290.x
- Chaudron, C. (1988). *Second language classrooms : research on teaching and learning*. Cambridge [Cambridgeshire] ; New York: Cambridge University Press.
- Chomsky, N. (1975). *Questions of form and interpretation*. Lisse: Peter De Ridder Press.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, Holland ; Cinnaminson, [N.J.]: Foris Publications.
- Cleland, A. A., & Pickering, M. J. (2003). The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *Journal of Memory and Language*, 49(2), 214-230. doi: 10.1016/s0749-596x(03)00060-3
- Comrie, B. (1976). *Aspect : an introduction to the study of verbal aspect and related problems*. Cambridge ; New York: Cambridge University Press.
- Conway, A., Kane, M., Bunting, M., Hambrick, D., Wilhelm, O., & Engle, R. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12(5), 769-786. doi: 10.3758/bf03196772
- Daneman, M. (1991). Working memory as a predictor of verbal fluency. *Journal of Psycholinguistic Research*, 20(6), 445-464. doi: 10.1007/bf01067637
- Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 9(4), 561-584.
- Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. [doi: DOI: 10.1016/0749-596X(86)90018-5]. *Journal of Memory and Language*, 25(1), 1-18.
- Daneman, M., & Merikle, P. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, 3(4), 422-433. doi: 10.3758/bf03214546
- Davies, A., & Elder, C. (2004). *The handbook of applied linguistics*. Malden, MA: Blackwell.

- de Graaff, R. (1997). The EXperanto experiment. *Studies in Second Language Acquisition*, 19(02), 249-276. doi: doi:null
- DeKeyser, R. (1993). The effect of error correction on L2 grammar knowledge and oral proficiency. *The Modern Language Journal*, 77(4), 501-514.
- DeKeyser, R. (2000). The robustness of Critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22(04), 499-533.
- DeKeyser, R. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning*, 55(S1), 1-25. doi: 10.1111/j.0023-8333.2005.00294.x
- DeKeyser, R. (2007a). *Practice in a second language : perspectives from applied linguistics and cognitive psychology*. Cambridge ; New York: Cambridge University Press.
- DeKeyser, R. (2007b). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition : an introduction*. Mahwah, N.J. ; London: (pp. 97-113). Mahwah, N.J. ; London: Lawrence Erlbaum Associates.
- Dörnyei, Z. (2005). *The psychology of the language learner : individual differences in second language acquisition*. Mahwah, N.J.: L. Erlbaum Associates.
- Doughty, C. (1991). Second language instruction does make a difference. *Studies in Second Language Acquisition*, 13(04), 431-469. doi: doi:10.1017/S0272263100010287
- Doughty, C. (2001). Cognitive underpinnings of focus on form. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 206-257). Cambridge: Cambridge University Press.
- Doughty, C., & Pica, T. (1985). Input and interaction in the communicative language classroom: a comparison of teacher-fronted and group activities. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 115-132). Rowley, Mass: Newbury House Publishers.
- Doughty, C., & Pica, T. (1986). Information Gap Tasks: Do They Facilitate Second Language Acquisition? *TESOL Quarterly*, 20(2), 305-325.
- Doughty, C., & Varela, E. (1998). Communicative focus on form. In C. Doughty & J. Williams (Eds.), *Focus on Form in Classroom Second Language Acquisition*. Cambridge: Cambridge University Press.
- Doughty, C., & Williams, J. (1998). *Focus on form in classroom : second language acquisition*. New York: Cambridge University Press.
- Edge, J. (1993). *Essentials of English language teaching*. London ; New York: Longman.
- Egi, T. (2007). Interpreting recasts as linguistic evidence: The roles of linguistic target, length, and degree of change. *Studies in Second Language Acquisition*, 29(4), 511-537.

- Ehrman, M. E., Leaver, B. L., & Oxford, R. L. (2003). A brief overview of individual differences in second language learning. *System*, 31(3), 313-330.
- Ehrman, M. E., & Oxford, R. L. (1995). Cognition plus: Correlates of language learning success. *The Modern Language Journal*, 79(1), 67-89.
- Ellis, N., & Collins, L. (2009). Input and second language acquisition: The roles of frequency, form, and function introduction to the special issue. *The Modern Language Journal*, 93(3), 329-336.
- Ellis, N. C. (1994a). Implicit and explicit language learning - An overview. In N. C. Ellis (Ed.), *Ellis, N. C. (1994). Implicit and explicit learning of languages* (pp. 1-31). London ; San Diego, Calif: Academic Press.
- Ellis, N. C. (1994b). *Implicit and explicit learning of languages*. London ; San Diego, Calif.: Academic Press.
- Ellis, N. C. (1996). Sequencing in SLA. *Studies in Second Language Acquisition*, 18(01), 91-126. doi: doi:10.1017/S0272263100014698
- Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and Implicit language knowledge. *Studies in Second Language Acquisition*, 27(2), 305-352.
- Ellis, N. C., & Schmidt, R. (1997). Morphology and longer distance dependencies. *Studies in Second Language Acquisition*, 19(02), 145-171. doi: doi:null
- Ellis, N. C., & Sinclair, S. G. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *The Quarterly Journal of Experimental Psychology A*, 49, 234-250.
- Ellis, R. (1990). *Instructed second language acquisition : learning in the classroom*. Oxford, UK ; Cambridge, Mass., USA: B. Blackwell.
- Ellis, R. (1992). Learning to communicate in the classroom. *Studies in Second Language Acquisition*, 14(01), 1-23. doi: doi:10.1017/S0272263100010445
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (1995). Interpretation tasks for grammar teaching. *TESOL Quarterly*, 29(1), 87-105.
- Ellis, R. (1999a). *Learning a second language through interaction*. Amsterdam ; Philadelphia: J. Benjamins.
- Ellis, R. (1999b). Theoretical perspectives on interaction and language learning. In R. Ellis (Ed.), *Ellis, R. (1999). Learning a second language through interaction* (pp. 3-31). Amsterdam ; Philadelphia: J. Benjamins.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford, U.K.: Oxford University Press.

- Ellis, R. (2004). Individual differences in second language learning. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics* (pp. 525-551). Malden, MA: Blackwell Publishing.
- Ellis, R. (2006). Researching the effects of form-focussed instruction on L2 acquisition. [doi:10.1075/aila.19.04ell]. *AILA Review*, 19, 18-41.
- Ellis, R. (2007). The differential effects of corrective feedback on two grammatical structures. In A. Mackey (Ed.), *Conversational interaction in second language acquisition : a collection of empirical studies*. Oxford: Oxford University Press.
- Ellis, R. (2012). *Second language research and language pedagogy*. Malden, MA: Wiley-Blackwell.
- Ellis, R., Basturkmen, H., & Loewen, S. (2001). Learner uptake in communicative ESL lessons. *Language Learning*, 51(2), 281-318. doi: 10.1111/1467-9922.00156
- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, 28(2), 339-368.
- Ellis, R., & Sheen, Y. (2006). Reexamining the role of recasts in second language acquisition. *Studies in Second Language Acquisition*, 28(4), 575-600.
- Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension, and the acquisition of L2 word meanings. *Language Learning*, 44(3), 449-491. doi: 10.1111/j.1467-1770.1994.tb01114.x
- Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research*, 9(2), 147-171. doi: 10.1191/1362168805lr161oa
- Farrar, M. J. (1990). Discourse and the acquisition of grammatical morphemes. *Journal of Child Language*, 17(03), 607-624. doi: doi:10.1017/S0305000900010904
- Farrar, M. J. (1992). Negative evidence and grammatical morpheme acquisition. *Developmental Psychology*, 28(1), 90-98.
- Field, A. P. (2009). *Discovering statistics using SPSS : (and sex, drugs and rock'n'roll)* (3rd ed.). London: SAGE.
- Foster, P., & Ohta, A. S. (2005). Negotiation for meaning and peer assistance in second language classrooms. *Applied Linguistics*, 26(3), 402-430. doi: 10.1093/applin/ami014
- French, L. M. (2006). *Phonological working memory and second language acquisition : a developmental study of francophone children learning english in Quebec*. Lewiston, N.Y.: Edwin Mellen Press.
- Gardner, R. C. (1985). *Social psychology and second language learning : the role of attitudes and motivation*. London ; Baltimore, Md., U.S.A.: E. Arnold.

- Gass, S. M. (1997). *Input, interaction, and the second language learner*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: L. Erlbaum Associates.
- Gass, S. M., & Varonis, E. M. (1994). Input, interaction, and second language production. *Studies in Second Language Acquisition*, 16(03), 283-302. doi: doi:10.1017/S0272263100013097
- Gathercole, S., & Baddeley, A. D. (1993). *Working memory and language*. Hove, [England]: Lawrence Erlbaum.
- Grimshaw, J., & Pinker, S. (1989). Positive and negative evidence in language acquisition. *Behavioral and Brain Sciences*, 12(02), 341-342. doi: doi:10.1017/S0140525X00048950
- Guará Tavares, G. (2008). *Pre-task planning, working memory capacity, and L2 speech performance*. (PhD thesis), Universidade federal de Santa Catarina, Florianopolis.
- Gülzow, I., & Gagarina, N. (2007). *Frequency effects in language acquisition : defining the limits of frequency as an explanatory concept*. Berlin ; New York: Mouton de Gruyter.
- Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly*, 36(4), 543-572.
- Harley, B. (1989). Functional grammar in french immersion: A classroom experiment. *Applied Linguistics*, 10(3), 331-360. doi: 10.1093/applin/10.3.331
- Harley, B. (1993). Instructional strategies and SLA in early French immersion. *Studies in Second Language Acquisition*, 15(02), 245-259. doi: doi:10.1017/S0272263100011980
- Harley, B., & Hart, D. (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition*, 19(03), 379-400. doi: doi:null
- Harley, B., & Hart, D. (2002). Age, aptitude and second language learning on a bilingual exchange. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 301-330). Amsterdam ; Philadelphia, PA: J. Benjamins Pub.
- Harley, B., & King, M. L. (1989). Verb lexis in the written compositions of young L2 learners. *Studies in Second Language Acquisition*, 11(04), 415-439. doi: doi:10.1017/S0272263100008421
- Harmer, J. (2001). *The practice of English language teaching* (3rd ed.). Harlow: Longman.

- Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, 14(01), 25-38. doi: doi:10.1017/S0272263100010457
- Havranek, G. (2002). When is corrective feedback most likely to succeed? *International Journal of Educational Research*, 37, 255-270.
- Havranek, G., & Cesnik, H. (2001). Factors affecting the success of corrective feedback. *EUROSLA Yearbook*, 1, 99-122.
- Hedge, T. (2000). *Teaching and learning in the language classroom*. Oxford: Oxford University Press.
- Hendrickson, J. M. (1978). Error correction in foreign language teaching: Recent theory, research, and practice. *The Modern Language Journal*, 62(8), 387-398.
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125-132.
- Iwashita, N. (2003). Negative feedback and positive evidence in task-based interaction: Differential effects on L2 development. *Studies in Second Language Acquisition*, 25(1), 1-36.
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. [doi: 10.1016/0010-0285(89)90003-0]. *Cognitive Psychology*, 21(1), 60-99.
- Juffs, A. (2004). Representation, processing and working memory in a second language. [Article]. *Transactions of the Philological Society*, 102(2), 199-225. doi: 10.1111/j.0079-1636.2004.00135.x
- Juffs, A. (2005). The influence of first language on the processing of wh-movement in English as a second language. *Second Language Research*, 21(2), 121-151. doi: 10.1191/0267658305sr255oa
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99(1), 122-149.
- Kaplan, M. A. (1987). Developmental patterns of past tense acquisition among foreign language learners of French. In B. VanPatten, T. Dvorak & J. F. Lee (Eds.), *Foreign language learning : a research perspective* (pp. 52-60). Cambridge [Cambridgeshire]; New York: Newbury House.
- Katz, S. L., & Blyth, C. S. (2007). *Teaching French grammar in context : theory and practice*. New Haven: Yale University Press.
- Keck, C. M., Iberri-Shea, G., Tracy-Ventura, N., & Wa-Mbaleka, S. (2006). Investigating the empirical link between task-based interaction and acquisition: A meta-analysis. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 91-131). Philadelphia: John Benjamins.

- Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition*, 11(02), 261-271. doi: doi:10.1017/S1366728908003416
- Krashen, S. D. (1981). *Second language acquisition and second language learning* (1st ed.). Oxford ; New York: Oxford University Press.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition* (1st ed.). Oxford ; New York: Pergamon.
- Krashen, S. D. (1985). *The input hypothesis : issues and implications*. London ; New York: Longman.
- Krashen, S. D., & Terrell, T. D. (1983). *The natural approach : language acquisition in the classroom* (1st ed.). Oxford [Oxfordshire] ; New York; San Francisco: Pergamon Press; Alemany Press.
- Lantolf, J. P., & Poehner, M. E. (2008). *Sociocultural theory and the teaching of second languages*. London ; Oakville, Ct.: Equinox Pub.
- Larsen-Freeman, D., & Long, M. H. (1991). *An introduction to second language acquisition research*. London ; New York: Longman.
- Leeman, J. (2003). Recasts and second language development: Beyond negative evidence. *Studies in Second Language Acquisition*, 25(1), 37-63.
- Lenneberg, E. H. (1967). *Biological foundations of language*. New York,: Wiley.
- Lenneberg, E. H., & Lenneberg, E. (1975). *Foundations of language development : a multidisciplinary approach*. New York: Academic Press.
- Leonard, L. B. (2011). The primacy of priming in grammatical learning and intervention: A tutorial. *Journal of Speech, Language and Hearing Research*, 54(2), 608-621. doi: 10.1044/1092-4388(2010/10-0122)
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60(2), 309-365. doi: 10.1111/j.1467-9922.2010.00561.x
- Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition*, 27(03), 361-386. doi: doi:10.1017/S0272263105050163
- Loewen, S., & Erlam, R. (2006). Corrective Feedback in the Chatroom: An Experimental Study. *Computer Assisted Language Learning*, 19(1), 1-14.
- Loewen, S., & Nabei, T. (2007). Measuring the effects of oral corrective feedback on L2 knowledge. In A. Mackey (Ed.), *Conversational interaction in second language acquisition : a collection of empirical studies*. Oxford: Oxford University Press.
- Loewen, S., & Philp, J. (2006). Recasts in the adult English L2 classroom: Characteristics, explicitness, and effectiveness. *The Modern Language Journal*, 90(4), 536-556.

- Long, M. H. (1981a). Input, interaction, and second language acquisition. *Annals of the New York Academy of Sciences*, 379(1), 259-278. doi: 10.1111/j.1749-6632.1981.tb42014.x
- Long, M. H. (1981b). Questions in foreigner talk discourse. *Language Learning*, 31(1), 135-157. doi: 10.1111/j.1467-1770.1981.tb01376.x
- Long, M. H. (1983a). Linguistic and conversational adjustments to non-native speakers. *Studies in Second Language Acquisition*, 5(02), 177-193. doi: doi:10.1017/S0272263100004848
- Long, M. H. (1983b). Native speaker/non-native speaker conversation and the negotiation of comprehensible input1. *Applied Linguistics*, 4(2), 126-141. doi: 10.1093/applin/4.2.126
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). San Diego: Academic Press.
- Long, M. H. (2007). *Problems in SLA*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Long, M. H., Inagaki, S., & Ortega, L. (1998). The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish. *The Modern Language Journal*, 82(3), 357-371.
- Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom : second language acquisition* (pp. 15-63). New York: Cambridge University Press.
- Lyster, R. (1998a). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, 48(2), 183-218.
- Lyster, R. (1998b). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, 20(1), 51-81.
- Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26(3), 399-432.
- Lyster, R., & Izquierdo, J. (2009). Prompts versus recasts in dyadic interaction. *Language Learning*, 59(2), 453-498.
- Lyster, R., & Mori, H. (2006). Interactional feedback and instructional counterbalance. *Studies in Second Language Acquisition*, 28(2), 269-300.
- Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19(1), 37-66.
- Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA. *Studies in Second Language Acquisition*, 32(2), 265-302.

- MacIntyre, P. D. (1999). Language anxiety: A review of research for language teachers. In D. Young (Ed.), *Affect in foreign language and second language learning* (pp. 24-46). Boston: MacGraw-Hill College.
- MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning, 44*(2), 283-305. doi: 10.1111/j.1467-1770.1994.tb01103.x
- Mackey, A. (1999). Input, interaction, and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition, 21*(4), 557-587.
- Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied Linguistics, 27*(3), 405-430. doi: 10.1093/applin/ami051
- Mackey, A. (2007). *Conversational interaction in second language acquisition : a collection of empirical studies*. Oxford: Oxford University Press.
- Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning, 60*(3), 501-533. doi: 10.1111/j.1467-9922.2010.00565.x
- Mackey, A., & Goo, J. (2007). Interaction research in SLA: a meta-analysis and research synthesis *Conversational interaction in second language acquisition : a collection of empirical studies*. Oxford: Oxford University Press.
- Mackey, A., Perdue, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition, 22*(04), 471-497. doi: doi:null
- Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *The Modern Language Journal, 82*(3), 338-356.
- Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback and L2 development. In P. Robinson (Ed.), *Individual differences and instructed language learning*. Amsterdam ; Philadelphia, PA,: J. Benjamins Pub. (pp. 181-209). Amsterdam ; Philadelphia, PA: J. Benjamins Pub.
- McDonough, K. (2005). Identifying the impact of negative feedback and learners' responses on ESL question development. *Studies in Second Language Acquisition, 27*(01), 79-103. doi: doi:10.1017/S0272263105050047
- McDonough, K. (2006). Interaction and syntactic priming: English L2 speakers' production of dative constructions. *Studies in Second Language Acquisition, 28*(02), 179-207. doi: doi:10.1017/S0272263106060098
- McDonough, K. (2007). Interactional feedback and the emergence of simple past activity verbs in L2 English. In A. Mackey (Ed.), *Conversational Interaction in Second Language Acquisition* (pp. 323-338). Oxford: Oxford University Press.

- McDonough, K., & Mackey, A. (2006). Responses to recasts: Repetitions, primed production, and linguistic development. *Language Learning*, 56(4), 693-720. doi: 10.1111/j.1467-9922.2006.00393.x
- Miyake, A., Carpenter, P. A., & Just, M. A. (1994). A capacity approach to syntactic comprehension disorders: making normal adults perform like aphasic patients. [doi: 10.1080/02643299408251989]. *Cognitive Neuropsychology*, 11(6), 671-717. doi: 10.1080/02643299408251989
- Miyake, A., & Friedman, N. P. (1998). Individual differences in second language proficiency: working memory as language aptitude. In A. F. Healy & L. E. Bourne (Eds.), *Psycholinguistic studies on training and retention* (pp. 339-364). Mahwah, N.J.: Lawrence Erlbaum.
- Miyake, A., & Shah, P. (1999). *Models of working memory : mechanisms of active maintenance and executive control*. Cambridge ; New York: Cambridge University Press.
- Mizera, G. J. (2006). *Working memory and L2 oral fluency*. (PhD Thesis), University of Pittsburgh.
- Muranoi, H. (2000). Focus on form through interaction enhancement: Integrating formal instruction into a communicative task in EFL classrooms. *Language Learning*, 50(4), 617-673. doi: 10.1111/0023-8333.00142
- Myles, F., Hooper, J., & Mitchell, R. (1998). Rote or rule? Exploring the role of formulaic language in classroom foreign language learning. *Language Learning*, 48(3), 323-364. doi: 10.1111/0023-8333.00045
- Myles, F., Mitchell, R., & Hooper, J. (1999). Interrogative chunks in French L2. . *Studies in Second Language Acquisition*, 21(01), 49-80. doi: doi:null
- Nabei, T., & Swain, M. (2002). Learner Awareness of Recasts in Classroom Interaction: A Case Study of an Adult EFL Student's Second Language Learning. *Language Awareness*, 11(1), 43-63.
- Nagata, H., Aline, D., & Ellis, R. (1999). Modified input, language aptitude and the acquisition of word meanings. In R. Ellis (Ed.), *Learning a second language through interaction* (pp. 133-149). Amsterdam ; Philadelphia: J. Benjamins.
- Nassaji, H. (2009). Effects of recasts and elicitations in dyadic interaction and the role of feedback explicitness. *Language Learning*, 59(2), 411-452. doi: 10.1111/j.1467-9922.2009.00511.x
- Nelson, K. E. (1987). Some observations from the perspective of the rare event cognitive comparison theory of language acquisition. In K. E. Nelson & A. Van Kleeck (Eds.), *Children's language* (Vol. 6, pp. 289-231). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

- Nicholas, H., Lightbown, P. M., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning*, 51(4), 719-758. doi: 10.1111/0023-8333.00172
- Nobuyoshi, J., & Ellis, R. (1993). Focused communication tasks and second language acquisition. *ELT Journal*, 47(3), 203-210. doi: 10.1093/elt/47.3.203
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50(3), 417-528. doi: 10.1111/0023-8333.00136
- Oliver, R. (1995). Negative feedback in child NS-NNS conversation. *Studies in Second Language Acquisition*, 17(4), 459-481.
- Oliver, R. (1998). Negotiation of meaning in child interactions. *The Modern Language Journal*, 82(3), 372-386.
- Oliver, R. (2000). Age differences in negotiation and feedback in classroom and pairwork. *Language Learning*, 50(1), 119-151. doi: 10.1111/0023-8333.00113
- Oliver, R., & Mackey, A. (2003). Interactional context and feedback in child ESL classrooms. *The Modern Language Journal*, 87(4), 519-533.
- Onwuegbuzie, A. J., Bailey, P., & Daley, C. E. (2000). The validation of three scales measuring anxiety at different stages of the foreign language learning process: The Input anxiety scale, the Processing anxiety scale, and the Output anxiety scale. *Language Learning*, 50(1), 87-117. doi: 10.1111/0023-8333.00112
- Panova, I., & Lyster, R. (2002). Patterns of corrective feedback and uptake in an adult ESL classroom. *TESOL Quarterly*, 36(4), 573-595.
- Philp, J. (2003). Constraints on "Noticing the gap": Nonnative speakers' noticing of recasts in NS-NNS interaction. *Studies in Second Language Acquisition*, 25(1), 99-126.
- Pica, T. (1983). Methods of morpheme quantification: Their effect on the interpretation of second language data. *Studies in Second Language Acquisition*, 6(01), 69-78. doi: doi:10.1017/S0272263100000309
- Pica, T. (1987). Second language acquisition, social Interaction, and the classroom. *Applied Linguistics*, 8(1), 3-21. doi: 10.1093/applin/8.1.3
- Pica, T. (1991). Classroom interaction, negotiation, and comprehension: Redefining relationships. [doi: 10.1016/0346-251X(91)90024-J]. *System*, 19(4), 437-452.
- Pica, T. (1994). Research on negotiation: What does it reveal about second language learning conditions, processes, and outcomes? *Language Learning*, 44(3), 493-527. doi: 10.1111/j.1467-1770.1994.tb01115.x
- Pica, T., Holliday, L., Lewis, N., & Morgenthaler, L. (1989). Comprehensible output as an outcome of linguistic demands on the learner. *Studies in Second Language Acquisition*, 11(01), 63-90. doi: doi:10.1017/S027226310000783X

- Pinker, S. (1989). *Learnability and cognition : the acquisition of argument structure*. Cambridge, Mass.: MIT Press.
- Poehner, M. E. (2008). *Dynamic assessment : a Vygotskian approach to understanding and promoting L2 development*. [Berlin?]: Springer.
- Poehner, M. E., & Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research*, 9(3), 233-265. doi: 10.1191/1362168805lr166oa
- Ranta, L. (2002). The role of learners' language analytic ability in the communicative classroom. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 159-180). Amsterdam ; Philadelphia, PA: J. Benjamins Pub.
- Revesz, A., & Han, Z. (2006). Task content familiarity, task type and efficacy of recasts. *Language Awareness*, 15(3), 160-178.
- Robinson, P. (1995). Attention, memory, and the "Noticing" hypothesis. *Language Learning*, 45(2), 283-331. doi: 10.1111/j.1467-1770.1995.tb00441.x
- Robinson, P. (1996). *Consciousness, rules, and instructed second language acquisition*. New York: P. Lang.
- Robinson, P. (1997a). Generalizability and automaticity of second language learning under implicit, incidental, enhanced, and instructed conditions. *Studies in Second Language Acquisition*, 19(02), 223-247. doi: doi:null
- Robinson, P. (1997b). Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning*, 47(1), 45-99. doi: 10.1111/0023-8333.21997002
- Robinson, P. (2002a). Effects of individual differences in intelligence, aptitude and working memory on adult incidental SLA: A replication and extension of Reber, Walkenfield and Hernstadt (1991). In P. Robinson (Ed.), *Individual differences and instructed language learning*. Amsterdam ; Philadelphia, PA,: J. Benjamins Pub. (pp. 211-266). Amsterdam ; Philadelphia, PA: J. Benjamins Pub
- Robinson, P. (2002b). *Individual differences and instructed language learning*. Amsterdam ; Philadelphia, PA,: J. Benjamins Pub.
- Robinson, P. (2002c). Learning conditions, aptitude complexes and SLA: A framework for research and pedagogy. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 113-133). Amsterdam ; Philadelphia, PA: J. Benjamins Pub.
- Robinson, P. (2003). Attention and memory during SLA. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 631-768). Malden, MA: Blackwell Pub.
- Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics*, 25, 46-73. doi: doi:10.1017/S0267190505000036

- Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar: A meta-analysis of the research. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133-164). Philadelphia, PA: John Benjamins.
- Sagarra, N. (2007). From CALL to face-to-face interaction: the effect of computer-delivered recasts and working memory on L2 development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition : a collection of empirical studies*. Oxford: Oxford University Press.
- Salaberry, M. R., & Shirai, Y. (2002). *The L2 acquisition of tense-aspect morphology*. Amsterdam ; Philadelphia, PA: J. Benjamins Pub.
- Sasaki, M. (1996). *Second language proficiency, foreign language aptitude, and intelligence: quantitative and qualitative analyses*. New York: Peter Lang.
- Sawyer, M., & Ranta, L. (2001). Aptitude, individual differences, and instructional design. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 319-353). Cambridge: Cambridge University Press.
- Saxton, M. (1997). The Contrast theory of negative input. *Journal of Child Language*, 24(01), 139-161.
- Saxton, M. (2010). *Child language : acquisition and development*. Los Angeles ; London: SAGE.
- Schachter, J. (1991). Corrective feedback in historical perspective. *Second Language Research*, 7(2), 89-102. doi: 10.1177/026765839100700202
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129-158. doi: 10.1093/applin/11.2.129
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.
- Schmitt, N., Dörnyei, Z., Adolphs, S., & Durow, V. (2004). Knowledge and acquisition of formulaic sequences: A longitudinal study. In N. Schmitt (Ed.), *Formulaic sequences : acquisition, processing, and use* (pp. 55-86). Amsterdam ; Philadelphia: John Benjamins Pub.
- Schwartz, B. D. (1993). On explicit and negative data effecting and affecting competence and linguistic behavior. *Studies in Second Language Acquisition*, 15(02), 147-163. doi:10.1017/S0272263100011931
- Seedhouse, P. (2004). *The interactional architecture of the language classroom : a conversation analysis perspective*. Malden, MA: Blackwell Pub.
- Sheen, Y. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research*, 8(3), 263-300.

- Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research*, 10(4), 361-392. doi: 10.1191/1362168806lr203oa
- Sheen, Y. (2007a). The Effect of Focused Written Corrective Feedback and Language Aptitude on ESL Learners' Acquisition of Articles. *TESOL Quarterly*, 41(2), 255-283.
- Sheen, Y. (2007b). The effects of corrective feedback, language aptitude, and learner attitude on the acquisition of English articles. In A. Mackey (Ed.), *Conversational interaction in second language acquisition : a collection of empirical studies*. Oxford: Oxford University Press.
- Sheen, Y. (2008). Recasts, language anxiety, modified output, and L2 learning. *Language Learning*, 58(4), 835-874.
- Sheen, Y. (2011). *Corrective feedback, individual differences and second language learning*. Dordrecht ; New York: Springer.
- Sheen, Y., & Ellis, R. (2011). Corrective feedback in language teaching. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (Vol. 2, pp. 593-610). New York: Routledge.
- Skehan, P. (1989). *Individual differences in second-language learning*. London ; New York: E. Arnold ; Distributed in the USA by Routledge, Chapman and Hall.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford ; New York: Oxford University Press.
- Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Robinson, P. (2002). Individual differences and instructed language learning* (pp. 69-93). Amsterdam ; Philadelphia, PA: J. Benjamins Pub.
- Sparks, R., & Ganschow, L. (1993). Searching for the cognitive locus of foreign language learning difficulties: Linking first and second language learning. *The Modern Language Journal*, 77(3), 289-302.
- Speidel, G. E., & Nelson, K. E. (1989). *The Many faces of imitation in language learning*. New York: Springer-Verlag.
- Sternberg, R. J. (2002). The theory of successful intelligence and its implications for language aptitude testing In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 13-43). Amsterdam ; Philadelphia, PA: J. Benjamins Pub.
- Swain, M. (1985). Communicative competence: some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, Mass.: Newbury House Publishers.

- Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom : second language acquisition* (pp. 64-81). New York: Cambridge University Press.
- Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal*, 82(3), 320-337.
- Swain, M., & Lapkin, S. (2000). Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan & M. Swain (Eds.), *Bygate, M., Skehan, P., & Swain, M. (2000). Researching pedagogic tasks : second language learning, teaching and testing. Harlow ; New York: Longman.* (pp. 99-118). Harlow; New York: Longman.
- Takashima, H., & Ellis, R. (1999). Output enhancement and acquisition of the past tense. In R. Ellis (Ed.), *Learning a second language through interaction* (pp. 173-188). Amsterdam Philadelphia: J. Benjamins.
- Tocalli-Beller, A., & Swain, M. (2005). Reformulation: the cognitive conflict and L2 learning it generates. *International journal of Applied Linguistics*, 15(1), 5-28.
- Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, 16(02), 183-203. doi:doi:10.1017/S0272263100012870
- Trofimovich, P., Ammar, A., & Gatbonton, E. (2007). How effective are recasts? The role of attention, memory, and analytical ability. In A. Mackey (Ed.), *Conversational interaction in second language acquisition : a collection of empirical studies*. Oxford: Oxford University Press.
- Trofimovich, P., & McDonough, K. (2011). *Applying priming methods to L2 learning, teaching and research : insights from psycholinguistics*. Amsterdam ; Philadelphia: John Benjamins Pub. Co.
- Ur, P. (1996). *A course in language teaching : practice and theory*. Cambridge [England] ; New York: Cambridge University Press.
- Vendler, Z. (1967). *Linguistics in philosophy*. Ithaca, N.Y.,: Cornell University Press.
- Vygotsky, L., Hanfmann, E., & Vakar, G. (1962). Thought and language Retrieved from <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&NEWS=N&PAGE=toc&SEARCH=2006-10268.dd&LINKTYPE=asBody&D=psbk>
- Wesche, M., Edwards, H., & Wells, W. (1982). Foreign language aptitude and intelligence. *Applied Psycholinguistics*, 3(02), 127-140.
- White, L. (1989). *Universal Grammar and second language acquisition*. Amsterdam ; Philadelphia: J. Benjamins Pub. Co.

- White, L. (1991). Adverb placement in second language acquisition: some effects of positive and negative evidence in the classroom. *Second Language Research*, 7(2), 133-161. doi: 10.1177/026765839100700205
- White, L. (2003). *Second language acquisition and universal grammar / Lydia White*. New York: Cambridge University Press.
- Widdowson, H. G., Cook, G., & Seidlhofer, B. (1995). *Principle & practice in applied linguistics : studies in honour of H. G. Widdowson*. Oxford: Oxford University Press.
- Williams, J. N. (1999). Memory, attention, and inductive learning. *Studies in Second Language Acquisition*, 21(01), 1-48. doi: doi:null
- Williams, J. N., & Lovatt, P. (2003). Phonological memory and rule learning. *Language Learning*, 53(1), 67-121. doi: 10.1111/1467-9922.00211
- Woodrow, L. (2006). Anxiety and speaking English as a second language. *RELC Journal*.
- Yang, Y., & Lyster, R. (2010). Effects of form-focused practice and feedback on Chinese EFL learners' acquisition of regular and irregular past tense forms. *Studies in Second Language Acquisition*, 32(2), 235-263.
- Young, D. J. (1990). An investigation of students' perspectives on anxiety and speaking. *Foreign Language Annals*(23), 539-553.

Appendix A Ethics Forms

Office of the Vice Chancellor
Ethics and Biological Safety Administration



The University of Auckland
Private Bag 92019
Auckland, New Zealand

UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE

Level 1, 79 Symonds Street
Telephone: 64 9 373 7500
Fax: 64 9 373 7600
Email: 64 9 373 7410

26 September, 2008

MEMORANDUM TO:

Nadia Mitka-Profozic
Applied Language Studies and Linguistics

Re: Application for Ethics Approval

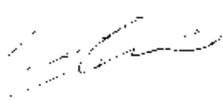
The Committee met on 17 September, 2008 and considered the application for ethics approval for your research titled "Corrective feedback, Individual differences and acquisition of L2 French." (Our Ref. 2008 / 360).

Ethics approval was given for a period of three years.

If the project changes significantly you are required to resubmit your application to the Committee for further consideration.

In order that an up-to-date record can be maintained, it would be appreciated if you could notify the Committee once your project is completed.

Please contact the Chairperson if you have any specific queries relating to your application. The Chair and the members of the Committee would be most happy to discuss general matters relating to ethics provisions if you wish to do so.



Lana Loi
Executive Secretary
University of Auckland Human Participants Ethics Committee

c.c. Head of Department / School, Applied Language Studies and Linguistics
Nadia Mitka-Profozic
137 Aukland Road
New Lynn

1. Please quote the reference number - (2008/360) in all future correspondence.

**PARTICIPANT CONSENT FORM
BOARD OF TRUSTEES**

THIS CONSENT FORM WILL BE HELD FOR A PERIOD OF SIX YEARS

Title: Individual Differences and Language Learning**Researcher/PhD Student: Nadia Mifka-Profozic**

We have been given and have understood the explanation of this research project. We have had an opportunity to ask questions and have them answered.

- We understand that the students' participation in this research is entirely voluntary.
- We understand that the students may withdraw from the research at any time up to December 1, 2008 without giving a reason.
- We understand that the students' grades will not be affected by either the refusal or agreement to participate in this research.
- We understand that all data gathered will be stored on the researcher's computer in case it is needed for publication or further research.
- We understand that the students will be audio-recorded and that after the data have been coded the recordings will be wiped.
- We understand that the information students provide in this research will be reported and published in a way that does not identify the students or the school as its source.
- We understand that we will have the opportunity to be given a report about the results of this research.
- We give permission to the researcher for the access to the school.

Signed:**Name:**
(please print clearly)**Date:**

**"APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS
ETHICS COMMITTEE
ON17 Sept 2008..... for(3)....years
Reference Number2008/360..."**



Fisher Building
18 Waterloo Quadrant, Levels 8 & 9
Auckland, New Zealand
Telephone 64 9 3737599 ext. 82001
Facsimile 64 9 308 2360

Private Bag 92019
Auckland, New Zealand

Department of Applied Language Studies and Linguistics

PARTICIPANT CONSENT FORM PRINCIPAL

THIS CONSENT FORM WILL BE HELD FOR A PERIOD OF SIX YEARS

Title: Individual Differences and Language Learning

Researcher/PhD Student: Nadia Mifka-Profozic

We have been given and have understood the explanation of this research project. We have had an opportunity to ask questions and have them answered.

- I understand that the students' participation in this research is entirely voluntary.
- I understand that the students may withdraw from the research at any time up to December 1, 2008 without giving a reason.
- I understand that all data gathered will be stored on the researcher's computer in case it is needed for publication or further research.
- I understand that students will be audio-recorded and that after the data have been coded the recording will be wiped.
- I understand that the information students provide in this research will be reported and published in a way that does not identify the students or the school as its source.
- I understand that we will have the opportunity to be given a report about the results of this research.
- I confirm that the students' grades will not be affected by either the refusal or agreement to participate in this research.
- I give permission to the researcher for the access to the school.

Signed:

Name:

(please print clearly)

Date:

**"APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS
ETHICS COMMITTEE**

**ON.....17 Sept 2008..... for...(3)... years
Reference Number2008 /360....."**



Department of Applied Language Studies and Linguistics

Fisher Building
18 Waterloo Quadrant, Levels 8 & 9
Auckland, New Zealand
Telephone 64 9 3737599 ext. 82001
Facsimile 64 9 308 2360

Private Bag 92019
Auckland, New Zealand

**PARTICIPANT CONSENT FORM
TEACHER OF FRENCH**

THIS CONSENT FORM WILL BE HELD FOR A PERIOD OF SIX YEARS

Title: Individual Differences and Language Learning

Researcher/PhD Student: Nadia Mifka-Profozic

I have been given and have understood the explanation of this research project. I have had an opportunity to ask questions and have them answered.

- I understand that students' participation in this research is entirely voluntary.
- I understand that students may withdraw themselves or any information traceable to them at any time up to July 15, 2009 without giving a reason.
- I understand that all data gathered will be stored on the researcher's computer in case it is needed for publication or further research.
- I understand that students will be audio-recorded and that after the data have been coded the recording will be wiped.
- I understand that the information students provide in this research will be reported and published in a way that does not identify the students or the school as its source.
- I understand that I will have the opportunity to speak with the researcher about the results of the research.
- I confirm that the students' grades will not be affected by either the refusal or agreement to participate in this research.

Signed:

Name:
(please print clearly)

Date:

**"APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS
ETHICS COMMITTEE
ON17 Sept 2008.... for.....(3)..... years
Reference Number...2008/360"**



THE UNIVERSITY OF AUCKLAND
NEW ZEALAND

Department of Applied Language Studies and Linguistics

Fisher Building
18 Waterloo Quadrant, Levels 8 & 9
Auckland, New Zealand
Telephone 64 9 3737599 ext. 82001
Facsimile 64 9 308 2360

Private Bag 92019
Auckland, New Zealand

PARTICIPANT CONSENT FORM PARTICIPANTS

THIS CONSENT FORM WILL BE HELD FOR A PERIOD OF SIX YEARS

Title: Individual Differences and Language Learning

Researcher/ PhD Student: Nadia Mifka-Profozic

I have been given and have understood an explanation of this research project. I have had an opportunity to ask questions and have them answered.

- I understand that my participation in this research is entirely voluntary.
- I understand that I may withdraw myself or any information traceable to me at any time up to July 15, 2009 without giving a reason.
- I understand that all data gathered will be stored on the researcher's computer in case it is needed for publication or further research.
- I understand that I will be audio-recorded and that after the data have been coded the recordings will be wiped.
- I understand that the information I provide in this research will be reported and/or published in a way that does not identify me as its source.
- I understand that I will have the opportunity to speak with the researcher about the results of the research.
- I understand that my grades will not be affected by either the refusal or the agreement to participate in this research.
- I agree to take part in this research.

Signed:

Name:
(please print clearly)

Date:

**"APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS
ETHICS COMMITTEE
ON 17 Sept 2008..... for(3)....years
Reference Number2008 /360"**

Appendix B References for illustrations used in the study

The following are the references for text-books from which the picture materials have been used in the present study.

Aries, Angela & Debney Dominique. Illustrations by John Plumb. (1987). *Facon de Parler*, 2. Sevenoaks, Kent : Hodder and Stoughton, LTD

Hingston, E. Illustrations by G. T. Martel. (1969). *Audio-lingual French. Supplement materials for the Fifth Form*. Wellington, London, Auckland: Reed Education

Recent School Certificate questions with suggested solutions: French, compiled by Simon B. Curnow. Auckland: ESA Publications

Sprake, D. (1982). *Comuniqué. Practice for the CSE and 16+ examination*. Oxford: Oxford University Press.

Permission to reproduce the illustrations was obtained from:

Simon Curnow and ESA Publications (NZ) Ltd., and Oxford University Press, Oxford.

All effort was made to obtain permission to reproduce all illustrations used in the study, however, with no success.

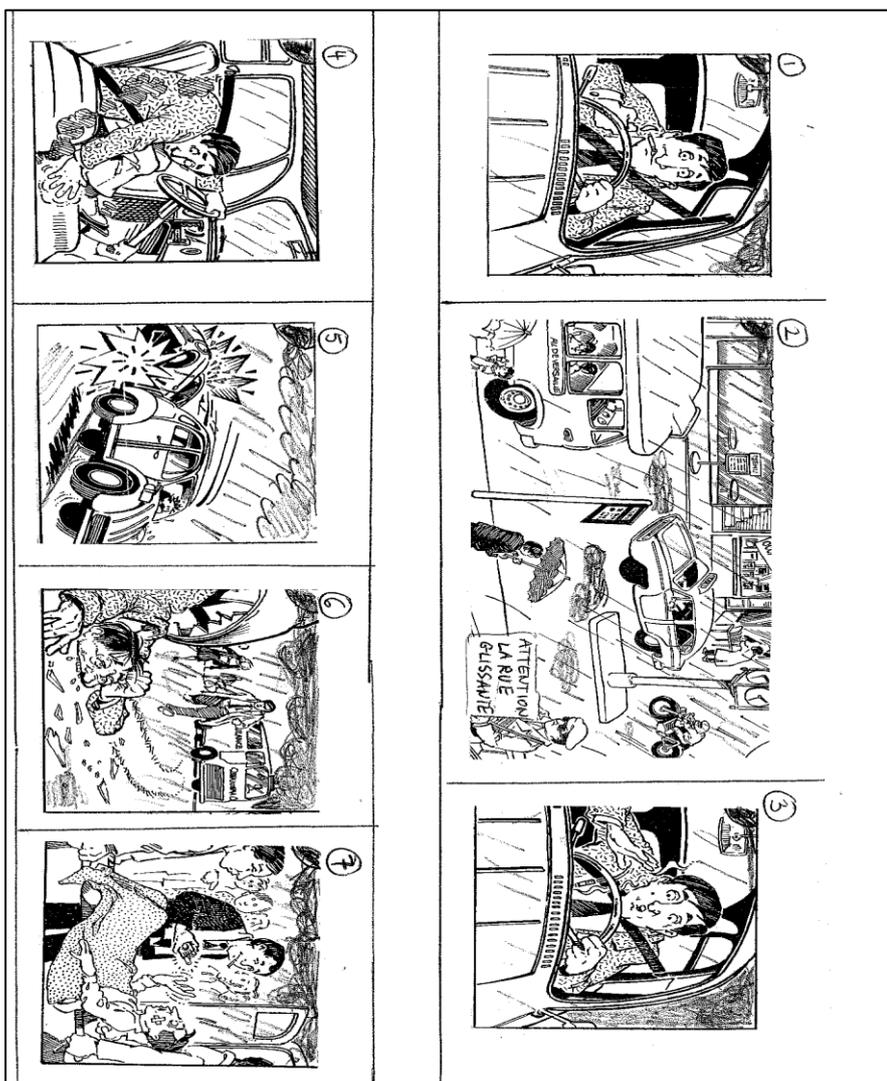
Appendix C Materials used for testing in oral production

Hier matin Professeur Maffard s'est levé trop tard. Il n'a pas entendu le réveille-matin...

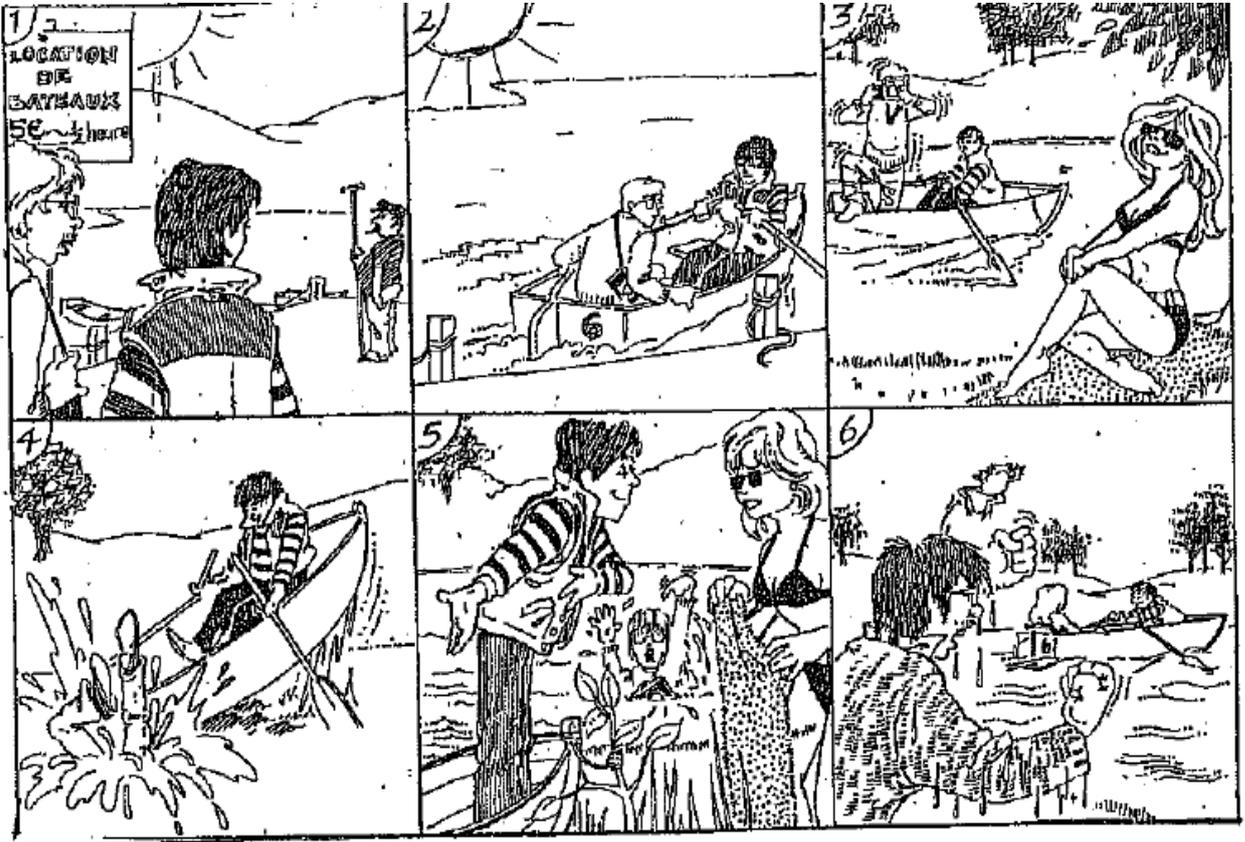


Reproduced with kind permission of ESA Publications (NZ) Ltd.

Appendix D Materials used for testing in written production



Cette histoire s'est passée il y a vingt ans. Mon oncle me raconte comment il a arrêté de fumer...



La semaine dernière mon cousin et son copain ont loué un canot au bord du lac. Voilà ce qu'il s'est passé

Reproduced with kind permission of Oxford University Press.

Appendix E Language analysis test

The list in the box below contains words/phrases from an imaginary language along with their English translation. Following this, there will be 14 short English sentences, each with four possible translations into the imaginary language. Based on the examples given in the box, we would like to ask you to try and work out which of the four options is the correct translation of each sentence. Thank you very much.

Imaginary language	English translations
kau	dog
meu	cat
kau meud bo	The dog is chasing the cat.
kau meud bi	The dog was chasing the cat.
so	watch
ciu	mouse
pa	we, us
xa	you
pasau meud bo	our dog is chasing the cat.
pa meud bo	We are chasing the cat.
paxbo	We are chasing you.
pa meud bor	We are chasing the cat.

1. The dog is watching the cat.

- a. kau meud so b.kau meud si
c. meu kaud so d. meu kaud si

2. The cat was watching the mouse.

- a. meud ciu so b. meu ciud so
c. meud ciu si d. meu ciud si

3. You are watching us.

- | | |
|----------|----------|
| a. paxbo | b. paxso |
| c. xapbo | d. xapso |

4. You are chasing the dog.

- | | |
|---------------|---------------|
| a. xa kaud bo | b. pa kaud bo |
| c. pa kaud bi | d. xa kaud bi |

5. We were watching you.

- | | |
|----------|----------|
| a. xapsi | b. paxso |
| c. paxsi | d. paxbi |

6. You are not watching the cat.

- | | |
|----------------|----------------|
| a. xa meud bor | b. xa meud sor |
| c. xa meud sir | d. xa meu sor |

7. You are not chasing us.

- | | |
|------------|-----------|
| a. paxbor | b. xapbo |
| c. xapabor | d. xapbor |

8. We were not watching the dog.

- | | |
|----------------|----------------|
| a. pa kaud sir | b. pa kau sir |
| c. pa kaud sor | d. pa kaud bir |

9. We were not chasing you.

- | | |
|-----------|-----------|
| a. xapbir | b. paxbir |
| c. paxbor | d. xapbor |

10. Your cat is chasing the mouse.

- a. xacu meud bo
- b. xaseu ciud bo
- c. meuxa ciud bo
- d. ciuxa meud bo

11. You are not watching our dog.

- a. xa paseud bor
- b. xa pasaud sor
- c. xa pasaud so
- d. xa pasaud bor

12. Our mouse was not chasing the dog.

- a. oasiu kaud bi
- b. xasiu kaud sir
- c. xasiu kaud bi
- d. pasiu kaud bir

13. Your mouse is chasing us.

- a. xa ciu pabo
- b. xasiu pbo
- c. xaciu pa bo
- d. xasiu pabo

14. Our cat was not chasing your dog.

- a. pseu xasaud bir
- b. pseu xsaud bir
- c. paseu xasaud bir
- d. paseu xsaud bir

Appendix F Speaking span test

The following are the instructions and the list of words used in the Speaking span test:

Instructions

Read each word out loud.

Remember each word in order.

At the prompt on the screen make a sentence for each word.

For example: you see: apple, train, bullet

You say: My apple is sour.

The train is late.

Which bullet was it?

Trial set

Fingers, noticed

Machine, results, compass

Set two

Kitchen, farmers

Signals, thirsty

Perfume, giraffe

Healthy, rewards

Biscuit, shampoo

Set three

Pencils, observe, journey

Nervous, quickly, younger

Trumpet, windows, believe

Earning, dentist, tallest

Parking, succeed, whisper

Set four

Butcher, wrinkle, ceiling, glasses

Certain warning, mittens, husband

Diapers, special, instant, plastic

Explain, stylish, garbage, request

Trouble, bending, advance, roasted

Set five

Teacher, stomach, foreign, cousins, quarter

Jealous, monthly, arrange, sweater, treated

Growing, surfing, ashamed, lettuce, cushion

Damaged, respect, private, clearly, witness

Appendix G Anxiety questionnaire

How do you feel about learning French?

There are some statements concerning your feelings about learning French. For each statement please tick the box that best describes how you feel. There is no good or bad answer – the researcher is only interested your personal feelings about learning a foreign language. So, please give your honest answers. Thank you very much for your cooperation.

	Strongly agree	Agree	Some what agree	Some what disagree	Disagree	Strongly disagree
I get flustered unless French is spoken very slowly and deliberately						
I get upset when French is spoken too quickly						
Learning new French vocabulary does not worry me, I can acquire it in no time						
I am anxious with French because, no matter how hard I try, I have trouble understanding it						
I do not worry when I hear new or unfamiliar words, I am confident that I can understand them						
I feel confident that I can easily use the French vocabulary that I know in a conversation						
I may know the proper French expressions but when I am nervous it just won't come out						
When I become anxious during a French assessment, I cannot remember anything I studied						

Participant's information:

Date of birth:

First language:

When did you start learning French:

Please answer the following three questions briefly:

1. What do you like most about learning French?

2. What is the most difficult about learning French?

3. What do you think was/were the aim/s of the activities we did today and last week?

THANK YOU VERY MUCH FOR YOUR TIME AND COOPERATION!

Appendix H Pilot Study : Results

12.1. Research question 1

12.1.1. Oral passé composé

Table 56 Descriptive statistics for oral PC tests

Group	Pre-test		Post-test 1		Post-test 2	
	Mean	SD	Mean	SD	Mean	SD
Recast (N=9)	50.3	29.7	60.2	28.2	62.8	22.7
Control (N=5)	58.2	25.6	62.0	20.1	72.4	17.8

The results of Mixed design Repeated measures ANOVA oral passé composé test revealed no statistically significant differences between the two groups, $F(1, 12) = 1.37, p = .64, > .05$. There was no significant time x group interaction but the tests of within-subjects effects revealed significant effects over time for both groups, $F(2, 24) = 6.16, p = .007, p < .05$. Pairwise comparison showed that statistically significant effect occurred between the pretest and the delayed post-test on oral passé composé tense.

12.1.2. Written passé composé

Table 57 Descriptive statistics for written PC tests

Group	Pretest		Posttest 1		Posttest 2	
	Mean	SD	Mean	SD	Mean	SD
Recast (N=10)	64.2	25.0	60.6	30.4	76.2	15.7
Control (N=5)	73.0	31.7	73.0	28.3	65.8	14.3

Mixed design Repeated measure ANOVA for written passé composé tests, again, revealed no statistically significant effects for group over time, $F(1, 13) = .114, p = .114, p > .05$. There was no significant time x group interaction and there were no significant effects for time found within the subjects, $F(2, 26) = .163, p > .05$.

12.1.3. Oral imparfait

Table 58 Descriptive statistics for oral IMP tests

Group	Pretest		Posttest 1		Posttest 2	
	Mean	SD	Mean	SD	Mean	SD
Recast(N=9)	35.5	21.6	51.7	23.3	46.7	22.6
Control (N=5)	58.2	7.6	48.6	21.9	55.4	17.8

Independent samples t-test on pre-test revealed that for oral imparfait there was a statistically significant difference on pre-test between the Recast and the Control group ($p = .035$, $p < .05$). On three other pre-tests there was no statistically significant difference between the treatment and the control group.

In order to account for the identified statistically significant difference between the Recast and the Control group on oral imparfait pre-test an ANCOVA, using the pre-test scores as covariate, was performed to compare the mean scores between the two groups at three different levels (pre-test, immediate post-test and delayed post-test) and hence to answer Research question 1. When the difference on the pre-test has been accounted for, the tests of between-subjects effects showed that there was no statistically significant difference between the two groups, on both the immediate post-test, $F(1,10) = .796$, $p = .393$, $p > .05$, and on the delayed post-test, $F(1, 10) = .219$, $p = .65$, $p > .05$. Also, no effect of time was found on tests within-subjects effects, $F(2, 24) = .181$, $p > .05$.

12.1.4. Written imparfait

Table 59 Descriptive statistics for written IMP tests

Group	Pre-test		Post-test 1		Post-test 2	
	Mean	SD	Mean	SD	Mean	SD
Recast (N=10)	61.9	26.1	59.4	22.4	70.6	25.8
Control (N=5)	65.0	30.4	68.6	21.7	54.2	25.0

Comparison of mean scores using the Mixed design Repeated measures ANOVA revealed no statistically significant differences between the two groups, $F(1, 13) = .019$, $p = .89$, $p > .05$.

Also, no evidence of significant effects over time was found on tests within-subjects effects, $F(2,26) = .025, p > .05$.

12.2. Research question 2

Uptake episodes were scored and calculated in percentage so that the number of episodes of uptake with repair for each participant was compared with the total number of recast episodes. The analysis was performed for each tense (imparfait and passé composé) separately.

Table 60 shows the total number of non-target like forms of the passé composé and imparfait, and instances of recasts followed by uptake with repair, for each tense separately.

Table 60 Number of recasts and uptake with repair

	Total non-target like forms	Recast episodes	Uptake with repair episodes
IMP	38	37	29
PC	41	41	28
Total	79	78	57

Table 61 presents the results of correlation analysis.

No significant correlation was found between uptake with repair and the gain scores on any of the post-tests for imparfait.

A statistically significant correlation was found between the gain scores and uptake with repair on oral passé composé immediate post-test ($r = .696, p < .05$), oral passé composé delayed post-test ($r = .800, p < .5$) and written passé composé immediate post-test ($r = .661, p < .05$).

Table 61 Correlation between uptake with repair and gain scores

	Short-term gains	Long-term gains
Oral PC	R=.696* p = .037	R= .800* p= .031
Written PC	R= .661* p= .053	R= .565 p=ns
Oral IMP	ns	Ns
Written IMP	ns	Ns

Note: R= Pearson correlation coefficient (uptake with repair – gain scores)

12.3. Research question 3

Table 62 Descriptive statistics for anxiety and LAA scores

Group	anxiety		LAA	
	Mean	SD	Mean	SD
Recasts	2.77	0.77	73.2	15.3
Control	2.89	0.03	85.2	6.45

In order to answer Research question 3 Pearson correlation coefficient was calculated, investigating the relationship between uptake episodes and individual difference factors: anxiety and language analytic ability (LAA).

No correlation was found between uptake with repair and anxiety. Also, there was no correlation evident between uptake with repair and language analytic ability.

12.4. Research Question 4

12.4.1. Results for anxiety

Pearson correlation was calculated between the anxiety scores and short-term gain scores (Gain 1) and long-term gain scores (Gain 2) for oral and written passé composé and imparfait.

Correlation analyses revealed a statistically significant negative correlation between anxiety and gain scores on delayed post-test for oral imparfait test in the Recast group:

$r = -.800^*$, $p < .031$.

No other significant correlations were found.

12.4.1.1. Results for language analytic ability

Pearson correlation analyses using analytic ability scores and short-term and long-term gain scores did not reveal any significant correlation between any of the gain scores and language analytic ability.

Appendix I Transcripts from the treatment tasks

The following are the excerpts from the treatment tasks. In one class recasts were provided as corrective feedback whereas the other class received clarification requests

Recast group

Task 1

001 T : uhm qui voudrait continuer ? hein ? qu'est qu' on peut voir sur votre =

002 = image ? ... dis- moi, dis-moi... bien, bien tu es ... ? =

003 T : = alors, ..., qu'est ce q'il c'est passé, qu'est-ce qu'il y a sur ton image ?

004 L2 : ummm les deux garçons un garçon acheté et

005 T : ils ont acheté ?

006 L2 : ils ont acheté une coca

007 T : bien bien

008 L2 : et et ils ont acheté seulement un coca

009 T : alors seulement un garçon avait de l'argent ? je ne sais pas ? =

010 = et qu'est ce que tu penses de cet image, qu'est ce qu' il s'est passé ?

011 A : errr un garçon devient devient jalous ?

012 T : oh, il est devenu jalous ?

013 L2 : ummm

014 T : il est devenu ?

015 L2 : jalous d'autre garçon et l'autre garçon dit

016 T : il a dit ?

017 L2 : il a dit je l'ai achetée attends

018 T : est-ce que tu penses qu'ils se sont disputés ?

019 L2 : uhm non

020 T : ummm ... non ?

021 T : bien alors ils ont acheté seulement une bouteille mais tous les deux avaient

022 soif ...et sur la troisième image ... moi j'ai la troisième image ... en fait ils ont

023 commencé à se disputer uhmm ils se sont disputés parce que l'autre garçon –

024 c'était Philippe ...

032 T : ils ont acheté une bouteille de coca ... um et plus tard qu'est ce qu'ils ont fait ?

033 L2 : ils sont ils sont

034 L1 : ils sont disputés ?

035 T : ils se sont disputés ?

036 L1 : ils se sont disputés et ils se sont battre

037 T : ils se sont battus ?

038 L1 : battus et et Paul a poussé son frère sur la rue et la voiture a frappé le garçon

039 T : oh la voiture s'approchait au même moment la voiture s'approchait

040 L1: ummmm... =

041 L1 : = a frappé a frappé le garçon

042 L1 : et la voiture s'arrête

043 T : la voiture s'est arrêtée ?

044 L1 : s'est arrêtée et le garçon était blessé et et e-et

045 T : alors, la voiture de police um les policiers qu'est ce qu'ils ont fait ? ils se sont =

046 = arrêtés et puis ?

- 047 L1 : ils ont appelé le la ambulance et et
- 048 T : pourquoi ?
- 049 L1 : parce que le garçon était blessé
- 050 T :... oui
- 051 L1 : umm il a perdu connaissance et le ambulance emmené à l'hôpital
- 052 T : uh l'a ammené à l'hôpital ?
- 053 L1 : uh oui l'a ammené à l'hôpital

Recast group

Task 2

- 021 T : Quelqu'un d'autre qui voudrait prendre part ? =
- 022 = sur la troisième image qu'est-ce qu'il y avait ? umm ?
- 023 L11: ils se sont ils sont arrivés au bord de la lac et umm ils ils mangent ils mangent le =
- 024 = déjeuner
- 025 T : tu peux les voir manger ? où ils ont déjà fini à manger ?
- 026 L11 : ils ont fini
- 027 T : alors ils ont mangé
- 028 L11: ils ont mangé le déjeuner et umm umm la sœur la sœur a la sœur a la sœur a se =
- 029 = faire les bronzer
- 030 T : elle se faisait bronzer ?
- 031 L11 : elle se faisait bronzer oui
- 032 T : et la mère la mère qu'est-ce qu'elle faisait ?

033 L11 : elle a se relaxé

034 T : elle se relaxait ?

035 L11 : se relaxait/é oui

036 T : et le père ?

037 L11 : le père il il a fait pêche

038 T : tu vois la photo a-t-il fini de pêcher ?

039 L11 : non je pense

040 T : alors il faisait de la pêche ?

041 L11 : oui

042 T : il faisait de la pêche

043 L11 : oui

Clarification Request group

Task 1

- 001 T : maintenant umm (.) comment vous racontez cette histoire hein ? =
- 002 = je vais vous aider
- 003 T : qui voudrait commencer de la première image ? (.) qui va commencer ?
- 004 L34 : je peux um il y a deux frères ... Philippe et Paul ummm il a joué le foot
- 005 T : hein ? qu'est-ce qu'ils faisaient ? les garçons tu peux les voir =
- 006 = qu'est-ce qu'ils faisaient ? (CR 1st with a hint indicating IMP 'faisait')
- 007 L34 : sorry I don't understand um (.) ils jouaient au foot ?
- 008 T: très bien. Ils jouaient au foot. (T recasts – by mistake)
- 009 L34 : et il faisait beau avec beaucoup de soleil... il faisait très chaud donc ils ont très soif
- 010 T : umm ? pardon ? (CR 1st)
- 011 L34 : ils étaient ...non il avaient ?
- 012 T : bien bien
- 013 L34 : très soif... et ils a acheté..
- 014 T : hein ? tu peux répéter ? (CR 1st)
- 015 L34 : ilssss ... sont ils sont
- 016 T : ummm comment ? (CR 2nd)
- 017 L34 : ils ont acheté
- 018 T : bien
- 019 L34 : un bouteille de boisson... et puis à la supermarché et puis ils est attend de =
- 020 = la bus, de bus
- 021 T : hein ? (CR 1st)
- 022 L34 : umm le bus

- 023 T : regarde ... moi j'attendais le bus (.) et ils ? hein ? (CR 2nd and a hint)
- 024 L34 : ils sont attend le bus ...
- 025 T : um
- 026 T : il y a quelqu'un d'autre qui voudrait continuer ? tu veux continuer ? nous =
- 027 T = nous sommes arrêtés quand ils sont sortis du supermarché (.) tu veux =
- 028 = continuer ?
- 029 T : qu'est-ce qu'ils ont fait quand ils sont sortis du supermarché ?
- 030 L20 : Paul s'est battu ?
- 031 T : hein? tu peux répéter? (CR 1st)
- 032 L20 : I don't know how to say ? ils ont?...
- 033 T : hein ? (CR 2nd)
- 034 L20 : parce que Phillipe a eu eu? la boisson
- 035 T : umm
- 036 L20 : mais umm il y avait une boisson seulement
- 037 T : oui bien
- 038 L20 : puis Paul a poussé Philippe et Philippe a il est tombé dans la rue
- 039 T : très bien
- 040 L20 : et les policiers est frappé il est (or : et) ensuite er er les policiers
- 041 T : pardon ? (CR 1st)
- 042 L20 : um
- 043 T : tu veux continuer (name) ?
- 044 L36: er er la voiture de police est s'est [rəversé] s'était [rəversé] um et er =
- 045 = Philippe est blessé
- 046 T : hein ? tu peux répéter comment tu as dit ? (CR 1st)
- 047 L36 : ils sont er il a blessé et une ambulance s'approchait

048 T : lentement um d'abord (.) les policiers qu'est-ce qu'il ont fait ? (*CR 2nd and a question referring back to first error*)

049 L36 : er er ils téléphoné ...

050 T : pardon ? tu peux répéter ça ? (*CR 1st*)

051 L36 : Paul a téléphoné à l'hôpital (*change of s ubject – but correct*)

052 T : bien

053 L36 : et une ambulance s'approchait et (or : est) arrivé et um et um

054 : T : bien

Clarification Request group

Task 2

064 T : um et tout d'un coup quelque chose s'est passée .. qu'est-ce qu'il s'est passé tout =

065 = d'un coup ?

066 L28 : le voiture tombé

067 T : comment tu as dit ? (*CR 1st*)

068 L28 : la voiture est tombé et umm...

069 T : uh oui

070 L28: le fille c'est très surprise (engl. pronunciation)

071 T : tu peux répéter ? hein ? (*CR 1st*)

072 L28: et le fille c'est très surprise et er

073 T : hein ? (*CR 2nd*)

074 L28: le voiture c'est nouvelle voiture et c'est très expensive

075 T : 'chère' c'est 'expensive' =

- 076 = mais comment la voiture est tombée dans le lac ? =
- 077 = quand la sœur s'est retourné =
- 078 = qu'est-ce qu'elle a vu ?
- 079 L34 : le voiture est glissé et
- 080 T : déjà ? qu'est-ce qu'elle a vu? (*CR 1st*)
- 081 L34 : er le voiture est glissé um et tombé dans le lac et le père il est très surpris
- 082 T : comment ? tu peux répéter ? (*CR 2nd*)
- 083 L34 : err le père est pêché
- 084 T : il pêchait bien (*T recasts*) mais quand il a vu la voiture dans le lac comment était-il ?
- 085 L34 : il a vu le voiture err tomber et il est très surpris
- 086 T : ummm comment ? (*CR 1nd*)
- 087 L34 : il est très surpris

Appendix J Transcripts from tests

The following are the excerpts from tests in oral production:

Learner 19 Pre-test

le professeur Maffard a dormi plus tard
parce que le (clock) réveille-matin ne sonné/er pas
er er parce que c'est très froid um
la femme de professeur Maffard errr um reveiller/é um le professeur Maffard se reveiller
um parce que le femme um son femme (how do you say come in) entrer um
à huit heures et demie um
la professeur Maffard um prendre le petit déjeuner um
et il sortier/ié
il pleut et il vent um
il ne manquer le bus err
il arriver le le college plus tard
et il entrer la classe um à neuf heures et demie err
err err huh (oh how do you say the students) les étudiants er er surprised um yeah

Learner 21 Post-test 1

c'est une histoire de Gérard et oui c' était umm c'était un matin très beau
err et Gérard a réveill-é et il il a pensé de sa err petit petite copine
et elle est elle est pensé aller en vacances avec sa umm sa petite copine
umm parce que il il c' était très soleil et
elle il il a téléphoné son sa petite copine elle s'appelle Magali

et elle est il est il a dit : tu... vous avez faire une rendez-vous
et elle est... était très contente et il s'habille son il s'habille son polo
et il a met son pantaloon et oui parce que il c'était très chaud et
il... il est il était en retard et il dépêché s'est dépêché dép-déêché
et il il il est il est allé descendre vite et son taxi a attendre attende
mais Gérard a tomb... est tombé et il il a cassé la jambe
et il était en trop tard et oui errr
son son ... sa mere appelé l'ambulance a appelé l'ambulance
et il il est vennn... venu en mons...en monn... (?)
et sa petit copine elle est très triste parce que elle elle a
elle a pensé ils sont allés en vacances et oui

Learner 47 Post-test 1

une matin il s'est il était dormir et c'était très tôt
et puis il umm c'était très froid et puis um
sa femme il s'est réveillé il s'est levé très tard
puis il s'est dépêché et puis il a prend une café et mangé sa petit déjeuner
er et puis errr il a couri dehors et il a il s'est de pris sa bus
mais il est manqué le bus
errr il faisait ne beau pas c'était plus très froid
alors parce que il a manqué le bus il a pas marché à l'école
et il a arrivé au lycée en retard et quand il a arrivé
les étudiants étaient attendé la classe et errr il s'excusé
et c'est ca

Learner 20 Post-test 1

ummm cette histoire d' un homme qui s'appelle professeur Maffard
um il errr c'était un jour de l'hiver et umm il est dormi umm

il faisait très froid et pendant la nuit et son reveille-matin n'est pas sonné
ummm donc la femme de professeur Maffard a ouvri la porte de chamber de professeur
Maffard
ummm et professeur Maffard s'est réveillé et s'est levé umm
et puis er il a pris le café umm et il a mangé et s'est dépêché umm
donc il est en retard il était en retard umm
il faisait pluie et ummm il n'avait pas de parapluie
donc il est sorti et il est sorti et il etait mouillé umm ummm
aussi il a manqué le bus donc il était très fâché uhhh ummm
quand il arrivé au lycée il était en retard umm
et son ses étudiants err sont err attend ... attendre err yeah ses étudiants attendant
ummm il a entré la classe et les étudiants étaient surprise umm
mais professeur Maffard s'est excusé

Learner 11 Post-test 2

Ummm le professeur Maffard a dormi a dormi umm plus à bonne heure ...
Le professeur Maffard umm n'a pas s'est umm s'est n'a pas se réveillé umm
parce que il a oublié err se faire
il a oublié se faire réveiller le matin
donc umm il y il y avait une grasse matinée
umm sa uh sa ff... sa femme ouvrié a ouvrié la porte
et il s'est il s'est lavé umm
il il pleuvait et il a prendre prené le café
et umm et il a mangé le petit oh il a prené la petit
il a prené la petit déjeuner
umm il a cour- cou- couri il a couri dans le dans la rue umm
mais il a manqué manqué le bus
umm il a arrivé au lycée et il était en retard
umm les étudiants umm les 'etudiants ont attendu pour il ummm
quand il a entré entré le la classe
les étudiants oh ils étaient surprise s'excuser/é

Learner 8 Post-test 2

Umm un matin Giselle se s'est réveillée
Et il faisait très beau le soleil brillait
Et puis elle a décidé de se promener
elle passait/é elle a passé devant la maison umm devant une maison
où il y avait un peintre et le il a peindre les volets
et puis il err l'échelle regardé à Giselle
um il était il n'est pas il n'a pas concentré et puis
il est tombé et la peinture renversé et la peinture était noire
elle est tombée err il est tombé il s'est cassé la jambe
puis Giselle a téléphoné au l'ambulance
l'ambulance est venue err
et le peintre s'est blessé et ummm l'ambulance est arrivée avec un infirmier
puis il lui ammenait/é à l'hôpital et
le chien aboyait/é err le chien a aboyé au le au l'infirmier umm
et Giselle était très stressée parce que c' était leur chien
er puis ummm le peintre était à l'hôpital

une jour la semaine derniere umm une fille umm
se réveillé ummm s'a réveillé
ummm dans le matin à six heures
quand le soleil umm était bri umm
et elle a décidé umm se promener umm dans la rue err
avec son chien err
quand elle err s'a passé devant la maison umm
il y avait ummerr un peintre
umm et il a tombé dans la rue et il s'est cassé la jambe umm
umm elle s'a shoqué ummm ummm
la femme a umm de lui umm ella err s'a shoqué aussi umm et
um elle a téléphoné
um l'ambulance umm de l'hôpital umm
et après ca l'ambulance umm a arrivé umm dans la rue et il a ammené à l' hôpital

