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 recognize 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.

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.
Second language development, language learning motivation and language learning opportunities

A longitudinal case study of German high school exchange students in New Zealand

Luzia Magdalena Sauer

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Applied Language Studies and Linguistics, the University of Auckland, 2017.
Abstract

This thesis reports longitudinal case studies of three female German high school exchange students in New Zealand. The students had advanced English proficiency levels at the beginning of their 5.5 month stay and were part of a large cohort of fee-paying international students who shared the same first language (L1) in the high schools where they chose to study.

The study combines a detailed analysis of the students' L2 development along with an analysis of the learners' language learning motivation, the social contexts in which they operated and the opportunities that these afforded for language learning.

The data collected consisted of weekly diary entries, monthly reports, and six individual audio-recordings of monthly semi-structured interviews. A qualitative data analysis was performed to scrutinize the students' motivation and language learning opportunities as evidenced in their self-reports. A quantitative data analysis was carried out to capture developmental patterns in speech performance, using various measures of L2 complexity, accuracy, and fluency.

Language learning opportunities were dynamically constructed between the students and their socio-cultural environment and were unique for each student. The students' involvement in their L1 communities presented a challenge to the creation of L2 learning opportunities. Each student's motivation was affected by a complex interaction between their goals, identities, and agency, and their perceptions of their L2 communities. The efforts they expended varied and were most clearly evident in social groups that gave them access and validated their sense of self. The students' L2 development was non-linear and differed individually. Only the results for fluency were consistent, pointing to overall improvements. Findings for accuracy and complexity varied. Some trends, such as a decrease in lexical complexity, have not typically been observed in previous SA studies. A key finding was that the students' L2 development was characterized by their adaption to the speech patterns of their native speaker interlocutors over time.

The thesis provides a detailed, longitudinal account of the motivational and linguistic processes that characterize study abroad, providing insight into how and why learners perform differently in seemingly identical contexts of learning, and in this way adds to the existing literature on study abroad.
Acknowledgements

I would like to thank the high schools for giving me access to the students who participated in this study and for clarifying so readily my questions that came up along the way.

Enormous thanks go to my ten participants, not only for volunteering to participate, but for making themselves vulnerable by sharing their unique experiences with me so openly. I cannot thank Jana, Chiara and Alia enough for keeping in touch and explaining the details that I required to close information gaps.

This thesis would not have been possible without the academic guidance of my supervisor, Distinguished Professor Rod Ellis. His invaluable input and feedback kept me on track and helped me see the forest for the trees. I greatly appreciate his incredible effort at reading and returning my work promptly despite his very busy schedule.

My interraters Julius Sugioka, James Heard and Harim Lee deserve a big thank you. I am grateful to Russell Greenwood for proofreading my thesis in accordance with the Third Party Editing and Proofreading of Thesis and Dissertations Guidelines of the University of Auckland.

My family, Barbara, Kaspar, Stephan and Reta Sauer, and Verena and Olivier Braun, supplied continued emotional support from Switzerland as well as the much-needed occasional parcels with Swiss chocolates and encouraging letters. Living in New Zealand and writing this thesis would not have been possible without them.

Very special thanks go to my partner, James Heard, who has been the most valuable source of encouragement and support. He has accompanied me on this journey over the past three years, listening to my ideas, and developing a real interest in the area himself. I am deeply grateful for his patience and endurance, not to mention for running all the daily chores over the past few months.

I also gratefully acknowledge the financial support I received for completing this thesis through the University of Auckland Doctoral Scholarship.

I am indebted to the contributions all the above-mentioned people and organisations made to my thesis, yet I acknowledge that I bear full responsibility for the contents of this thesis.
# Table of Contents

1 Introduction ................................................................................................................................. 1
  1.1 Personal motivation..................................................................................................................... 1
1.2 Theoretical frameworks .............................................................................................................. 3
  1.2.1 Motivation.............................................................................................................................. 3
  1.2.2 Second language development.............................................................................................. 4
1.3 This thesis ..................................................................................................................................... 5
1.4 Thesis overview ........................................................................................................................... 6

2 Literature Review and Theoretical Background ............................................................................. 8
  2.1 Introduction ............................................................................................................................... 8
  2.2 Literature review on study abroad ............................................................................................... 8
    2.2.1 The study abroad context .................................................................................................... 8
    2.2.2 Language learning activities during study abroad ................................................................. 10
    2.2.3 Studies of language learning motivation ............................................................................ 14
    2.2.4 Linguistic outcomes of study abroad .................................................................................. 16
      2.2.4.1 Fluency ............................................................................................................................ 17
      2.2.4.2 Accuracy .......................................................................................................................... 19
      2.2.4.3 Lexical complexity .......................................................................................................... 20
      2.2.4.4 Grammatical complexity................................................................................................. 21
      2.2.4.5 Initial L2 proficiency level .............................................................................................. 21
    2.2.5 Longitudinal case studies ..................................................................................................... 23
    2.2.6 Conclusion ............................................................................................................................ 27
  2.3 Theoretical background ............................................................................................................ 28
    2.3.1 Second language development ............................................................................................ 28
      2.3.1.1 CAF .................................................................................................................................. 28
      2.3.1.2 Limited resources and speech production ....................................................................... 33
      2.3.1.3 Dynamic Systems Theory .............................................................................................. 34
    2.3.2 Motivation and language learning opportunities ................................................................. 36
      2.3.2.1 Introduction .................................................................................................................... 36
      2.3.2.2 The concept of affordances ........................................................................................... 37
      2.3.2.3 Constructions of the self ................................................................................................. 39
      2.3.2.4 Social identity ............................................................................................................... 40
      2.3.2.5 Community of practice ................................................................................................. 41
      2.3.2.6 Conclusion ....................................................................................................................... 42

3 Methodology .................................................................................................................................... 43
  3.1 Introduction ............................................................................................................................... 43
  3.2 Pilot study .................................................................................................................................... 43
  3.3 Participants .................................................................................................................................. 44
  3.4 Design ......................................................................................................................................... 47
4.4.3.4 Smallwords………………………………………………………………………117
4.4.4 Accuracy …………………………………………………………………………120
  4.4.4.1 Error-free clauses ……………………………………………………………120
  4.4.4.2 Correct finite verb phrases …………………………………………………121
4.4.5 The relationship between CAF …………………………………………………123
  4.4.5.1 Supportive growers …………………………………………………………124
  4.4.5.2 Competitive growers ………………………………………………………126
4.4.6 Summary ………………………………………………………………………128

4.5 Learning opportunities and L2 development ……………………………………129

5 Case Study 2: Chiara ………………………………………………………………134
  5.1 Background information ………………………………………………………134
  5.2 Social settings……………………………………………………………………135
    5.2.1 The host family ………………………………………………………………135
    5.2.2 The high school ………………………………………………………………138
      5.2.2.1 Inside the classrooms …………………………………………………138
      5.2.2.2 Outside the classrooms ………………………………………………..142
    5.2.3 Outside of the high school …………………………………………………142
      5.2.3.1 Social activities and trips ………………………………………………142
      5.2.3.2 Hobbies …………………………………………………………………143
    5.2.4 Summary ……………………………………………………………………144
    5.2.5 Perception of L2 progress …………………………………………………144
  5.3 L2 development……………………………………………………………………147
    5.3.1 Grammatical complexity ……………………………………………………147
      5.3.1.1 The subordination index ………………………………………………147
      5.3.1.2 Mean length of clause …………………………………………………150
    5.3.2 Lexical diversity ………………………………………………………………152
    5.3.3 Fluency ………………………………………………………………………153
      5.3.3.1 Breakdown fluency ……………………………………………………153
      5.3.3.2 Speed fluency …………………………………………………………155
      5.3.3.3 Repair fluency …………………………………………………………157
      5.3.3.4 Smallwords ……………………………………………………………….160
    5.3.4 Accuracy ………………………………………………………………………162
      5.3.4.1 Error-free clauses ………………………………………………………162
      5.3.4.2 Correct finite verb phrases ……………………………………………163
    5.3.5 The relationship between CAF ……………………………………………165
      5.3.5.1 Supportive growers ……………………………………………………166
      5.3.5.2 Competitive growers …………………………………………………169
    5.3.6 Summary ……………………………………………………………………171

5.4 Learning opportunities and L2 development …………………………………172

6 Case Study 3: Alia …………………………………………………………………177
  6.1 Background information ………………………………………………………177
6.2 Social settings.................................................................................................................. 178
   6.2.1 The host family........................................................................................................... 178
   6.2.2 The high school......................................................................................................... 181
      6.2.2.1 Inside the classrooms ......................................................................................... 182
      6.2.2.2 Outside the classrooms ..................................................................................... 184
   6.2.3 Outside of the high school ......................................................................................... 186
      6.2.3.1 Social activities and trips .................................................................................. 186
      6.2.3.2 Hobbies ............................................................................................................. 187
   6.2.4 Summary ................................................................................................................... 188
   6.2.5 Perception of L2 progress ......................................................................................... 189

6.3 L2 development .............................................................................................................. 191
   6.3.1 Grammatical complexity ......................................................................................... 191
      6.3.1.1 The subordination index ..................................................................................... 191
      6.3.1.2 Mean length of clause ...................................................................................... 194
   6.3.2 Lexical diversity ........................................................................................................ 196
   6.3.3 Fluency ..................................................................................................................... 198
      6.3.3.1 Breakdown fluency ............................................................................................. 198
      6.3.3.2 Speed fluency .................................................................................................... 200
      6.3.3.3 Repair fluency ................................................................................................... 202
      6.3.3.4 Smallwords ........................................................................................................ 205
   6.3.4 Accuracy ................................................................................................................... 207
      6.3.4.1 Error-free clauses ............................................................................................... 207
      6.3.4.2 Correct finite verb phrases ............................................................................... 208
   6.3.5 The relationships between CAF .............................................................................. 210
      6.3.5.1 Supportive growers ........................................................................................... 211
      6.3.5.2 Competitive growers ......................................................................................... 213
   6.3.6 Summary ................................................................................................................... 215

6.4 Learning opportunities and L2 development ............................................................... 216

7 Synthesis: Key Issues in Case Studies ............................................................................ 221
   7.1 Introduction .................................................................................................................. 221
   7.2 Motivation .................................................................................................................... 221
      7.2.1 Motivational orientation ....................................................................................... 221
      7.2.2 Behavioural motivation ......................................................................................... 223
      7.2.3 Attributional motivation ......................................................................................... 225
   7.3 Language learning opportunities ............................................................................... 227
      7.3.1 The host families .................................................................................................... 228
      7.3.2 Inside the classrooms ............................................................................................ 229
      7.3.3 Outside the classrooms and outside of the high school ......................................... 231
      7.3.4 Hobbies ............................................................................................................... 232
      7.3.5 Concluding comments .......................................................................................... 233
   7.4 L2 development .......................................................................................................... 236
      7.4.1 Interview factors affecting CAF ............................................................................ 237

vii
## 7.4.2 CAF results

7.4.2.1 Similarities

7.4.2.2 Differences

## 7.4.3 CAF relationships

7.4.3.1 Supportive relationships

7.4.3.2 Competing relationships

## 7.4.4 Concluding comments

## 8 Conclusions and Implications

8.1 Introduction

8.2 Motivation and language learning opportunities

8.2.1 Limitations and future research

8.3 L2 Development

8.3.1 Limitations and future research

8.4 Pedagogical implications

## References

## Appendix 1 – Ethics Approval

## Appendix 2 – Interview Template

## Appendix 3 – Diary Entry Guidelines

## Appendix 4 – Monthly Report Guidelines

## Appendix 5 – Transcription Excerpt

## Appendix 6 – Error Coding Principles
List of Figures

Figure 1: Parent and child nodes identified in this study .................................................. 82
Figure 2: Subordination Index for Jana ............................................................................. 101
Figure 3 Subordinate clauses for Jana ............................................................................. 102
Figure 4: Mean length of clause for Jana ........................................................................ 104
Figure 5: Lexical diversity for Jana ................................................................................ 106
Figure 6: Empty pauses for Jana .................................................................................... 109
Figure 7: Filled pauses for Jana ..................................................................................... 110
Figure 8: Speech Rate for Jana ....................................................................................... 111
Figure 9: Mean length of run for Jana ............................................................................ 112
Figure 10: Repetitions for Jana ....................................................................................... 114
Figure 11: Reformulations for Jana ................................................................................ 115
Figure 12: Frequency of smallwords for Jana ................................................................. 117
Figure 13: Error-free clauses for Jana ............................................................................. 120
Figure 14: Correct finite verb phrases for Jana ............................................................... 121
Figure 15: CAF relationships for Jana ............................................................................ 123
Figure 16: Relationship between mean length of clause and lexical diversity for Jana .... 124
Figure 17: Relationship between speech rate and mean length of run for Jana .......... 125
Figure 18: Relationship between fluency and accuracy for Jana .................................... 125
Figure 19: Relationship between fluency/accuracy and complexity for Jana ............... 126
Figure 20: Subordination Index for Chiara ..................................................................... 147
Figure 21: Subordinate clauses for Chiara ..................................................................... 148
Figure 22: Mean length of clause for Chiara .................................................................. 150
Figure 23: Lexical diversity for Chiara .......................................................................... 152
Figure 24: Empty pauses for Chiara ............................................................................. 153
Figure 25: Filled pauses for Chiara .............................................................................. 154
Figure 26: Speech rate for Chiara ................................................................................. 155
Figure 27: Mean length of run for Chiara ...................................................................... 156
Figure 28: Repetitions for Chiara ................................................................................. 157
Figure 29: Reformulations for Chiara .......................................................................... 158
List of Tables

Table 1: Participant profiles...........................................................................................................46
Table 2: Dates and lengths of the semi-structured interviews (item numbers in brackets) ....51
Table 3: Submission dates and length of online diary entries (item numbers in brackets) .....53
Table 4: Submission dates and length of the monthly diaries (item numbers in bracket) ......54
Table 5: Transcription conventions ................................................................................................55
Table 6: Word boundaries – transcription principles......................................................................56
Table 7: Unpruned and pruned versions of the same speech sample ........................................57
Table 8: Clause type definitions and examples.............................................................................58
Table 9: Intercoder agreement ........................................................................................................58
Table 10: Handling clause problematic features of oral productions .........................................62
Table 11: Discourse and pragmatic markers that were not given clausal status........................65
Table 12: Grammatical complexity measures used in this study..................................................67
Table 13: Classification of non-finite verb structures......................................................................68
Table 14: Lexical complexity measure used in this study..............................................................69
Table 15: Accuracy measures used in this study ............................................................................69
Table 16: Fluency measures used in this study ................................................................................72
Table 17 Types of repetitions used in this study .............................................................................74
Table 18 Examples of reformulations in this study .........................................................................75
Table 19: Words that were not counted as smallwords in this study:............................................77
Table 20: Interpretations of correlation coefficients, based on Mukaka, 2012 ..............................79
Table 21: Student constellations in the different subjects ...............................................................90
Table 22 Finite complement clauses and direct speech for Jana ..................................................103
Table 23: Range of smallwords for Jana........................................................................................118
Table 24: Results of all CAF measures for Jana ............................................................................128
Table 25: Student constellations in the different subjects ...............................................................138
Table 26: Finite complement clauses and direct speech for Chiara ............................................149
Table 27: Range of smallwords for Chiara ....................................................................................161
Table 28: Results of all CAF measures for Chiara ........................................................................171
Table 29 Student compositions in Alia’s classes ..........................................................................182
Table 30: Finite complement clauses and direct speech for Alia ................................................. 193
Table 31: Range of smallwords for Alia .................................................................................. 206
Table 32: Results of all CAF measures for Alia ...................................................................... 215
Table 33: Similarities in development of fluency .................................................................. 239
Table 34: Similarities in the use of clause types .................................................................. 240
Table 35: Differences in development of fluency ................................................................. 241
Table 36: Differences in the development of accuracy ........................................................ 242
Table 37: Differences in the development of grammatical complexity ......................... 243
Table 38: Differences in the development of lexical diversity ........................................ 244
Table 39: Supportive CAF relationships .............................................................................. 245
Table 40: Competing CAF relationships .............................................................................. 247
1 Introduction

1.1 Personal motivation

At the age of 16, I travelled to New Zealand for 10 months as an exchange student in the hope of immersing myself in a new culture, having new experiences, and improving my English skills. The thought of going away like this one day had intrigued me since I was a child. Hearing my mother’s stories about her own experiences as an exchange student in America in 1966 sparked my interest in studying abroad (SA), as did the trips that we made to the USA as a family to visit the friends she had made during her exchange year.

In July of 1999, I embarked on my adventure with only one year of formal English schooling (but some informal English learning experiences through my travel experiences) and a thirst for learning more. I was placed in a host family with three children in the Auckland region and went to a local private high school, where I remained the only international student for my entire sojourn. For ten months, I was, with very few exceptions, fully immersed in the New Zealand culture and English language learning. Very high costs for overseas calls and the use of the internet made regular contact with family and friends at home impossible. In fact, besides a few emails and letters, I only called my family three times. There were very few exchange students in Auckland that I was aware of and meeting was difficult given the poor public transport system. I knew that when I arrived in New Zealand I would be on my own and have to do what it took to fit into my new home if I was to enjoy and profit from the experience.

There were ample English language learning opportunities during my SA. I made several very close friends at high school. All of them were local students and native English speakers as there were no other potential friends. We regularly spent weekends with each other’s families. I spent many afternoons with a group of local teenagers who did not attend my school. My relationship with my host family was strong and we spent a lot of time together. Through their extensive circle of friends, I became acquainted with a large number of New Zealanders who regularly visited us or we them. My host family also took me away for weekends and we spent the long summer holidays away camping.

However, though the experiences I had in New Zealand became highly enjoyable, to reach the point of being able to enjoy them I had to overcome enormously challenging obstacles. Many
social interactions did not come ‘for free’, at least not at the beginning, but had to be actively sought, over and over again. Failure was initially a regular experience. I often had to overcome feelings of inferiority and doubt which accompanied these interactions. Dealing with and negotiating cultural and linguistic differences presented immense intellectual and emotional difficulties.

At the end of my sojourn, these authentic cultural encounters, which constantly pushed me outside my comfort zone, really made my sojourn a life-changing experience; one that required me to communicate exclusively in English for ten months. This set of factors is what made the experiences I had so rewarding and helped me to develop my English so rapidly. My experience was similar to the experience my mother had in her exchange year. However, it is vastly different to the experiences that many European students appear to have today in their exchange programmes in New Zealand.

Today, New Zealand’s international education industry is thriving. It has grown to a level such that at many high schools over 100 international students are enrolled and they are often from the same few countries. These students are typically full fee-paying students, rather than ‘exchange students’ in the traditional sense of the term. In 2014, 14,112 fee-paying international students were enrolled in New Zealand high schools. From these, 1,758 were Germans (Ministry of Education, 2014). Thus, instead of a student being the only international student, or perhaps one of very few international students at a school, s/he is now often surrounded by his/her compatriots.

Nevertheless, SA experiences are still considered the ‘sine qua non’ of language education and thousands of students worldwide spend time and money abroad on international high school exchange programmes, often assuming that they will become ‘fluent’ English speakers as a result. There appears to be little understanding of the situation that the students will find themselves in before they arrive in their chosen country to study abroad. It appears that there is a pervasive assumption that the experience will in and of itself improve the students’ English skills and lead to diverse and interesting experiences. However, studies have shown that opportunities for second language (L2) learning depend on the students’ conscious learning efforts in combination with factors that may be beyond their control, such as the qualities of their socio-cultural environment (Willis Allen, 2010; Amuzie & Winke, 2009; Badstübner & Ecke, 2009; Churchill & DuFon, 2006; Hernandez, 2010; Regan, Howard, & Lemée, 2009). The more possibilities an SA student has to engage with the L2, the more likely s/he will
improve the L2 (Willis Allen, 2002). Thus to develop, SA students must actively seek and engage in language learning activities.

This study was motivated by a desire to understand whether or not during high school study abroad in New Zealand, (where large numbers of German international students attend the same high school) learners are motivated to create and take advantage of English language learning opportunities, what the nature of these opportunities is, and what changes occur in the learners’ L2 development.

1.2 Theoretical frameworks

1.2.1 Motivation

Language learning motivation is concerned with “the effort that learners put into learning an L2 as a result of their need or desire to learn it” (Ellis, 2008, p. 972) and is a crucial variable in language learning (Willis Allen, 2010; Churchill & DuFon, 2006; Dörnyei, 2009a; Ellis, 2008; Kim, 2009b; Norton, 1995). In the past, motivation has been treated as a monolithic and stable construct. The research, typically carried out by means of questionnaires, implied that learners possess different types of motivation in a stable manner regardless of environmental influences. This view of motivation has recently been challenged. According to Dörnyei (2009b), motivation is not stable, but dynamic. It is not monolithic, but complex. And it is not just learner internal but is dependent on external factors of the environment.

Motivation researchers have more recently adopted a socio-dynamic perspective on L2 motivation research (Dörnyei and Ushioda, 2011b). This approach acknowledges two issues in particular: the learners’ socio-cultural environment, and the dynamic and complex nature of motivation. It treats language learning students as individuals with multiple identities which are shaped in real-life contexts. The learning context is viewed as something that is mutually constructed between the learners and members of their social environment (Dörnyei & Ushioda, 2011a; Ushioda, 2009). Learning is considered an inherently social process (Dörnyei, 2009b). The focus of L2 motivation is on the interaction between the agent and the “complex system of social relations, activities, experiences and multiple micro- and macro contexts in which the person is embedded, moves, and is inherently part of” (Ushioda, 2009, p. 220).

Consequently, there is substantial variation regarding the extent to which individual learners in the same context acquire an L2 and group averages are no longer believed to be able to account
for these dynamics. In order to investigate the detailed and complex nature of L2 motivation, Dörnyei and Ushioda (2011b) suggest conducting dense, multidimensional longitudinal case studies which can shed light on developmental transitions of a few, rather than many learners.

1.2.2 Second language development

Second language acquisition (SLA) literature has mainly been concerned with learners’ linguistic progress (see 3.6.5.1 for a distinction of L2 performance, development and progress in this study). Earlier studies typically examined L2 proficiency holistically, based on the learners’ subjective ratings or on general language tests, such as the OPI. These studies generally found that SA is beneficial for the improvement of oral and aural skills, but less so for reading and writing skills (Regan et al., 2009). More recent studies have scrutinised specific areas of the students’ L2 proficiency, such as their fluency, their sociolinguistic knowledge, lexical sophistication, and specific grammatical features. Conclusive findings have been made only in regard to the most examined area, oral fluency, which has been found to significantly improve during SA (Kinginger, 2009; Llanes, 2011). More research is needed, however, to show the effect that SA has on the remaining L2 areas (Llanes, 2011).

To date, most SA studies of second language acquisition (SLA) have examined ‘fixed’ linguistic outcomes and produced generalizable findings (e.g. SA fosters fluency). More recently, researchers have voiced a concern that averaged group data often “has no validity for any individual” (Larsen-Freeman, 2006, p. 598) and that if individual development were tracked, a rather different picture of L2 acquisition processes would emerge (Freed, 1995; Götz, 2013; Larsen-Freeman, 2006; Milton & Meara, 1995; Towell, 2002). While group averages may be desirable for pedagogical purposes, case-study designs are much more powerful when it comes to demonstrating language use. Given the scarcity of such designs, little is known about the dynamics inherent in individual learners’ language development during SA.

In the past, most SLA studies have adopted a pre-/post-test methodology to capture L2 growth by means of experimental tasks or tests. While such a design may show averaged group outcomes, it does not depict the incremental changes in the students’ interlanguage systems. Recently, researchers have begun to study language learning as a dynamic process. They no longer strive to know the extent to which a learner ‘acquires’ a language, but they are interested in the variability of language learners’ subsystems and how they change over time. In fact, they argue that L2 proficiency is never complete and that linguistic skills continually grow and
In comparison to mainstream SLA research that typically investigates a few linguistic structures, dynamic systems theory (DST) aims to “paint a full picture of the development of an individual’s proficiency across all components of their linguistic repertoire” (Schmid, Verspoor, & MacWhinney, 2011, p. 40) and to use data that is produced “under relatively natural conditions” (p. 39). To research dynamic L2 systems, multiple dimensions of a learners’ proficiency must be considered (Verspoor & van Dijk, 2011). Language complexity, accuracy and fluency (CAF) are considered “distinct and competing areas of L2 performance” (Housen, Kuiken, & Vedder, 2012a, p. 3) and examining these domains therefore lends itself well to a comprehensive investigation of L2 proficiency (Larsen-Freeman, 2006; Norris & Ortega, 2009). Due to learners’ limited attentional resources, CAF components are furthermore believed to compete with each other, resulting in ‘trade-offs’ between them (Schmid et al., 2011). DST scholars acknowledge that a system cannot be investigated in its totality but must be broken up into focal points of interest.

From a DST perspective, a researcher who wishes to investigate the incremental changes in language learners’ subsystems needs to collect longitudinal, naturalistic data and to investigate as many different linguistic variables as possible (Schmid et al, 2011). The research may not be able to produce generalizable findings but can show what is true for the individuals under investigation.

1.3 This thesis

This research examines the dynamic interplay of the L2 development, the language learning motivation, and the language learning opportunities of three German high school international students who spent 5.5 months in New Zealand. This study addresses the following research questions:

1. What motivations did the participants display during their study abroad?
2. What language learning opportunities did the participants experience during their study abroad?
3. What language development occurred in the participants’ oral L2 performance over the study abroad period? What were the factors that influenced their development?

To conduct my research, I used a case study design with qualitative and quantitative data analysis methods. I used interview data to understand the idiosyncrasies of the individuals’ oral L2 development and introspective methods (interviews, learner diaries and monthly reflective
reports) to examine the learners’ evolving motivation as it is co-constructed in interaction with target language (TL) members. The learners’ oral language samples were assessed by means of multiple measures of complexity, accuracy and fluency. I took an inductive approach, aiming to arrive at an understanding of the learners’ changing patterns of motivation as evidenced in their self-reports using content analysis. This thesis is exploratory.

An investigation of the motivation, language learning opportunities, and linguistic development of German high school exchange students abroad merits careful attention for a number of reasons. These students were placed in a context which differed from their home context in significant ways, resulting in unique L2, social and cultural experiences at school, in their host families, and during their free time.

This study is unique because the SA experiences of European students in general, and German high school exchange students in particular, have rarely been examined (Kinginger, 2009). Due to their often high initial L2 proficiency and knowledge of other foreign languages, they can be expected to have markedly different language learning experiences than, say, American SA students, who have been the subjects of most studies in this area.

Both the perception of ‘motivation’ as a dynamic and socially-constructed concept and the understanding of language growth as an ongoing dynamic development have only recently attracted interest among researchers. Both concepts require further investigation. Due to their complex and dynamic nature, the study of motivation and language development is considered to be most fruitful when conducted over a longer period of time in order to make changes visible. As the learners' interlanguage development and motivation are inherently emergent, variable, and non-linear, it can only be thoroughly understood through longitudinal studies (Ellis, 2008).

1.4 Thesis overview

This thesis is divided into eight chapters. Chapter Two provides a discussion of the current study abroad literature, with a particular focus on findings about language development, the types of language learning activities that exist abroad, and language learning motivation. In the second part, Chapter Two focuses on the theoretical frameworks that were used in this study for the examination of the students’ language development and motivation. Chapter Three provides information about the methodology used. It introduces the participants and explains the procedures and instruments that were used for data collection and data analysis processes.
Chapters Four, Five and Six are detailed descriptive reports of each individual participant. These chapters describe the language learning opportunities that existed for the learners as evident in their self-reports, and the changes that were observed in their interlanguage development as elicited in semi-structured interviews. Each chapter also attempts to explain the relationship between the learners’ L2 development and their language learning opportunities. In Chapter Seven, the thesis moves on to synthesize the results that were obtained in the three case studies. In this chapter, the general factors that shaped the students’ motivation, language learning opportunities, and second language development are discussed and the research questions are answered. Chapter Eight considers the theoretical and practical implications of the research and gives suggestions for future research.
2 Literature Review and Theoretical Background

2.1 Introduction

The purpose of this chapter is twofold. I will first provide an overview of the study abroad (SA) literature that focuses on language learning activities, language learning motivation, and linguistic outcomes. I will then outline the theoretical frameworks and concepts that framed the study – both for second language development, and for motivation and language learning opportunities.

2.2 Literature review on study abroad

2.2.1 The study abroad context

Second language acquisition (SLA) researchers often distinguish between different types of learning contexts: 1) the foreign language context where students learn an L2 among other subjects at school (i.e. EFL setting); 2) domestic immersion programmes (IM) where the L2 constitutes the medium of instruction for all or many subjects; 3) the study abroad (SA) context where students learn the L2 in a naturalistic setting for a limited period of time and often in conjunction with formal instruction, and 4) the naturalistic context where non-native speakers (NNSs) have taken up permanent residency and may or may not receive formal L2 instruction (Llanes, 2011). This study is only concerned with the SA context.

Kinginger (2009) defined the term *studying abroad* as a “sojourn of pre-defined duration, undertaken for educational purposes” (p. 11). In this study, the term ‘educational’ includes learning via formal school work and learning through interactions outside formal school work. Studying abroad is therefore defined in this study as a sojourn of pre-defined duration, undertaken for the purpose of academic, cultural and/or language learning.

There are two broad options available to German students intending to study abroad at a foreign high school: the students can privately organise their sojourn; or they can have their sojourn organised for them by an SA programme. In this study, a high school SA programme is considered to be an organisation that offers adolescents (in return for a programme fee) the
possibility of residing in a foreign country with a host family, where they receive formal education at a local high school for a predetermined period of time.

A further distinction can be made between classical and modern types of German high school SA programmes. In classical types of SA programmes, places are often limited to a few students, which creates fierce competition. Students can choose the country for their SA, but not where they will live in that country. In modern types of SA programmes, places are generally unlimited and the students can choose the region where they will live and the school where they will study. Exchange organisations of the second type are usually referred to as ‘agencies’. Agencies occupy an intermediary position between the SA students and their high schools in the host countries. They are advisory bodies and are typically responsible for aspects related to the students’ travel between the home and the host country, communication with the overseas high school (including student enrolments at the host schools), and student visa-related issues. Agencies do not deal directly with host families. The agency ensures that the participants meet the programme criteria. Criteria for programme participation are typically based on the students’ school achievements (grade average of at least 3 – the equivalent of ‘achieved’). Students of agency based programmes tend to be referred to as ‘international full fee-paying students’ rather than ‘exchange students’ in New Zealand.

The SA context is considered particularly fruitful for an investigation of L2 acquisition given the apparent plethora of language learning opportunities (Churchill & DuFon, 2006). The study of L2 acquisition in SA contexts is still in its infancy despite the large amount of research that has been carried out, due to the complexity of the SA context (Churchill & DuFon, 2006, p. 27). One of the difficulties is the idiosyncratic nature of individual students’ SA experience, in combination with individual differences (IDs), which complicates making generalisations about SA students’ overseas experiences and their linguistic outcomes. Additionally, researchers have measured IDs and L2 progress using different instruments and analyses, which yielded incomparable results. Findings are often controversial, as they are only attributable to the specific setting, instruments, and students in question (Churchill & DuFon, 2006; Llanes, 2011).

Various studies have documented the ways learner external as well as learner internal factors influence students’ interlanguage development during SA. In the following sections, I will present the key findings of prior SLA studies during SA. Firstly, I will describe SA process variables and their impact on L2 learning. Secondly, I will examine the impact of learner
motivation on L2 acquisition. Then, I will discuss the effects SA has on learners’ oral L2 proficiency. Finally, I will provide a more detailed review of three longitudinal case studies which examined the relationship between learner variables, contextual influences, and L2 development.

2.2.2 Language learning activities during study abroad

The way a learner uses and acquires a second language (L2) correlates with the setting (Ellis, 2008). The SA context (and naturalistic contexts in general) is associated with informal (implicit) rather than formal (explicit) learning of the target language (TL) (Ellis, 2008; Grieve, 2010; Schauer, 2006). Conversely, in the EFL setting students typically experience formal learning and pay attention to rules (Ellis, 2008; Ife, 2000). In reality, however, learning opportunities vary just as much within each setting as they do between them. During study abroad, some learners have regular opportunities to practice the L2, while others have fewer opportunities. Learning settings are immensely complex arrangements of L2 learning opportunities.

Often, students, teachers and programme coordinators believe that spending time in the target language (TL) country automatically leads to higher language proficiency. Yet, studies have shown that ‘just being abroad’ is not enough (Amuzie & Winke, 2009; Badstübner & Ecke, 2009; Churchill & DuFon, 2006; Hernandez, 2010; Regan et al., 2009; Willis Allen, 2002). In the following, I will review what the existing literature reveals about the actual language learning activities of L2 learners abroad, both inside and outside the classroom, and their impact on L2 growth.

L2 learning inside the classroom

SA research typically focuses on the examination of informal language learning outside the classroom (Kinginger, 2009). High school classroom settings and their effect on language attainment have rarely been investigated. Generally, high school exchange students are believed to be “afforded increased opportunity to interact with native speakers” through their participation in mainstream classrooms (Spenader, 2005, p. 382). Ward and Kennedy’s (1993) study of high school AFS students showed that more contact with host nationals related to increased fluency, less homesickness and fewer socio-cultural problems. Conversely, they showed that students who “reside primarily in an ‘expatriate bubble’” (p. 143) experienced psychological wellbeing but not socio-cultural adaption.
Churchill’s (2006) study of Japanese high school students in America provides a more nuanced picture of the foreign classroom. He showed that the sociocultural context of the high school classroom has a significant impact on interactions among classroom attendees, shaping not only language learning opportunities but also the students’ identities, social relations, and well-being. Churchill attributed this to influences such as the time students were given to interact with each other, classroom organisation, and the adjustments made by the local students and teachers in response to their arrival. The students felt involved and accepted when they were invited to participate as competent members. They felt excluded from classroom participation during teacher-centred classes and during individual tasks. Churchill’s research demonstrates that learning in the foreign classroom is heavily influenced by the host community’s treatment of the exchange students but does not account for the fact that students have agency and are capable of changing the conditions of their learning (Menezes, 2011; Shotter & Newson, 1982).

**L2 Learning Outside the Classroom**

Informal out-of-class communication with NSs presents an especially attractive opportunity for SA students to practice their L2, because this type of communication is unavailable in the home country (Kinginger, 2009). Social interaction may occur on an irregular basis during service encounters (e.g. at the post office, in a shoe shop) or on a more regular basis with friends and host family members.

Forming friendships with L2 speakers constitutes an important goal for many SA participants (Hernandez, 2010; Willis Allen, 2010). However, building such networks and engaging in regular, let alone quality, interaction with NSs is, according to the SA literature, extremely challenging (Churchill & DuFon, 2006; Hernandez, 2010; Ife, 2000).

Making friends does not happen overnight, but depends on a number of factors related to student agency and host community attitudes and behaviour. Churchill and DuFon (2006), Isabelli-Garcia (2006), and Hernandez (2010) showed that the students who succeeded in creating networks distinguished themselves from others owing to their positive attitudes, openness towards the host culture, and their sustained efforts to gain access to NS interaction. These students also tended to show a tolerance and desire for unmodified NS input during interaction and did not shy away from linguistically challenging conversations with NSs. According to Isabelli-Garcia (2006), extended networks “may be the driving force behind language acquisition in the SA context” (p. 257). In Ginsberg and Miller’s (2000) study, the students’ opportunities to mingle with NSs depended on their interests. Yet despite the quantity
of time spent with NSs and the intimacy of the relationships, only a few participants yielded linguistic gains from these interactions. The authors argued that language gains do not result from the opportunities students are provided with to learn a language but from the “complex interactions among learning strategies, the students’ ideas about language and how it is learned, motivation, and the learning support provided by their [host] contacts” (p. 237).

As a result of the difficulties that students perceive between their own and the host culture, many SA students choose to socialise in the secure environment of their L1 peers (Churchill & DuFon, 2006; Doyle, 2009; Grieve, 2010, Ife, 2000; Kinginger, 2009; Tanaka, 2004; Ward & Kennedy, 1993). This tendency to stick with L1 peers can be explained by the students’ “degree of reluctance to leave the safety of the national fold” (Ife, 2000, p. 33). Pellegrino Aveni (2005) also found that NNSs who struggle to handle the linguistic challenge of L2 dialogues often develop a self-image that is inferior to the one at home. Instead of risking embarrassment, they choose not to participate in NS conversations.

In short, the act of establishing (deep) friendships abroad is complicated. Being motivated to make friends does not suffice. Students must also demonstrate great willingness and self-confidence to participate not only in linguistically demanding NS interaction but also in a culture whose norms may deviate considerably from the home culture norms. Additionally, social networks are not created unless with reciprocal action by members of the host culture.

**L2 Learning in the Host Family Setting**

In theory, with its abundance of authentic input, opportunities for interaction, and possibilities for output, the homestay environment is an ideal site for language acquisition during SA. Living in host families can also be culturally and socially rewarding. However, homestays have received very little attention in the SA literature (Schmidt-Rinehart & Knight, 2004) and not every scholar agrees they are advantageous.

The studies indicate that linguistic and/or cultural interaction in host families does not just happen. It is largely determined by the nature of the relationship between the student and the host family, their personalities, the amount of time spent together, and the level of the students’ initial L2 proficiency (Churchill & DuFon, 2006; Kinginger, 2009). In the most favourable case, students converse with their host parents and siblings on an everyday basis and are thus provided with unique language learning opportunities (Kinginger, 2009; Tanaka, 2004). Ideal host parents involve their host children in their social activities which give them access to social
and linguistic practices otherwise unavailable. Schmidt-Rinehart and Knight (2004) found that the mealtime constituted the most important opportunity for students to engage in daily interaction and exchange information about local customs and traditions. Adjustment to host family practices was most successful when the students’ initial L2 proficiency allowed them to be involved in extended conversations or when host family members perceived their host students as open and mature.

Student-host family relationships can also prove problematic. Students in unfortunate relationships have lamented that their host parents either talked too much and listened too little, or interacted too little, or that they did not include the student enough in the host family’s social activities (Schmidt-Rinehart & Knight, 2004). Tanaka (2004) found that some SA students’ felt unsupported in their L2 learning because their host parents did not play the role of ‘language teacher’. Consequently, exchange students and hosts often only exchanged routine short dialogues in simple, everyday English and students regularly retreated into their rooms or went out with L1 friends. At their worst, homestay arrangements may lead to isolation and humiliation (Kinginger, 2009). Studies have shown that it is not uncommon for SA students to feel rejected by their host families, and unfavourable homestays may influence the student’s attitudes towards the host culture in general (Churchill & DuFon, 2006; Grieve, 2010).

Conversely, the more contact students have with their host family members, the more they are likely to improve their fluency (Hernandez, 2010), their written accuracy (Llanes, Tragant, & Serrano, 2012), and their socio-pragmatic knowledge (Kinginger, 2009). Thus, living with host families constitutes an important, but not necessarily positive, component of the students’ SA experience and their language development.

NON-INTERACTIVE LANGUAGE LEARNING ACTIVITIES

Not many studies have investigated the extent to which high school SA students actively engage in analytical language learning activities. The few studies that do exist suggest that learners’ contact with the written language is very restricted (Ife, 2000; Tanaka, 2004). Ife’s participants typically pursued activities that did not “involve to any great extent . . . analytical engagement with the language” (e.g. checking grammar and word meanings, keeping vocabulary lists and translating texts into the L1) (p. 34, emphasis in original). Tanaka (2004), whose study involved learners in New Zealand language schools, reported similar results. Only a few of his participants engaged in private reading activities (reading books and newspapers)
during short-term SA. However, it remains somewhat doubtful to what extent students themselves want to develop their analytical knowledge while studying abroad.

2.2.3 Studies of language learning motivation

Despite the impact that the factors mentioned above may have on L2 acquisition abroad, researchers repeatedly point out the degree of variation of development that occurs among students who are exposed to the same (or very similar) external factors. In fact, it has been shown that the extent to which L2 learning occurs depends largely on individual differences (IDs) (Churchill & DuFon, 2006; Freed, 1995; Isabelli-Garcia, 2006; Kinginger, 2009; Regan, 1998) and “a full account of how learners differ with regard to how, how much, and how fast they learn an L2 will need to take account of both social and psychological factors, and how these interact” (Ellis, 2008, p. 641).

Different types and quantities of IDs have been identified (e.g. identity, intelligence, personality, anxiety, learner beliefs, and learning strategies) but the most thoroughly researched ones are language aptitude and motivation. This study is concerned with the latter. Language learning motivation has been shown to correlate significantly with L2 success (Dörnyei, 2009). A modern view on motivation reveals that it is multi-faceted and dynamic in nature, and involves sub-constructs such as learner agency and learner identity (Dörnyei & Ushioda, 2009).

Language learning motivation is a highly complex phenomenon and depends on various factors both relating to the language learner and to the community the learner engages with. An especially problematic aspect of motivation research concerns its conceptualisation. Some researchers focus on the stable and more general aspects of motivation that are typically related to the students’ attitudes and beliefs (Dörnyei & Ushioda, 2011a). They thereby fail to address the dynamic nature of motivation; namely that it is bound to action and that it emerges as a result of the characteristics of specific learning contexts.

For example, Badstübner and Ecke (2009), in a study of short-term SA students (Americans in Germany), examined the relationship between students’ motives for studying abroad and their anticipated language and culture gains. Their findings suggest that establishing contacts with L2 members and the opportunity for cultural enrichment (both relating to integrative motivation) as well as being able to travel, were reported as the highest rated motives for going abroad. Students’ self-reports (questionnaires) revealed that their integrative motivation
correlated significantly with expectations of gains in speaking and listening abilities, “two skills that are necessary for interaction and integration in the culture” (p. 47).

However, according to Churchill and DuFon’s (2006) review, pre-departure language learning motivation does not automatically lead to L2 acquisition, let alone to immersion during SA. Rather, maintaining motivation to learn the L2, and to actively participate in L2 conversations, requires not only a high level of motivation, but also an extrovert personality, substantial language proficiency, and the willingness to communicate. Hence SA students’ initial or even pre- SA ‘motivation’ to learn a language is distinct from the type(s) of motivation students continually re-develop (or do not develop) as a result of social interactions abroad.

In their review of motivation studies, Churchill and DuFon (2006) identified two factors that are likely to mitigate the interplay between motivation and L2 acquisition during SA: length of stay, and previous language learning experience. Students who only spend a few weeks abroad are likely to be less motivated to integrate into the host community, and more likely to develop less positive attitudes towards the host culture than long-term students. Given the advantage of social networks that might develop as a result of integration, some researchers have doubted the value of short-term programmes (Willis Allen, 2002). However, other studies challenge the advantage of long-term programmes, showing that short-term programmes can lead to an increase in learner motivation and that, on the other hand, “some learners did not develop a positive attitude towards the host community even after a year abroad” (Churchill & DuFon, 2006, p. 15).

Willis Allen (2010) and Kim (2009b) conducted their studies from an activity theory perspective. Both studies suggest that agency has a very important influence on motivation and SLA. They found that motivation to learn an L2 is not just an integral and stable characteristic of SA students but is shaped by their capacity for goal setting and self-regulation. Students who were able to set specific and effective goals for what they wanted to learn during their SA experience also possessed the ability to reflect on the communication they experienced, which delivered more substantial benefits and enhanced their language learning (Willis Allen, 2010). Being the agent of their own actions was a decisive factor for these students in strengthening their language-learning motivation. Additionally, these students typically possessed what Willis Allen called “higher-level cognitive motives”, arising from an integrative motivation to learn the L2 (2010, p. 45).
In Willis Allen’s study, the students with “lower-level cognitive motives” (p. 45), arising from an instrumental motivation (e.g. earning a French minor) typically developed unfocused goals (e.g. improving fluency), and invested less effort in learning the L2, hoping that increased language skills would develop as an inevitable consequence of living in the target country. They took little responsibility for their learning, blamed their limited interaction with host nationals on others, and were unable to enhance or maintain their language learning motivation during SA. Willis Allen (2010) concluded that language-learning motivation is dynamic in nature and shaped by self-regulatory strategies (learner internal) as well as relationship with others (learner external). Believing that SA automatically leads to L2 proficiency (the ‘osmosis’ myth) may even prove detrimental to students’ linguistic development during SA. Kim (2009b) concluded that without internalizing reasons for learning and integrating motivation with specific learning goals and a sense of participation, motivation cannot emerge and L2 learning is unlikely to be facilitated.

Isabelli-Garcia’s (2006) and Hernandez’s (2010) studies of SA students’ language also exemplify the correlation between motivation, attitudes towards the host culture, meaningful interaction, and L2 learning. Both studies show that the higher the motivation of the students in learning the L2 and becoming part of the community is, the more extensive their social network becomes (more detail on Isabelli-Garcia’s study is provided below). Depending on the types of motivation for L2 learning, attitudes towards the host culture, quality of input, and strength of social network, then, the language development of SA students in the same programme is promoted or impeded. Thus it is not just the opportunities SA students are provided with to communicate, which fosters L2 acquisition during SA, but what the students do with them.

2.2.4 Linguistic outcomes of study abroad

In the past two decades, a large number of studies have investigated learners’ linguistic developments. Earlier studies tended to examine learners’ general holistic proficiency, which provided global scores but failed to take into account “changes that occur in real language use” and “progress on specific aspects of the learner’s linguistic repertoire” (Regan et al., 2009, p. 21). These studies suggest that SA is beneficial for the improvement of oral and aural skills but less so for reading and writing skills (Regan et al., 2009).
Due to researchers’ continuing desire to “define and document the emergence of proficiency” (Kinginger, 2009, p. 40), the analytical focus of SA studies has gradually expanded (Churchill & DuFon, 2006). More recent studies scrutinize specific areas of the students’ L2 proficiency (e.g., lexical acquisition). Furthermore, the focus on oral/aural skills of earlier studies has shifted to an investigation of students’ literacy as well as their sociolinguistic and sociopragmatic knowledge (Churchill & DuFon, 2006; Grieve, 2010; Schauer, 2006). Current studies are increasingly conducted from a sociolinguistic perspective, moving beyond the mere documentation of linguistic structures to examine communicative competence (Regan et al., 2009). Besides a change in focus, Churchill and DuFon (2006) detected a “renewed interest in comparative studies” (p. 27). Typically, such studies investigate potential differences of L2 processes in the ‘at home’, the immersion and the SA contexts. According to Llanes (2011), there is still too little literature showing the respective impact of the SA context on the development of each area.

In the following, I will give an account of what is known to date about the linguistic oral L2 development of SA students.

2.2.4.1 Fluency

Probably the most commonly held belief about SA is that time spent abroad is especially beneficial for improving ‘fluency’. Rather than ‘fluency’, students typically refer to the global construct of ‘oral proficiency’ (i.e. Lennon’s broad sense of fluency, 1990). Researchers agree that fluency develops most out of the four language skills during SA and that the SA context is superior to any other context for fluency development (Kinginger, 2009; Llanes, 2011). Rather than becoming entirely native-like, though, L2 learners are believed to increase “some type of global fluency; the ability to ‘sound good’” (Freed, 1995, p. 10) which has been found to be linked to increased speech rate and to the use of fillers, modifiers, formulae, and the use of compensation and communication strategies in a native-like way (Freed, 1995).

Features that have been addressed when investigating fluency development during SA include speech rate, the frequency and nature of pauses, the length of fluent speech runs, repairs, clusters of dysfluencies, and naturalness of rhythm and intonation (Freed, 1995; Kinginger, 2009). Researchers have used a number of different instruments and analytical methods to assess learners’ fluency.
Early research used native speaker (NS) judgements of learner performance to measure fluency. The first standardized procedure to measure oral proficiency was implemented by the American Council on the Teaching of Foreign Languages (ACTFL) in the 1980s. The test consists of an Oral Proficiency Interview (OPI) which has had a considerable influence on establishing assessment standards (Kinginger, 2009, p. 42). Ever since its implementation, however, the OPI has been criticized for its inability to discriminate progress (it only measures holistic scores at a particular point in time) (Freed, 1995), its unequal treatment of distinct-level students (it favours lower-level students) (Llanes, 2011), its gender and interviewer bias, as well as its inability to test short-term development (Kinginger, 2009). Despite these flaws, the OPI (and other similar interviews) is still a popular and widely used to test L2 learners’ oral proficiency.

Using the OPI, Hernandez (2010) found that L2 learners either maintained or improved global oral proficiency during one semester abroad. Davidson’s (2010) large scale study indicated that chances to make “cross-threshold gains” in the OPI test (i.e. significant gains) were highest after one academic year abroad (p. 14). In his study, less than a third of one-semester abroad students crossed the threshold, while more than half of the one academic-year-students moved from an intermediate to an advanced fluency-level (Davidson, 2010). According to Davidson (2010), reaching an advanced fluency level during SA is essential for successful interaction with NSs and “appropriate levels of rapport building” (p. 19).

Other language elicitation instruments include controlled experimental tasks (e.g. narrative retellings), and semi-controlled or uncontrolled methods such as role plays and interviews, or recordings of naturally-occurring speech. These tasks often afforded more specific fluency measures. Studies based on specific fluency measures have shown that an SA period is likely to increase learners’ speech rate and mean length of run, and to decrease dysfluent pauses (Freed, 1995; Lennon, 1990; Llanes et al., 2012; Mora & Valls-Ferrer, 2012; Tonkyn, 2012; Towell, Hawkins, & Bazergui, 1996; Wood, 2012). In contrast, the use of reformulations has been shown to increase during SA as a sign of advanced monitoring skills (Freed, 1995; Lennon, 1990; Riggenbach, 1991). Many researchers, however, noted that there were considerable individual differences between participants (e.g. Larsen-Freeman, 2006; Riggenbach, 1991). Additionally, the results obtained in studies that used more than two data collection points (pre- and post-test) showed considerable variation in the development of proficiency variables (Mora & Valls-Ferrer, 2012; Wood, 2012).
Testing oral proficiency must be handled with caution, as general definitions of fluency, against which L2 learners are measured, may not converge with the actual fluency of NSs. That is, NSs’ oral performance has been shown to exhibit the very features that would be considered dysfluent when uttered by NNSs (Kinginger, 2009). In essence, Kinginger argues that “there will always be questions about the ecological validity of research in this [proficiency-related] domain” until tests are actually based not on academic norms but on the “realities of natural language use” (p. 67).

2.2.4.2 Accuracy

Testing SA learners’ grammar is rare and has yielded mixed results. Though “research on grammar … has begun to look at discrete features, … this research has been largely restricted to a few study abroad contexts and to relatively few target languages” (Churchill & DuFon, 2006, p. 27). While some studies emphasise the superiority of the SA context over other contexts (AH and IM) for grammatical progress (Guntermann, 1995, cited in Llanes, 2011; Mora & Valls-Ferrer, 2012), most studies report limited or no grammatical progress during SA (Collentine, 2004) or significant individual differences in learner development (Larsen-Freeman, 2006). In general, the existing literature questions whether the SA context, which typically presents more opportunities to engage in informal than formal language learning activities, is as favourable to the development of grammatical features as it is to the acquisition of fluency. Grammatical features are typically taught in EFL or ESL classrooms, but not explicitly referred to during informal daily interaction in the SA context (Skehan, 1998). Due to the advantage of classroom instruction for developing grammatical knowledge, at home students have often been found to manifest very similar or even higher improvement in their grammatical competence than SA students (Llanes, 2011). Gains in accuracy during informal conversations in the SA context are questionable because NS communication is “elliptical and incomplete in surface form” (Skehan, 1998, p. 25). During participation in L2 interactions, the need for precise syntax diminishes, while that for meaning increases (Skehan, 1998).

However, some researchers who have examined the acquisition of particular grammatical features during SA (e.g. such as gender agreement, past time markers or the subjunctive) found evidence of the superiority of the SA context (Howard, 2001, 2005, 2006, cited in Isabelli-Garcia, 2006; Llanes, 2011; Isabelli-Garcia, 2006). As with the other domains of L2 proficiency, the SA length appears to play a major role regarding the enhancement of
grammatical proficiency. Longer-term SA has been associated with greater gains in grammar acquisition (Kinginger, 2009) than shorter-term residence abroad.

The problem associated with tracking learners’ grammatical development overseas concerns the application of traditional measurement instruments. In Kinginger’s (2009) view, such instruments fail to measure what is learned during SA (e.g. variations of standard forms) and, instead, focus on aspects learnt in classroom contexts. She claims that as long as instruments are not adapted to testing language acquired during SA, changes in students’ interlanguage system cannot be adequately captured. Thus future research “will need to come to terms with variation in language in general and particularly with non-pedagogic language use” (Kinginger, 2009b, p. 82).

2.2.4.3 Lexical complexity

Vocabulary acquisition during SA has not been thoroughly researched (Llanes, 2011). What has been investigated so far has shown the SA context promotes lexical growth. Findings demonstrate that studying abroad leads to an enriched lexicon (approximating NS norms), to a reduction of lexical errors, and to an improvement of lexical richness (Kinginger, 2009; Wood, 2012).

Foster and Tavakoli’s (2009) study on Iranian’s vocabulary growth showed that the group that studied in a naturalistic context used much more diverse vocabulary than the at-home group. Larsen-Freeman’s (2006) study of Chinese learners in the USA showed an increase in group scores for lexical diversity, but large differences in individuals. Daller and Xu (2007) found that Chinese learners of English in ESL contexts displayed speech of higher lexical diversity on average than the Chinese learners in EFL contexts. Mora and Valls-Ferrer (2012) measured advanced L2 learners’ lexical sophistication in oral speech during three informal interviews, twice during their studies in a Spanish university EFL context and once during their obligatory three-months sojourn in an English speaking country. Their averaged scores increased with each interview and in particular after studying abroad.

Fitzpatrick (2012), whose interest was in the dynamics of vocabulary acquisition, found that it was not the learners’ general vocabulary knowledge, but specific aspects of their vocabulary knowledge that accelerated during SA. By implementing a written word association task (completed every six weeks by one SA student), she detected a steady increase in some aspects of vocabulary knowledge (e.g. number of collocations, NS-like associations) yet striking
inconsistencies in others (e.g. word form, form-meaning connections and orthography) (p. 81). Fitzpatrick (2012) concluded that learners do not increase their knowledge in predictable ways but that the acquisition process is chaotic and that “we should pay attention to the ‘messy little details’ of the process” (p. 92).

Due to the beneficial effect SA has on vocabulary acquisition, Llanes (2011) strongly recommends that further research be conducted in this area.

### 2.2.4.4 Grammatical complexity

The effect that SA has on learners’ levels of oral grammatical complexity has rarely been investigated. The few studies that exist have drawn attention to the difficulty of assuming that complexity in oral speech grows in a linear fashion or grows at all (particularly compared to the written mode where complexification is a sign of more advanced language use) (Brown & Yule, 1983; Ferrari, 2012). Typical characteristics of oral production are: non-hierarchical clause constructions (i.e. parataxis), unspecific vocabulary; simple common clause conjunctions, incomplete sentences, a lack of complex noun phrases (e.g. adjectives rarely modify nouns) and generally less densely packed information (Brown & Yule, 1983). Generally, nominalisations and other phrasal or clausal complexification strategies are indicative of scientific texts but not of conversation (Biber, 1991). Brown and Yule (1983) emphasised that only after mental rehearsal, or reproduction of the same output on various occasions does oral speech become grammatically more complex for native speakers.

The complexity of SA students is dynamic and shows great variation in clausal density in individual learners (Ferrari, 2012; Larsen-Freeman, 2006; Polat & Kim, 2014) but not averaged group data (Larsen-Freeman, 2006; Llanes et al., 2012). Clause length for the NNS in Polat and Kim’s (2014) study was also highly variable with no clear trend over a period of one year (fortnightly measurements).

### 2.2.4.5 Initial L2 proficiency level

The findings reported above are generalisations. Ultimately, every SA student’s learning trajectory is unique. SA researchers have detected variables that are said to predict individual students’ L2 development during SA, such as age, gender, length of stay, and initial L2 proficiency level. I will briefly discuss the findings that have been made with regard to the
aspect that is most relevant to this study: the relationship between L2 development and initial L2 proficiency level.

During SA, in-depth L2 practice is often only achieved when students have overcome linguistic limitations. Beginners are easily bewildered by “the amount, delivery rate, and complexity of the language that surrounds them”, particularly in conversations where NSs do not fine-tune their speech to the level of the learners (Segalowitz & Freed, 2004, p. 174). As a consequence, novice learners may be unable to participate in language activities other than very simple ones (Rivers, 1998). If linguistic inadequacy cannot be overcome by students, they may choose to avoid interaction with L2 speakers during their SA by withdrawing from contact or spending time with L1 peers (Tanaka, 2004).

Some studies have concluded that only students with a certain pre-departure L2 level (known as the ‘threshold level’) are able to substantially benefit from the input-rich context of the naturalistic environment and it is advantageous for students to have “a somewhat good command of the L2” prior to going abroad (Llanes, 2011, p. 208). More experienced learners have been considered better able to manage the “ceaseless flow of TL input” during conversations with NSs (Rivers, 1998, p. 492) and to be “better capable of attending to subtler aspects of the L2” than learners in less-advanced stages (Regan et al., 2009, p. 39). Given this advantage, some studies hypothesize higher level students are also more likely to actively seek interaction with NSs, while lower level students “may shy away from such situations, because of the greater communicative difficulties they entail” (Regan et al., 2009, p. 39).

The body of research on students’ initial L2 proficiency level and L2 gains during SA reports mixed findings, however. Much of the research indicates that advanced level language learners do not profit more linguistically from SA than lower level students (Llanes, 2011; Regan et al., 2009). In fact, studies have shown that lower-level students typically show more progress than advanced level students after studying abroad (Regan et al., 2009). This finding may not, however, be surprising, as “lower-level learners have the most to gain by definition” (Churchill & DuFon, 2006, p. 26).

Furthermore, there is evidence indicating that learners benefit from different kinds of learning opportunities. Informal conversations, for example, do not always accelerate L2 growth for advanced L2 learners. Non-interactive language activities (e.g. reading) have been found to deliver more benefits for high-level students (Freed, 1995). Regan et al. (2009) found that the development uncovered by research is dependent on the aspect under investigation. For
example, advanced learners have been shown to be better able to attend to pragmatic aspects than grammatical aspects (p. 39).

The majority of studies have investigated American students abroad. European students have not received much attention (Kinginger, 2009; Llanes, 2011). Research on Americans is somewhat one-sided, especially when it comes to the length of SA programmes and pre-programme language proficiency. Most of the research findings relate to language learners of low proficiency. Llanes (2011) therefore suggests reducing this imbalance by increasing studies of European SA students. She claims that examining European exchange students abroad is especially interesting as bi-/multilingualism has been positively related to greater success in L2 acquisition. European students are likely to have passed the threshold level of L2 proficiency prior to SA and are arguably better able to meaningfully engage in language-related activities than less advanced students (Churchill & DuFon, 2006; Llanes, 2011).

2.2.5 Longitudinal case studies

Longitudinal SA case studies which combine L2 assessment with in-depth qualitative investigations are uncommon. Exceptions are Kinginger’s (2008), Spenader’s (2005; 2011) and Isabelli-Garcia’s (2004; 2006), studies which examined the SA experiences of a small cohort of American learners in France, Sweden, and Argentina, respectively. These three studies will be reviewed in detail.

Kinginger (2008) investigated 24 American university students on a semester-long SA in France. She took a sociocultural perspective with a particular focus on L2 learners as “persons with hearts, bodies, and minds, with memories, fantasies, loyalties, identities” (p. 104). Various instruments were used for language elicitation. Once before and once after their SA, a language awareness interview was held to assess the students’ knowledge of sociolinguistic variation, a French test was administered to assess the learners’ reading and listening skills, and a role-play task was conducted to provide speaking samples on the speech acts ‘inviting’ and ‘requesting’. A focus group of students was also asked to keep a biweekly journal and a logbook to note L2 contact hours. They participated in two semi-structured interviews (pre-SA and mid-SA) in English about language learning motives and L2 history. The L2 results showed SA had a generally positive effect on the students L2 skills. After a semester abroad, the group increased their French listening skills, their ability to distinguish between informal and formal forms of address, and their awareness of colloquial forms for leave-taking. Fewer gains were registered
for reading skills despite the students’ engagement in academic work. Considerable individual
differences were observed for all measures.

Kinginger’s (2008) analysis of six focal students demonstrated that studying abroad is not a
generalizable experience, but unique to every student, and that its quality is largely determined
by environmental influences (i.e. the level of access provided by L2 communities) and the way
learners make sense of the sojourn, using their own “interprettive tools recruited from among
the narrative and ideological resources available to them” (p. 104). Kinginger showed that these
relationships are mediated by “power differentials in which age, gender, nationality, race, or
ethnicity are highlighted” (p. 21). She also teased out a link between the focal students’ L2
gains and their language learning experiences. One student showed a real desire to immerse
himself in the host community by doing volunteer work and was on “constant lookout for
insights about the language” (p. 86). He developed networks with French nationals which
increased his already very high L2 skills during SA. Another student made an effort to immerse,
yet real engagement in L2 social networks, and hence L2 progress, was impeded by her serious
L2 comprehension difficulties. Students who spent most of their SA with compatriots showed
few L2 gains.

Students who felt challenged in their identities as Americans and who frequently
communicated with family back home had fewer opportunities for L2 interaction. Another
impediment for American students abroad was the desire of French interlocutors to practice
their English with them. Kinginger concluded that the consequences of living in a globalized
world restricted L2 learning opportunities during SA, requiring of students “more profound
and durable commitment than has been needed in the past” (p. 105).

Kinginger’s (2008) study provides valuable insights into a few individuals’ SA journeys, their
successes and difficulties. The drawback of the study was the diversity of the instruments she
used for L2 data collection (which renders generalisation and interpretation difficult,) as well
as the lack of NS baseline data. Kinginger (2008) emphasized, however, that the primary focus,
and hence the strength of the study, was the qualitative exploration of the learners’ experiences.
She advised future researchers to undertake “interprettive projects integrating narrative study
and formal assessment” (p. 114).

Spenader (2005; 2011) investigated the year-long SA experiences of four US high school
exchange students in Sweden. Her study was ethnographic with a particular focus on the
interplay between cross-cultural adaptation (framed in terms of cultural identity), the
relationships with host nationals, and language learning at the beginner level. Spenader collected data using observation (on two occasions), interviews with students, teachers and host parents (on two occasions), monthly e-mail questionnaires, and the Acculturation Index to examine the quality of the students’ overseas experiences. She tested the learners’ global proficiency and fluency levels (judged impressionistically by a NS) by means of standardised interviews (OPIs) five and ten months into their sojourn.

All participants made L2 gains both for fluency and for global proficiency, albeit to different degrees. One student achieved an Intermediate High level, and three students reached the ‘Superior’ ACTFL level after 10 months abroad. Spenader concluded that studying abroad is highly beneficial even for students with no previous L2 experience. Spenader’s analysis also suggested a link between the students’ L2 gains and the extent to which they identified with Swedish nationals. Those who desired a Swedish identity reported higher L2 gains. The student who was less successful at language learning rejected the TL culture and removed himself from L2 communities.

Spenader’s case study analysis also revealed that language learning was affected by “a supportive host environment and meaningful relationships with host nationals” (2011, p. 392). The closer the friendships, the more opportunities for the use of the L2. However, the student with the closest L2 networks did not report the highest L2 gains of the group – instead, highest L2 gains were measured for the only student who maintained a close relationship with an L1 friend. Spenader concluded that “there is evidence in support of a more complex relationship between sojourn and co- and host-nationals” (2005, p. 185) than has been assumed before and that contact with compatriots is only detrimental to language learning when it replaces contact with L2 communities. Compatriots fulfil an important role when it comes to psychological support.

Positive influences of language learning were also found for students who exhibited risk-taking behaviour, assertiveness, and a sense of humour. Some impediments were also outside the control of the students, such as their place of residence (in isolated, rural areas) or programme variables (e.g. being placed in L2 rather than mainstream courses). Furthermore, all learners reported that learning opportunities were not created automatically, but required considerable effort.

Like Kinginger (2008), Spenader’s research provides comprehensive in-depth insights into the sojourners experiences in Sweden, taking into account multiple people’s perspectives and
aspects related to the learners’ identity and motivation. The study’s scope is probably most obviously limited in the area of L2 assessment. Only global proficiency and impressionistic fluency was examined and L2 progress was unambiguously linked to higher scores. The study failed to account for differences in language scores attributable to discourse context and changes in overall proficiency. Making comparisons with a NS baseline could have been more meaningful.

Isabelli-Garcia (2004; 2006) studied the relationship between the SA environment and the development of communicative ‘functions’ (i.e. the ability to narrate ‘simple’ vs. ‘complex’ stories and opinions), fluency, and accuracy of four university-level American learners of Spanish during a semester-long sojourn in Argentina. She examined a variety of aspects of the sojourners’ SA experience, including language learning motivation, learner identity, acculturation, attitudes towards host nationals, and the social networks the students were able to build with host nationals.

Data was collected by means of pre- and post- SA simulated oral proficiency interviews (SOPI) (for global proficiency measures), six informal interviews during the learners’ SA (to quantitatively measure accuracy development), weekly diary entries (for qualitative assessments of the students’ experiences), daily social network logs, and a pre-programme questionnaire to investigate the learners’ motivational intensity and orientation (low vs. high; instrumental vs. integrative, or intrinsic).

Isabelli-Garcia found that the extent to which the students were able to form L2 networks depended on their attitudes towards the host culture, the type of motivation they displayed, and the extent to which they were able to acculturate to the Argentinian lifestyle. Motivation, in turn, was symbiotically influenced by the extent to which they were able to form friendships with locals. Two of the four students were able to build extended “second order zone” networks (forming relationships with friends’ friends) by making considerable effort and showing a tolerance for “annoying behaviors of the members the [sic] host culture” (2006, p. 255). The other two students were less successful at building complex relationships due to lower willingness to interact and negative attitudes towards the host culture.

Of the four participants, three students improved their ACTFL scores by one level, and one did not show progress. For the accuracy measures, Isabelli-Garcia noted much variation, including periods of ‘backsliding’ and periods of improvements, with only a few learners showing “a steady increase with little to no regression” (2004, p. 62). Fluency increased (the quantity and
flow of their utterances), although development was not linear. Interpreting the development of ‘communicative function’ proved problematic, as it was found to be mediated by personality and confidence. Some learners kept using simpler functions, while others progressed to using more complex narrative functions during their SA.

Isabelli-Garcia detected a direct relationship between the qualitatively measured variables and L2 development for three students (i.e. high motivation, positive attitudes, and minimal cultural differences were linked to increased L2 scores). However, one student improved L2 levels despite displaying negative attitudes towards and having poor relationships with host nationals. Isabelli-Garcia interpreted this outcome as a consequence of his literary engagement (reading the newspaper). Isabelli-Garcia (2006) concluded that there is a “complex relationship between motivation, acculturation and the development of social networks that ultimately provide opportunities for exposure to the target language and extended interactions that may be the driving force behind language acquisition in the SA context” (p. 257).

Although affording some rich insights, a limitation of this research is the examination of social networks by means of their ‘strands’ (i.e. number of connectors) rather than the quality of these relationships. Isabelli-Garcia assumed that learners are pushed to speak about a larger range of topics and to use more communicative functions when interacting with extended networks, rather than ‘first order’ networks. It is questionable, however, to what extent this is true, since in-depth conversations with a single host national may afford ideal opportunities for the use of a range of linguistic and discourse structures.

2.2.6 Conclusion

Despite significant advances that have been made in the past two decades, the study of L2 acquisition during SA is still in need of further, in-depth research. The variety of programme types, languages involved, instruments used, and individual learner differences make generalisations across SA studies problematic (Churchill & DuFon, 2006; Llanes et al., 2012). According to SA researchers, future research should combine in-depth qualitative analysis of SA processes with an examination of L2 development (Churchill & DuFon, 2006; Kinginger, 2009; Regan et al., 2009). Although such SA studies are likely to pose considerable methodological challenges, Churchill and DuFon (2006) claim that “a series of such studies conducted in different programmes and target language contexts could go a long way towards
improving our understanding of what is learned, by whom and under what conditions” (Churchill & DuFon, 2006, p. 27).

2.3 Theoretical background

2.3.1 Second language development

2.3.1.1 CAF

Complexity, accuracy, and fluency (CAF) measures have been widely used for investigations of L2 performance and proficiency (Housen et al., 2012a). The three dimensions are theorized to have independent status in L2 proficiency/performance and are therefore considered suitable for examining it (Ellis, 2005; Ellis & Barkhuizen, 2005; Housen et al., 2012a; Larsen-Freeman, 2006; Polat & Kim, 2014; Skehan, 2009). According to Housen et al. (2012a), complexity is believed to mirror “internalisation of new L2 elements”, accuracy is believed to show “modification of L2 knowledge”, and fluency is believed to reflect “consolidation and proceduralization of L2 knowledge” (p. 3). It has been argued that “all three must be considered if any general claims about learners’ L2 performance and proficiency are to be made” (Housen et al., 2012a, p. 3). Each CAF dimension is considered “multi-layered, multifaceted, and multidimensional” (Housen et al., 2012a, p. 5). It is this nature of the CAF components that makes their use varied but also controversial.

The exact conceptualisation of CAF measures, their validity and reliability as measures of L2 performance, and the extent to which they reflect more advanced language use have been debated (e.g. Ferrari, 2012; Lambert & Kormos, 2014; Norris & Ortega, 2009; Pallotti, 2009). Any assessment of learner performance must determine the factors that influence the CAF components since they “may be differentially manifested under different conditions of L2 use, and … differentially developed by different types of learners and under different learning conditions” (Housen et al., 2012a, p. 3).

Below, I will describe how each CAF dimension has been defined and operationalised, and to what extent it can account for language progress. I will also describe what external factors have been shown to influence CAF, and the theorized interdependency of the CAF components as suggested by Skehan’s (1992) Trade-off Hypothesis. Lastly, I will show how these views connect to those of dynamic systems theory and briefly describe the core assumptions of this theory.
GRAMMATICAL COMPLEXITY

Complexity is considered the most controversial construct of the CAF triad (Housen et al., 2012a). Little agreement exists regarding the definition of the term (Bulté & Housen, 2012). A first distinction must be made between using the term ‘complexity’ for tasks (i.e. the objective demandingness of a task) and for language (i.e. the quantitative and qualitative components of a language feature). Grammatical complexity measures are concerned with the latter, in particular with the “state of [L2 learners’] declarative linguistic IL knowledge” (Housen et al., 2012a, p. 5).

Bulté and Housen (2012) and Norris and Ortega (2009) emphasized that when measuring grammatical complexity, it is crucial to take into account different sources of complexity and to avoid complexity measures that compute the same phenomenon. Many measures are available; “[e]ach measure has its own strength but also presents challenges in terms of reliability, validity, sensitivity and discriminatory power, and, not in the least, its practical feasibility” (Bulté & Housen, 2012, p. 35). Grammatical complexity measures usually aim to quantify the length of a unit, or the amount and type of coordination, subordination and embedding (Bulté & Housen, 2012).

Norris and Ortega (2009) advised measuring the following three dimensions of complexity for both written and oral data: 1) the level of coordination as an indication of progress at the beginner level (e.g. coordination index); 2) the level of subordination as an indication of complexification at an intermediate and upper-intermediate level (e.g. subordination index), and 3) the level of phrase complexification or the length of clause as an indication of complexification at the most advanced level of second language proficiency (e.g. mean length of clause) (p. 563-564). They also recommend using not only general measures that tap global complexity but also more fine-grained measures to locate specific features of the speaker’s performance (e.g. types of subordinate clauses).

Pallotti (2009) argued that there are three notions of complexity that must be distinguished from each other to account for language development even though they are often related to one another. Language features can be structurally complex (i.e. referring to entities that are composed of at least two elements), they can be ‘difficult’ (i.e. cognitively demanding), or they can be ‘acquired late’ in the developmental process (i.e. advanced language structures).
Another key issue that must be borne in mind when using grammatical complexity measures to examine L2 development or proficiency is that complexity “is always measured for particular purposes in particular settings and with particular developmental targets in mind” (Norris & Ortega, 2009, p. 575). Although this is also true for accuracy and fluency, factors related to the communicative activity, such as discourse genre, are crucially important when it comes to interpreting complexity measures. Ferrari (2012), for example, showed that different contexts require different levels of complexification for language production to appear appropriate. In highly interactional spontaneous speech (e.g. telephone calls), grammatical complexity is low. In monologic speech (e.g. retelling tasks), utterances are more complex. Spontaneous interactional speech is generally characterised by its simplicity and non-hierarchical clause constructions (Carter & McCarthy, 2006). Hence the assumption that clauses become ‘longer’ or utterances more hierarchical in more advanced language in informal conversations is invalid.

**Lexical Complexity**

Lexical complexity does not figure independently in the CAF triad and is typically assumed to form a part of general complexity. The construct does not feature prominently in CAF studies. Skehan (2009) argued that lexical complexity should be treated as a separate construct from the triad given its prominent status in speech production models.

Lexical complexity has been said to consist of separate but interrelated components: lexical density (the ratio of lexical words to tokens), lexical sophistication (proportion of advanced words in a text), lexical diversity (the range of a learner’s vocabulary), and the number of errors in vocabulary use (Read, 2000). Most studies in SLA have measured lexical diversity, which has been identified as the component that best correlates with NS ratings of the quality of NNS output (Lu, 2012). Lexical diversity indicates a learner’s productive lexical resources (Milton, 2009).

Lexical diversity has typically been measured by means of the type-token ratio (TTR). The value of TTR as a valid measure of lexical diversity has been questioned due to its sensitivity to text length. That is, “samples containing larger numbers of tokens give lower values for TTR and vice versa” (MacWhinney, 2013, p. 127). Despite the fact that the operationalization of lexical diversity is still somewhat unsatisfying, some measures have proven more reliable than others. Typically, these are Guiraud’s Index (Daller, Turlik, & Weir, 2013; Levkina & Gilabert, 2012), Uber-Index (Dewaele & Pavlenko, 2003), and vocd (Jarvis & Daller, 2013).
The extent to which lexical diversity per se is an indicator of language progress is questionable. As for fluency, external factors such as topic are likely to account for changes in the construct (Pallotti, 2009).

**Accuracy**

Accuracy is considered “the most straightforward and internally consistent construct of the CAF triad” (Housen et al., 2012a, p. 4). Like complexity, the construct is believed to be linked to learners’ current interlanguage knowledge (both procedural and declarative). The correctness of L2 performance can be defined as “the extent to which an L2 learner’s performance . . . deviates from a norm” (Housen et al., 2012a, p. 4). However, what exactly the “relative nature of deviation and error” is, must be clearly determined before speech analysis is conducted (p. 4). To evaluate accuracy, one must decide whether to compare deviations from prescriptive target language (TL) norms or from “non-standard and even non-native usages fully acceptable in some social contexts or some communities” (p. 4). According to Housen et al. (2012a), in the case of the latter, it might be more accurate to refer to speech performance accuracy as ‘appropriateness’ or ‘acceptability’. Similarly, before analysing data, one must choose to include in the error count either only errors that are absolutely wrong or also features that are dis-preferred. However, such a distinction is not always straightforward and subjectivity plays a role in this decision-making process (Ellis & Barkhuizen, 2005). Lambert and Kormos (2014) argued that researchers must find a principled way of identifying errors to ensure reliability.

There are both general and specific indices of language accuracy. General measures include counting the total number of errors per clause, per unit, or per hundred words. Specific measures are applied when researchers have an interest in particular types of errors and their development over time. Most researchers claim that more general error measures are better able to distinguish levels of proficiency than specific measures (Albrechtsen, Henriksen, & Faerch, 1980; Skehan & Foster, 1999; 2012; Tonkyn, 2012). However, more recently, there have been calls for “more distinct linguistic features, as a complement to the use of more global measures” (Housen et al., 2012a). One possibility is to focus on one specific type of error that is generally considered serious or reflective of general accuracy.

When interpreting the construct in speech performance, Pallotti (2009) warns that accuracy must be considered separately from interlanguage progress. She claimed that accuracy is not a direct or valid indicator of progress since “we may find texts that are very accurate but scarcely
developed . . . and texts containing many errors but exhibiting several traits of evolution” (p. 592).

**Fluency**

Lennon (1990) explained that there are two senses of fluency: a broad sense and a narrow sense. The broad sense is used as a cover term for oral proficiency in general. It is the type of fluency that is considered “a mark of social accomplishment” (p. 189). SLA researchers consider fluency in its narrow sense whereby fluency constitutes only one component of overall proficiency and is measured in terms of various fluency-related phenomena. Compared to accuracy and complexity, fluency is considered a pure performance phenomenon. Its use is not linked to linguistic knowledge or a memory ‘store’ but to speech processing mechanisms related “to the manner in which linguistic information has been stored and can be recalled from memory systems” (Towell, 2012). Fluency can be defined as the “ability to fill time with talk without unnatural hesitations” (De Jong, Steinel, Florijn, Schoonen, & Hulstijn, 2012, p. 123).

Fluency is a multi-dimensional construct and different subcategories have been identified. One of the most influential classifications of fluency variables is that by Tavakoli and Skehan (2005), who distinguished between breakdown fluency (empty and filled pauses), repair fluency (repetitions and reformulations), and speed fluency (speech rate, articulation rate and mean length of run). According to Towell (2012), the former two sub-constructs relate to the degree to which a learner feels s/he has reliably stored information and to which s/he has established procedures for repair operations in case of communication breakdown. Speed fluency is “reliant on procedures for storage and recall” (p. 55). Many researchers have argued that fluency measures are good indicators of language development/progress. However, as for accuracy and complexity, fluency may be “subject to too many variables to reflect development directly” (Lambert & Kormos, 2014, p. 610).

Fluency can also be enhanced through the use of formulaic sequences. Formulaic sequences are ‘prefabricated units’ which can be “stored and retrieved as whole units from the mental lexicon” (Götz, 2013, p. 23). In informal oral conversation, formulaic sequences are commonly multi-word sequences such as ‘and something like that’ (Biber, Conrad, & Cortes, 2004). ‘Smallwords’ are considered a subcategory of formulaic sequences (Hasselgren, 2002). Hasselgren (2002) defined smallwords as “small words and phrases, occurring with high frequency in the spoken language, that help to keep our speech flowing, yet do not contribute essentially to the message itself” (p. 150) (e.g. ‘like’). Hasselgren believes that native-like
fluency is only achieved with the acquisition of a large enough range and varied use of smallwords. Götz (2013) claimed that smallwords can also increase speech rate when used as lexical fillers.

2.3.1.2 Limited resources and speech production

Skehan’s (1992) ‘Trade-off Hypothesis’ claims that L2 learners’ cognitive resources (i.e. attentional capacity and working memory) are limited during speech performance. It states that all three CAF components compete with each other for attentional resources and that learners are compelled to focus their attention on one CAF construct or another when producing language (but not on all three at the same time). The Trade-off Hypothesis proposes that different conditions can affect different aspects of speech production processes, thus leading to differences in output.

Skehan (2009) drew on Levelt’s (1989) Model of Speech Production to explain how trade-offs occur. A more recent model of speech processes that has specifically been developed for the study of L2 production is Kormos’s (2011) Bilingual Model of Speech Production. It is based on Levelt’s (1989) Speech Production Model but additionally accounts for the existence of a “fourth and L2 specific knowledge store: a declarative memory of syntactic and phonological rules in L2” (p. 41). Both Skehan (2009) and Kormos (2011) postulate that attention during speech processes is limited and has to be shared between multiple resources, affecting speech performance.

The speech production models assume that three components or processes are responsible for language production: speech conceptualization, encoding, and articulation. In the ‘conceptualization’ stage, speakers form the plan or propositional content of their output. In the ‘formulation’ stage, this ‘pre-verbal plan’ is syntactically and lexically encoded by retrieving lexis and grammar from memory stores. In the ‘articulation’ stage, the plan is verbalised.

In L1 speech, encoding mechanisms are believed to operate largely automatically. Only planning (conceptualization) and monitoring are subject to conscious attentional control. In L2 speech, particularly for learners of lower proficiency levels, all three speech processing stages are believed to be subject to attentional control and to proceed in a serial rather than a parallel fashion (Kormos, 2011). This deficit in attention results in performance trade-offs.
In L2 speech, when the conceptual demands are high (e.g. dynamic and abstract information must be expressed, the topic at hand is unfamiliar or formulaic language is not available), then less attention can be given to encoding processes. This, according to Skehan (2009) and Kormos (2011), positively affects grammatical and lexical complexity, but leads to less fluent and accurate output. Kormos (2011) believes that in particular the L2-specific declarative knowledge store accounts for slow retrieval in the formulator stage. Only advanced L2 learners, who have acquired rules in the form of procedural knowledge are able to retrieve them quickly and unconsciously, even when conceptual demands are high.

Conversely, if less attention is devoted to planning (e.g. because the topic is familiar and no complex ideas need to be expressed), then more resources are available for message formulation, resulting in more accurate and fluent but less complex output. This is because conceptually ‘simpler’ tasks do not demand that speakers activate complex schemata. In tasks of this kind, learners make use of frequently-used and well-known (i.e. centrally stored) words whose syntactical information is well proceduralised and retrieval is not a laborious process.

2.3.1.3 Dynamic Systems Theory

Larsen-Freeman (2009) stated that investigating the causal relationships between task-related factors and CAF is simplistic and does not advance the way the triad is understood. She suggests that ‘the way forward’ is to study CAF interactions, and to do so from a Dynamic Systems Theory (DST) perspective. DST lends itself well to an interpretation of the longitudinal study of learners’ oral capacity since it is takes account of time and variability. That is, DST is "a general theory that explains how any complex system, which consists of a set of interrelated variables that continually mutually affect each other, may change over time" (Verspoor & van Dijk, 2011, p. 25).

DST assumes, like Skehan’s (1992) Trade-Off Hypothesis, that learners have to make choices during language production regarding what aspects they want to allocate their attentional resources to. Given this tension, trade-offs are expected between the different language components (Schmid et al., 2011). According to DST, a meaningful approach for describing developmental systems involves considering the interaction between variables (i.e. ‘growers’), rather than just describing a single variable’s development. Vespoor, Wander, and de Bot (2009) defined ‘growers’ as “a variable that changes quantitatively on the basis of a principle of increase (e.g. learning) or decrease (e.g. forgetting or suppressing)” (p. 189). However,
‘increase’ and ‘decrease’ are relative concepts and depending on the variable, they can refer to both ‘learning’ and ‘forgetting/suppressing’. For an accurate description of the relationship between variables, their developing patterns can be compared (i.e. whether or not they both increase or decrease at any one timeiii). Four kinds of relationships between ‘growers’ have been defined by Verspoor and van Dijk (2011): neutral, supportive, competitive and conditional/precursor relationships. Relationships between variables without meaningful interactions are considered neutral. A supportive relationship between two variables exists when the growers develop in the same way. It indicates that the two variables support each other in their development (e.g. fluency may support accuracy). A competitive relationship exists between growers that develops in alternating patterns (e.g. complexity may compete with accuracy). Conditional or precursor interactions between two growers signal that one variable is dependent on the growth of the other variable (e.g. subordination may condition nominalization). In a dynamic system, relationships between variables can also be asymmetrical (i.e. they change over time). For instance, a formerly competitive interaction may develop into a supportive relationship.

This thesis employs DST as a conceptual metaphor for understanding the way systems (i.e. the three CAF domains) change over time rather than as a mathematical model to examine language variation. From a DST perspective (as from other ‘usage-based’ or ‘emergentist’ theories), intra-individual variability is considered an important source of information about the underlying developmental process. A few assumptions underlying a DST view of development are considered particularly important to this study (based on de Bot & Larsen-Freeman, 2011; Larsen-Freeman, 2006, 2009; Larsen-Freeman & Cameron, 2008; Verspoor et al., 2011). These are:

1. Language is a complex dynamic system that shows nonlinear development. The development of linguistic skills can grow and decline, resulting in a waxing and waning of language patterns. Language is never completely acquired and development is ongoing.
2. Language is a complex dynamic system consisting of components or subsystems that are completely interconnected. Linguistic subsystems interact in ways that are neutral, supportive, competitive, or conditional.
3. There is constant change and systems only temporarily settle into attractor states. An attractor state is the state which the system prefers to be in over other states at a
particular point in time. It is the long-term behaviour of the system, a stable pattern of the system that is not easily perturbed by small changes in the conditions.

2.3.2 Motivation and language learning opportunities

In this thesis, a particular focus was given to the language learning opportunities that the participants experienced during their SA as a result of contextual affordances and the learners’ agency. Language learning motivation was considered to significantly contribute to the creation of these language learning opportunities and therefore given special attention in this thesis. Below I will introduce the theoretical backgrounds that were used to interpret the findings of this study.

2.3.2.1 Introduction

Over the past few years, the previously static, linear, and one-dimensional conceptualisation of motivation has been considered insufficient to “do justice to idiosyncrasies of personal meaning-making in social context” (Ushioda, 2009, p. 218). From a socio-dynamic perspective, motivation is conceptualised as a multi-dimensional, non-linear, dynamic concept (Dörnyei & Ushioda, 2009). Language learners are viewed as individuals with multiple identities that are shaped in real-life contexts. This view calls for a qualitative approach to investigating motivation that takes into account interaction processes and contextual phenomena.

One of the most ardent advocates of this view, Ushioda (2009), claims that to gain knowledge about the emergence of L2 motivation, examining cognitive psychological processes in isolation from the sociocultural and socio-historically situated background is insufficient. In order to truly grasp how students develop, maintain or lose motivation in any context, the interaction between the individual and the social environment must be examined. From her perspective, a “person-in-context relations view” (p. 220) does not need to rely on a particular theoretical framework to account for the emergence of motivation, but may draw insight from different frameworks which emphasise “a more conceptually embedded relational view of motivation and identity” (p. 220), particularly those that “highlight the dynamic interplay between agency and social structure” (p. 221).

In my study I will selectively draw on the following frameworks in support of my conclusions: the concept of affordances (Menezes, 2011; Shotter & Newson, 1982), communities of

2.3.2.2 The concept of affordances

‘Affordances’ is a term that refers to how an individual perceives the environment and responds to it. It is borrowed from ecology where it describes the relationship between an organism and other elements in the ecosystem. Affordances can be described as “demands and requirements, opportunities and limitations, rejections and invitations, enables and constraints” (Shotter & Newson, 1982, p. 34) in the environment. Defining characteristics of the concept of affordance are ‘perception’, ‘interpretation’, and ‘action’. Before conditions in the environment become affordances and can be acted upon, they must not only exist in the environment but must also be perceived and interpreted as affordances (Menezes, 2011). An example of how the same environmental property can serve different affordances is a fallen leaf in the forest. A worm might interpret the fallen leaf as food and break it down into smaller pieces. A frog might interpret the leaf as an affordance for shelter and sit underneath it. In all instances, before making use of the leaf, the leaf must exist and the animals have to perceive it in the forest, and then interpret it as useful to their specific purposes.

In SLA, the concept of affordance is used as a metaphor to describe the inter-relatedness of an individual and the social environment s/he is embedded in. From this perspective, the individual and the environment are viewed as “mutually constituted and mutually defining” (Shotter & Newson, 1982, p. 33). A person’s environment contains human beings and cultural artefacts. An individual acts in response to what s/he apprehends the environment “offers, demands, or affords” (p. 33); i.e. the environments’ ‘affordances’. Affordances for language learning are therefore not simply “properties of the environment” but they are “restricted by the user’s perceptions” (Menezes, 2011, p. 61). The ability to notice affordances is decisive for language learning: “language learners living in similar ‘niches’ . . . can have different perceptions, which afford them different experiences, and, consequently, differences in language development” (p. 62). Once affordances have been noticed, they must be responded to. Affordances only ‘emerge’ when humans interact with their social world; they are “action in potential” (p. 61).

Within this framework of affordances, the term ‘niche’ refers to the specifics of a person’s “relational position” in the world (p. 62). A niche is a “set of affordances” (p. 62) that the
environment offers and which a learner can take advantage of. When language learners share a niche with other learners (e.g. the classroom environment), resources may be restricted and their use becomes a matter of competition. In these cases, affordances have to be found in other niches (e.g. outside the classroom). One means of ‘enhancing’ or ‘enlarging’ a niche is by travelling abroad. SA niches offer more opportunities to be in contact with English than non-English speaking environments. However, these contexts do not guarantee language learning opportunities as affordances only emerge “from language use, from the participation of the learners in a niche and from those with whom they interact in social practices” (Menezes, 2011, p. 67). Not all students succeed in perceiving the opportunities offered by their environments as affordances, and thus fail to take advantage of them (Menezes, 2011).

Besides acting in their world, individuals can also act upon their world. They can create new conditions by exerting their own agency. Agency is linked to the significance a learner attributes to his/her learning goal, and depending on the extent of learner agency, individuals approach activities with different levels of effort and persistence. Some humans employ carefully selected strategies to assist their learning process, while others take on less responsibility for their learning – thereby taking the risk of failing to reach their goals. However, agency is not determined by the individual alone but is constantly “co-and renegotiated with those around the individual and with the society at large” (Lantolf & Pavlenko, 2001, p. 148).

According to sociocultural perspectives, agency is also not formed on the spot but socially and historically constructed. A learner’s behaviour is a reflection of ideologies about language learning that have been “appropriated in childhood [and] that incline us to act and react in specific ways” (Lantolf & Pavlenko, 2001, p. 145). However, behaviour can also be changed because humans are capable of critical analysis of their social interactions and because they are able to determine their goals in a dynamic way.

Learning a language is therefore “a matter of acting in the world, a matter of autonomy and agency” (Menezes, 2011, p. 71). In order to understand learning outcomes, especially in naturalistic environments, one must understand the degree of agency taken by the learner and the extent of access the learner is given in a community. In Menezes’ (2011) view, the relationship between environmental affordances (i.e. what the learner perceives the environment offers) and learner agency can best be understood when we “pay attention to what learners have to tell us” (p. 60).
2.3.2.3 Constructions of the self

According to Pellegrino Aveni (2005), language use in the naturalistic setting cannot be thoroughly understood without an examination of the learner’s self in that context. She postulates that language learning success abroad depends not only on learners’ motivation but is likely to be the consequence of the complex relations between learners and their social environment, the access they are granted to participate in the community and the possibilities they are given to portray an ideal image of themselves.

Social interaction, and thus language use, is hence inextricably linked to the construction of the self (Pellegrino Aveni, 2005). She postulates that when individuals are not in full control of their language production, such as during a stay in a foreign country, the conveyance of the self becomes problematic. Grammatical errors, wrong lexical choices and insufficient sociopragmatic competence, for example, may result in the learner perceiving his/her L2 self as inferior to his/her L1 self. When sojourners use the L2 in a foreign environment, they calculate the risks they take in creating an undesirable image of themselves.

According to Pellegrino Aveni (2005), language learners are believed to constantly evaluate whether the benefits of speaking in a certain context at a certain time with a certain group of people will outweigh the potential harm. Calculations of risks are made based on the communicative goals of the learners. According to Pellegrino Aveni, L2 learners engage in communication in order to pursue three types of goals: to exchange information, to build social networks and to practice the L2. To minimise the threat to the self in interaction, learners tend to strive for goals which they believe they can achieve. However, the evaluation of cost and benefit students make for each communicative situation also depends on the communicative need they have for using the L2. Certain situations make a cost-benefit evaluation redundant simply because the communicative need is paramount. In such situations, students simply must use the L2 despite the potential threat it causes. Typically, the more welcome a student feels in social interaction, the more relaxed s/he will act and the more likely s/he will take linguistic risks. Conversely, the more disturbing the L2 learner perceives his/her presence to be in communicative situations, the more likely s/he will expect his/her self to be threatened, and the fewer risks s/he is willing to take in conversation. In interaction involving potential threat, learners may also feel that their presence is unwelcome and that their status within the group of interlocutors is inappropriate.
Feeling uneasy during social interaction may arouse anxiety – this may manifest itself in various ways: learners may deliver a qualitatively poorer language performance; they may lessen the extent to which they take risks in using new linguistic structures, and they may reduce the degree to which they use language spontaneously (Pellegrino Aveni). The effects of anxiety may suppress the students’ desire to participate in social interaction to such an extent that they feel unable to form future relationships in the foreign country. In extreme instances, the fear of threat to the self may result in the learner’s complete retreat from communicative interaction (Pellegrino Aveni). Needless to say, avoiding participation in the community may not only be emotionally detrimental, but also linguistically.

Language learning outcomes, as has been made rather clear, must be assessed in relation to an individual’s position vis-à-vis his/her social environment.

### 2.3.2.4 Social identity

A term that is often used interchangeably with ‘self’ is ‘identity’. However, while self relates only to perceptions that an individual has of his/her own behaviour and characteristics, identity may refer both to personal specifications and to group affiliations. Various different types of identities exist, such as national identity, cultural identity, ethnic identity, personal identity or social identity. Norton, one of the leading figures in identity research, prefers to use the expression ‘social identity’ to refer to “identity as multiple, a site of struggle, and subject to change” (Norton, 1995, p. 9). Social identity is linked to people’s desire “for recognition, the desire for affiliation, and the desire for security and safety” (Norton, 1997, p. 410). Desires, in turn, are related to material and symbolic sources, to which people have more or less access. Hence the concept of social identity indirectly relates to notions of power, gender, race and other relations that arise out of the complex relationship between learners and their language learning contexts. Social identity theory accounts for the inequitable relations of power that learners encounter and that may limit their opportunities to use the L2 (Norton, 1995).

Norton argues that it is not ‘motivation’, but ‘investment’ that best captures the complex relationship between learners and NSs and can shed light on learners’ “sometimes ambivalent desire to learn and practice” a language (2012, p. 411). Investment, according to Norton and Toohey (2005), can be understood in terms of ‘cultural capital’, that is the “knowledge, credentials, and modes of thought that characterize different classes and groups in relation to specific sets of social forms” (p. 122). Hence, if a learner invests in the L2, s/he presumes that
s/he acquires of a range of symbolic and material sources which will later have the potential to increase the value of his/her own cultural capital. Norton and Toohey (2005) furthermore argue that as “value of their cultural capital increases, so learners’ sense of themselves and their desires for the future are reassessed” (p. 122).

Like Pellegrino Aveni (2005), Norton and Toohey (2005) argue that individuals are diverse, complex beings and they change dynamically and constantly over time and space. Thus social identity, like the self, has a changing quality. Identity is inherently subjective as the learner understands him/herself and his/her relation to the world through thoughts and emotions. Norton and Toohey (2005) further argue that in dialogue with others, that is through the use of language, the individual’s selves and identities are manifested and created. This view of social identity as a textured construct that is shaped during social interaction is influenced by the poststructuralist stance that Norton and Toohey take.

2.3.2.5 Community of practice

Since social interaction is a determining force when it comes to the creation of identity and the self, it is pivotal to understand the degree of access to participation that learners have in the new community. Lantolf and Pavlenko (2001) argue that the ability to find a place in a new culture largely depends on the nature and extent of access that members of the host community offer the learner. Participation in given communities of practice may range from ‘non-participation’ to ‘full participation’. Newcomers to a community interact with ‘old-timers through a process of “legitimate peripheral participation” (Lave & Wenger, 1991). That is, with the assistance of long-term members and by engaging in activities of a community, newcomers become increasingly experienced with the practices of that particular community and thus steadily move towards fuller participation in that community. As “learning is about mediated participation” (Lantolf & Pavlenko, 2001, p. 148) and because participation impacts on the learner’s construction of the self, there is a close relationship between the degree of participation, acquiring new knowledge through participation, and the construction of the individual’s’ identity. Individuals both define themselves through the practices they are engaged in and those they are not. Hence, in the process of positioning themselves in a new community, “there is a dialectic struggle between the learner and the community out of which emerges the learner’s position and identity” (Lantolf & Pavlenko, 2001, p. 149). When learners change their social environments and thus when (or if) they move from the periphery to full participation they can encounter numerous difficulties which “may prove shocking and result
in a long, difficult and isolated transition phase” (Lantolf & Pavlenko, 2001, p. 149). During phases of “linguistic, social and cultural transitions . . . the learners’ multiple identities . . . become sites of contestation and renegotiation” (Lantolf & Pavlenko, 2001, p. 149).

Although a community has tremendous power in determining the learners’ position in that community, learners also have agency in determining the extent of their participation. Learners may either opt for a position (and thereby often struggle) in the community or may deliberately resist (or partly resist) access to the community by remaining on its margins. Unquestionably, learners who are denied access to the host community or choose to remain at its margins do not engage in social interactions and are therefore deprived of language learning opportunities.

2.3.2.6 Conclusion

In this study, the participants’ engagements in social activities with host nationals in the different social settings are considered opportunities for language learning. This belief rests on the view that language learning is a “fundamentally social, cultural, and temporal activity” (Morita, 2004, p. 575) and that language learning processes “emerge from relationships between individuals and their sociocultural environments” (Kurata, 2011, p. 138). It is acknowledged that the extent to which students participate in social activities is situationally dependent and dynamic and, that the characteristics of these activities determine possibilities for interlanguage development and progress.

In particular, it appears from the above review of social-oriented accounts of L2 learning that what individual learners do in terms of finding and making opportunities depends on their perceptions and interpretations of environmental ‘properties’, the degree to which they feel validated and safe during communication, the abilities they have in finding an identity that allows for participation and the right to speak in a community, and the extent to which they are given access to groups ‘of power’ by members of the host community. Hence the language learners of this study will be viewed as “self-reflective intentional agent[s]” (p. 218) who are “located in particular cultural and historical contexts” (Ushioda, 2009, p. 216).
3 Methodology

3.1 Introduction

This study adopted a longitudinal case-study approach to implement in-depth investigations of three German high school SA students’ language-related and social experiences in New Zealand. The study focused on the students’ language learning motivation and the language learning opportunities they experienced. It also set out to obtain naturally sounding oral speech samples collected on multiple occasions to account for L2 development and progress. Data analysis was both descriptive (for the language data) and interpretative (for the analysis of motivation and social contexts).

This examination was intended to tie second language development to the context in which it occurred (i.e. high school SA in New Zealand) and to show to what extent and how the different settings changed the nature of the learner’s motivation, their language learning opportunities, and, possibly, their L2 proficiency (and vice versa). The context of interest incorporated both social and non-social types of L2 interactions in the host family, inside and outside the classroom, and in the students’ free time.

The following research questions were addressed in this study:

1. What motivations did the participants display during their study abroad?
2. What language learning opportunities did the participants experience during their study abroad?
3. What language development occurred in the participants’ oral L2 performance over the study abroad period? What were the factors that influenced their development?

To answer these three research questions, I examined how and why language learning motivation, language learning opportunities, and L2 development arose and changed over time.

3.2 Pilot study

In September 2013, I carried out a one-month pilot study with seven German international students. The pilot study was conducted to reflect on the suitability of the data collection
instruments. I did one interview with each student, collected diary entries twice a week and collected one reflective report.

I wanted to establish the best questions to employ in interviews. I was also concerned that open questions would not elicit fulsome data and closed questions might narrow the focus too much and result in ‘thin’ data (Brown and Yule, 1983). The pilot study established that my open interview questions stimulated long, fulsome answers resulting in rich data. Consequently, no changes were made to the questions.

During the pilot study, the diary entries were written inconsistently and varied in length and detail. All seven participants reported that writing two diary entries a week was unfeasible. The quantity and quality of their diary entries was negatively affected by demotivation rather than the comprehensibility of the guidelines (see section 3.5.3). I therefore reduced the number of diary entries that the students had to write in the main study from two per week to one per week.

The reflective reports resulted in clear, informative texts, and the students reported enjoying the activity. I did not change this data collection method.

3.3 Participants

For this longitudinal study German students were selected because the linguistic and social SA experiences of students from European backgrounds and with high L2 proficiency levels have rarely been investigated (Kinginger, 2009). German speakers were chosen because I and the participants shared a language (High German) which would enable them to write journal entries in German if they desired. Many German students undertake their SA at New Zealand high schools. Targeting German students not only limited individual differences but also ensured participant numbers.

I sought assistance from New Zealand high schools and requested introductions to German international fee-paying students to complete my study. The high schools were informed of the nature of the research, the requirements for the students, and the fact that all data would be treated confidentially. The high school principals of the students that participated in this study signed a consent form agreeing to inform the students of my study. With their signatures, they also confirmed that the students’ participation in the project would not in any way affect their relationship with high school staff members or their school grades. The high schools forwarded
my letter of interest to German students who were enrolled at their school for one or two semesters in 2014. This letter included a participant information sheet that described the project and procedures, their rights to withdraw at any time, the ‘koha’ (see below), information on data storage (i.e. data would be destroyed after six years) and that all the collected data would be treated confidentially. The letter informed the students that participation was voluntary. The students who were still interested in participating then signed a consent form. All participants were provided with the contact details of the university and the ethics approval number. All information sheets and consent forms were approved by the Ethics Committee of the University of Auckland (see Appendix 1). Every effort was made to proceed ethically and empathically, considering the vulnerable position these young people occupied being alone in a foreign country and sharing their personal stories with me.

Originally, I collected data from ten German students, but I only analysed data from three students in this study. The large cohort ensured that I had rich, varied data and could complete collection of the number of desired data points. All ten participants completed the study, producing copious data. Since the scope of this study was limited, data analysis processes were confined to the investigation of the small sample of three.

I selected three German girls as my focus participants. The three students had completed one semester of SA and were the only one-semester students who had completed six interviews (rather than five). All following descriptions will be in regard to these three participants only. To protect the participants’ and their associates’ identities, I have used pseudonyms. The three students will be referred to as Jana, Chiara and Alia.

The three participants studied in New Zealand for 5.5 months between late January and early July 2014 (Table 1). Chiara and Alia were recruited while they were still in Germany and expressed interest in the study via email. I met the two girls together a week after their arrival in New Zealand, explained the research in more detail, and encouraged them to clarify any points that were unclear. I conducted the first interview with them a week after this informal meeting. I met Jana for our first interview on 24 February 2014, a month into her sojourn, and a week after meeting her for the first time at her high school, as part of the recruitment process.

Besides differences in age and length of time studying English, Jana’s (17 years), Chiara’s (15 years) and Alia’s (16 years) profiles were relatively similar. They came from middle-class German speaking families with two parents and one younger sibling and lived in suburban areas of Germany. All three had travelled inside and outside Europe. Jana and Chiara had had
previous experiences with language and/or cultural exchanges (see Chapters 4, 5, and 6 for more detail).

Table 1: Participant profiles

<table>
<thead>
<tr>
<th></th>
<th>Jana</th>
<th>Chiara</th>
<th>Alia</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Germany</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age during sojourn</td>
<td>17/18</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>School Year in Germany</td>
<td>11</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Years of English language learning</td>
<td>10</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td><strong>New Zealand</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SA period in 2014</td>
<td>24/01-08/07</td>
<td>25/01-06/07</td>
<td>25/01-06/07</td>
</tr>
<tr>
<td>German speakers at New Zealand high school</td>
<td>16</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>School Year in New Zealand</td>
<td>12/13</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>

The participants were enrolled at a ‘Gymnasium’ school in Germany. Jana was in Year 11, Alia in Year 10 and Chiara in Year 9. Gymnasium is one of three types of high schools in Germany (the others being ‘Hauptschule’ or ‘Realschule’) and the only type that allows students direct admission to universities upon successful completion. Gymnasia therefore have a strong academic focus.

Jana, Chiara and Alia used German agencies to set up their SA. They selected a New Zealand state high school as their host institution. Alia and Chiara attended the same high school (which will be referred to as high school 1; while Jana’s high school will be referred to as high school 2). The two high schools were near the sea and a city. They were also similar with regard to the fees that they charged (roughly 13,000 Euro), the subject choices and the extra-curricular activities offered. Both high schools required the students to wear a uniform. 15 German-speaking international students attended high school 1, and 16 German-speaking international students attended high school 2.

The three students were placed in three different school years based on their age and time table management issues (Table 1): Jana took both Year 12 and Year 13 classes, Alia was placed in Year 12, and Chiara in Year 11. Both high schools required the students to select six subjects. Some subjects were compulsory and others could be chosen by the students, depending on their year level. Chiara was required to study English, Mathematics and Science. Jana and Alia were required to study English and Maths. High schools 1 and 2 followed different policies regarding English classes for international students after assessing their English on the first school day. High school 1 placed high achievers in Mainstream English classes, and students with lower L2 proficiency levels in ESOL classes. Chiara and Alia were placed in Mainstream English
classes. At high school 2, international students with good English proficiency levels could choose between mainstream English classes and EAP classes. Jana chose EAP classes. At both schools, the number of international students that attended the same class depended on the students’ subject choices. International students gathered in EAP and ESOL classes and in ‘fun’ subjects (e.g. PE, Drama, and Photography), see Chapters 4, 5, and 6 for more details. At high school 1, some classes were specifically designed for international students and not attended by domestic students (e.g. PE).

Host family placements were organised by the New Zealand high schools. Home-stay coordinators at high schools undergo a process of selecting suitable local families through interviews and home visitations. The international students are then matched with families based on their mutual interests (as stated in the enrolment process) and availability. Host families receive a weekly fee of approximately $260 NZ in return for providing the students with accommodation and ‘full board’ (i.e. breakfast, lunch and dinner). Host families are supposed to enforce the rules of the high schools (e.g. a nightly curfew and rules regarding sleep-overs).

3.4 Design

A case-study approach was adopted to provide detailed accounts of the three German students’ unique motivational and linguistic processes during their high school SA experiences in New Zealand. Data was collected over a period of 5.5 months – from the onset to the end of the participants’ SA.

Three data collection instruments were used to elicit data: monthly semi-structured interviews to obtain English oral language samples; and weekly diary entries and monthly reflective reports to obtain written self-reports. The students showed different levels of reliability and diligence when it came to submitting their diary entries and reflective reports. None of the participants missed an interview.

Quantitative methods and a descriptive approach were employed to analyse the participants’ oral language samples. Multiple measures of language complexity, accuracy, and fluency were used for this purpose. The students’ interview data, weekly online learner diaries, and monthly online learner reports were scrutinized for the students’ behaviour and experiences abroad, using qualitative analysis methods (content analysis) and an interpretative approach.
This study is exploratory in nature and the findings do not allow for broad generalisations. However, the students’ experiences were also characterized by commonalities, which might provide more generalizable results, at least for German high school exchange students in New Zealand.

In the following sections, I will describe in detail the data collection instruments, and the quantitative and qualitative data analysis methods that were used in this study.

3.5 Data collection

3.5.1 Background questions

I used a specific set of questions about the participants’ biographical and learning background during the first interview in addition to the general interview template (e.g. questions about their age, their family situation in Germany, their language learning background etc.) (Appendix 2). These questions were asked not only for the purposes of acquiring valuable information (a questionnaire could have been sent out for that) but also to build rapport and to give the students the opportunity to warm up to the interview setting (e.g. the use of a microphone).

3.5.2 Semi-structured interviews

It is well established that the relationship between researchers and learners, “can have a profound effect on the language produced by the learners” (Ellis & Barkhuizen, 2005, p. 25). Building close rapport with the participants was of central importance to this study, not only to ensure continued participation, but also to encourage them to “talk volubly” (Milroy, 1987, p. 48). Two factors were essential to ensuring the success of the interviews: having things in common with the participants, and establishing “exchange relationships” (Milroy, 1987, p. 48).

Previous studies have shown that having things in common with participants can help the researcher better understand the students’ perspectives and needs, as well as create a sense of trust in the researcher. These factors can enhance detailed and open interactions (Labov, 1984; Milroy, 1987; Morita, 2004). I and the three participants were the same gender, had a similar linguistic and cultural background (Swiss-German and German origins), and had studied (or were studying) in New Zealand during SA. Although I am older and held a position of power in relation to the participants, these differences seemed to be of little concern to the participants.
who reported that they considered me like a ‘friendly senior’ or ‘big sister’. These commonalities assisted in building trust and rapport.

The second aspect for creating successful interviews, the provision of “tokens of exchange” (Milroy, 1987, p. 49), had to be carefully planned. Milroy (1987) claimed that through the exchange of goods and services “free of charge”, obligations are created and “applied to influence an individual’s behaviour. If the individual wishes to protect social relationships, these constant obligations must be honoured” (p. 49). Following Milroy’s (1987) advice, I made what I perceived as “tokens of exchange equal in value to those [I] wished the informant to give” (p. 49). These included ‘material’ tokens: a ‘koha’, or gift, of $50 NZ per month upon completion of the study requirements, and a printed journal of the students’ blog entries at the end of their sojourn. I provided transport to and from the interviews. I bought coffee and snacks during the interviews. I included ‘symbolic’ tokens by offering compliments, expressions of gratitude, sympathy and fostering genuine interest in the participants’ experiences.

This study collected spontaneous oral speech data both to examine changes in the second language (L2) learners’ speech performances and to collect information about the students’ language learning motivation and experiences. For purposes of the latter, interviews constituted a superior instrument to questionnaires (Dörnyei & Ushioda, 2011a), particularly because questionnaires do not accept superficial or ambiguous responses as data (Ratner, 2002). Semi-structured interviews were also considered superior to other types of language elicitation techniques for a successful examination of language development.

Tasks and standardized interviews have been the standard methods for language elicitation of SA participants because they elicit ‘tidier’ data than less structured interviews do. Interviews are often shunned for social, practical and theoretical reasons. If ‘rich’ data is to be collected from ‘strangers’, successful interviews depend on significant personal involvement between researcher and participants and the maintenance of good personal rapport. Practically, interview data can be strenuous to analyse due to different tone qualities and interruptions. Theoretically, interviews have been criticized for not producing valid results (given the tendency for topic variation etc.). Wood (2012), for example, stated that picture and film retellings produced generally “clearer results” (p. 103) compared to results obtained in interviews or discussions.

However, while interview data is messy, it is also “exciting” (Haastrup and Phillipson in Faerch & Kasper, 1983, p. 142) as well as more reliable (if the goal is to measure authentic and
spontaneous language). Compared to tasks and standardized interviews, data elicited in semi-structured interviews reflects what speakers do while they are abroad: they share experiences with each other in a friendly way (Brown & Yule, 1983; Collentine, 2004; Grieve, 2010).

Semi-structured interviews were considered the most appropriate method for eliciting unrehearsed speech. Like conversations, which are considered “the basic form of human interaction”, the interviews used in this study were believed to be an “appropriate and egalitarian vehicle for the display of oral proficiency” (van Lier, 1989, p. 495). The interviews were characterised by features typical of conversations, such as “face-to-face interaction, unplannedness (locally assembled), unpredictability of sequence and outcome, potentially equal distribution of rights and duties in talk, and manifestation of features of reactive and mutual contingency” (van Lier, 1989, p. 495). They allowed the learners to produce language under naturalistic conditions where “all aspects of the linguistic production process . . . [were], as far as possible, fully under the control of the learner” (Schmid et al., 2011, p. 39). By using semi-structured interviews, I had control over the topics, which made comparisons between individuals possible. What distinguished these interviews most significantly from naturally occurring conversations was the level of asymmetry in the exchanges (van Lier, 1989). That is, my interest in obtaining comprehensive data from the participants resulted in me largely controlling the interviews according to that plan.

Van Dijk, Verspoor, and Lowie Wander (2011) suggested collecting dense, longitudinal, and individual data was critical for a successful examination of language development. In this study, oral language data was collected six times over a period of 5.5 months, implementing a time-series design “frequent enough to capture the relevant properties underlying the developmental process” (Larsen-Freeman, 2006a, p. 595) for a dynamic description of the learners’ language development. A large collection of measurement points over long periods of time is recommended to study development of this kind. Large numbers of data points are often impractical, if not impossible, and very few studies in the field of applied linguistics can draw from data collected at shorter intervals (e.g. Polat & Kim, 2014; Spoelman & Verspoor, 2010; Verspoor, Lowie Wander, & van Dijk, 2008). Six data collection points were sufficient for a broad examination of the learners’ dynamic language development. The aim was to have regular intervals of approximately one month between each interview session, but achieving this depended on the students’ abilities and willingness. The interviews were audio-recorded with a high-quality Sony recorder (ICD-AX412F), and lasted between 32 minutes and 1 hour 25 minutes (Table 2).
Table 2: Dates and lengths of the semi-structured interviews (item numbers in brackets)

<table>
<thead>
<tr>
<th>2014</th>
<th>February</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
<th>July</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jana</td>
<td>24 (1)</td>
<td>20 (2)</td>
<td>14 (3)</td>
<td>14 (4)</td>
<td>09 (5)</td>
<td>02 (6)</td>
</tr>
<tr>
<td></td>
<td>01:02:54</td>
<td>01:00:48</td>
<td>00:44:21</td>
<td>01:02:48</td>
<td>00:53:44</td>
<td>01:16:02</td>
</tr>
<tr>
<td>Chiara</td>
<td>12 (1)</td>
<td>12 (2)</td>
<td>09 (3)</td>
<td>09 (4)</td>
<td>02 (5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>00:54:45</td>
<td>00:43:21</td>
<td>00:50:45</td>
<td>00:38:37</td>
<td>00:32:40</td>
<td>01:13:19</td>
</tr>
<tr>
<td></td>
<td>30 (6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alia</td>
<td>12 (1)</td>
<td>13 (2)</td>
<td>15 (3)</td>
<td>15 (4)</td>
<td>12 (5)</td>
<td>01 (6)</td>
</tr>
<tr>
<td></td>
<td>01:07:46</td>
<td>01:08:05</td>
<td>00:58:12</td>
<td>01:01:54</td>
<td>00:48:01</td>
<td>01:25:52</td>
</tr>
</tbody>
</table>

All interviews were conducted in English. The students were asked questions about their lives, their attitudes, and their language use, as well as the members of their social networks. These questions stimulated their interest and developed conversations (Labov, 1984). My interview template was inspired by Isabelli-Garcia’s (2006) and Kim’s (2006) sample interview questions. Their studies also concerned language learning motivation and language learning opportunities, and monthly interviews over the period of six and ten months, respectively.

Three broad types of questions were used in the semi-structured interviews: general guiding questions, prepared follow-up questions, and ad hoc probing questions. The guiding questions were posed during each interview (Appendix 2). They were open-ended and general in nature. They required the students to reflect on their study-abroad experiences, specifically with regard to the New Zealand culture (e.g. ‘is there anything in particular that you liked or didn’t like about the New Zealand culture?’), their social networks (e.g. ‘who are you spending your break times with?’), their hobbies (e.g. ‘how is football practice going?’), their English language learning progress (e.g. ‘how is it going with your English?’), and perceptions about themselves (e.g. ‘how do you feel about yourself here in New Zealand?’) - without specifically asking about their motivation.

The second type of questions, the follow-up questions, changed with each interview as they were formulated on the basis of the students’ diary entries and reflective reports (e.g. ‘tell me more about your trip to Australia’). Their purpose was to elicit discussion of experiences the participants had described in their self-reports. Thirdly, when I felt a need for more insight during the interviews I asked ad hoc probing questions, such as ‘why do you think that happened?’. According to Ratner (2002), follow-up questions of this kind can penetrate beneath the superficial responses and help “comprehend true motives, perceptions, attitudes, emotions and personality traits” (p. 146).
As suggested by Labov (1984), the interviews were hierarchical both as a whole and with regard to the different topic ‘modules’. Each interview started with the very general question ‘has anything special happened in the past month?’ to give the students the opportunity to speak freely, and to emphasize the importance of the conversational nature of the interviews. Subsequently, more detailed questions were posed. The order of the interview template questions was not strictly adhered to but was adapted to the course that each interview took. The extent to which the students answered the questions varied with each interview, depending on their needs and interests.

3.5.3 Online diaries

Learner diaries allow researchers to gain access to the students’ lives in a way that might be otherwise impossible (Dörnyei, 2007). I considered personal records of their daily lives appropriate for investigating their dynamic learning experiences.

The students were given guideline questions, developed to elicit language-related events experienced, other events, differences between Germany and New Zealand, and how they felt about their experiences. The website wordpress (http://wordpress.com) was ideal for the journals. The participants’ wordpress accounts were set up in a way that the blogging instructions were displayed in a side bar, making the writing task more user-friendly (Appendix 3). The students were asked to write diary entries at least once a week. They were permitted to write the entries in English or German. The three students chose to write all diary entries in English. Jana and Chiara wrote a total of 15 diary entries, and Alia wrote 14 entries (Table 3).

The quality of the diary entries varied considerably. Some diary entries provided an elaborate account of the participants’ experiences, including the ways they felt during a particular event. Other entries were brief and factual and did not generate rich data. However, as mentioned above, in case of the latter, the interviews afforded opportunities for me to probe their entries more deeply.
### Table 3: Submission dates and length of online diary entries (item numbers in brackets)

<table>
<thead>
<tr>
<th></th>
<th>2014</th>
<th>February</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
<th>July</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Jana</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>02 (1)</td>
<td>01 (5)</td>
<td>01 (9)</td>
<td>01 (11)</td>
<td>12 (14)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>385 w.</td>
<td>412 w.</td>
<td>371 w.</td>
<td>352 w.</td>
<td>355 w.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>09 (2)</td>
<td>07 (6)</td>
<td>18 (10)</td>
<td>11 (12)</td>
<td>16 (15)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>391 w.</td>
<td>329 w.</td>
<td>388 w.</td>
<td>372 w.</td>
<td>374 w.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>16 (3)</td>
<td>16 (7)</td>
<td></td>
<td>22 (13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>309 w.</td>
<td>387 w.</td>
<td></td>
<td></td>
<td>431 w.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>25 (4)</td>
<td>22 (8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>384 w.</td>
<td>399 w.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Chiara</strong></td>
<td></td>
<td>02 (1)</td>
<td>18 (5)</td>
<td>06 (10)</td>
<td>18 (12)</td>
<td>22 (14)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>176 w.</td>
<td>181 w.</td>
<td>315 w.</td>
<td>191 w.</td>
<td>237 w.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>14 (2)</td>
<td>23 (6)</td>
<td>15 (11)</td>
<td>25 (13)</td>
<td>29 (15)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>118 w.</td>
<td>259 w.</td>
<td>280 w.</td>
<td>234 w.</td>
<td>90 w.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>23 (3)</td>
<td>23 (7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>92 w.</td>
<td>384 w.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>23 (4)</td>
<td>23 (8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>254 w.</td>
<td>155 w.</td>
<td>30 (9)</td>
<td>251 w.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Alia</strong></td>
<td></td>
<td>01 (1)</td>
<td>02 (4)</td>
<td>08 (6)</td>
<td>11 (7)</td>
<td>02 (10)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>285 w.</td>
<td>439 w.</td>
<td>165 w.</td>
<td>311 w.</td>
<td>555 w.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>09 (2)</td>
<td>11 (5)</td>
<td>17 (8)</td>
<td>09 (11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>320 w.</td>
<td>423 w.</td>
<td>317 w.</td>
<td>192 w.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>16 (3)</td>
<td>463 w.</td>
<td>29 (9)</td>
<td>18 (12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>159 w.</td>
<td></td>
<td></td>
<td>169 w.</td>
<td>22 (13)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>375 w.</td>
<td>27 (14)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>294 w.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

w.: words

### 3.5.4 Reflective reports

It was originally my aim to also collect written speech samples. So I asked the students to write monthly reflective reports about their perceived second language (L2) and cultural development during the previous month. The students were asked to write a minimum of 500 words per report and they were instructed to reflect on both their L2 performance and their experiences with the host culture (Appendix 4). Because these reports were intended to serve as language assessment data, they had to be written in English. As with the journal entries, reminders were often necessary as a result of the students’ forgetfulness or laziness. The three learners submitted all required reports and their lengths varied between 353 and 1301 words (Table 4).

There was too much data generated to allow for the analysis of both oral and written language development in this study. However, the monthly reports provided relevant additional insights into the students perceived language progress, their attitudes, and experiences in New Zealand.
Table 4: Submission dates and length of the monthly diaries (item numbers in bracket)

<table>
<thead>
<tr>
<th></th>
<th>2014</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>March</td>
<td>April</td>
<td>May</td>
<td>June</td>
<td>July</td>
<td>August</td>
<td>Sept.</td>
</tr>
<tr>
<td>Jana</td>
<td>20 (1)</td>
<td>15 (2)</td>
<td>20 (3)</td>
<td>15 (4)</td>
<td>16 (5)</td>
<td>15 (6)</td>
<td>15 (6)</td>
</tr>
<tr>
<td></td>
<td>556 w.</td>
<td>468 w.</td>
<td>668 w.</td>
<td>616 w.</td>
<td>654 w.</td>
<td>489 w.</td>
<td></td>
</tr>
<tr>
<td>Chiara</td>
<td>18 (1)</td>
<td>14 (2)</td>
<td>08 (3)</td>
<td>02 (4)</td>
<td>07 (5)</td>
<td>24 (6)</td>
<td>379 w.</td>
</tr>
<tr>
<td></td>
<td>564 w.</td>
<td>528 w.</td>
<td>516 w.</td>
<td>368 w.</td>
<td>475 w.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alia</td>
<td>14 (1)</td>
<td>13 (2)</td>
<td>11 (3)</td>
<td>10 (4)</td>
<td>353 w.</td>
<td>31 (6)</td>
<td>880 w.</td>
</tr>
<tr>
<td></td>
<td>822 w.</td>
<td>817 w.</td>
<td>1301 w.</td>
<td>583 w.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

w.: words

3.6 Quantitative data analysis

The learners’ oral language samples were measured quantitatively. I used multiple measures of complexity, accuracy, and fluency (CAF). The procedures resulted in numerical data. Before analysis, the learners’ data had to be transcribed, segmented into units, and coded. These three processes will be described below.

3.6.1 Transcription

The audio-recorded interviews were transcribed using the transcription software ‘InqScribe’ ("InqScribe"). Each speech sample was transcribed in full. However, only a portion of each sample was used for the linguistic analysis of the learners’ L2 performance. The interviews were transcribed orthographically and verbatim. All participant repetitions, reformulations, filled and silent pauses were noted. Features not noted included intonation and stress patterns, laughter, coughing and other non-verbal sounds or background noises, and the researchers’ back channelling sounds that were uttered during the speech of the participants (e.g. ‘oh’, ‘really?’) unless they triggered a response in the participants (as minimal as ‘yeah’). Only the participants’ turns were carefully transcribed. Non-lexical filled pauses were restricted to three realizations in the data (‘ehm’, ‘eh’, ‘hmm’) for reasons of consistency (Retherford, 2000).

The interview samples were formatted based on the CHAT transcription system outlined in MacWhinney’s (2013) manual of the CHILDES project. This procedure is common for researchers of child L1 acquisition, and has been adopted by researchers of the dynamic systems paradigm (e.g. Schmid et al., 2011). Adhering to the CHAT transcription standards is indispensable as only properly formatted transcripts can be read by the CLAN software – the language analysis tool of the CHILDES project which was used in this study. Table 5 shows the transcription conventions that were used in this study.
Two different transcription procedures were adopted for fluency and complexity word counts. Unpruned versions of the transcript were used for fluency and pruned versions for complexity. It was important for this study to avoid counting multi-word compounds (e.g. ‘New Zealand’) as one word as that would have affected the fluency measures. However, these multi-word compounds were counted as single units for analysing complexity to recognise their semantic integrity and to represent the learners’ lexical diversity reliably. It was also important to exclude non-lexical fillers (e.g. ‘ehm’) and repaired material from the complexity count, as suggested by Foster, Tonkyn, and Wigglesworth (2000).

Word count decisions also had to be made for verb-subject contractions, semi-auxiliaries, and alphabetisms (Brown, 1973; Götz, 2013; Retherford, 2000). Calculating word counts of these types for language analysis varies among researchers. Justifications for the different counts are typically founded on how researchers believe the features are represented in speakers’ memory (i.e. as unanalysed chunks or creatively assembled units). For example, Biber (1999) believes that “[i]t is likely that [subject-predicate] contractions are processed by the speaker and hearer as single words, and therefore, for the purposes of studying dysfluency phenomena, that they should be treated as such” (p. 1061). The count of compound nouns was particularly difficult in this study. Biber (1999) stated that "practice varies as to whether to represent a compound as two orthographic words, one unbroken orthographic word, or a hyphenated word. Partly this is because there is no clear dividing line between compounds and free combinations" (p. 326).

I adhered to the rules displayed in Table 6 for the word counts in my study following multiple sources, my own decisions, and the CHAT transcription standards:

Table 5: Transcription conventions

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>(.3)</td>
<td>pauses in milliseconds</td>
</tr>
<tr>
<td>xxx</td>
<td>unidentifiable speech</td>
</tr>
<tr>
<td>zehn@g</td>
<td>German words</td>
</tr>
<tr>
<td>A@1 S@1 B@1</td>
<td>spelt letters</td>
</tr>
<tr>
<td>[?]</td>
<td>best guess</td>
</tr>
</tbody>
</table>

Two different transcription procedures were adopted for fluency and complexity word counts. Unpruned versions of the transcript were used for fluency and pruned versions for complexity.
<table>
<thead>
<tr>
<th>Feature</th>
<th>Sources and/or applied principles</th>
<th>Examples for unpruned version (fluency)</th>
<th>Examples for pruned version (complexity)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alphabetisms</td>
<td>(Goetz, 2013).</td>
<td>PE (2 words)</td>
<td>PE (1 word)</td>
</tr>
<tr>
<td>Subject-verb / verb-negation contractions</td>
<td>(Rhetherford, 2000)</td>
<td>it’s / don’t (1 word)</td>
<td>it’s / don’t (1 word)</td>
</tr>
<tr>
<td>Semi-auxiliaries</td>
<td>(Brown, 1973, Retherford, 2000).</td>
<td>wanna / gonna (1 word)</td>
<td>wanna / gonna (1 word)</td>
</tr>
<tr>
<td>Non-lexical fillers</td>
<td>(Retherford, 2000)</td>
<td>&amp;hmm /&amp;ehm / &amp;eh (1 word each)</td>
<td>&amp;hmm /&amp;ehm / &amp;eh (not counted)</td>
</tr>
<tr>
<td>Repair phenomena</td>
<td>various sources</td>
<td>normally I we &amp;eh go home by bus (8 words)</td>
<td>normally &lt;I&gt; [///] we &amp;eh go home by bus (6 words)</td>
</tr>
<tr>
<td>Proper multi-word nouns traditionally spelt as two orthography words / hyphenated compounds</td>
<td>Written and counted as separate words for the fluency counts</td>
<td>New Zealand / self confident (2 words)</td>
<td>New_Zealand / self_confident (1 word)</td>
</tr>
<tr>
<td>Open compounds</td>
<td>Written and counted as separate words for the fluency and the complexity counts (the combinations were less restricted and more interchangeable than the other types of compound nouns)</td>
<td>credit card / host dad (2 words)</td>
<td>credit card / host dad (2 words)</td>
</tr>
<tr>
<td>Closed compounds</td>
<td>Written and counted as one word for the fluency and the complexity counts</td>
<td>airport (1 word)</td>
<td>airport (1 word)</td>
</tr>
<tr>
<td>Numbers</td>
<td>Written and counted as separate words for the fluency and the complexity counts</td>
<td>fifty seven (2 words)</td>
<td>fifty seven (2 words)</td>
</tr>
</tbody>
</table>

Note that words in this study were defined as types and tokens, not lemmas. Hence different inflections of a word (e.g. *student*, *students*) were counted as different word types. This approach was appropriate given that English “contains only a minimum of inflections compared to other languages, e.g. French or Czech, where using lemmas would be more appropriate” (Siskova, 2012, p. 27).

The following two examples in Table 7 from Alia’s data show a pruned and an unpruned version of the same speech sample:
Table 7: Unpruned and pruned versions of the same speech sample

<table>
<thead>
<tr>
<th>Unpruned version</th>
<th>Pruned version</th>
</tr>
</thead>
<tbody>
<tr>
<td>(.97) and &amp;ehm (1.59) and then James gave me &amp;ehm because I said on Friday :: that I really want like a singlet with New Zealand</td>
<td>(.88) because it it wou it would make me so proud :: (.78) to wear it in Germany</td>
</tr>
<tr>
<td>(.97) and [///] &amp;ehm (1.59) and then &lt;James gave me&gt; [///] &amp;ehm because I said on Friday :: that [///] that I really want like a singlet with New_Zealand</td>
<td>(.88) because it [///] &lt;it wou&gt; [///] it would make me so proud :: (.78) to wear it in Germany</td>
</tr>
</tbody>
</table>

Once the audio-recorded interviews were transcribed, they had to be coded (see under each construct how I coded the measures in the transcripts). To avoid the danger of mistyping a code, and to make the coding process faster, I created in Word a macro for each fluency, complexity, and accuracy code used in this study (Verspoor et al., 2011). After coding, CLAN software was used to perform frequency counts on the CAF-coded documents. All 100 AS-units that were used for analysis from each interview were copied from a Microsoft Word document (saved as plain text) into a new file of the CLAN editor. Files in CLAN were then checked for accuracy (the transcripts must follow the standards of the CHAT transcription format) and adjusted where necessary. Each measure then underwent a frequency count. The codes that were used to mark fluency and complexity features are described in the respective CAF sections. An example of a transcript including all codes appears in Appendix 5.

3.6.2 AS-units

To compare results from one speech sample with another, “a principled way of dividing transcribed data into units” (Foster et al., 2000, p. 354) must be used. I segmented data using “AS-units” (Foster et al., 2000). The AS-unit is widely used to segment oral data into units and recommended as a valid unit of analysis for spoken language (see Ellis & Barkhuizen, 2005; Foster et al., 2000).

The AS-unit is mainly a syntactic unit. This makes spoken language more easily analysable than with intonational and semantic units. AS-units may incorporate multiple clauses and may thus reflect units of planning. Foster et al. (2000) defined the AS unit as “a single speaker’s utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clause(s) associated with either” (emphasis in original, p. 365). The definitions of these clause types are shown in Table 8 and are supplemented with my own examples. Note that the symbol ‘::’ is used to separate clauses from each other.
Table 8: Clause type definitions and examples

<table>
<thead>
<tr>
<th>Clause type</th>
<th>Definition by Foster et al. (2000)</th>
<th>Examples from this study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent clause</td>
<td>A clause that minimally includes a finite verb</td>
<td>it’s a little bit yeah (.97) difficult sometimes with my homestay</td>
</tr>
<tr>
<td>Independent sub-clausal unit</td>
<td>A unit that consists of “either one or more phrases which can be elaborated to a full clause by means of recovery of ellipted elements from the context of the discourse or situation . . . or a minor utterance, which will be defined as one of the class of ‘Irregular sentences’ or ‘Nonsentences’”</td>
<td>and on Wednesday ehm (.34) just a short training. urgh (.76) oh my god (.70) yeah.</td>
</tr>
<tr>
<td>Subordinate clause</td>
<td>A clause that has at least a finite or non-finite verb element and minimally one more clause element. Exception: non-finite verbs that follow modal verbs or that are part of complex verb phrases</td>
<td>and (.50) I didn’t ate anything :: because I had didn’t had like ehm (.47) dessert, ehm I can do some of my competitions. I don’t know why. (1 AS-unit, 1 clause)</td>
</tr>
</tbody>
</table>

**INTERCODER RELIABILITY**

Based on Levkina and Gilabert (2012, p. 183), 10% of the data was double-coded by an inter-rater for AS-unit segmentations. The inter-rater was a PhD student in the field of applied linguistics. The interrater worked with the interview transcriptions, not the audio-files. When intonation patterns were used for analysis procedures, only my interpretations counted. One training session with in-depth discussion of the segmentation rules was carried out. In two subsequent steps, the rater segmented an equivalent of 20 AS-units for further training purposes. The results were discussed and discrepancies accounted for. The interrater was then given the equivalent of 10 AS-units per interview per participant (as interpreted by me), which made a total of 60 AS-units per participant. Intercoder reliability figures were calculated as simple percent agreement (i.e. total coding decisions minus discrepancies divided by the total coding decisions). The interrater and I came to 92% agreement on the analysis of the data (Table 9).

Table 9: Intercoder agreement

<table>
<thead>
<tr>
<th>Segmenting AS-units</th>
<th>Intercoder reliability</th>
<th>Double-coded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jana</td>
<td>93%</td>
<td>10% (60 AS-units)</td>
</tr>
<tr>
<td>Chiara</td>
<td>98%</td>
<td>10% (60 AS-units)</td>
</tr>
<tr>
<td>Alia</td>
<td>86%</td>
<td>10% (60 AS-units)</td>
</tr>
<tr>
<td>average</td>
<td>92%</td>
<td></td>
</tr>
</tbody>
</table>
3.6.3 Speech samples for analysis

3.6.3.1 General approach

Unless researchers conduct a full analysis of their data, they must determine the nature and size of the speech samples to analyse. I made three main choices for selecting data for analysis regarding: 1) the amount of data to be analysed (i.e. the number of AS-units); 2) the positions of the selected AS-units to be analysed from the larger set of the data, and 3) whether or not I needed to exclude data from the selected passages to ensure a coherent analysis.

Regarding the amount of data, analysing 100 utterances is a common and accepted procedure to obtain a representative sample of speech (Brown & Yule, 1983; Retherford, 2000; van Geert & van Dijk, 2002). Retherford (2000) explained that although more utterances would produce more representative samples, “in an effort to be realistic, practical, and efficient, 100 utterances” (p. 11) are sufficient.

No established guidelines exist regarding the choice of the position of the selected AS-units. In first and second language research, detailed explanations of sample size are rare and standard procedures do not appear to have been established. It appears that the choice of sample position is dependent on the data at hand and the purposes of the research. It seems that unless researchers analyse the first 100 units of their data (van Geert & van Dijk, 2002), or 100 consecutive units after “the first page of transcription” (Brown, 1973, p. 54) or the first ten minutes from the beginning of the recording (Oliveira, 2012), the position of the language sample to be analysed in interactive language performance research is determined on the basis of the discourse topic (e.g., the number of AS-units per topic, as in Tonkyn, 2012) or turn length (e.g. turns of at least 100 words, as in Polat and Kim, 2014).

Decisions on sample size in my own study were influenced by the setting for the data collection (cafes) and the speech elicitation technique (semi-structured interviews). I decided to analyse 100 AS-units rather than speech of a certain period of time. Analysing units during the first 30 minutes of speech, for example, would have resulted in an enormous variation in the number of AS-units to be analysed, not only from one participant’s interview to the next, but also between participants. Secondly, although the amount of data gathered with the shortest interview (32 minutes) would have allowed me to analyse more than 200 AS-units per interview, I limited the number of AS-units to 100, since this appeared to be the most appropriate way to successfully manage my data and keep the size representative.
For developmental data to be comparable it must share critical features (Götz, 2013). This made selection of the position of the 100 utterances difficult. Oliveira (2012) found that prompted narrative differs from unprompted narrative in terms of length and elaboration (the former are short, the latter elaborate). Polat and Kim (2014) argued that longer (uninterrupted) passages may show different levels of accuracy and complexity than shorter passages. Brown and Yule (1983) furthermore argued that “what is required of a speaker in a long turn is considerably more demanding than what is required of a speaker in a short turn” (p. 17).

The data collected for this study consisted of both long uninterrupted turns and short, interactive turns. Longer turns were produced when the participants felt that their story was “worthy of being heard” (Oliveira, 2012, p. 628). I decided to adopt an approach inspired both by Oliveira (2012) and Polat and Kim (2014). Basing selection on topics like Tonkyn (2012) was rejected not only because Polat and Kim (2014) demonstrated that topic does not necessarily have an impact on speech performance, but also because the topics that were addressed in this study often blended in with each other so that taking them apart would have been an impossible task. In this study, I selected from each interview transcript a number of participant turns of at least 50 uninterrupted words (including non-lexical fillers) – and at most as many words as they produced in one turn - there was no maximum upper limit to the word count. I collected as many consecutive 50+word turns that were required to obtain a corpus of 100 AS-units starting from the beginning of each interview. If 50+word turns contained whole utterances that were produced in German, these turns were not considered for analysis. Turns with isolated German words were not excluded from analysis. If the participants were interrupted (either by me or waiting staff) after they had uttered a minimum of 50 words but before they had finished their turn, the words after the interrupted utterance were eliminated from analysis, unless the utterances contained more than 50 words, in which case they constituted a new turn. Uninterrupted speech is defined in this study as speech that is in no obvious way influenced by the interviewer’s participation in the conversation. Back channelling comments (e.g. ‘yeah’, ‘really’) were not considered interruptive unless they evoked a response in the participant (as short as ‘yeah’).

The background questions of the first interview (see 3.5.1) were not considered for the selected language samples. The maximum time used to answer these questions was twelve minutes. To make comparisons of speech samples meaningful, I eliminated the first twelve minutes (plus my next question) of each participants’ first interview for data analysis processes.
I chose 50+word turns for three reasons. Examining the data showed that real elaboration of a topic tended to occur at the 50-word boundary. All three participants produced enough 50+word passages to obtain 100 AS-units. Choosing long turns for analysis allowed me to compare the data in a meaningful way.

Last, a Level Three analysis was applied to the 50+word turns to exclude “uncharacteristically short units” (Foster et al., 2000, p. 371). According to Foster et al. (2000), a Level Three analysis is to be used “where analysis of non-fragmentary AS-units is required” and recommended for researchers who are interested in studying “what the performer can do in the production of relatively ‘complete’ units” (p. 371). The following items, which sometimes occurred at the beginning of a 50+word turn, were eliminated from the data in accordance with a Level Three analysis (examples adapted from Foster et al., 2000, p. 371):

- One-word minor utterances (e.g. ‘ok’)
- Verbatim echo responses (A: I think two years B: two years).
- V-less elliptical units involving ellipsis of elements of the interlocutor’s speech (A: What is your mother tongue then? B: Arabic Arabic)
- Units involving “substitution of clause, predicate, or predication level units of interlocutor’s speech” (A: do you think that…. B: yes I think so).

In cases where the conjunctions ‘but’, ‘or’, and ‘because’ (Example 1) linked the short initial units with the subsequent more complete ones, they were also eliminated from the analysis unless these conjunctions were not connected to the researcher’s questions, but to the participants’ previous turn – for example, when the researcher interrupted them (Example 2, the underlined words were eliminated from the data set). The conjunction ‘and’ in this position was not eliminated from the data as it was unclear whether the conjunction was dependent on the short previous answer or simply reflected strung together thoughts typical of unplanned speech (Example 3) (Huddleston & Pullum, 2002, 87a). I analysed all data beginning with the first ‘complete’ unit that followed these short responses.

Example 1:
Luzia: … you don’t have the sweets anymore afterwards in your room?
Chiara: I try but it's really hard if you're hungry in your room...

Example 2
Chiara: it's just stupid I don't know
Luzia: ok that's hard yep
Chiara: yes because I was like ‘ok when they think I use too much internet…

Example 3
Jana: no no no I don't know I think so yeah and then yeah I think
Luzia: let’s clarify that later
Jana: yeah and then ehm I don't know I was just I saw him but he didn't say hi to me…

3.6.3.2 Addressing specific issues

The procedures outlined by Foster et al. (2000) often lack precision and do not include procedures for all speech features. I used other sources and/or devised my own procedures to fill gaps. Table 10 lists how I handled clause coordination and subordination, topicalisation, asides, and discourse/pragmatic markers. My decisions were based on a number of sources (indicated in brackets). Where I had no guidance, I developed a procedure that best ensured consistency. The treatment of direct speech complements and discourse/pragmatic markers is discussed separately below.

Table 10: Handling clause problematic features of oral productions

<table>
<thead>
<tr>
<th>Clause coordination</th>
<th>Definition</th>
<th>Handled</th>
<th>Examples from this study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clause coordination</td>
<td>Coordinated clauses with their own subjects and finite verb phrases</td>
<td>Treated as separate units.</td>
<td>she works the whole day</td>
</tr>
<tr>
<td></td>
<td>(Foster et al., 2000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>In the coordinated clause, the (identical) subject or subject + auxiliary/modal are dropped, and the coordinated clause is not divided from the first clause by a pause of more than 0.5 seconds (Foster et al., 2000)</td>
<td>Treated as coordinate clauses</td>
<td>it’s just you sit there :: and watch. (1 AS-unit, 2 clauses)</td>
</tr>
<tr>
<td></td>
<td>Postposing coordinates (a non-adjacent element of the previous clause. (Huddleston and Pullum, 2002)</td>
<td>Treated as belonging to previous clause</td>
<td>and so you (0.35) can’t ehm (0.92) ehm (0.36) talk with them :: or go going to the next room. (1 AS-unit, 2 clauses)</td>
</tr>
<tr>
<td></td>
<td>Additional elements added to the independent clause. (Huddleston and Pullum, 2002)</td>
<td>Treated as coordinated clauses</td>
<td>she thinks (.41) ehm James is funny (.33) but not Lorenz. (1 AS-unit, 1 clause)</td>
</tr>
<tr>
<td></td>
<td>Apposition (Noun or prepositional phrases that refer to the same entity in the main clause, but are added to the matrix clause to clarify or exemplify the contents of the</td>
<td>Treated as belonging to previous clause</td>
<td>I’m better than her :: but not in every event. (2 clauses, 1 AS-unit)</td>
</tr>
</tbody>
</table>

Table 10: Handling clause problematic features of oral productions
Clausal subordination

| Coordination of two or more subordinate clauses that function as constituents of the same main clause | Treated as coordinate clauses regardless of whether they had the same subject or not. Key criterion: dependence on the same main clause | and our hostmom was very interested to know how it was and that we’re (1.2) ok after that. | (1 AS-unit, 4 clauses) and (1.33) then the other day (0.76) I woke up and was all (1.25) still relaxed. | (2 AS-units) |
| Coordinate clauses without overt conjunction | Treated as independent AS-units | he don’t get money because he don’t work. | (1 AS-unit, 2 clauses). |
| Optional because- and which-clauses within the same tone unit as at least one of the other preceding clause elements of the AS-unit. | Treated as subordinate clause | and otherwise she will send an email to us which is really ridiculous. | (2 AS-units). |
| (Foster et al., 2000) | | so my family in Germany we eat all together in the kitchen on the table. | (1 AS-unit, 1 clause) |

Optional because- and which-clauses not within the same tone unit as at least one of the other preceding clause elements of the AS-unit. (Foster et al., 2000)

Topicalization of noun phrases (Foster et al., 2000)

Treated as belonging to the same AS-unit as the following clause unless separated from it by falling intonation and a pause of at least 0.5 seconds.

Topicalization of noun phrases (Foster et al., 2000)

Treated as belonging to the same AS-unit as the following clause unless separated from it by falling intonation and a pause of at least 0.5 seconds.

Topicalization of noun phrases (Foster et al., 2000)

Treated as belonging to the same AS-unit as the following clause unless separated from it by falling intonation and a pause of at least 0.5 seconds.

Asides

brief side comments that interrupt the flow of an utterance

Treated as belonging to previous clause

ehm it w was (0.70) like (1.28) how is it called (0.41) pants and (0.63) like a jumper like a full (0.69) training suit (0.63) &ehm with (0.51) NZL on the back. | (1 AS-unit) |

| unit separator :: clause separator |

**DIRECT SPEECH**

Foster et al. (2000) do not explain how to treat dialogue complements in direct speech. As the participants of this study used these, I had to determine whether to treat dialogue complements as part of the previous unit or as a new unit. In direct speech reports, the complement takes the form of an independent clause (i.e. it could be coded as an independent clause) but performs the function of a dependent clause (i.e. it could be coded as a finite complement clause). According to Romaine and Lange (1991), in spoken language “both direct and indirect quotes
can be considered complements of the verb of saying and thus form a subordinate clause which is the direct object of the verb of saying” (p. 233). In line with Romaine and Lange (1991), as well as the rules outlined for segmenting data into c-units (“Segmenting utterances into c-units”), I treated direct speech complements as dependent finite complement clauses (Example 4). All successive independent clauses that occurred as part of direct speech quotes, however, were counted as separate units (Example 5), unless they fulfilled the criteria outlined above for coordinated clauses (Example 6).

Example 4: (0.38) and she said :: ‘oh yeah it’s not a problem’. (1 AS-unit, 2 clauses)

Example 5: (0.68) and she said :: ‘no you just have to be careful with the water | and (1.38) your showers are ok :: but Luana's are too long’. (2 AS-units, 2 clauses each)

Example 6: and I'm always thinking that :: ‘ok I'm gonna eat :: and watch tv :: and just be quiet’. (1 AS-unit, 4 clauses)

Reporting clauses that lacked a reporting verb (a ‘zero quotative’) were treated like quotative frames with a reporting verb (Example 7). Treating dialogue complements as subordinate clauses led to the unsatisfactory consequence that verb-less direct speech complements such as ‘oh’ received clausal status (which, according to Foster et al., 2000, would not fulfil the criterion of subordinate clauses (Example 8). I did not see another way of dealing with this problem. Indirect speech complements were treated in the same way as direct speech complements (Example 9).

Example 7: and then :: ‘ok it’s good’. (1 AS-unit, 2 clauses)

Example 8: and I was like :: ‘oh’. (1 AS-unit, 2 clauses).

Example 9: she thought :: he was right. (1 AS-unit, 2 clauses).

DISCOURSE AND PRAGMATIC MARKERS

Discourse and pragmatic markers (DPMs) were also not dealt with by Foster et al. (2000) and present difficulties when it comes to segmenting oral data into AS-units. Difficulties arise with DPMs that contain a finite verb (e.g. ‘I mean’, ‘you know’), which, according to Foster et al.’s (2000) principles, would be counted as separate clauses. However, DPMs do not add to the propositional content of an utterance and function structurally “outside the boundaries of the clause” (Carter & McCarthy, 2006, 90). It would be unreasonable to give them clausal status. In this study, DPMs of formulaic nature containing a finite verb were not given clausal status, in line with C-unit analysis rules (“Segmenting utterances into c-units”).

64
Table 11 lists all DPMs (containing a finite verb) that were not given clausal status in this study. Two steps were followed to arrive at this list: First, I tested whether or not a phrase was used as a DPMs or a non-formulaic clause by deleting it from the utterance. If no change to the propositional content of the utterance occurred and if the utterance was still grammatically complete, then I categorised the phrase as a DPMs (Example 10). If the phrase had a literal meaning and could not be eliminated from the utterance without making it incomplete or changing its content, then I did not categorise the phrase as a DPMs (Example 11).

Example 10: (0.54) but ehm (1.13) I don’t know it was so funny :: because eh we are we came here with the same organisation. (1 AS-unit, 2 clauses)

Example 11: but (0.38) sometimes (0.49) I don’t know really :: what to think about that :: because (0.48) she could do more for that. (1 AS-unit, 3 clauses)

Secondly, the DPMs that I identified in my data were then compared to the examples of DPMs provided in Diessel’s (2004), Carter and McCarthy’s (2006) and Hasselgren’s (2002) works.

Table 11: Discourse and pragmatic markers that were not given clausal status

<table>
<thead>
<tr>
<th>Markers</th>
<th>Examples from my data</th>
</tr>
</thead>
<tbody>
<tr>
<td>I don’t know</td>
<td>but ehm (1.13) I don’t know it was so funny because eh we are we came here with the same organisation.</td>
</tr>
<tr>
<td>I think (that)</td>
<td>and it was I think one year ago / I think it was two weeks ago.</td>
</tr>
<tr>
<td>I would say</td>
<td>ehm yeah I think some of them are I would say normal not boring but just normal.</td>
</tr>
<tr>
<td>I mean</td>
<td>I mean I I’m a person who thinks a lot.</td>
</tr>
<tr>
<td>that means</td>
<td>that means (0.45) they’re getting like the exact amount of food they need to cook some (0.37) meals.</td>
</tr>
<tr>
<td>You know</td>
<td>and I wanted to try something different than the everyday life you know.</td>
</tr>
<tr>
<td>it’s just (that)</td>
<td>and ehm (1.00) I don’t know it’s just (0.63) I have the feeling I need some friends</td>
</tr>
<tr>
<td>The thing is</td>
<td>yeah and oh yeah the thing was ehm she asked me where I live</td>
</tr>
<tr>
<td>It’s like</td>
<td>and then s in PE it’s like we are one group and I belong to that group and I was like ‘yes’.</td>
</tr>
<tr>
<td>to be honest</td>
<td>to be honest, I didn’t go to school on (0.43) Friday.</td>
</tr>
<tr>
<td>(I’m) sorry</td>
<td>and then (0.47) when’s ten ten thirty I’m really sorry but I’m too tired to doing anything else</td>
</tr>
<tr>
<td>that’s (the reason)</td>
<td>that’s why he’s always late.</td>
</tr>
<tr>
<td>why</td>
<td></td>
</tr>
</tbody>
</table>

3.6.4 CAF measures

In this study, multiple global and specific measures of complexity, accuracy and fluency (CAF) were employed to measure L2 development. Findings of previous studies regarding the difficulties inherent in the operationalisation and measurement of CAF have been taken into account to make an informed choice about the measures to be used (see Chapter 2). The CAF measures in this study constitute the dependent variables, as descriptors of L2 performance, measured to demonstrate the effect of time spent studying abroad. For the assessment of these
measures, a frequency count was performed with each component for each interview and the changing numbers were plotted on a graph to depict development.

To potentially account for the factors that might influence the development of CAF, where possible, I carried out an examination that went beyond studying numbers to considering “learner performance from a qualitative standpoint, concerned with the details of how the use of language change[d] to yield new performances” (Larsen-Freeman, 2006, p.597). Attention was also given to CAF “as a dynamic and inter-related set of constantly changing subsystems” (Housen et al., 2012a, p. 9). CAF interactions over time were examined using dynamic description (Larsen-Freeman, 2006) - a ‘mild’ version of describing dynamic language development, yet one that still honours its complex and dynamic nature. The measures used in this study will be described in the following sections.

3.6.4.1 Grammatical complexity

This study used two of the most common types of grammatical complexity measures: the length-based measure mean length of clause (measuring complexity on the sub-sentential level) and the subordination measure subordination index (measuring complexity on the sentential level). This study further examined the frequency and range of subordinate clauses for a more sophisticated analysis of grammatical complexity. Subordinate clauses were divided into finite complement, adverbial and relative clauses, and non-finite complement and adverbial clauses (in line with Schmid et al., 2012) (Table 12). Clause coordination was not examined in this study given the high initial L2 proficiency levels of the participants. Table 12 lists the measures used in this study and provides examples for the subordinate clauses. It also provides the codes used for the transcriptions. Independent clauses and independent sub-clausal units were coded as [% indcl] and [% indsc1], respectively (see examples below).
Table 12: Grammatical complexity measures used in this study

<table>
<thead>
<tr>
<th>Complexity</th>
<th>Measures and codes</th>
<th>Definition</th>
<th>Examples from my data with codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub-sentential complexity</td>
<td>Mean length of clause</td>
<td>Average number of words per (independent or subordinate) clause</td>
<td>I don't know [% indcl] :: if I mentioned her [% scl:co].</td>
</tr>
<tr>
<td></td>
<td>Subordination Index</td>
<td>Average number of (independent or subordinate) clauses per AS-unit</td>
<td></td>
</tr>
<tr>
<td>Sentential complexity</td>
<td>Finite complement clauses [% scl:co]</td>
<td>Clauses functioning as the subject or object of the matrix clause (also a clause after a preposition or a comparison clause)</td>
<td></td>
</tr>
<tr>
<td>Syntactic sophistication</td>
<td>Finite adverbial clauses [% scl:adv]</td>
<td>Clauses functioning as an adverbial</td>
<td>but then I wasn't sure [% indcl] :: because I didn't want to come to that group (.52) alone you know [% scl:adv].</td>
</tr>
<tr>
<td></td>
<td>Finite relative clauses [% scl:r]</td>
<td>Clause functioning as post modifiers of a noun</td>
<td>normal things like you would also tell your mother at home or someone [% indcl] :: (.41) who is picking you up from school [% scl:r].</td>
</tr>
<tr>
<td></td>
<td>NonFinite complement clauses [% scl:co: nf]</td>
<td>Non-finite clause functioning as complements</td>
<td>I'm not used [% indcl] :: to have (.73) ten weeks of school [% scl:co: nf].</td>
</tr>
<tr>
<td></td>
<td>Non-Finite adverbial clauses [% scl:adv: nf]</td>
<td>Non-finite clause functioning as adverbials</td>
<td>but I just bought them for them [% indcl] :: (.36) to make them a little (.47) bit happy [% scl:adv: nf].</td>
</tr>
</tbody>
</table>

:: clause separator

When measuring clause structures it is essential to determine how one defines a clause (Bulté & Housen, 2012). The key issue concerns infinitive and participial forms and whether or not to treat them as belonging to the previous finite structure or to count them as separate clauses. This decision has a significant impact on clause length and clause subordination. In this study, infinitive and participial forms were divided into two types: formulaic/simple structures (e.g. ‘I wanna go’) and complex structures (e.g. ‘I wanted her to do the washing’) (Diessel, 2004). Only complex non-finite structures were given clausal status. Formulaic/simple non-finite structures were counted as belonging to the previous clause. The classification of the verb structures into simple or complex types depended both on the type of the complement-taking verb (e.g. modality verb or manipulative verb), and whether or not an additional element (e.g. noun phrase or wh-marked) stood between the complement-taking verb and the non-finite structure. Table 13 shows the types of non-finite clauses that were included in my analysis. It was hoped that this type of analysis would “respec[t] the linguistic integrity of advanced learner
samples” (Bulté & Housen, 2012, p. 39) (i.e. by acknowledging their ability to form complex non-finite clauses) and, at the same time, not significantly lower MLC values (i.e. by treating formulaic non-finite verb structures as phrases, not clauses). This procedure also controls for “item-based usage” of subordination (Lambert & Kormos, 2014).

Table 13: Classification of non-finite verb structures

<table>
<thead>
<tr>
<th>Simple utterances</th>
<th>Complex utterances</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-finite complement as object</strong></td>
<td><strong>Non-finite complement as subject</strong></td>
</tr>
<tr>
<td>modality verbs + (to) infinitive</td>
<td>complement-taking verb + noun phrase + (to) infinitive/gerund</td>
</tr>
<tr>
<td>want, (would) have, (would) like, got, try, need, start, begin, stop</td>
<td>(e.g. ‘I want to see Tauranga’).</td>
</tr>
<tr>
<td>manipulative/cognition-utterance verbs + (to) infinitive</td>
<td>complement-taking verb + wh-marker + (to) infinitive/gerund</td>
</tr>
<tr>
<td>make, force, forget, decide, say, know, see</td>
<td>(e.g. ‘I don’t know how to do this’).</td>
</tr>
<tr>
<td>(e.g. ‘I forgot to buy the chocolate’).</td>
<td>anticipatory ‘it’ (/this) + v + adjective phrase + noun phrase + non-finite complement</td>
</tr>
<tr>
<td></td>
<td>(e.g. ‘it’s difficult for me:: to ignore it’)</td>
</tr>
<tr>
<td>Finite verb + adjective + preposition + nonfinite complement</td>
<td>Finite verb + adjective + preposition</td>
</tr>
<tr>
<td>(e.g. ‘I’m keen on:: going to the movies’)</td>
<td>(to) infinitive as subject + minimum a verb phrase</td>
</tr>
<tr>
<td></td>
<td>(e.g. ‘getting there:: is easy’)</td>
</tr>
<tr>
<td><strong>Non-finite adverbial clauses</strong></td>
<td>(e.g. ‘He smiled:: after answering the question’)</td>
</tr>
</tbody>
</table>

:: clause separator

Decisions must also be made with regard to the placement of linkers that cannot be unambiguously assigned to one AS-unit or another. In this study, single-word discourse/pragmatic markers between AS-units (e.g. ‘yeah’), which could be part of either unit, were positioned in the transcripts according to linguistic context and prosodic features (i.e. sometimes at the end of a unit, Example 12, sometimes at the beginning of a unit, Example 13).

Example 12: and maybe it’s in your head and (.55) all these things (.63) yeah | (3.39) and I need it for my abitur (1.05) so yeah |.

Example 13: so she knew from that | (1.04) and yeah (.50) it’s (.98) I can talk to her (.42) so |.

3.6.4.2 Lexical complexity

Most SLA studies have measured lexical diversity, which has been identified as the component that best correlates with NS ratings of the quality of NNS output (Lu, 2012). Lexical diversity is an indication of a learner’s productive lexical resources (Milton, 2009).
The language data (pruned version) of all three participants in this study consists of a minimum of 790 and a maximum of 1001 tokens per interview and 100 AS-units. Since Guiraud’s Index has been found to be reliable for learner data below 1000 (Daller et al., 2013), and because it is considered more reliable, and more stable than other commonly used measures (e.g. TTR), using Guiraud’s Index was considered the best measure of lexical diversity for this study (Table 14).

A shortcoming for measuring lexical diversity is that normalised scores are as yet unavailable in the literature (Milton, 2009). Milton (2009) explains that a particular score (say 6) on Guiraud’s Index does not indicate whether the speaker has “huge productive vocabulary resources or average resources or poor ones” (p. 130). This shortcoming is overcome by comparing the results of this study with results from previous studies of native-speaker and non-native speaker performances (such as Daller & Xu, 2007; Levkina & Gilabert, 2012; Llanes et al., 2012; Lu, 2012; Mora & Valls-Ferrer, 2012).

Table 14: Lexical complexity measure used in this study

<table>
<thead>
<tr>
<th>Complexity</th>
<th>Measure</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical diversity</td>
<td>Giraud’s index</td>
<td>The number of types divided by the square root of the number of tokens</td>
</tr>
</tbody>
</table>

3.6.4.3 Accuracy

In my study, I used one standard global and one specific measure to count accuracy. Table 15 lists the measures and provides definitions based on Ellis and Barkhuizen (2005).

Table 15: Accuracy measures used in this study

<table>
<thead>
<tr>
<th>Accuracy</th>
<th>Measure</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global accuracy</td>
<td>Percentage of error-free clauses</td>
<td>The number of error-free clauses divided by the total number of independent clauses, sub-clausal units and subordinate clauses multiplied by 100.</td>
</tr>
<tr>
<td>Specific accuracy</td>
<td>Percentage of correct finite verb phrases</td>
<td>The number of correct finite verb phrases divided by the total number of verb phrases multiplied by 100</td>
</tr>
</tbody>
</table>

For the global measures, I counted the percentage of error-free clauses (Skehan & Foster, 2012). Counting error-free clauses rather than the total number of erroneous words has the advantage that it enables the researcher to allow for ambiguous error coding. This is particularly the case when the ambiguity involves strings of words (i.e. when the error domain is broad, see Lennon, 1991). For this measure, each learner utterance was scrutinized for grammatical and
semantic inaccuracies. To enable consistency and reliability, I only considered errors in this count that were absolutely wrong. Word order errors were excluded from analysis since in spontaneous speech, “items often appear in positions that are dictated by communicative needs and . . . positioning is generally more flexible” (Carter & McCarthy, 2006, 88).

To complement the global accuracy measure I selected the percentage of correct finite verb phrases as a specific measure. Errors in verb phrases have been shown to impair communication and Tonkyn’s (2012) study revealed that “[o]verall Error density, and Error frequency in the VP seem to be promising indices of progress, and to be aligned with judges’ views of level” (p. 239, emphasis in the original). For purposes of analysis, all finite verbs were coded with the symbol [% fv] and a frequency count was conducted with the CLAN software. When students left out the auxiliary in a complex verb phrase, (e.g. ‘when you always joking with others . . . ’), the non-finite verb was coded as a finite verb (i.e. with the symbol [% fv]) to ensure that the phrase was counted. The finite verb phrases were subsequently scrutinised for their accuracy (in complex verb phrases, errors on both the finite and the non-finite verb were considered) and the total number of finite verb errors per total number of finite verb phrases was calculated. For this specific measure, only morphological and/or tense errors were counted. Semantic errors (i.e. wrong vocabulary choice but correct inflection) were only considered in the global accuracy count. Note that finite verbs that occurred as part of discourse/pragmatic markers were not included in the count.

The norm against which the learners’ speech was measured was that of informal spoken English used by New Zealand teenagers. Inaccuracy in this study refers to the extent to which the participants’ speech performances deviated from the speech performance of their same-age NS peers. Non-standard uses of language were considered acceptable if their usage was common by NSs (e.g. ‘ehm yeah and then so the weekend started really bad’ - instead of ‘badly’). Isolated German words that occurred in the speech samples were not coded as errors. When the same error was repeated, it was counted every time it occurred. Non-target-like word choices were considered accurate when the participants made use of lexicalisation strategies (e.g. circumlocution) to generate L2 items which were not considered absolutely wrong. For example, Alia was unable to produce the word ‘bubble wrap’ and instead described the material as “air paper with bubbles” (int. 6). Identifying tense errors in a consistent and reliable manner caused difficulties. In informal speech, speakers can “exercise considerable liberty in tense and aspect choice” (McCarthy & Carter, 2001). I adopted a few principles that I applied in a consistent manner in cases of ambiguity (Appendix 6).
Two inter-raters were used for the identification of errors: a New Zealand adult ESL/EFL teacher and a 17-year-old New Zealand boy. In a first step, the adult NS and I (in separate processes) identified all speaker utterances that contained grammatical or semantic errors. All erroneous utterances were then given to and assessed by the 17-year-old NS boy who determined if the errors were definitely wrong or if they passed as standard uses of the model population (i.e. adolescent New Zealand English). The raters were also consulted to establish the principles that were applied for the identification of tense errors, as described above. Although only one adolescent native speaker was consulted in this matter, he was considered a reliable person to identify errors of this kind not only because of his age and nationality, but also because we were able to thoroughly discuss the nature of the errors and solve disagreements. This procedure was also adopted by Lennon (1990) who asked a local group of same-age people to identify errors.

This study acknowledges that “accuracy per se is not a direct indicator of interlanguage development” (Pallotti, 2009, p. 592, emphasis in original), particularly when performance is assessed with global measures. After all, the measure encompasses all errors, and does not distinguish between developmental stages of different types of errors. The scope of this study did not allow for a discussion of the learners’ developmental stages with regard to global error measures, but an attempt was made to illustrate what appeared to be the most difficult areas for the students with regard to finite verb phrases. This may be significant, as “the study of learner error remains of practical significance to language pedagogy” (Ellis & Barkhuizen, 2005, p. 70).

3.6.4.4 Fluency

Fluency is a multi-dimensional construct and different subcategories have been identified. Fluency can be defined as the “ability to fill time with talk without unnatural hesitations” (De Jong et al., 2012).

In analogy to Tavakoli and Skehan (2005) a distinction was made in this study between breakdown fluency, repair fluency, and speed fluency (Table 16). Additionally, the frequency and range of “smallwords” (Hasselgren, 2002), a category of formulaic sequences, were investigated.
Table 16: Fluency measures used in this study

<table>
<thead>
<tr>
<th>Fluency</th>
<th>Measure</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed fluency</td>
<td>Speech rate</td>
<td>The number of unpruned words articulated per minute, including all repair phenomena and all pauses.</td>
</tr>
<tr>
<td></td>
<td>Mean length of run</td>
<td>The number of unpruned words between pauses of 0.3 seconds</td>
</tr>
<tr>
<td>Breakdown fluency</td>
<td>Empty pauses</td>
<td>The number of silences longer than 0.3 seconds per hundred words</td>
</tr>
<tr>
<td></td>
<td>Filled pauses per hundred words</td>
<td>The number of non-lexical fillers (ehm, hmm, eh) per hundred words</td>
</tr>
<tr>
<td>Repair fluency</td>
<td>Repetitions per hundred words</td>
<td>The number of repeated words per hundred words</td>
</tr>
<tr>
<td></td>
<td>Reformulations per hundred words</td>
<td>The number of reformulations of form and content per hundred words</td>
</tr>
<tr>
<td>Formulaics</td>
<td>Smallwords</td>
<td>“small words and phrases, occurring with high frequency in the spoken language, that help to keep our speech flowing, yet do not contribute essentially to the message itself” (Hasselgren, 2002, p. 150).</td>
</tr>
</tbody>
</table>

**SPEED FLUENCY**

Speech rate (SR) measures the time it takes a speaker to produce speech, including pause time. It is a global measure of speech production that lends itself well to an objective analysis of fluency and that is likely to “show fluency developments over time in a single speaker” (Lennon, 1990, p.26). The measure is widely used and has been designated a “fairly sound indicator” of fluency (Wood, 2012, p.20). SR has been analysed in different ways and it remains “debatable which of the measurements is to be preferred” (Götz, 2013, p. 15). In this study I calculated SR in unpruned words per minute (i.e. the number of all words the speakers uttered, including all hesitations, repairs and all filled pauses). A word count rather than a syllable count was adopted, based on the fact that informal English speech predominantly consists of one-syllable words (Gut, 2009) and that there is no statistically significant difference in SR when measured in syllables per minute or words per minute (wpm) for unscripted speech (Götz, 2013, p. 15). Counting the number of words was also chosen for reasons of reliability. Compared to syllables, counting allows for an automatic word count and hence absolute accuracy.

For the SR measure in my data, I counted the minutes and seconds of the duration of each 50+word turn from the last word uttered by the researcher to the last word of the 50+word turn uttered by the participant (to ensure the pauses before the participant turn was counted, following Goldman-Eisler, 1968). If the participants initiated their turn before I finished mine,
the analysis began with the participant’s first word of their turn. I stopped measuring the time after the last word of the 100th unit, whether or not the speaker’s turn continued. Following Riggenbach (1991), empty pauses longer than 3 seconds were not counted “to give a more representative figure for the rate of speech” (p. 428) (e.g. a pause of 4 seconds counted as a pause of 3 seconds). Such long pauses tended to occur in my data when the participants were involved in behaviour unrelated to speech processes, such as taking a sip of their drink.

Articulation rate was not considered in this study because as a measure of speed, it is “not strongly related to the proceduralization of lexical and syntactic knowledge” (deJong & Perfetti, 2011, p. 539). Furthermore, in some interview settings (restaurants), background noise featured on the recordings, making exact pause measurements difficult. While it was possible to give an approximation of the measured pauses (and certainly possible to allocate a pause as below or above the cut-off point), it was too risky to reliably calculate the total amount of silent time necessary for the calculation of articulation rate.

Mean length of run (MLR) is defined as the size of the runs of speech that occur between pauses (Wood, 2012). Towell et al. (1996) considered MLR as a ‘higher-order’ measure of fluency since it indicates the degree to which language performance is automatized. Like SR, MLR can be quantified in different ways (in syllables, phonemes, or words per run). In this study, a run was considered as the number of words uttered between pauses, and calculated by dividing the total number of tokens by the total number of runs. As for SR, all words between runs were counted as tokens, including non-lexical filled pauses (‘ehm’, ‘eh’, ‘hmm’) and all hesitation phenomena (Götz, 2013). Although some researchers choose to exclude filled pauses and repair phenomena from an MLR count, these phenomena were considered fluency enhancing strategies and therefore words (see Götz, 2013). For purposes of calculating MLR, a ‘pause’ was considered to exist when it was 0.3 seconds or longer (i.e. cut-off point).

REPAIR FLUENCY

This study counted the frequency of repetitions and reformulations. A repetition is defined as an “immediate and verbatim repetition of a word or phrase” (Skehan and Foster, 1999, p. 107). Conducting a systematic analysis of repetition positions was beyond the scope of this study, although such an analysis can shed light on the ‘naturalness’ with which repetitions are used. Table 17 lists the types of repetitions that were considered for the frequency count in this study. Combining a predicate with a predicate-negation contraction (e.g. ‘I should shouldn’t go’) was considered a reformulation, not a repetition. Only involuntary repetitions were counted
(reflective of planning) - deliberate repetitions that were used for emphasis (‘I was so so tired’) were not counted, when they could be unambiguously identified as such (using the linguistic context and intonation phenomena for reference).

In the CHAT transcription format, the repeated material (if longer than one word) was enclosed in angle brackets and the code [/] marked the repetition. Apart from empty pauses and non-lexical pauses (‘ehm’, ‘eh’, ‘hmm’), as well as the fillers ‘yeah’ and ‘like’, no material could stand between the earlier material and the repetition (MacWhinney, 2013).

Table 17 Types of repetitions used in this study

<table>
<thead>
<tr>
<th>Types of repetitions</th>
<th>Examples from this study</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-word repetitions</td>
<td>and she ehm (.36) change [/] (.32) ehm (1.05) change the group.</td>
</tr>
<tr>
<td>Multi-word repetitions</td>
<td>but so &lt;when I&gt; [/] (.32) when I ask her something she gonna (.78) answer it and (.34) gonna help me as good as she can.</td>
</tr>
<tr>
<td>Repetition of syllables</td>
<td>but ehm (.78) they are person from Brazil and Ch [/] Chile.</td>
</tr>
<tr>
<td>Expanded repetitions</td>
<td>and I want to ehm (.44) learn more vocabulary like in Bio [/] Biology or (.96) Health.</td>
</tr>
</tbody>
</table>

Reformulations occur when the speaker identifies an error during or immediately following production, stops, and reformulates the speech. Reformulations come in two forms: one that involves a change of content or sentence structure (including false starts) and one that involves a change of form. Some researchers distinguish between the two forms (e.g. Foster et al., 2000; MacWhinney, 2013), while others measure them together (e.g. Biber, 1999; Götz, 2013; Lennon, 1990). Distinguishing between the two types can be difficult and it is doubtful if reliable means can be found to do so consistently. Given these difficulties, I did not discriminate between the two for the frequency count. However, a more qualitative analysis was conducted to illuminate the participants’ behaviour regarding unambiguous cases of changes of form and content. Table 18 lists the types of reformulations that were considered for the frequency count in this study, based on MacWhinney’s (2013) CHAT transcription format.

In this study, a trail-off was not counted as a false start but as a complete unit. A trail-off is a grammatically complete utterance (i.e. consisting of a subject and a finite verb) that is abandoned (usually with the signal ‘yeah’) without being repaired (Biber, 1999; MacWhinney, 2013). Trail-offs are usually followed by a completely new utterance (MacWhinney, 2013) (e.g. Jana: ‘I was just a bit yeah | (0.42) I accepted that because I thought ‘oh yeah I’m new
here”’. 2 AS-units). Participant clarifications that included the conjunction ‘or’ to introduce another element to be considered as an alternative were not treated as reformulations.

In the CHAT transcripts, the reformulated material (if longer than one word) was enclosed in angle brackets and the code [///] marked the reformulations. If the lexical fillers ‘yeah’ or ‘like’ occurred between repairs and the earlier repaired material, they were not considered part of the repair (e.g. ‘because the bus <is> [///] like (0.55) need this time’).

Table 18 Examples of reformulations in this study

<table>
<thead>
<tr>
<th>Types of reformulations</th>
<th>Examples from this study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change of form</td>
<td>I see [///] saw him on the football field. so that was &lt;very&gt; [///] (0.31) a very short time before I left.</td>
</tr>
<tr>
<td>Change of lexical item</td>
<td>I read [///] wrote that in the blog.</td>
</tr>
<tr>
<td>Addition of elements</td>
<td>During the lesson she was constantly with [///] on her mobile phone. &lt;He went&gt; [///] I think he went to the soccer field.</td>
</tr>
<tr>
<td>Elimination of elements</td>
<td>but &lt;he’s&gt; [/] &lt;he’s really&gt; [///] (1.54) he is friendly and good.</td>
</tr>
<tr>
<td>Replacement of elements</td>
<td>and now &lt;she is&gt; [///] (0.38) she talks with me and yeah</td>
</tr>
<tr>
<td>False start</td>
<td>&lt;and they would&gt; [///] and I was like ‘oh no’.</td>
</tr>
</tbody>
</table>

**BREAKDOWN FLUENCY**

According to Wood (2012), the study of pause times and frequencies “tells us a great deal about speech fluency” (p. 24). This study examined the frequency of empty and filled pauses. Although an investigation of pause positions might reveal interesting results, constraints of time and space did not allow for such an analysis. Cut-off points for empty pauses (EPs) have been debated exhaustively. The lower end cut-off point for pauses has traditionally been between 0.25 seconds and 0.3 seconds (Götz, 2013). In line with Wood (2012), a lower end cut-off point of 0.3 seconds was chosen because “[a]nything less than 0.3 seconds is easily confused in a spectrogram with other speech phenomena such as the stop phase of a plosive sound” (p. 107). In line with Riggenbach (1991), a higher end cut-off point of 3 seconds was chosen “to give a more representative figure for the rate of speech” (p. 428).

EPs were measured using the PRAAT software. The PRAAT programme converts audio files into waveforms and spectrograms. It reacts to zooming and scrolling which makes finding pauses and sounds easy. The programme allows users to highlight selected parts of the waveform or the spectrogram (e.g. the silent periods between two sounds) which are attributed
by start times, end times, and time steps (in seconds). Highlighting the area between two waves (i.e. EPs) provides users with the length of the pause in seconds. Pauses in this study were counted in seconds and milliseconds (e.g. 1.25). They are defined as the length of complete silence between two sounds. Vowel elongation was not counted as part of a pause. Due to the setting of the interviews (cafes), background noise also featured on the recordings. In some instances, this created challenges when measuring pauses.

EPs were marked by brackets in the transcripts, with a number representing the pause length in seconds (Example 14).

Example 14: and (1.02) it’s something special :: when we eat in front of the TV | (.97) and (1.23) yeah (1.96) it’s (.46) it’s not really (.46) personal.

For filled pauses (FPs), I conducted a frequency count of the non-lexicalised fillers ‘ehm’, ‘eh’, and ‘hmm’. In the data, filled pauses were coded with the symbol & (MacWhinney, 2013) for the purposes of the CHAT analysis (Example 15) (see 3.6.1).

Example 15: &ehm yeah (.42) it was &ehm my last chance (.51) to (.47) go away during my school time.

**SMALLWORDS**

Given the spatial and temporal limits of my study, formulaic patterns were not scrutinized in this study. However, I counted the range and frequency of smallwords, which occurred in the three participants’ speech. The range of smallwords under investigation was based on Hasselgren’s (2002) list of smallwords and on smallwords specific to my own corpus. Identification procedures were as follows: I first scanned the learners’ data for smallwords that had already been identified by Hasselgren. Secondly, I incorporated into the list of smallwords the discourse/pragmatic markers (DPMs) listed in 3.6.3.2. Last, the single-word smallwords ‘yeah’ and ‘so’ were identified as smallwords in this study according to Hasselgren’s working definition: “Smallwords: small words and phrases, occurring with high frequency in the spoken language, that help to keep our speech flowing, yet do not contribute essentially to the message itself” (Hasselgren, p. 150). ‘Yeah’ and ‘so’ occurred particularly frequently in my participants’ data, yet did not feature in Hasselgren’s list of smallwords.

In analogy with Hasselgren’s (2002) procedures, I only included in my study smallwords that occurred with an overall frequency of 6 times and more (i.e. in the entire corpus). Similarly, I adapted Hasselgren’s approach to form a group of smallwords with relative low frequency but
with similarity in form and function: the smallwords ‘and everything’, ‘and (all this) stuff’, and ‘and things’ were considered parts of the same group. The smallwords investigated in this study contain single-word DPMs (e.g. ‘like’), and fixed multi-word sequences (e.g. ‘you know’). In total, the identification process generated the following twelve smallwords (or groups of smallwords):

Like, yeah, so, just, and everything/and (all this) stuff/and things, you know, and/or something like, I think, I don’t know, a bit, I mean, ah/oh

Polysemous expressions that did not conform to Hasselgren’s working definition of smallwords were excluded from the smallword count in this study (Table 19):

Table 19: Words that were not counted as smallwords in this study:

<table>
<thead>
<tr>
<th>Words</th>
<th>Uses of these words that did not count as smallwords</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>ah/ oh</td>
<td>when used as an exclamation (including sighing)</td>
<td>‘and I was so ah’</td>
</tr>
<tr>
<td></td>
<td>when part of a fixed phrase</td>
<td>‘oh my god’</td>
</tr>
<tr>
<td></td>
<td>when used as a preposition</td>
<td>‘he is like her’</td>
</tr>
<tr>
<td></td>
<td>when used as a verb</td>
<td>‘I like him’; ‘I would like to go’</td>
</tr>
<tr>
<td></td>
<td>when used as an element of another discourse marker/smallword</td>
<td>‘or something like that’; ‘it’s like’</td>
</tr>
<tr>
<td></td>
<td>when used for direct speech</td>
<td>‘he was like ‘turn the volume down’</td>
</tr>
<tr>
<td>just</td>
<td>when adverbial in function</td>
<td>‘I just got there when he closed the store’</td>
</tr>
<tr>
<td></td>
<td>when ‘just’ can be replaced by ‘only’</td>
<td>‘I just spent one break with them’</td>
</tr>
<tr>
<td></td>
<td>part of a larger discourse marker where removal would change the meaning of the utterance</td>
<td>‘just to let you know’</td>
</tr>
<tr>
<td>so</td>
<td>when it fulfills a consecutive function</td>
<td>‘and so we went there’</td>
</tr>
<tr>
<td></td>
<td>when it fulfills an adverbial function</td>
<td>‘it was so good’</td>
</tr>
</tbody>
</table>

3.6.5 Describing and interpreting L2 data

3.6.5.1 Terminology

In this study, I make a distinction between L2 performance, L2 development, and L2 progress. L2 performance refers to the participants’ oral L2 production during any one interview. It is indicative of what learners know implicitly and as such provides a source of information about their L2 competence. L2 performance is believed to be variable in this study. L2 development refers to the changes that take place in L2 production from one performance to the next. L2
development thus describes variation in learner production. This use of the term is in line with Dynamic Systems Theory which views development as variable, unstable and ongoing (i.e. there is no linear relationship between development and advanced language use) (see 2.3.1.3). L2 progress refers to clear advances in the participants’ L2 systems. L2 progress can only be determined through in-depth analyses of the development of each L2 measure.

I acknowledge that my data does not lend itself to an unproblematic interpretation of L2 progress. This thesis used relatively naturally occurring data (spontaneous speech) which is variable and heterogeneous, even for L1 speakers. When interpreting learners’ development of fluency, accuracy and complexity, not-controlled-for variables (e.g. topic, speech mode) can change the outcome of the measure, without suggesting changes in proficiency levels. I dealt with this difficulty by describing as clearly as possible whether the changes in the measure (i.e. ‘development’) account for progress or simply variation. I also applied correlation coefficients for an objective measure to assist in that interpretation. A smooth, linear trajectory was interpreted as more indicative of progress than a jumpy nonlinear trajectory (see 3.6.5.2).

3.6.5.2 Visual analysis

To describe and interpret longitudinal single-case studies, data are traditionally presented in time-series line graphs. Visual inspections are carried out to determine if change has happened and what the trend is (Barkaoui, 2014).

For my study, visual analysis (i.e. a visual inspection of the data patterns) was considered a valid technique to present L2 development. For each variable, I used a time-series line graph. Development for all three participants was shown using identical scales. The lower and upper limits of the y-axis correspond to the lowest and highest numbers that any of the participants in this study reached. The y-axis represents the dependent variables (i.e. the CAF measures). The x-axis on the graph represents the independent variable ‘time’ (i.e. the six interviews).

The danger that comes with describing changes in time-series graphs is ‘subjectivity’. To show change beyond pure visual description, Pearson’s product-moment correlation coefficient (CC) (denoted by ‘r’) was used in this study. CCs provide an objective measure to determine how close paired data is to being linear (Taylor, 2016). CCs were standardised to enable meaningful comparison among the measures (Reifman, 2007). CCs have been used to indicate the level of variability or stability of variables in single-case studies in applied linguistics, albeit rarely (Isabelli-Garcia, 2004). To interpret the size of CCs, I adopted the cut-off points used in past
research and as demonstrated in Mukaka (2012). Positive ‘r’ values indicate positive development and negative values negative development. The closer ‘r’ is to 0, the more variation there is between the data (i.e. there is little or no linear relationship between time and the dependent variable). ‘R’ values close to +1 or -1 represent development with little variation (Table 20).

<table>
<thead>
<tr>
<th>Size of Correlation</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>.90 to 1.00 (-.90 to -1.00)</td>
<td>Very high positive (negative) correlation</td>
</tr>
<tr>
<td>.70 to .90 (-.70 to -.90)</td>
<td>High positive (negative) correlation</td>
</tr>
<tr>
<td>.50 to .70 (-.50 to -.70)</td>
<td>Moderate positive (negative) correlation</td>
</tr>
<tr>
<td>.30 to .50 (-.30 to -.50)</td>
<td>Low positive (negative) correlation</td>
</tr>
<tr>
<td>.00 to 0.30 (.00 to -.30)</td>
<td>Negligible correlation</td>
</tr>
</tbody>
</table>

When interpreting the participants’ changing performance, I compared their results to native and non-native English speaker data of previous studies that measured spontaneous oral speech. Comparing results across studies is problematic due to differences between the studies in, for example, topic, sample size, and language elicitation methods. There is in particular some debate as to whether learner language should be compared to NS norms. I took the position that the participants’ language is not a ‘deficient’ version of NS systems (Schmid et al., 2011), and that it would be difficult to interpret their performance without comparisons to NS performance. Using previous studies was “a helpful touchstone” (Polat & Kim, p. 189) for the discussion of the participants’ oral L2 changes, as it has been in previous L2 research (e.g. Polat & Kim, 2014; Skehan, 2009; Verspoor et al., 2008). Comparisons were predominantly made with SA studies, but, given the scarcity of SA studies using the CAF triad (Juan-Garau & Pérez-Vidal, 2007; Mora & Valls-Ferrer, 2012; Tonkyn, 2012), or any one CAF dimension (e.g. Freed, 1995; Isabelli-Garcia, 2006; Lennon, 1990; Riggenbach, 1991; Wood, 2012), I occasionally used studies outside this research domain for comparisons.

In this study, I also described and interpreted the relationships between CAF variables. To compare measures of different scales (e.g. some are percentages, others are frequency counts), scores were standardised. I used Z-scores for this purpose in the manner set out by Larsen-Freeman (2006). Z-scores allow the comparison of raw scores from different bunches of data. Absolute values of Z-scores denote how many standard deviations a score is away from the mean. Thus they provide information regarding the relationship between a bunch of scores and their mean. Z-score values range from -2 to +2. A Z-score value of 0 indicates the value is
equal to the group mean. Positive Z-score values refer to a score being above the group mean, while negative Z-scores refer to values below the group mean. The ‘group’ refers to the group of measures that are being compared with each other. Larsen-Freeman (2006) demonstrates that although Z-scores report averaged data, they “provide a true description of the behaviour of the individual” (p. 601).

Verspoor and van Dijk (2011) emphasised that comparing different variables “makes sense only if the variables that are studied in combination have a meaningful relationship to each other” (Verspoor and van Dijk, 2011, p. 86) based on a “clear theoretical or empirical motivation” (p. 88). The variables that were compared in this study have been shown to be meaningfully related to each other in previous CAF studies (e.g. Robinson, 2007; Skehan, 2009). They were speech rate and mean length of run for fluency, mean length of clause and lexical diversity for complexity, and error-free clauses for accuracy. The remaining variables were excluded from comparisons in this part. The subordination Index was considered a poor measure for complexity in oral speech (read Chapters 4 to 6 for more details) and finite verb phrase was considered too specific to be compared to other variables. Only supportive and competitive relationships were considered, as the methods used in this study were not rigorous enough to determine precursor relationships.

### 3.7 Qualitative data analysis

This study adopted an interpretivist paradigm, which assumes that the researcher is part of the research project. More particularly, it was informed by the fact that “experiences and values of both research participants and researchers substantially influence the collection of data and its analysis” (Wahyuni, 2012, p. 71). The interpretations of the findings of the learners’ social experiences and language learning motivation were considered “the product of the researcher’s subjective interpretation of the data” (Dörnyei, 2007, p. 38) and knowledge was generated not from implementation of previous theories but “based on participants’ unique perspective and grounded in the actual data” (Hsieh, 2005, p. 1280). Credibility of the analysis was established by means of my prolonged engagement with the participants and data (Hsieh, 2005). Additionally, my specific experiences (EFL/ESL teacher, former international student) and cultural and linguistic background (European German-speaker) assisted in interpreting the significance of the discovered patterns in connection with the research questions (Ellis & Barkhuizen, 2005). Given the interrelationship between data generation and researcher-participant-relationships, as well as my unique experiences, external validity was not tested as
it was not considered an essential procedure when analysing data of this type within this theoretical paradigm.

The participants’ subjective accounts of the world were considered acceptable knowledge in constructing truth. The objective truth of their accounts was not confirmed and no evidence regarding the justification of their claims exists. This approach is in line with a relativist or interpretivist approach, which considers the participant’s knowledge and interpretation of the world to be “reality” (Braun & Clarke, 2013, p. 27). Reality was understood as socially constructed, subjective, dynamic, and multiple (Wahyuni, 2012). This research is thus value-bound and emic.

I used the software NVivo (QSR International PTy Ltd., 2012) for data management and consistent content analysis. For the analysis, I systematically worked through the entire data set (i.e. the participants’ responses to my interview questions, their blog entries and monthly reports), coded for themes and identified patterns within these themes that revealed how language-learning opportunities and language learning motivation came about. I integrated the results into “a set of conclusions” (Ellis & Barkhuizen, 2005, p. 271) and subsequently described and explained them. No attempt was made at building or developing theory.

The unit of analysis for each theme ranged from short phrases to extended discourse. Units represented matters such as the participants’ behaviour (e.g. chatting during classes), goals and desires (e.g. to go out with a Kiwi friend; to become a better L2 speaker), descriptions and explanations (e.g. seating arrangement in classrooms) and perceptions (e.g. about others’ behaviour; about their L2 progress). Evaluative comments were coded together with the matters they referred to. For instance, in Example 16, Chiara describes where she spent the break (not with Kiwis but with Germans) and how she felt about that (it was nice with Germans and she felt bad with Kiwis).

Example 16: I'm not really spending time my lunch breaks with them [Kiwis] because I don't know you don't really have things to talk about and yeah yeah they always have friends and I don't know them and then I always feel bad if they just talk to me or I feel bad if they just talk to them so yeah at the moment we just the Germans and yeah but that's nice. (Int. 6)

I followed three steps for this analysis. First, the data was analysed deductively. I worked through the data set to find references about social interactions/L2 that occurred in three major settings. The types of settings were based on the interview questions: the host family, the school, and activities that occurred independently of the former two contexts outside of school.
I placed each relevant ‘chunk’ from the data into the respective node (Figure 1). The students’ explicit comments relating to language learning which could not be tied to a specific setting were placed into the parent node ‘L2 learning’. These nodes constituted the highest level of organisation (parent nodes).

After these broad coding processes, I divided data relating to the parent nodes into more specific themes (child nodes). The parent node ‘school’ was divided into the child nodes ‘inside the classrooms’ and ‘outside the classrooms’. The parent node ‘outside the school’ was divided into the child nodes ‘trips/social activities’ and ‘hobbies’ (Figure 1). The parent node ‘L2 learning’ was divided into the two child nodes L2 learning reasons and goals (e.g. wishing to improve fluency to be able to contribute to group discussions) and perceptions of L2 progress (e.g. being dissatisfied with making mistakes).

Figure 1: Parent and child nodes identified in this study

In a third step, I analysed the data inductively. I delineated fine-grained themes within the child nodes. These themes constituted the smallest unit of analysis. For example, narrative units coded in the child node ‘inside the classrooms’ contained information about classroom activities (e.g. preparing for an English speech), classroom compositions (i.e. seating arrangements), and the participants’ performance in the subjects. Steps 2 and 3 required multiple re-readings of the data and analysis was ongoing and recursive until saturation was reached (Ellis & Barkhuizen, 2005).

After coding, I looked for patterns in the data that revealed how the participants’ different social contexts afforded opportunities for L2 use, both as a consequence of the learners’ own proactive behaviour (e.g. deliberately sitting next to Kiwi students) and the opportunities provided by L2 communities (e.g. classroom group activities).

For the discussion of language learning motivation, I adopted an approach that does not rely on any particular theory but addresses motivation in terms of the three central aspects that any general theory of motivation must address (Ellis, 2015):
1. The reasons a learner has for needing or wanting to learn an L2 (i.e. motivational orientation)

2. The effort a learner makes to learn the L2 and the learner’s persistence with the learning task (i.e. behavioural motivation).

3. The effect that learners’ evaluations of their progress have on their subsequent learning behaviour (i.e. attributional motivation).

For this purpose, I scrutinized the data for patterns relating to these three aspects of motivation. This study adopts a socio-cultural perspective on language learning motivation. This perspective acknowledges that learners are inseparable from their socio-cultural environments and that the socio-cultural environment affects the three central aspects of motivation. While separate analysis of each central aspect was necessary to gain insight into the effect of each aspect, I acknowledge that aspects of motivation are in reality part of the same complex, dynamic system.

In this thesis, each case is examined in a separate chapter (Chapters 4, 5 and 6). In the case study chapters, I first provide background information for each participant and for each social context and then go on to describe in a more analytical manner the kinds of opportunities that existed for language-related experiences. These descriptions are supported by examples that capture social activities involving L2 use or the participants’ evaluations of these. Following this section, I describe in detail the learners’ L2 development (and whether or not it accounts for progress), and finish the chapters with potential explanations of the relationship between L2 learning opportunities and L2 development. In Chapter 7, I analyse the learners’ motivation, language learning opportunities, and L2 development, pointing to commonalities and differences that occurred among the three participants.
4 Case Study 1: Jana

4.1 Introduction

This chapter has three separate sections. The first one provides an account of Jana’s social settings during her study abroad. In the second section I describe in detail her L2 development as measured over six data collection points. I finish the chapter with potential explanations regarding the relationship between L2 learning opportunities and L2 development. The structure of this chapter is identical for all three case study chapters.

4.2 Background information

JANA AND HER FAMILY IN GERMANY

Jana grew up in a medium-sized German city with her two parents (a stay-at-home mother and an engineer father) and younger brother. Her family exclusively speak German and they had never lived outside of Germany prior to Jana’s sojourn. Jana’s family had vacationed in many different European countries. Jana had also accompanied her father on business trips to Northern Africa, London and Iran, during which she made contact with people from different nationalities with whom she had to speak English:

when I travel with my dad or he had some some customers there or something like that ehm and I talk with them [in English] or in Iran I had to talk to to some people and also in my age and I'm still in contact with them. (Int. 1)

Jana arrived in New Zealand at the age of 17 and turned 18 in May 2014 which made her older than the average German international student (16-17). She considered herself an open minded person.

GERMAN SCHOOL AND LANGUAGE LEARNING HISTORY

Before her sojourn, Jana attended a private Gymnasium with a different curriculum than that of public schools. The school emphasized intellectual as well as creative, social, and critical thinking skills. The school also required the students to do a four-week internship abroad, which Jana spent in Switzerland. So, she had already experienced living abroad prior to going to New Zealand.
Jana left for New Zealand mid-way through Year 10, which was the last year in which it was possible for her to study abroad while attending Gymnasium (German students graduate in Year 12). She had studied English for ten years before her sojourn. She had three lessons per week in EFL classes (at all levels). The school also staged a drama performance in English in Year 8. Jana professed a passion for learning languages in general and for English in particular. She reported having been one of the best performing students in her English class. She also reported having read books in English while in Germany, as she put it, “just for fun” and because it constituted “a bit more challenge to read something in English” (int. 1).

**Reasons for Studying Abroad in New Zealand and General English Language Learning Motives**

Jana’s motivation for studying abroad was mainly a desire to have new experiences away from familiar surroundings. Before going abroad, Jana had grown tired of the sameness of the school and her friends. She wanted to study in an English speaking country because her English skills would allow for better communication than her French skills, which she thought were not “really good” (int. 1). New Zealand was her chosen destination due to a gut feeling and because friends recommended the country:

> I wanted to go to Australia or New Zealand I don't know why but I was not interested to go to America and then I heard from friends and yeah that New Zealand is so that the people are so friendly and open and it it was my feeling. (int. 1)

Jana’s language learning motives included personal reasons, instrumental reasons (e.g. school examinations, English for a job or studies), and social networking reasons (see 4.3.4 for more details). The length of Jana’s sojourn was not negotiable. As a Year 10 student, she was required to return after 5.5 months.

**4.3 Social settings**

As mentioned above, this section gives an account of the social settings that afforded Jana opportunities for language-related experiences during her SA period, and her reflective responses about them. Social settings in this study are defined as “arrangements for learning, involving one or more learners in a particular place, who are situated in particular kinds of physical, social or pedagogical relationships with other people (teachers, learners, others) and material or virtual resources” (Benson, 2011, p. 13). Three social settings were identified in
this study: the host family, the school (inside the classroom and outside the classroom) and out-of-school settings (trips/social activities and hobbies) (see Chapter 3).

For these descriptions, I took into account Jana’s full data set (interviews, diary entries and monthly reports). In the quotes from Jana’s oral interview data, I added quotation marks to indicate the beginning and end of direct speech passages for purposes of readability. Data from reports and blog entries were left unchanged (including errors).

4.3.1 The host families

HOST FAMILY 1

Jana’s first host family included a host mother, Scarlett, a host father, Andi, and a South American host sister, Luana. Andi had two children who lived with their mother and who occasionally visited the family. Apart from mentioning that these kids were very shy, Jana never referred to them. Both host parents worked full time. Andi worked as an IT specialist from home and Scarlett had an administrative job. They were experienced at hosting international students.

Generally, Jana’s social activities in the host family were limited to dinner, some conversations with her host mother, and a few social activities outside the home. The host father rarely interacted with the girls and spent most of the time in his room. Jana and Luana usually got on well with each other and often spent time together home. While conversations with her host mother became more frequent and pleasant over time for Jana, she generally did not perceive the opportunities for social interaction in her host family setting as enriching or conducive to learning.

The ‘double-placement’ arrangement brought relief to Jana as she did not wish to be ‘alone’ with her host family. Jana and Luana were required to do dishes once a week, to clean their rooms and to do their washing. Often, Jana and Luana did their chores together. In the beginning, the girls sometimes socialised. They visited each other’s international friends or went to town together. After a time, Jana disliked parts of Luana’s personality and felt less inclined to spend time with her outside the house. The girls communicated in English, yet Jana did not consider the level of interaction, as with other South American international students, facilitative of L2 learning. Luana attended the same college as Jana. The girls walked to school together and attended the same Maths class. They did not interact during breaks at school.
Jana, Luana and Scarlett regularly got together during dinner. To Jana’s dislike, dinner was eaten in front of the TV on the couch. She reported that they did not eat ‘as a family’ given her host father’s absence from the meal. The three women watched TV series while eating and did not talk much. However, they sometimes chatted before or after dinner. Although Jana mentioned getting used to this arrangement, she disapproved of it because it was not ‘personal’. Jana was also dissatisfied because she was accustomed to lengthy conversations with her family in Germany who always ate together at the table.

At the beginning of her homestay, Jana had some experiences that affected her rapport with the family. Minor incidents, such as the host mother forbidding the girls to snack after school to ensure they had an appetite for dinner made Jana feel ‘unsure’ about the rules in the home. Other negatively interpreted events resulted in Jana choosing to spend little time at home. For example, Jana felt that her host parents did not show an interest in her person beyond shallow enquiries about her day. Jana furthermore described the conversations with her host mother as superficial, and complained that she failed to respond to Jana’s openness by sharing her own experiences. It also appeared difficult for Jana to interpret her host mother’s changing day-to-day moods, which manifested in the subtleties in her host mother’s language use:

\[\text{ehm she's she changes her her mood so sometimes she she comes home and she's like 'hi how are you?' and that's it and I think I did something and the other time she comes in 'hi my love how are you?' and you are like 'oh hi I'm fine thanks' ehm yeah sometimes it's a bit difficult. (Int. 1)}\]

In response to these circumstances, Jana sought conversations in English with locals at school who she believed were more interested in her and with whom she could engage in conversations that were more conducive to L2 learning:

\[\text{I prefer to speak with someone I want to speak you know or it's it's so to speak with her is not really deep and ehm yeah and what I talk with her is not ehm so difficult that I can learn from it but ehm yeah it's at the moment it's enough for me to speak in in the school but if if I eh yeah I don't really miss it to talk to her but ehm I would talk to her if I if I ehm yeah it's hard to say but if I wouldn't talk that much in English I think I would talk with her but I don't miss it so actually. (Int. 1)}\]

After living with the host family for approximately one month, Jana felt more comfortable. She got involved in a number of activities with her host mother that she enjoyed and which made her feel at home. These included meaningful conversations about her feelings and social achievements, occasional cooking for the host mother, a few movie nights, and going for walks.
One experience in particular positively affected Jana’s feelings towards the host family and made her feel part of it. Scarlett was unexpectedly taken to hospital and so Jana, together with Andi and his brother visited her at the hospital. This situation brought with it the host father’s attention and care. For the first time, Jana and Luana shared a meal with Andi which Jana thought was “really cute” (int. 3).

Although Jana became happier and more at home in the host family over time, her host parents’ behaviour in the third month re-invoked negative feelings. Scarlett regularly criticized Jana and Luana for behaving improperly, such as showering too long or not washing clothes correctly. Andi, on the other hand, acted inappropriately in Jana’s opinion, for example, when he made her serve dinner for him without showing appreciation. Despite being aware that she was the ‘new’ family member and therefore had to accept cultural differences, her host parents’ behaviour left Jana feeling estranged. Jana was encouraged to change host families by Luana and a staff member at her school. She left with the uncomfortable feeling that they had never desired to build a relationship with her but only hosted international students to “fix their problems” (int. 4) and, particularly, for financial reasons. Jana was eventually also happy about parting from Luana because she felt that Luana demanded constant attention.

**HOST FAMILY 2**

After spending 3.5 months in New Zealand, Jana changed host families. Her second host family was Anna, a European host mother, Jamie, a Kiwi host father, and their one-and-a-half-year-old daughter. Both parents worked full-time (receptionist and IT professional) and typically returned home just after Jana. Compared to the first host family, Jana had ample opportunities for close day-to-day interaction with the family members both during regular activities (e.g. dinner) and activities in the evenings and at weekends. Pre-dinner rituals in the second host family involved all family members who regularly gathered in the kitchen and living room area after work to talk. Dinner was not eaten in front of the TV, which further allowed for meaningful conversations.

Jana’s second host family appears to have accepted her and treated her as a “person of consequence, deserving of access to everyday communicative practices” (Tan & Kinginger, 2013, p. 156). It seems that this attitude in particular, coupled with their genuine efforts to socialize with her and mutual feelings of affection, fostered a close relationship. For example, Jana mentioned that their conversations went beyond ‘shallow’ dialogue to extended interactions about her experiences, reflecting a true interest in her person:
they really ask me 'what did you do at school?' ‘do you have homework?’ ehm ‘oh how was that?’ and they remember things you know they are interested and so I have to talk more with them and I like it. (Int. 4)

Besides in-depth conversations, Jana’s second host family regularly invited her to join them for outings at the weekends, such as walks on the beach, visiting the markets, cafes, dining out, and even a baby shower. These activities often involved the company of family friends who Jana believed were “really nice and easy to talk to” (int. 6).

The host family also seems to have created welcome linguistic challenges for Jana. Compared to the first host family, Jana felt that she could learn English in conversations with Anna and Jamie. The following excerpt shows that she used these opportunities to negotiate for meaning:

There is a difference of their language use. It is another league of conversation, they really care about me. I see this as a good opportunity to learn and to ask for words I don’t know. (Report 3)

She also considered family evenings in front of the TV after dinner an opportunity to develop her language skills.

This close involvement with the second host family also increased Jana’s general sense of wellbeing, her knowledge about places and customs, and her confidence. Being with them resulted in a reduced need and desire for interaction with L1 peers.

Overall, Jana met the second host family’s inviting behaviour with much enthusiasm herself and believed that moving in with them was “the best decision” (int. 4). After only two months in the homestay, she developed genuine affection for them and left with the comfortable feeling that they hosted her not for money but because they truly cared.

4.3.2 The high school

Before Jana officially started college in New Zealand, she had already established contacts and become friends with the German-speaking international students at her college. She met Mona, who later became her best friend, on the journey to New Zealand through her exchange organisation/agency. Jana met all the other German speaking international students during the ‘international student orientation programme’ which was organised by the school a week prior to the official start of the semester. Jana first met the local students in the classrooms at the beginning of the first semester.
4.3.2.1 Inside the classrooms

At Jana’s college, international students were required to choose six subjects. Jana took Drama, Mathematics, English for Academic Purposes (EAP), Physical Education (PE), French, and Tourism. Jana chose to attend EAP classes because she believed that she would need the skills the class promised to teach for EFL classes in Germany (e.g. writing essays and making presentations). Unlike most other international students at the college, Jana was placed in Year 13 for PE and French classes due to her age (she turned 18 while in New Zealand) and timetable conflicts. Jana studied her other subjects alongside Year 12 students (like most international students). The student composition of different classes varied (Table 21). Only local students attended her Tourism class whereas only international students (mostly Germans due to their similar L2 proficiency level) attended EAP. Maths, Drama, and French classes contained a mix of local and international students, including German speakers. PE classes were attended by local students and one South American international student.

<table>
<thead>
<tr>
<th>Year</th>
<th>PE</th>
<th>Tourism</th>
<th>Maths</th>
<th>French</th>
<th>Drama</th>
<th>EAP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>√</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>12</td>
<td>√</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Level 3</td>
<td>x</td>
<td>√</td>
<td></td>
<td></td>
<td>√</td>
<td></td>
</tr>
</tbody>
</table>

The different subjects afforded Jana different opportunities for L2 use. Possibilities for initial contact and subsequent interaction with local students in the classrooms appeared to depend on 1) student composition in the classrooms (e.g. classes with or without other international students), 2) seating arrangements (i.e. sitting next to a local student, a German speaker, or another international student) and 3) the level and type of interaction the subject demanded of the students (e.g. playing badminton vs. studying at the table quietly). Interactions with teachers were rare and depended on Jana’s perception of their attitudes and behaviour towards her. Jana’s level of engagement in classroom activities appeared to be determined by her level of motivation for each subject and for schoolwork in general.

During the course of her sojourn, Jana successfully built close social networks with local students from a few classes. Some of these friendships extended beyond the classroom into
break times and after-school activities (see 4.3.3). Other subjects did not afford Jana with equal opportunities for engagement with local students, and, subsequently, were not facilitative of establishing friendships with Kiwis. Jana’s study efforts were generally minimal and she tended to prioritize social relationships over school.

In Drama, Maths and EAP, Jana did not develop relationships with local students. This seems to have been mainly the result of the student composition of the classroom and seating orders. Jana attended Drama classes with several German-speaking international students including her best friend Mona. Although she initially described Drama classes as very relaxed and thus conducive to establishing contacts with local students, she did not end up making L2 friendships. Jana’s Math classes consisted of local students, German-speaking international students, and other international students. Jana sat next to her South American host sister Luana and did not interact with Germans during this class. However, like in Drama, Jana did not report developing friendships with local students in the Maths class. Jana and Luana sometimes used class time to chat in English about their host family. Mostly, however, the girls day-dreamed or did creative activities and did not pay attention to classroom procedures:

> I'm sitting in the ehm lesson and I'm just drawing pictures I don't know really I do nothing and I'm sitting next to my ehm host eh sister and ehm I think she’s nearly all the time eh with her on her phone. (Int. 3)

EAP classes were specifically designed for international students and opportunities for interactions with local students were therefore unavailable. Jana was frustrated that the EAP classes compelled her to interact with German-speaking people whom she tried to avoid to become closer to the Kiwis. This prompted her to suggest that future SA students not do EAP English and choose mainstream English instead.

In these three subjects, Jana also devoted minimal effort to classroom work. In Drama, for example, Jana had to prepare for a theatre performance with her classmates. Practice mainly included learning lines and reciting them on stage or in front of a camera. Although Jana reported feeling challenged by the dynamic nature of the play’s dialogue which made her insecure, she expressed an unwillingness to put effort into the subject:

> I have ehm a drama performance tomorrow and I still don't know my lines cause I'm so busy and . . . I'm just not in the mood to learn my lines cause I think it's my last week and I'm I have better things to do than to learn my lines. (Int. 6)
Jana’s attitude towards the EAP classwork was similar. She initially said that she would take English very seriously but appeared to keep study efforts at a minimum. For example, Jana only ever learnt the words for the weekly vocabulary tests a few minutes before class which she thought was enough. At the end of her sojourn, Jana said she was disappointed with the subject:

they said you're doing essays and you know things like that and I thought ‘oh it's cool because I need it’ but then in the end it was just not really good. (Int. 6)

Jana did not consider Maths important and did not feel the need to study for the tests or do much homework. She felt that her social activities outside of school were much more important.

Unfavourable attitudes towards the Maths and Drama teachers inhibited Jana’s desire to interact with them beyond what was necessary. She felt that these teachers considered international students a “disruptive inconvenience” (Tan & Kinginger, 2013, p. 160). For example, in Maths, Jana initially showed an interest in handling the classroom demands for which she required the teacher’s help. However, when the teacher did not provide her with satisfactory assistance, Jana said she had given up doing work and refrained from further interaction with the teacher.

Jana also harboured a dislike towards her Drama teacher because he made jokes about the European international students. This did not result in withdrawal, but threatened Jana’s identity as a German national. For example, the teacher made fun of a German girl’s use of the taboo word ‘fucking’:

my drama teacher annoys me so much cause he always makes his jokes with the internationals and I don't know it's a joke but sometimes I take it serious because I think there's a little bit of truth in it or he thinks like that I don't know why but ehm today he said oh oh yeah ehm the bride had to say 'I need a fucking miracle' and she wasn't supposed to say 'fucking' but she didn't know it and because she's international as well and ehm then one girl said ‘ah that's so rude’ and he said ‘oh that's so Europe’ and I was like ‘are you kidding me?’ you know it's supposed like a joke but it just yeah … and most of them they laugh about it but Mona and I we are always like ‘hmm funny’. (Int. 6)

Compared to EAP, Maths, and Drama classes, the conditions of the French, Tourism, and PE classes were facilitative of developing friendships with local students. In all three subjects the determining factor that encouraged the formation of social networks appeared to be the student composition of the classroom, particularly the seating arrangements. In French, Jana became
close friends with Kate, a Kiwi girl. Jana sat next to Kate despite several other German-speaking internationals attending the class. This presented ample opportunities for lengthy English conversations during French, for example about the school ball or ‘boys’. Choosing not to be part of the ‘international group’ was important to Jana not only because it allowed her to connect with local students but also because she considered the international students’ behaviour during classes as inappropriate and did not want to be associated with it:

I think you're in another school and you're there as a guest you know and you have to respect their rules even if it's so different to Germany but most of them they don't and in French they are so loud and I am so embarrassed cause I'm not part of their group so I'm sitting with Kate and they are so loud and it's so rude to be loud and then to talk in German. (Int. 6)

Jana considered Kate trustworthy and an important source of knowledge about social relationships at the school. She also found in Kate an ideal partner for stress-free English conversations. Kate was understanding of Jana’s situation as a second language speaker because she herself spoke two other languages and was therefore aware of the challenges involved in speaking a foreign language. After roughly four months into Jana’s sojourn, Jana and Kate began to meet outside the school for coffee and parties (see 4.3.3).

In Tourism, Jana was the only international student. She sat next to the Kiwi girls Ella and Claire with whom she conversed a lot during class due to the ‘un-demandingness’ of the subject and because it was enjoyable. Jana was particularly fond of Claire and the friendship appears to have been characterized by a “two-way exchange of learning opportunities” (Tan & Kinginger, 2013, p. 159). Jana profited from the friendship with Claire because she knew “really good stories” (int. 1), she could practice her English in a non-threatening way, and was provided with potential access to further social networks outside the school. Claire, in turn, seemed to show an interest in knowing about life abroad or life as an international student. Ultimately, the girls’ friendship did not evolve beyond classroom interaction. Because Jana interpreted Claire’s behaviour toward her as oscillating between a desire for proximity and distance, she felt that her status as a close friend was not sufficiently validated to develop it further:

she is really nice but it's ehm I don't know she changed a bit in the last last week maybe I don't know and sometimes she comes to me and said ‘ah Jana what's going on?’ and she tells me everything and the next day she just said ‘hi’ so I don't know it's just we are friends but just in the lesson but that's fine. (Int. 3)
PE classes also consisted of local students as well as one South American international student. In this setting, Jana built increasingly closer relationships with the PE girls which involved social activities outside the classroom (see 4.3.2.2) and outside the school context (see 4.3.3). Practical activities and sitting next to local students appear to have played significant roles in creating opportunities for social interactions. After sitting next to the South American student for approximately a month, Jana decided to change seats and sit next to a Kiwi girl. In retrospect, she considered this action as the key moment that provided her with access to the PE girls:

I remember ehm we were sitting in the classroom ehm I don't know why I was always sitting with ehm the one [South American] guy and then one day I was just I thought ‘oh I could just sit there’ and I sit next to Marie and she said ‘ah you're so pretty’ and then we talked and then it was my seat there. (Int. 6)

At the same time, the class started to practice different sports in the gym and went on an outdoor kayaking trip. Being involved in these sports activities, as opposed to doing theory in class, helped Jana forge friendships. Jana interpreted the PE girls’ inviting behaviour during the practical activities as very positive and felt that this enabled her to become part of that group.

4.3.2.2 Outside the classrooms

Jana spent her school breaks (morning tea and lunch) with both German speakers and local students. However, the nature and intensity of these contacts changed considerably over time both because of her pro-active behaviour and as a result of friendships that developed from classrooms contacts (see 4.3.2.2). Jana actively reduced interactions with her German speaking peers while increasing her interactions with local students. Three broad phases could be noted in this process: the first phase was marked by close interaction with other German speakers and a lack of Kiwi friendships. The second phase was marked by separating from the main German speaking group with another German friend (Mona) and participating in local social networks. In the third phase, Jana almost exclusively interacted with Kiwi friends.

During the first 1.5 months, Jana spent all school breaks and her leisure time with the German speaking international students. The German group consisted of both new arrivals (including Jana) and second-semester students. Two local students (international students’ friends) often also joined the group. During the school breaks, the German-speaking group physically sat apart from the local and other international students. According to Jana, this group constellation, despite the presence of a few English native speakers, was not conducive to
speaking English. She reported that speaking English to Germans was uncomfortable because it made her conscious of her L2 shortcomings:

it's ehm it's easier for me to talk in English to natives because if I talk in English to Germans I'm always like ‘ah they kn they maybe recognize my mistakes because it's stupid’ but it's my feeling so I it's I don't like it to talk eh in English to other Germans cause I start to think about what I'm saying and if I'm together with ehm New Zealanders or ehm then I have to say it and I've no choice and it's more easy so I just do it. (Int. 2)

After 1.5 months at the college, Jana realised that spending all breaks and leisure time with other German speakers not only limited English language learning opportunities but also the possibility of establishing social networks with locals. For these reasons, Jana decided to leave the group of internationals together with Mona. The process of leaving the internationals was gradual. First, Jana and Mona began to sit separately from the German-speaking group during some breaks, which was very successful as they were immediately approached by Kiwis. In mid-April, Jana and Mona decided to cut all ties with the Germans and the pair completely ‘split up’ from the group of internationals despite the group’s disappointment. They subsequently spent the breaks with different groups of local students. Jana and Mona also attempted to increase their L2 practice by speaking English with each other:

we try [speaking English] we really try it sometimes if some people can hear us and we don't want that they can understand us then we speak German and so yeah maybe half of the time we speak German and half of the time ehm English. (Int. 2)

Soon after, Jana was asked by her French classmate, Kate, to spend morning tea or lunchtimes with her and her friends. Both Kate and some of the PE girls were prefects and spent the breaks together at the ‘prefects’ table’. Despite being nervous to join this group, Jana felt she could not jeopardize this chance and accepted the invitation. Spending the breaks with the Year 13 students constituted ideal circumstances for Jana to cultivate her friendships with the girls from French and PE and make new friendships. Among the Year 13 boys, Jana attracted considerable interest and became popular. One boy, Connor, apparently fell in love with Jana and regularly asked her to spend time with them in the breaks and to join them for parties. Although Jana did not reciprocate his feelings, their friendship facilitated her access into the prefects group:

through Connor everybody knew me so and then I wasn't a part of that group then but then the people talked about me and then it started yeah. (Int. 6)

Jana’s increasingly fuller membership in the Kiwi groups changed her relationship with Mona. Although Jana was initially grateful for Mona’s companionship, which empowered her to leave the German group, she desired to spend less time with her over time. In the second school term,
Jana and Mona agreed to spend only one of the two school breaks with each other and the other one individually with Kiwi friends. Jana felt that this was necessary to avoid giving others the impression that she preferred spending time with Mona. She was also afraid that speaking German with others would keep locals from establishing relationships with her. Jana was also not prepared to give Mona access to her group of friends as she feared that Mona’s presence in the newly-established group of Kiwi friends would threaten her own membership status:

she doesn't know the girls and the girls doesn't know her and I was afraid I really like to hang with her but I'm afraid when she's here when she's there with me and the girls that you know it's hard enough to talk with them to be part of their group and I don't want to have somebody else … it's not safe you know. (Int. 4)

In June, as Jana’s social networks with Kiwis became stronger and she met with them outside of school, her relationship with Mona continued to weaken. The change of host family, who lived further away from school, also resulted in less contact with Mona. Although Jana remained friends with Mona, she ultimately felt so closely integrated into her Kiwi networks that she desired to study in New Zealand without other international students.

4.3.3 Outside of the high school

The social activities Jana engaged in with her friends outside school mirrored those in the break times. During the initial two months, when her friendships with Kiwis were still new, she spent her free time with German speakers only. Mid-way through her exchange semester, Jana’s increasingly closer relationships with Kiwis, the change of host families, and her participation in the football team resulted in increased contacts with her Kiwi friends outside school, and reduced outings with the Germans.

4.3.3.1 Social activities and trips

In her first month abroad, Jana and her German-speaking friends undertook various activities, such as playing tennis after school, daily visits to the beach, celebrating each other’s birthdays, and going to concerts and festivals at weekends. During these times, Jana said that she did not talk in English. After Mona’s and Jana’s complete separation from the international student group at school, the two girls spent most of their leisure time together. Jana reported having felt “very blessed” to have found such a good friend in Mona and could not imagine doing the exchange without her (blog 01/04/14).
After roughly four months, Jana also began to meet the New Zealand girl Kate (the friend she made in French classes) outside of school for coffee. A month later, they went out for pizza nights at friends’ places, and parties. The latter two activities often also involved Kate’s Kiwi friends and Jana reported that spending time with Kiwis was invaluable to her.

Towards the end of Jana’s sojourn, social activities outside the school also began to take place with the PE girls because they turned eighteen and were legally allowed to drink alcohol and enter nightclubs. Jana was invited to a PE girl’s eighteenth birthday dinner and she went clubbing twice. Going out not only assisted in becoming closer with the PE girls but constituted opportunities to make further contacts with friends that joined them for the night out or the people that offered lifts to these events (i.e. parents and other friends). Some events also became the topic of the next day’s ‘gossip’ at school for which Jana was approached by strangers:

that school is just they only care about rumours and gossip … and ehm it was just so funny because some people looked at me and said ‘oh hey I heard about town’. (Int. 6)

The school ball in Jana’s final week in New Zealand constituted the most important and last event that she attended with her PE friends. The girls asked Jana to join them for the main event, and invited her to a pre-ball and an after-ball which was exclusive to Year 13 students. The latter included a sleep-over at the host’s house and breakfast the next morning where she had to say goodbye. The PE girls’ invitations to these events signalled to Jana that she was finally considered a full member of the group:

I mean I’m invited to the pre-ball and after-ball and I’m part of them especially with the PE girls ehm I didn’t realise that before but I saw this when we went to town when they invited me you know and I s realised ‘oh I’m part of their group’ you know. (Int. 6)

Jana, unlike all the other German speaking internationals, decided against going on the South Island trip during the Easter holiday (organised through the college) but instead to visit family friends (a German-Australian couple with a four-year-old daughter) in Australia. Jana spent most of the week exploring the near-by city on her own, or went for walks with her German friend. She was also provided with some opportunities for speaking English with her friend’s Australian husband and daughter, as well as during service encounters in the city. Jana greatly enjoyed the family’s company and was proud having travelled alone.

4.3.3.2 Hobbies

In March, Jana joined the football team which included the two PE girls Rachel and Eleanor, some South American international students and another German-speaking girl. During
training. Jana consciously decided against spending time with the German-speaking girl because she said she would prefer being with Kiwis. In particular, Jana enjoyed spending time with Rachel and Eleanor outside of school. She got to see them three days after school for football and was invited to drive to games in Rachel’s car. Jana’s connection with the locals also helped her ‘adopt’ her desired Kiwi identity:

[Our] coach said that we are without internationals for the game and I corrected her, she looked at me and told me that I’m not a real exchange. I was so happy to hear that. (Blog 11/06/14)

4.3.4 Summary

Jana’s opportunities for English language interactions varied between and within all three social contexts. She continually reduced the extent to which she spoke German to her L1 peers, and increased the extent to which she communicated in English with her L2 peers and host family members.

4.3.5 Perception of L2 progress

Besides analysing Jana’s data for instances of language interaction, I also scrutinized it for references to her perceptions about her own English learning progress, goals and motivations. In the following, I will summarise and comment on Jana’s own perceptions of her language learning progress during her study abroad.

Jana’s language-related goals for the SA period mainly concerned improving speaking and listening abilities and developing a greater lexicon. Jana did not report a desire to improve her grammatical abilities. She generally believed that her grammar was good and that the occasional difficulties she encountered with tenses were not important enough to pay attention to. Rather, her comments indicated that she took a functional approach to language learning:

I think my grammar is ok . . . tenses and those things I want to yeah they are sometimes a bit mixed but ehm yeah I'm not focused on my grammar . . . it's more about ta talking and speaking listening. (Int. 1)

Jana’s perceptions of her linguistic progress were mixed. In interview 2, she felt she had become more confident but generally was dissatisfied with her oral performance and felt a need “to improve it more” (int. 2). She experienced difficulties when using language spontaneously due to a lack of planning opportunities:
I was really good in English in Germany so in my subject English hmm because you have time to think about everything and I ehm now I'm here and I really have to talk spontaneously so it's not bad I think but it's I'm not satisfied. (Int. 2)

In report 1, Jana mentioned that understanding and listening to English had become easier and connected this progress to watching movies in English.

In interview 3, Jana believed that ‘splitting up’ from the German-speaking group helped her English speaking skills a lot. In particular, she reported getting used to talking spontaneously through her interactions with Kiwi friends. Despite making progress, she admitted that she was seldom satisfied with her performance, particularly her language accuracy (see below for a discussion of this apparent contradiction):

I'm still the little perfectionist who thinks ‘ah is that correct? no. (Int. 3)

In interview 4, Jana stated that she spoke English on a ‘higher level’ since she changed families. While still believing that her grammar could improve, Jana declared that she could understand everything without actively listening. She also expressed surprise at the apparent improvement in her vocabulary:

In the past month I sometimes was surprised about my vocabulary and I wondered where that knowledge came from I guess this is because of the natural learning I’m not studying the words I just hear and keep them without knowing it. (Report 3)

Despite these achievements, Jana once again referred to her perfectionist nature in interview 4, claiming that she would never be satisfied with her progress. In interview 5, Jana repeated that conscious attention was no longer needed to follow the teachers in class. She also mentioned that she had stopped paying conscious attention to her use of English and that she had lost her ‘fear’ of saying something wrong. She also mentioned caring less about others’ judgements of her L2 proficiency level. However, Jana also reported feeling uncomfortable when native speakers noticed errors in her speech. Although she appreciated their feedback, she perceived the act as face-threatening:

I talk more in English since I started dating some boys. With some of them it is really easy to talk and others always try to correct me. I can’t say what I prefer, I’d rather know about the mistakes I make but on the other hand it is a bit frustrating. I think the best way to learn is to face the mistake, that’s why I try not to be offended if they correct me and see it as a chance to improve my English. (Report 4)
Jana finished the sojourn feeling “not really happy” (int. 6) about her English progress. She particularly lamented the fact that despite the establishment of her social networks she still spent much time speaking German:

> even if I was friends with Kiwis maybe more than anybody else from our group I talk so often in German ehm I thought when I came here that I would only talk in English for six months you know and I didn't. (Int. 6)

She was disappointed that speaking English did not come more naturally or sub-consciously. Jana left New Zealand in the belief that further improvements may have occurred given more time in New Zealand and if all the internationals had left.

It is interesting to note that Jana’s approach to language learning and her feelings about it did not appear to align. On the one hand, she asserted she was a risk-taker who prioritized a functional approach to language learning (i.e. she did not want to focus attention on grammar during speaking but ‘just speak’). On the other hand, she reported being anxious of making mistakes in front of other Germans and of having errors pointed out by native speakers. A possible explanation of this contradiction may lie in the tension between Jana’s genuine desire to focus on meaning during speech and the “necessary consequence of educational systems putting emphasis on ‘correctness’” which inevitably makes students feel like “failures since only relatively few, exceptional, individuals will achieve this ability to hold conversations in which they produce exclusively ‘correct’ and ‘complete’ forms” (Brown & Yule, 1983, p. 22).

### 4.4 L2 development

The following section provides a detailed account of Jana’s L2 development. Various measures of complexity, accuracy and fluency were employed to assess her L2 performances. Jana’s results were compared to those of previous studies to allow for a meaningful discussion of her L2 development. This section’s argument is organised starting with complexity, followed by fluency and accuracy. I adopted the same structure for all three case study chapters.

#### 4.4.1 Grammatical complexity

To examine changes in grammatical complexity, I investigated the amount and type of subordination (subordination index and types of subordinate clauses) and the mean length of clause in the participants’ data. Pruned measures were used for the latter to ensure that repairs and filled pauses did not count towards measurements of complexity (Foster et al., 2000).
4.4.1.1 The subordination index

The number of clauses Jana used per AS-unit (cpu) varied with no clear trend \((r=.185)\) (see Figure 2). There was a small overall decrease in the subordination index (SI) of 5\%\textsuperscript{ix} between interview 1 (1.46 cpu) and interview 6 (1.39 cpu). With the exception of an atypically high SI score in interview 4 (1.74 cpu), the scores varied within a small range of 1.29 and 1.46 cpu.

Compared to previous results for SI obtained in informal interviews for native speakers (NSs) and non-native speakers (NNSs), Jana’s scores (≥1.44) were higher on average or within the range of some NSs (Nippold, Hesketh, Duthie, and Mansfield, 2005: max ≥1.39 cpu; range: 1.08-1.82). But they were lower than those measured on average for the NNSs and NSs in Mora and Valls-Ferrer’s (2012) study (NNSs ≥1.60-1.62, NSs ≥1.83), in Llanes et al.’s (2012) study (NNSs ≥1.70/1.87) and in Polat and Kim’s (2014) study (NSs ≥1.97; range: 1.82-2.22). I suspect that there are differences in clause definitions between these studies (e.g. the treatment of non-finite structures), making comparisons difficult.

The comparisons suggest that Jana’s utterances were relatively simple. However, this characteristic is typical of spontaneous speech since “[t]he needs of real-time communication do not allow the speaker time to construct over-elaborate patterns of main and subordinate clauses” (Carter and McCarthy, 2006, 87a). Previous work (Ferrari, 2012; Lambert & Kormos, 2014; Pallotti, 2009) has also emphasized that changing SI scores may simply reflect variations in the speech activity (not progress). In Jana’s case, the production of more complex utterances in interview 4 appeared to reflect more direct speech reports – a variation in the speech activity (see 4.4.1.1).
RANGE AND FREQUENCY OF SUBORDINATE CLAUSES

To obtain a more comprehensive picture of the types of clauses Jana used, the range and number of subordinate clauses was counted (Figure 3).

![Figure 3 Subordinate clauses for Jana](image)

As can be seen from Figure 3, Jana produced subordinate clauses of all kinds and, with the exception of non-finite adverbial clauses, did so right from the beginning. The majority of subordinate clauses in her oral production were finite complement clauses (between 15 and 48 clauses per 100 AS-units). Finite adverbial clauses constituted the second most frequently used type of subordinate clauses (3-14 clauses). Non-finite complement clauses ranked third in terms of their frequency (1-7 clauses), and finite relative clauses and non-finite adverbial clauses hardly ever occurred in Jana’s data (a total of 4 and 3 times, respectively).

Compared to previous findings, Jana’s distribution of finite subordinate clause types appears appropriate to the communicative activity. Nippold et al.’s (2005) study showed that the most frequently used types of finite subordinate clauses in conversational discourse are complement clauses, followed by adverbial and relative clauses. With regard to the frequency with which Jana used each subordinate clause type, it seems that she overused finite complement clauses (particularly in interview 4) and underused finite relative clauses (Nippold et al., 2005). I will show below that the heavy focus on finite complement clauses often took the form of direct speech complements. The non-finite structures that were given clausal status in this study are considered later acquired features and less common in spontaneous speech than formulaic non-
finite structures (e.g. ‘I wanna go’) (Diessel, 2004). Hence the scarcity with which Jana used non-finite clauses is considered normal.

The findings provide evidence that Jana was able to form all types of subordinate clauses, indicating advanced levels of English. It is unclear, however, whether the production of relative clauses and non-finite adverbial clauses, for example, was difficult for Jana and she therefore used them infrequently, or whether this was simply a response to the communicative activity.

**DIRECT SPEECH COMPLEMENTS**

The case of direct speech complements deserves extra attention. Section 4.4.1.1 showed that interview 4 contained a particularly high SI and high numbers of finite complement clauses. A closer examination revealed that a large number of complement clauses in this interview were direct speech complements (Table 22). They accounted for 73% of all finite complement clauses and for 33% of the total number of clauses (including independent clauses). Direct speech reports (both in the form of complement clauses and subsequent independent units) occurred in half (48) of all utterances in interview 4, in a third (31) in interview 6 but only in approximately 13 utterances in the remaining interviews.

<table>
<thead>
<tr>
<th>Interview</th>
<th>Raw number of reporting clauses per 100 AS-units (and number of dependent direct speech complements)</th>
<th>Raw number of following AS-units with direct speech clauses (and number of clauses involved)</th>
<th>Total number of utterances containing direct speech per 100 AS-units</th>
<th>Ratio of direct speech complements / total number of finite complement clauses</th>
<th>Total number of direct speech complements (ratio direct speech complements / total number of clauses)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9 (9)</td>
<td>6 (6)</td>
<td>15</td>
<td>9/24 (38%)</td>
<td>15/147 (10%)</td>
</tr>
<tr>
<td>2</td>
<td>7 (9)</td>
<td>5 (6)</td>
<td>12</td>
<td>8/15 (53%)</td>
<td>15/129 (12%)</td>
</tr>
<tr>
<td>3</td>
<td>8 (8)</td>
<td>5 (5)</td>
<td>13</td>
<td>8/16 (50%)</td>
<td>13/130 (10%)</td>
</tr>
<tr>
<td>4</td>
<td>33 (37)</td>
<td>15 (20)</td>
<td>48</td>
<td>35/48 (73%)</td>
<td>57/173 (33%)</td>
</tr>
<tr>
<td>5</td>
<td>11 (12)</td>
<td>3 (3)</td>
<td>14</td>
<td>11/27 (41%)</td>
<td>15/145 (10%)</td>
</tr>
<tr>
<td>6</td>
<td>24 (25)</td>
<td>7 (10)</td>
<td>31</td>
<td>24/34 (71%)</td>
<td>35/139 (25%)</td>
</tr>
</tbody>
</table>

Direct speech, as part of constructed dialogue, is used for storytelling and enables narrators “to take a stance on the represented words” (Lampropoulou, 2013, p. 20). It is considered a preferred form of speech for adolescents, particularly females, to convey their views and feelings. This speech mode has an expressive, dramatic, and evaluative function (Romaine
& Lange, 1991). The figures in Table 22 demonstrate that direct speech served Jana as a particularly effective means to report and evaluate the events in interviews 4 and 6, possibly because they were emotional to her. As will be seen in the following sections, Jana’s use of direct speech also significantly impacted on the scores she achieved for mean length of clause and mean length of run. I believe that Jana’s varying use of direct speech reports explains the ‘waxing and waning’ (Larsen-Freeman, 2006) patterns of these measures.

4.4.1.2 Mean length of clause

Jana started her sojourn with a mean length of 6.13 words per clause (wpc) and finished it with 5.63 wpc (Figure 4), which accounts for an overall decrease of 0.5 words or 8%. Only in interview 3 (6.46) did she use more wpc than in interview 1. In interview 4, she produced the overall shortest clauses (5.20 wpc). The mean length of clause (MLC) development is marked by much fluctuation, as mirrored in a low negative correlation coefficient of -.385.

In comparison to Polat and Kim’s (2014) study, Jana’s scores were within the ranges measured both for the NNS (4-6.67 wpc) and the three NSs (5.05-6.76 wpc). Since even NSs vary in their performance, Jana’s fluctuations can be interpreted as normal; possibly even the overall MLC decrease. Fluctuating levels of MLC did not appear to be indicative of rising or falling proficiency levels but of different narration styles.

Below I will provide a few explanations concerning the nature of fluctuation measured for Jana’s MLC levels. A particular focus will be placed on the impact of direct speech on complexity levels.

In utterances containing direct speech reports, the reporting clauses were typically formulaic and lacked complexity (i.e. they were short). They often only involved a coordinator (e.g.
'and’), a subject, plus the reporting verbs ‘said’ (Example 17 and Example 18), ‘thought’ (Example 19 and Example 20), or ‘be like’ (Example 21 and Example 22). Sometimes, they included an object. To fulfil their function as indicators of direct speech reports, no additional elements were absolutely necessary in the reporting clause. Typically, the elements of the reporting clause added up to a length of three (e.g. ‘and I thought’) to four (e.g. ‘and we were like’) words. The length of the direct speech complements, on the other hand, varied greatly, ranging from very short clauses consisting of exclamations only (Example 18) to long clauses containing prepositional phrases and complex verb phrases (Example 21).

Example 17 (int. 4): (.65) we said :: ‘ok why not’.

Example 18 (int. 4): and she said :: ‘ah shh’.

Example 19 (int. 4): (.36) and I thought :: ‘ok something's wrong’.

Example 20 (int. 4): (.55) and I thought :: ‘oh I should ask her’.

Example 21 (int. 4): and we were like :: ‘ok but she just came back from her holidays’.

Example 22 (int. 4) I was like :: ‘ok I didn't know :: that it (.33) would work that easy’.

Conversely, when Jana made little use of direct speech, she reported events and thoughts by means of more factual descriptions. Often, they were relatively detailed, involving background information and evaluations intended to make a ‘moral’ point (Lampropoulou, 2013). Utterances of this type included, for example, adverbial phrases (‘after school’ in Example 23, ‘the next time’ in Example 24), prepositional phrases (‘in the last weeks’ in Example 23), adverbs (‘really tired’ in Example 23), and more complex sentence constructions (e.g. conditional clauses in Example 24). Evidently, reports like this went hand in hand with grammatically more complex language (i.e. longer clauses).

Example 23 (int. 3): (.49) and (.58) I have to say (.99): in the last weeks I was really tired after school.

Example 24 (int. 3): because if somebody doesn’t explain it to you :: you won’t ask the next time.

Different narration modes perform different communicative functions and require different levels of linguistic precision (Brown & Yule, 1983). Describing a feeling or action comprehensibly calls for the use of more adjuncts than reporting past events by means of direct speech reports. Isabelli-Garcia (2004) argued that detailed narratives reflect higher levels of L2 proficiency than simple narratives. In the case of direct speech reports, simpler (i.e. shorter)
utterances may be a sign of more appropriate communicative competence. It is possible that the direct speech mode is acquired late(r) in the developmental process of L2 speech for students whose education is heavily based on the acquisition of academic discourse genres, as implied by Renouf and Kehoe (2011). In this regard, direct speech would reflect complexity as a sign of a later acquired language construct.

To determine exactly the relationship between narrative modes and clausal complexification, a thorough investigation is required, which was beyond the scope of this study. The findings above, however, appear to support the conclusion that the use of direct speech reports significantly contributed to lower MLC scores in interview 4.

4.4.2 Lexical diversity

In this study, lexical diversity was measured using Giraud’s Index. Giraud’s Index is a sophisticated type-token measure, calculated by dividing the number of types by the square root of the number of tokens. It is denoted by ‘R’.

![Figure 5: Lexical diversity for Jana](image)

Figure 5 shows that Jana began her sojourn with R= 8.49 and ended it with R=8.08 (-5%). Her lexical diversity (LD) scores fluctuated as implied by a low negative correlation coefficient (r=-.429). Jana reached the lowest score in interview 4 (R=7.38).

Compared to previous studies who measured LD of NNSs spontaneous oral speech by means of Guiraud’s Index, Jana’s values were higher at all times (Levkina & Gilabert, 2012: R= 5.84-6.24; Daller & Xu, 2007: R= 5.03 -6.18; Lu, 2012: R= 3.3-6.9; Mora & Valls-Ferrer, 2012: R= 6.76-6.97). Jana’s average LD value (R= ております7.97) was even above the average scores measured for the NSs in Mora & Valls-Ferrer (2012) (ております7.86).
While her linguistic repertoire was advanced from the beginning, it appears remarkable that the measures following interview 1 were below Jana’s first score. Previous studies have shown that “learners who live in the target language community, [sic] increase their performance in this variable to the extent that there are no more significant differences compared to native speakers” (Götz, 2013, p. 66). Another study has shown a close correlation between an increase in speech rate (which Jana demonstrated, see 0) and greater LD (Milton, 2009).

However, changes in LD reflect not only L2 progress but also factors related to discourse choices. As for MLC, the fluctuations in Jana’s LD levels appear to have been mainly influenced by her choice of narration modes. Sharing experiences predominantly by means of storytelling (i.e. using direct speech) seems to have resulted in the production of simple and repetitive vocabulary (i.e. lower LD scores). Narratives in which Jana gave more factual and detailed accounts seem to have triggered the use of lexically more diverse items (i.e. higher LD scores).

Two excerpts from interviews 3 (high LD scores) and 4 (lowest LD scores) demonstrate the difference that narration modes can make on LD scores. In Example 25, Jana explained why she believed her Maths teacher ‘hated’ all the internationals. To build a convincing argument, she referred to the different people involved in the situation; the activities that unfolded as a lead up to the situation and those that happened during the situation; the teacher’s attitudes and behaviour; and Jana’s evaluation of the situation. To do this, she used relatively infrequent and context-specific nouns (e.g. ‘assessment’, ‘dean’), collocations (e.g. ‘drawing pictures’), adverbs (e.g. ‘nearly’), and adjectives (e.g. ‘ridiculous’).

Example 25: I’m sitting next to my ehm host eh sister (1.34) and ehm (2.19) I think (.65) she’s (1.59) nearly all the time eh with her (.59) on her phone (.77) and I think our teacher she (1.03) noticed that already (.50) but she says nothing because she really she doesn’t care about us (.62) but if we don’t do the homework (.47) or (1.68) with the assessment (.56) she cares about us so we have to do the homework and otherwise she will send an email to us (.73) to our host parents (.71) which is really ridiculous (1.19) ehm (.30) I talked with eh our dean (.43) about that. (Int. 3)

In contrast, Jana’s use of the direct speech mode to report the events that led up to her host family change in interview 4 resulted in the use of formulaic (see 4.4.1.2) and repetitive (e.g. ‘and’, ‘said’, ‘ask’) language.

Example 26: (.93) and then I (1.05) thought ‘ok (.43) maybe it’s better if I ask her’ (.48) and I (.65) ehm (.30) said to her ‘oh I really need it’ (.52) and she said ‘ah you don't have to ask this is your home I don't want you to ask and I said ‘yeah but last time you
were a bit angry’ (.68) and she said ‘no you just have to be careful with the water (.77)
and (1.38) your showers are ok but Luana’s are too long’ (1.00) and ehm (.49) yeah
things like that. (Int. 4)

The discussion above suggests that there is a strong relationship between discourse choices and
LD. Evidently, fluctuations deriving from such decisions can only emerge when learners’
conversational language is assessed (i.e. when they have complete freedom in their language
choices). It is important to note that at the advanced learner level, less complex language is not
necessarily indicative of L2 ‘attrition’. Rather, Jana showed an ability to adapt her language
to the situation at hand. Sometimes, she may have perceived it as more appropriate to be
‘entertaining’ and thus ‘perform stories’ rather than provide me with complex descriptions or
analyses of her experiences in New Zealand. This duality of attention devotion has been
described in Kormos (2011). She claimed that what speakers pay attention to depends not only
on ‘task-demands’ but also on the context of social interaction (including power relations,
politeness norms of a given culture, social status, the stake of the given interaction etc.).

4.4.3 Fluency

Below I will present the results obtained from the frequency counts of three types of fluency:
breakdown fluency, speed fluency and repair fluency. The cut-off point for empty pauses was
set at 0.3 seconds. I also counted the range and number of smallwords.
4.4.3.1 Breakdown fluency

EMPTY PAUSES

Figure 6: Empty pauses for Jana

Jana started her sojourn with 18.76 empty pauses (EPs) per hundred words (phw) and finished it with a slightly lower number (16.74 phw), making for an overall reduction of 11%. Interview 4 contained the lowest number of EPs (15.19 phw). The development of EPs was moderately linear (correlation coefficient = -.696).

Jana’s (⌀16.95 phw) performance resembled those measured for the highly advanced German speakers of English in Götz’s (2013) study both with regard to their average and range (⌀15.13 phw; range: 8-30 phw). Her results are also comparable to those found for the ‘fluent’ participants in Riggenbach’s (1991) study (10-22 EPs phw). They were well below the range within which her ‘nonfluent’ group performed (33-73 EPs phw). Relative to Götz’s NSs (⌀ 3.99 phw), however, Jana used over four times as many EPs and performed outside their range (1.5 - 9 EPs phw) at any given time.

The comparisons suggest that Jana’s performance was advanced but did not come close to NS norms. An overuse of unfilled pauses in NNS English speech has been considered a “major and crucial problem area even for highly advanced German learners of English” (Götz, 2013, p. 98). It is likely that in interview 4, the heightened use of direct speech reports, which involved formulaic (i.e. well memorised) clauses, alleviated the cognitive burden of online processing and enabled her to pause less.
Figure 7: Filled pauses for Jana

Figure 7 shows that Jana initially used 4.26 filled pauses (FPs) per hundred words (phw) and finally produced 3.64 FPs phw, making for an overall reduction of 15%. Jana used the lowest number of FPs in interview 4. The development fluctuated without a clear trend. This is confirmed by a “negligent” (Mukaka, 2012, p. 71) correlation coefficient of -.097.

On average, Jana used 2.98 pauses phw. Her scores were well below those scored by adult NNSs (Götz’s, 2013: 5.12 phw; range: 1-14) and adolescent NNSs (Hasselgren: 5.98/7.78phw). They were very similar to the averages scored by Hasselgren’s (2002) adolescent English NS group (2.81) and slightly above Götz’s (2013) adult NS group (2.27; range: 0.5-8 phw). In contrast to Götz who found that “even advanced learners need more planning devices [i.e. filled pauses] than native speakers” (p. 113), it appears that Jana’s FPs were a non-problematic area of her English. All of her scores lay within the range documented for NSs, suggesting that the variation in FPs reflects native-like spontaneous speech behaviour.
### 4.4.3.2 Speed fluency

**Speech Rate**

Figure 8: Speech Rate for Jana

<table>
<thead>
<tr>
<th></th>
<th>121.02</th>
<th>144.61</th>
<th>128.51</th>
<th>148.23</th>
<th>152.65</th>
<th>144.18</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>120</td>
<td>130</td>
<td>140</td>
<td>150</td>
<td>160</td>
<td>170</td>
</tr>
</tbody>
</table>

Figure 8 shows that Jana’s unpruned Speech Rate (SR) increased from 121.02 words per minute (wpm) in interview 1 to 144.18 wpm in interview 6. She made an overall improvement of 19%. Her SR development took a relatively sharp dip in interview 3 and a slight dip in interview 6. A correlation coefficient of .693 implies that the SR development was moderately linear.

Compared to Lennon’s (1990) intermediate-level German speakers of English, Jana spoke much faster than three of them (74-99 wpm) but slower than the fastest speaker at the end of her sojourn (162 wpm). Compared to Götz’s (2013) highly advanced NNSs, Jana spoke slower on average (≥139.86 wpm) but within the range measured for all individuals (≥160 wpm; 117-190 wpm). Jana spoke considerably slower than the NSs in Götz’s (2013) study on average (≥213 wpm) and none of her scores featured within their range (175-280 wpm). According to Tauroza and Allison’s (1990) categories of British English SR (i.e. NS scores), Jana would have been classified as a ‘moderately slow’ (120-160 wpm) speaker of English.

Faster SR indicates that the processes involved in speech production take place more automatically (i.e. less consciously) (Lennon, 2000; Kormos, 2011; Wood, 2012). SR becomes higher when words are retrieved more easily (i.e. quicker) as a result of practice and when syntactical and morphological rules are applied more automatically (i.e. quicker). Speech rate can also become faster when learners make use of formulaic sequences, since they can be retrieved in one step as “readily-assembled constructions” (Kormos, 2011, p. 46).
Jana’s increased SR suggests that the SA context was beneficial to the development of more proceduralized knowledge. However, higher or lower SR scores reflect not only L2 proficiency levels but also ‘task’ differences. The formulaic nature of direct speech reporting clauses did not appear to strongly affect her SR levels. However, the low score in interview 3 may reflect laborious processes in the encoder stage during what appeared to be a cognitively demanding interview (as evidenced by high complexity scores).

**MEAN LENGTH OF RUN**

Mean length of run describes the extent to which a speaker is able to plan an utterance before s/he needs to pause again (i.e. before an EP). This measure has implications for proceduralization skills of the speaker.

![Figure 9: Mean length of run for Jana](image)

Figure 9 shows that between interviews 1 and 6, Jana increased the mean length of run (MLR) by 12% from 5.19 wpr to 5.81 wpr. The development did not follow a linear pattern, which is confirmed by a moderately positive correlation coefficient ($r=.645$). There was a peak in MLR in interview 4 (6.54 wpr). This score gave way again to lower scores in interviews 5 and 6.

Jana’s runs ($\approx 5.80$ wpr) were relatively similar in length compared to the results obtained for NNSs of upper-intermediate /advanced proficiency levels (Mora & Valls-Ferrer, 2012: $\approx 6.17/5.93/7.50$ wpr; Götz, 2013: $\approx 5.88$ wpr). Her overall rate of change (+0.62 wpr) was significantly lower than that found in Mora and Valls-Ferrer (+1.33 wpr over 9 months). Compared to NSs, who have been shown to be able to encode between 8 to 10 words (Pawley & Syder, 1983), 10.3 words (Götz, 2013) or even 12.49 words (Mora & Valls-Ferrer, 2012) in one operation during spontaneous speech, Jana’s performance was less native-like. Note that
these comparisons must be taken with caution since measures differed with regard to the cut-off points for empty pauses.

Typically, it has been proposed that changes in MLR can be explained by the proceduralization of different types of knowledge, such as “knowledge of syntax and of lexical phrases” (Towell, 2002, p. 84). Higher MLR values have also been explained by the use of more formulaic language (Götz, 2013). Like for the complexity measures, direct speech reports seem to explain the fluctuating nature of the MLR development.

Three explanations are proposed which might account for the interrelationship between direct speech and long runs. The first explanation concerns the formulaic nature of the quotatives in Jana’s speech. Jana’s data shows that she made repeated use of the same three quotatives: 1) ‘pronoun + said’, 2) ‘pronoun + was like’, and 3) ‘pronoun + thought’. It is likely that these quotatives functioned as memorized clauses, and were not creatively constructed with each utterance. This would be supported by Tannen (1986, cited in Romaine & Lange, 1991). It appears that the formulaic reporting clauses that Jana used enabled her to “attend to planning tasks while leaving the production of the current clause ‘on automatic pilot” (Pawley & Syder, 1983, p. 222), leading to the production of longer runs.

With regard to Kormos’ (2011) speech production model, the relationship between direct speech quotatives and MLR can be explained by undemanding processes in the encoder stage. Kormos (2011) explains that when a communicative context calls for the application of a particular formula, the conceptual chunk activates “corresponding readily-assembled linguistic constructions stored as one lemma” (p. 46). As formulas are “produced faster and with less conscious effort than creatively-constructed elements of the message […] [t]he process of routinization lessens the cognitive load involved in planning the message” (p. 46). It appears that the production of direct speech was conceptually less demanding than the use of other speech modes (e.g. descriptions, opinions) because it allowed for the activation of formulaic chunks, which can be retrieved quickly and easily, thus making speech faster (or runs longer).

The second reason concerns the level of syntactical embedding of direct speech clauses into utterances. Romaine and Lange (1991) stated that the use of direct speech does not pose a challenge but “may be a simplifying device for speakers […] because it allows them to avoid some of the more problematic aspects of syntactic and semantic incorporation, such as deictic shifting required in the indirect mode” (p. 268).
The third reason concerns morpho-syntactical properties of direct speech complements. The examples above showed that most direct speech reports were expressed in the present simple tense (arguably to draw the interlocutor closer to the happenings of the story). The use of present tense has been associated with early stages of L2 (and L1) learning and can thus be considered cognitively less demanding than other tense forms (Ellis, 2008). This would also make it quicker for retrieval (i.e. knowledge has been proceduralized).

Clearly, the use of direct speech reports, though powerful in its explanatory power, cannot account for all MLR changes. The production of longer or shorter runs can also be significantly influenced by other kinds of ‘demands’ on speaking processes, such as the macrostructure of a speaker’s narration (i.e. number and sequence of elements), differences in topics, or epistemological memory searches. Jana may have also placed pauses intentionally before certain words for stylistic effects (i.e. drawing attention to what is coming). For a conclusive explanation of factors affecting Jana’s MLR levels, more in-depth analyses would be required. It must also be accepted that not everything that impacts speech can be explained: “we can never be sure about what goes on in the mind of the speaker during pausing nor about the real reason for it. We must always allow for the possibility that the mind is preoccupied with things that are not directly related to the ongoing production” (Schilperoord, 1996, p. 10, cited in Götz, 2013).

4.4.3.3 Repair fluency

**Repetitions**

Figure 10: Repetitions for Jana
Figure 10 shows that Jana reduced the number of repetitions from 2.03 per hundred words (phw) in interview 1 to 0.91 in interview 6. This makes for a dramatic decrease in repetitions of 55%. The development is highly linear, as confirmed by a correlation coefficient of -.856.

Repetitions are relatively rare both in native and non-native speech. Biber (1999) mentioned that “in every hundred occurrences of a word, it is exceptional for more than one repeat to occur” (p. 1056). Compared to the averages obtained for Götz’s (2013) NSs (≥0.82 phw) and advanced NNSs (≥0.69 phw), Jana (≥1.35) overused repetitions. However, all her scores were within the range of NS scores (0.05-2.07 repetitions phw) and became increasingly more native-like.

Götz’s (2013) study revealed that NNSs used fewer repetitions than NSs on average, but that performance varied in particular regarding the word types the groups repeated. According to her, an underuse of repetitions may imply that non-native speakers have not internalized the feature as a “fluency-enhancing strategy” (p. 109) or simply prefer to use other features for that purpose, such as discourse markers, or filled pauses. It is possible, according to this interpretation, that Jana may have internalized the use of repetition as a speech planning strategy.

**Reformulations**

Figure 11: Reformulations for Jana

Figure 11 shows that Jana increased the number of reformulations from 1.42 per hundred words (phw) in interview 1 to 2.51 phw in interview 6 (+77%). Apart from an atypically large increase in interview 3, her development was almost linear as confirmed by a high correlation coefficient of .899.
Like Lennon’s (1990) participants, Jana increased her reformulations during SA. Previous literature has considered this trend to be indicative of improved levels of proficiency. Freed (1995) concluded that more reformulations tend to be accompanied by the expression of more complex thought. Riggenbach (1991) explained that through the process of speech monitoring, language learners show an awareness of their own linguistic shortcomings by way of self-correcting. Wood (2012) mentioned that reformulations of discourse are features of NS performance and therefore do not make for perceived dysfluency.

Reformulations in this study include reformulations of content and of form (‘retracings’) since not all reformulations could be unambiguously assigned to one or the other type. However, unambiguous cases were investigated in more depth with the aim of gaining insight into the participants’ monitoring skills.

The majority of Jana’s reformulations concerned false starts (Example 27). This type of reformulation has been considered a native-like speech management strategy and unobtrusive to the listener (Götz, 2013; Riggenbach, 1991; Wood, 2012). In other instances, Jana seems to have reformulated her speech to accommodate the listener’s comprehension of the narration – for example, by providing background information (Example 28). In Example 29, Jana reformulated her utterance for reasons of precision, possibly because she believed that the term ‘friends’ was too vague a reference for the interlocutor to understand (In the examples below, reformulations were placed in angle brackets).

Example 27 (int. 6): (.35) and (1.50) <Friday was> I was so frustrated on on Friday.

Example 28 (int. 6): (1.16) <and then I saw on Facebook> they did a Facebook group.

Example 29 (int. 6): (.82) hmm (3.20) I went to town with (.37) ehm three other <friends> PE friends ehm with Rachel (.99) and Tess and Brittany.

Modifications for grammatical inaccuracy were extremely rare in Jana’s data. In Example 30, she corrected a verb tense; in Example 31, she modified a preposition; and Example 32 involves the retracing of an idiomatic expression.

Example 30 (int. 6): and then (.47) after dinner I (1.40) ehm (.34) went home and <get> (.30) got changed.

Example 31 (int. 3): (1.34) and ehm (2.19) I think (.65) she’s (1.59) nearly all the time eh <with her> (.59) on her phone.

Example 32 (int. 3): <most the time> most of the time it’s about New Zealand but.
Jana’s grammatical modifications demonstrate awareness of (a major portion of) errors and a capacity for correction. The small number of corrections appears to have little to do with proficiency levels but a lack of opportunities for corrections. As will be seen below (see 4.4.4), Jana displayed particularly high levels of language accuracy.

4.4.3.4 Smallwords

Figure 12: Frequency of smallwords for Jana

Figure 12 shows the frequency with which Jana used a selected number of smallwords during her sojourn. She began her sojourn with 4.97 smallwords per hundred words (phw) and ended it with 6.72 smallwords phw. This accounts for a total increase of 35%. The upward trend is moderately positive (r= .619). Interview 5 exhibits an unusual peak in the development (9.15 smallwords phw).

In terms of frequency, Jana appears to have ‘overused’ smallwords on average (≈6.68 smallwords phw). Hasselgren (2002) measured a greater range of smallwords than I did (see Chapter 3), but all of her participants used significantly fewer smallwords than Jana (NS: ≈4.46; the more fluent NNS group: ≈2.79; the less fluent NNS group: ≈2.31).

In a more qualitative analysis, the types of smallwords that Jana used in her speech were investigated (Table 23). Her range of smallwords varied between 8 (int. 6), 9 (int. 1, 2 and 3) and 10 (int. 4 and 5) at any one time. She used all twelve smallwords measured in this study over the course of her sojourn. The vast majority of her smallwords consisted of the single-word smallwords ‘yeah’, ‘so’, ‘just’, and ‘oh/ah’. Jana made limited use of multi-word smallwords, particularly ‘you know’, ‘I think’, ‘I don’t know’, and ‘a bit’:
Hasselgren (2002) identified three stages of acquisition of smallwords, each stage representative of increasingly higher L2 proficiency levels. With regard to this classification, Jana successfully used the stage 1 smallwords ‘I think’, ‘just’, ‘or something’, ‘a bit’ and, ‘you know’ and the stage 3 (native speaker level) smallwords ‘ah’, ‘oh’, ‘and everything’, and ‘like’ (Hasselgren’s stage 2 smallwords only include ‘well’). However, she did not incorporate the stage 3 smallwords ‘and stuff’ and ‘and things’ into her speech.

Hasselgren (2002) further discovered that NNSs tend to overuse (e.g. ‘I think’, ‘or something’) or underuse certain smallwords (e.g. ‘just’, ‘like’, ‘a bit’, ‘and things’). She hypothesized that the overuse of “smallwords may be an area of vocabulary where learners like to draw on a stock of familiar words and phrases – the ‘lexical teddy bears’ of speaking” (p. 155). As mentioned above, Jana frequently used the smallwords ‘yeah’, ‘so’, ‘just’, ‘I think’ and ‘ah/oh’. The smallword ‘I think’ is typically overused by NNSs, and ‘so’ (which has not been identified as smallwords by Hasselgren) appears to be a marker associated with German speech (Grieve, 2010). In this sense, it could be argued that with the exception of ‘just’ and ‘ah/oh’, Jana tended to overuse smallwords that are atypical of adolescent native English speech. Also, note that Jana made use of these smallwords right from the first interview, which implies that she had either acquired them prior to her SA period or during the first month abroad. Possibly, these smallwords occur more frequently in NS input or Jana simply noticed them more easily. In any case, Jana seemed to hold on to and use the same few ‘teddy bear smallwords’ as they may have provided her ‘islands of safety’ during online speech performance.

Table 23: Range of smallwords for Jana

<table>
<thead>
<tr>
<th>Interview</th>
<th>Like</th>
<th>yeah</th>
<th>so</th>
<th>just</th>
<th>And everything/and (all this stuff)/and things</th>
<th>You know</th>
<th>And/or-something like</th>
<th>I think</th>
<th>I don’t know</th>
<th>a bit</th>
<th>I mean</th>
<th>Ah/oh</th>
<th>Raw total</th>
<th>Total Phw</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>18</td>
<td>11</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>8</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>49</td>
<td>4.97</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>20</td>
<td>16</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>58</td>
<td>6.91</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>8</td>
<td>12</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>10</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>50</td>
<td>5.47</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>13</td>
<td>9</td>
<td>12</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>18</td>
<td>67</td>
<td>6.83</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>24</td>
<td>7</td>
<td>14</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>0</td>
<td>6</td>
<td>88</td>
<td>9.15</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>16</td>
<td>7</td>
<td>8</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>11</td>
<td>59</td>
<td>6.72</td>
</tr>
</tbody>
</table>
Table 23 further shows that what appear to be ‘newly’ acquired smallwords had their first appearance in interview 2 (‘I mean’) and 4 (‘like’, and ‘and everything’). Interestingly, the latter two are considered ‘stage 3’ smallwords by Hasselgren. It can thus be assumed that 4 months of residing abroad were necessary for Jana to arrive at the stage where she was able to produce more native-like, or later-acquired, smallwords.

Smallwords can have a variety of pragmatic and discourse functions (Grieve, 2010) (e.g. stance marker, emphasizer etc.). Götz (2013) argued that they can serve as “elegant fillers for planning phases instead of a filled or unfilled pause” (p. 40). In this function, they can also increase speech rate.

Overall, Jana already began her sojourn with a good range and high frequency of smallwords, which appear to have “served as elegant fillers for planning phases” (Götz, 2013, p. 40). Possibly, Jana substituted some empty pauses by smallwords over time, which may have made her speech faster (possibly, the high use of smallwords in interview 5 may, to some extent, explain the high speech rate scores during that time). Some of the smallwords Jana used were rather atypical of native speech and were potentially the result of L1 transfer (e.g. ‘so’). Jana also did not broaden the range of smallwords much, but appeared to generally ‘cling on’ to a few familiar ones during speech processes. It is doubtful whether 5.5 months in the naturalistic environment are enough for learners to notice smallwords in NS input, and make use of them themselves. Götz stated that the acquisition of smallwords is difficult even for native speakers, as they are hardly noticed and seem to be used in an “unpredictable manner” (p. 40). Ultimately, the usage of smallwords is optional, and Jana may not have desired to use phrases such as ‘and stuff’. These possibilities would need more thorough investigation than was feasible within the limits of this thesis.
4.4.4 Accuracy

4.4.4.1 Error-free clauses

Figure 13: Error-free clauses for Jana

Figure 13 shows that Jana’s accuracy scores were very high overall. She began her sojourn with a score of 88% and finished it with 94%, resulting in an overall improvement of 7%. Interview 2 represents a peak in her development (99%) and interview 5 a dip (87%). Jana’s development of error-free clauses fluctuated considerably, as reflected in a very low correlation coefficient ($r = -0.048$).

With an average of 92% error-free clauses, Jana scored well above the mean values reported for the intermediate-level students in Skehan and Foster (1999) (approximately $\pm 60\%$) and Kormos and Dénes (2004) ($\pm 53\%$) and above the levels achieved by the intermediate/advanced-level groups in Mora and Valls-Ferrer (2012) ($\pm 70.50-78.91$) and Kormos and Dénes (2004) ($\pm 86\%$). In comparison to Mora and Valls-Ferrer’s (2012) NS group ($\pm 99.60$), Jana performed less well.

The figures show that Jana performed at a highly advanced level with regard to language accuracy, if not a native-like level. After all, even native speakers make errors during spontaneous speech (Götz, 2013, p. 47). It is unclear, however, to what extent global accuracy levels can shed light on language progress. Pallotti (2009) explained that “accuracy per se is not a direct indicator of interlanguage development” (p. 592, emphasis in original). Conversely, Norris and Ortega (2009) suggested that accuracy measures may appropriately describe advanced learners’ language progress.
It was beyond the scope of this study to investigate the levels of advanced language use as reflected by late acquired features such as conditionals, passive voice and subjunctives. To see whether progress took place in one specific area, the more specific measure of correct finite verb phrases was employed (see 4.4.4.2).

### 4.4.4.2 Correct finite verb phrases

As can be seen from Figure 14, Jana hardly made finite verb phrase errors (fvphs). She began her sojourn with an accuracy rate of 98% and ended it with 97% (-1%). As for the global accuracy measure (see 4.4.4.1), Jana’s development experienced a peak in interview 2 (100%) and a dip in interview 5 (94%). Despite a moderately negative correlation coefficient ($r = -0.623$), her scores indicate that she had largely mastered the morphological rules and tense use of the verbs she used in her output.

During her sojourn, Jana produced fvphs that were on average accurate to 98%. Not many previous SLA studies have examined this specific feature of accuracy but Jana was well above the scores obtained for the NNS high-proficiency group in Wigglesworth (1997) (max. 85.51% correct fvphs).

To give a short account of the demandingness of the verb forms that Jana used (i.e. to account for language progress), I applied a more fine-grained analysis of the types of errors she made. Overall, Jana committed 6 verb form errors and 16 tense-choice errors (note that combined tense and morphological errors were counted as morphological errors). Regarding the former, Jana once did not correctly inflect the copula ‘be’ (Example 33), twice overgeneralised the regular past form ending ‘−ed’ to the irregular verb ‘split’ (Example 34), twice made an error
on the main verb of a complex verb phrase (Example 35) and once omitted an auxiliary (Example 36).

   Example 33 (int. 1): (.75) eh French is a bit (1.6) a bit weird because they (.36) they ehm (1.09) the the lessons is very (.40) different to them in Germany. (< are)

   Example 34 (int. 2): (.54) so (.85) we completely splitted up. (< split)

   Example 35 (int. 6): (1.26) and then I said saw on ehm (.31) facebook that one girl said ‘oh does anybody needs a ride?’ (< need)

   Example 36 (int. 6): (.30) but I couldn’t know. (< couldn’t have known)

Tense errors mainly concerned failure to use tense backshift for verbs in indirect speech complements (12 errors) (Example 37). Other errors include using the present tense for a future event (twice) (Example 38), and two instances of using the present tense for a past event (Example 39).

   Example 37 (int. 6): (.96) and oh yeah the thing was (.84) ehm (.69) she asked me where I live. (< lived)

   Example 38 (int. 5): (.48) she just said ‘no (.49) you’re not in school (.55) I go to the internationals’. (<will go)

   Example 39 (int. 6): I just (.47) you know <and ask> (.48) and I ask her. (< asked)

Backshifting errors indicate that Jana stored information about the formation of this feature in a way that was deviant from target language norms. It is possible that she overgeneralized the rules that apply to the formation of direct speech reports, and applied them to those of indirect speech reports. Her incorrect use is reflective of early developmental stages of this construct, yet the construct itself probably constitutes a later acquired, demanding form. This assumption would be supported by Goodell and Sachs (1992).

The fvph error analysis suggests that Jana predominantly made errors typical of advanced learners (Götz, 2013; Lennon, 1991). Lennon (1991) argued that at the advanced level, “highly localized morphological errors is less prevalent than error involving usage” (p. 184). It is unclear to what extent Jana’s errors were processing based or knowledge based. Possibly, the systematicity with which Jana failed to backshift tense in indirect speech could suggest that these errors were knowledge-based, not processing-based (Lennon, 1991). Given that “in the actual spoken output of a native speaker, there is a great number of erroneous sentences or word forms, or at least deviations from the norm” (Götz, 2013, p. 47), Jana’s abilities with regard to form during spontaneous speech can be considered highly advanced.
4.4.5 The relationship between CAF

Figure 15: CAF relationships for Jana

Figure 15 displays the nature of the relationship of five CAF dimensions (using Z-scores). I plotted the development of the global measures of fluency (speech rate and mean length of run), accuracy (error-free clauses), lexical diversity (Guiraud’s Index) and grammatical complexity (mean length of clause).

According to this figure, the following relationships could be detected: 1) A supportive relationship between the two complexity measures lexical diversity (LD) and mean length of clause (MLC). 2) A supportive relationship between the two fluency measures speech rate (SR) and mean length of run (MLR) and error-free clauses (EFC). 3) There is a competitive relationship between the complexity measures (LD/MLC) and the fluency (SR/MLR)/accuracy (EFC) measures. Not all relationships are fully supportive or competitive (i.e. the trend of the relationship was not always reflected in all movements - but growers were considered supportive or competitive when at least 3 out of 5 movements were identical or distinct, respectively). However, the relationships appear to point to trends in a manner that appears to explain overall speech behaviour. In the following, these relationships will be considered in more detail.
4.4.5.1 Supportive growers

**Grammatical complexity and lexical diversity**

Figure 16: Relationship between mean length of clause and lexical diversity for Jana

In Jana’s speech, MLC and LD developed in unison (i.e. they have a supportive relationship) most of the time (Figure 16): the longer the clauses, the higher the lexical diversity, and vice versa. Only at the end of the 6-month sojourn did this relationship change: Jana’s output was more lexically diverse in interview 5 than interview 6, while the opposite was true for MLC.

With regard to Kormo’s Speech production model, this supportive relationship can be explained by the fact that “lexical items are assumed to drive syntactic encoding processes” (p. 51). That is, complex concepts require the activation of less frequent words, which, in turn, are realised by more complex syntax.

Jana’s output fluctuated greatly in terms of the two complexity measures, exhibiting an increased performance in interview 3 and a dip in interview 4. I suspected that the use of direct speech largely influenced complexity measures. Direct speech reports typically consisted of very short and formulaic matrix clauses and non-embedded direct speech complements expressed in the present tense. Particularly the nature of the matrix clauses (short and formulaic) appears to have resulted in overall shorter clauses and less diverse lexical output. In contrast, I found that the interview with the highest complexity levels (interview 3) contained detailed descriptions of past events which required Jana to use more diverse lexis and longer clauses. The latter is especially true in cases where she felt she had or wanted to provide her interlocutor with extra information to make a point or to be comprehensible, as expressed in the form of phrases and adjuncts.
**Speech Rate and Mean Length of Run**

Figure 17: Relationship between speech rate and mean length of run for Jana

![Graph showing relationship between speech rate (SR) and mean length of run (MLR) for Jana.](image)

Figure 17 shows that SR and MLR exhibited a supportive relationship most of the time. Longer runs went hand in hand with faster SR (int. 2 and 4), while shorter runs went hand in hand with slower SR (int. 1, 3 and 6). Only in interview 5 did SR levels increase but MLR levels decrease. The analysis above demonstrated that changes in the two fluency measures can be related to levels of speech automatization (i.e. L2 progress). Jana’s ability to retrieve lexis and syntax appear to have been determined in the same way by whatever external forces were at play (e.g. different speech modes, number of elements, different topics, or variations in speaker style).

**Accuracy and Fluency**

Figure 18: Relationship between fluency and accuracy for Jana

![Graph showing relationship between fluency (MLR and SR) and accuracy (EFC) for Jana.](image)

Figure 18 shows that EFC developed in unison with fluency (MLR and SR) during the first 4 interviews. That is, higher accuracy levels went hand in hand with faster speech (int. 2 and 4) while lower accuracy levels were accompanied by slower speech (int. 1 and 3). Between interviews 4 and 6, accuracy competed with SR, and formed a partially supportive/competitive
relationship with MLR. In terms of allocation of attention, this relationship suggests that Jana simultaneously focused on fluency and accuracy during interviews 1 to 4. In interviews 5 and 6, SR and EFC developed a competitive relationship with each other, suggesting that her cognitive resources could no longer be distributed equally among the two dimensions. A focus away from accuracy in interview 5 might also suggest that Jana emphasised meaning over controlled output, which led to errors.

The relationship between accuracy and fluency can best be explained by findings from task-based research that focuses on influences of task types on all three CAF dimensions. These issues will be discussed in the following section.

4.4.5.2 **Competitive growers**

**FLUENCY/ACCURACY VERSUS COMPLEXITY**

Figure 19: Relationship between fluency/accuracy and complexity for Jana

![Graph showing the relationship between fluency/accuracy and complexity for Jana](image)

Figure 19 shows that a perfectly asymmetrical relationship exists between the fluency measures (MLR and SR)/ the accuracy measure (EFC) and the complexity measures (MLC and LD) during the first four interviews: when fluency/accuracy levels went up (int. 2 and 4), complexity levels went down (int. 1 and 3), and vice versa. Between interview 4 and 6, this relationship was less straightforward. SR and MLC competed with EFC. LD competed with MLR.
The general competition between fluency (MLR and SR)/accuracy and complexity (LD and MLD) suggests that when Jana spoke fast (int. 2, 4, 5), she produced more accurate but less complex output. Conversely, when fluency levels were lower (int. 1 and 3), Jana’s output was less accurate but more complex.

Given that there were competitive relationships between fluency and complexity as well as accuracy and complexity, it can be concluded that attention during online speech processes was limited for Jana and that she typically had more or fewer resources available (or focused on) either for fluency and accuracy (combined) or complexity. This trade-off can be explained by Kormos’ (2011) and Skehan’s (1992) account of the role of attention during speech production. They claim that when conceptual demands are low, speakers have more resources available for processes in the encoder stage (i.e. raised fluency and accuracy). When conceptual demands are high, fewer resources are available for processing, which leads to errors and less fluent output (but higher levels of complexity). Jana’s trade-off between complexity and accuracy/fluency also supports the findings made in task-based research (Skehan, 2009).
## 4.4.6 Summary

Table 24: Results of all CAF measures for Jana

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Average</th>
<th>Overall development</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subordination Index</strong></td>
<td>1.46</td>
<td>1.29</td>
<td>1.3</td>
<td>1.74</td>
<td>1.45</td>
<td>1.39</td>
<td>1.44</td>
<td>-5%</td>
<td>.185</td>
</tr>
<tr>
<td><strong>Mean length of clause (wpc)</strong></td>
<td>6.13</td>
<td>5.88</td>
<td>6.46</td>
<td>5.20</td>
<td>6.08</td>
<td>5.63</td>
<td>5.89</td>
<td>-8%</td>
<td>-.385</td>
</tr>
<tr>
<td><strong>Lexical diversity (Guiraud’s Index)</strong></td>
<td>8.49</td>
<td>7.85</td>
<td>8.25</td>
<td>7.38</td>
<td>7.78</td>
<td>8.08</td>
<td>7.97</td>
<td>-5%</td>
<td>-.429</td>
</tr>
<tr>
<td><strong>Empty Pauses (phw)</strong></td>
<td>18.76</td>
<td>17.28</td>
<td>17.83</td>
<td>15.19</td>
<td>15.90</td>
<td>16.74</td>
<td>16.95</td>
<td>-11%</td>
<td>-.696</td>
</tr>
<tr>
<td><strong>Filled Pauses</strong></td>
<td>4.26</td>
<td>2.15</td>
<td>2.95</td>
<td>1.83</td>
<td>3.01</td>
<td>3.64</td>
<td>2.98</td>
<td>-15%</td>
<td>-.097</td>
</tr>
<tr>
<td><strong>Speech Rate (wpm)</strong></td>
<td>121.02</td>
<td>144.61</td>
<td>128.51</td>
<td>148.23</td>
<td>152.65</td>
<td>144.18</td>
<td>140</td>
<td>+19%</td>
<td>.693</td>
</tr>
<tr>
<td><strong>Mean Length of Run (wpr)</strong></td>
<td>5.19</td>
<td>5.59</td>
<td>5.51</td>
<td>6.54</td>
<td>6.17</td>
<td>5.81</td>
<td>5.80</td>
<td>+12%</td>
<td>.645</td>
</tr>
<tr>
<td><strong>Repetitions (phw)</strong></td>
<td>2.03</td>
<td>1.79</td>
<td>1.09</td>
<td>1.02</td>
<td>1.25</td>
<td>0.91</td>
<td>1.35</td>
<td>-55%</td>
<td>-.856</td>
</tr>
<tr>
<td><strong>Reformulations (phw)</strong></td>
<td>1.42</td>
<td>1.43</td>
<td>2.08</td>
<td>2.04</td>
<td>1.98</td>
<td>2.51</td>
<td>1.91</td>
<td>+77%</td>
<td>.899</td>
</tr>
<tr>
<td><strong>Smallwords (phw)</strong></td>
<td>4.97</td>
<td>6.91</td>
<td>5.47</td>
<td>6.83</td>
<td>9.15</td>
<td>6.72</td>
<td>6.68</td>
<td>+35%</td>
<td>.619</td>
</tr>
<tr>
<td><strong>Error-free clauses</strong></td>
<td>88%</td>
<td>99%</td>
<td>91%</td>
<td>94%</td>
<td>87%</td>
<td>94%</td>
<td>92%</td>
<td>+7%</td>
<td>-.048</td>
</tr>
<tr>
<td><strong>Correct finite verb phrases</strong></td>
<td>98%</td>
<td>100%</td>
<td>98%</td>
<td>98%</td>
<td>94%</td>
<td>97%</td>
<td>98%</td>
<td>-1%</td>
<td>-.623</td>
</tr>
</tbody>
</table>

wpc: words per clause; phw: per hundred words, wpr: words per run

The development of lexical diversity (LD), mean length of clause (MLC) and the subordination index was marked by much fluctuation (all correlation coefficients were below 0.50, implying
low linearity) and a slight overall decrease (between -5% and -8%). Only one score for MLC (int. 3) and one score for SI (int. 4) was higher during the sojourn than at its beginning (int. 1). All LD scores dropped below the first one.

For fluency, the biggest rate of change occurred with reformulations (+77%), followed by smallwords (+35%), repetitions (-33%), speech rate (SR) (+19%), filled pauses (FPs) (-15%), mean length of run (MLR) (+12%), and empty pauses (EPs) (-11%). Reformulations and repetitions developed in a highly linear way (i.e. between r= .70 and .90). SR, MLR, EPs, and smallwords developed in a moderately linear way (i.e. between r=.50 and 0.70). FPs showed no linearity in their development.

Jana’s overall and specific accuracy levels were very high during all six interviews (≥92% and ≥97 respectively) and so showed very few overall changes (+7% for error-free clauses; -1% for correct finite verb phrases). Error-free clauses displayed a lot of fluctuation (r=-.048) while the development for finite verb phrases was moderately negatively linear (r=-.623). With the exception of tense backshift, where Jana made most of her errors, she mainly made errors typical of advanced L2 speakers (Götz, 2013; Lennon, 1991).

4.5 Learning opportunities and L2 development

The discussion above showed that the nature of the semi-structured interviews, as well as the data analysis procedures used in this study affected CAF measures (in particular the scores that were measured for complexity). However, since the semi-structured interviews resembled friendly conversations in style, they are believed to have captured language similar in production to conversations the participants had with their L2 peers and host families during their sojourn. I will in this section attempt to explain whether there was also a relationship between Jana’s L2 development and the types of language activities she experienced during her study abroad. Potential reasons for Jana’s changes in grammatical complexity, lexical diversity, accuracy and fluency will be given.

It was shown in section 4.4 that Jana’ speech became more fluent, less grammatically complex, less lexically diverse, and less accurate. However, the complexity measures and the lexical diversity measure ranged within NS norms at all times (with the exception of SI), and Jana’s accuracy levels were particularly high (almost native-like) at all times. It is unreasonable, therefore, to assume that the decreases that did occur were reflective of language ‘attrition’. Rather, they appear to have been reflections of normal variations in spontaneous speech.
Jana’s L2 activities predominantly involved informal, ‘uncontrolled’ conversations in all three social settings (i.e. the host family, the school, and outside of school). Analytical or ‘controlled’ English language learning activities only seemed to occur in the EAP classroom. Informal language learning activities changed with regard to their frequency and, to some degree, their quality during her sojourn. Growing social networks with Kiwis from school went hand in hand with increasingly more social interactions during the break times and in Jana’s free time. Joining the football club and changing host families provided additional opportunities for informal language learning.

Simultaneously, Jana’s use of German decreased over time, from speaking the language during all break times and in her free time during the first two months to speaking it only rarely with a few German friends in the second half of her sojourn. Jana’s accounts further revealed that the quality of interactions remained relatively stable with her native speaker peers (i.e. informal conversations about boys, the school ball etc.) but improved in the host family setting as a consequence of her change of families in the fourth month.

The opportunities for language learning that Jana was provided with during informal conversations with NSs were unlikely to induce increases in her complexity and accuracy scores. After all, being absolutely accurate and highly complex does not reflect NS norms for informal speaking situations (Brown & Yule, 1983; Ferrari, 2012; Skehan, 1998). Skehan (1998) in fact claimed that learning to participate in NS conversations means “learning not to use complete and well-formed sentences but learning how to make well-judged interventions which one’s conversational partners will judge as furthering the conversation” (p. 25). Furthermore, participating in conversations of large groups, which Jana was part of, does not usually provide speakers with “substantial floor time” which could allow for the increase in accuracy and complexity (Isabelli-Garcia, 2006, p. 248).

However, it seems that as a result of Jana’s high initial L2 proficiency levels, the L2 input she was exposed to during informal conversations enabled her to pay attention to more subtle aspects of English, such as the speech styles of her L2 communities. This would support Regan et al. (2009) who proclaimed that advanced learners are better able to attend to pragmatic aspects of the L2 than lower level learners. One speech style typical of female adolescent speech that Jana made frequent use of is direct speech reports (Romaine & Lange, 1991). It is possible that Jana’s desire to increasingly adapt to the norms of her Kiwi friends, including their pragmatic norms, encouraged her to use this structure more frequently in the second half.
of her sojourn when relationships became stronger (note that Jana accurately used direct speech reports from the first interview). By conveying opinions and feelings through this speech mode, Jana performed a social role typical for native speaker girls of her age (Romaine & Lange, 1991), adapting to their speech patterns (Giles & Coupland, 1991). In this sense, Jana’s slight decrease in the complexity measure MLC would indicate more native-like speech behaviour. Also, the more direct speech reports Jana used, the more native-like her subordination index became (see section 4.4.1.1).

Likewise, the lexical diversity in Jana’s output was potentially ‘shaped’ by the speech situations she was exposed to with her NS friends and family. Her scores decreased over time, with no measure ranking higher than the first interview. As the comparisons showed, this behaviour indicated an approximation to NS norms. Perhaps, Jana’s higher initial lexical diversity level was reflective of the more formal and academic classroom practices in her German high school which required her to produce lexically diverse language. During her sojourn, she may have adjusted these more formal speech patterns to more casual ones to match her NS interlocutors’ conversation style. From Jana’s data, it appears that she typically engaged in communication that did not require the use of specific vocabulary but rather ‘generalised’ vocabulary to talk about topics such as ‘boys’ or the school ball with her Kiwi friends, or daily activities with the host family. If this is true, it was neither necessary nor appropriate for her to produce language that was complex or more diverse. This interpretation would support Brown and Yule’s (1983) finding that spoken informal language is not marked by great lexical diversity but “general, nonspecific” vocabulary, that is often repeated to “save[e] the speaker from constantly having to hunt up a different word” (p. 9). Particularly interactions with the main purpose of maintaining social contact (e.g. casual chats, greetings), rather than delivering an important message (i.e. explaining an illness to a doctor), demand the use of “‘generalised’ vocabulary and sparse information-packing” (p. 12) to accommodate the needs of the listener.

Alternatively, lexical diversity scores may have been shaped by Jana’s language learning orientation towards correct output. Götz (2013) explained that producing new lexical items can be a challenging task for L2 speakers, especially if they favour being in control rather than risking errors through experimentation. To avoid the potential pitfalls of making errors, Jana may have preferred to use “the same, easily retrieved and ‘safe’ words […] over and over again” (Götz, 2013, p. 65). This language learning approach is likely to come at the expense of a lexically more diverse performance. The greater likelihood considering Jana’s reports of her
communicative interactions is that the communicative activities that Jana was involved in simply did not encourage or require the use of complex or diverse language.

The analysis of Jana’s speech production revealed that her accuracy levels were very high overall (despite some fluctuations in the development). There seems to have been little room for improvement in her accuracy – a finding which has been made before for advanced level learners (Churchill&DuFon, 2006; Isabelli-Garcia, 2004). Additionally, Jana may have been aware that in her social settings, using simple language structures was more appropriate. Repeating the same familiar structures would have allowed her to remain a highly accurate speaker as well as to speed up automatization of already existing knowledge (hence increase in fluency). Alternatively, the nature of the language learning activities she engaged in (predominantly informal conversations) may simply not have been conducive to improvements in accuracy. Skehan (1998) claimed that “language use, in itself, does not lead to the development of an analytical knowledge system since meaning attracts attention from form” (p. 27). It does appear that Jana may have profited linguistically from more analytical language learning, as suggested by Freed (1995).

Higher speed fluency (i.e. automatized L2 knowledge) has been shown to result from substantial exposure to L2 input and a learner’s own productions in social interactions (Ellis, 2008; Mora & Valls-Ferrer, 2012; Skehan, 1998; Wood, 2012). Kormos (2011) claimed that attending to “frequently-heard constructions can raise the activation level of the learner's representation of linguistic items and can help accessing and automatization of existent knowledge of the L2” (p. 50). Frequent opportunity to speak is considered crucial for students to “go beyond carefully constructed utterances and achieve some level of natural speed and rhythm” as “only by frequent use is the fluency side of speech likely to be improved” (Skehan, 1998). Jana’s increasingly extensive contact with native speakers appeared to have provided her with good opportunities to gradually automatize her declarative linguistic knowledge and use syntactical structures and lexical items with more ease and speed. More proceduralized knowledge in Jana’s speech is reflected in L2 productions with fewer pauses and higher speech rates. Jana also demonstrated an increased ability to make native-like speech reformulations to ensure listener comprehension, and to use smallwords typical of adolescent native speech. MLR did not increase dramatically and the peak performance in interview 4 was shown to largely depend on Jana’s use of direct speech.
In sum, increasingly frequent opportunities for informal interactions resulted in more automatized speech performances (i.e. as reflected in higher fluency scores), and the use of speech patterns typical of female adolescents – generalised vocabulary and direct speech mode - as reflected in lower complexity and lexical diversity scores. However, these informal conversations did not lead to high accuracy scores either because this was not necessary (given the use of simple and already accurate linguistic structures) or because it was not possible (focus on meaning during NS interactions).
5 Case Study 2: Chiara

5.1 Background information

CHIARA AND HER FAMILY IN GERMANY

Chiara grew up in a German city with her two parents (a real estate agent and an engineer) and younger sister. They spoke German at home and the family had never lived outside of Germany. The family had travelled to different destinations in Europe, as well as North America and Canada. Chiara arrived in New Zealand at 15 years old, which made her one of the youngest German international students at her school (she was placed in Year 11). Chiara reported a lack of self-esteem and that she was overly well behaved in Germany. It was her desire to change these traits in New Zealand so that she could return to Germany a different person.

GERMAN SCHOOL AND LANGUAGE LEARNING HISTORY

From school Year 5, Chiara attended a German public school that specialises in English immersion classes. Most students are not German and, according to Chiara can speak three or four languages. As the students get older, a greater number of subjects are taught in English at the school:

in the fifth grade I have PE in English in the sixth grade I had one extra English lesson in the seventh grade I had Geography in English in the seventh I had History in English now in in the eight I had gosh what did I have another subject in English ehm this year we have Social studies in English. (int. 6)

Given that many of Chiara’s classmates in Germany were English native speakers, she felt pressured to produce “quite good” English (int. 6). Failure to do so made her anxious:

I had a problem that people laughed at me about my th ehm in Germany so I tried not to say the word ‘I think’ and I hope that it got better but I notice that I still ehm make my sentence that I don't have to say that word because I'm so used to and I really try to don't say any word with th because I know then they will laugh. (int. 1)

Chiara explained that she was a very diligent and ambitious student. Compared to most other subjects, she did not achieve top marks in English classes, unlike all the native speakers apparently, which disappointed her.
Chiara left for New Zealand half-way through Year 9, just before entering Senior School (Year 10). She mentioned she had to comply with one educational requirement from Germany while abroad – to pass New Zealand exams with ‘achieved’. This would then allow her to attend the Senior School level.

Chiara studied English in EFL classes for seven years prior to going abroad. She also studied French and Japanese. Chiara had already completed two short-term exchanges as part of the German school curriculum in Europe prior to arriving in New Zealand. She had spent a week in the Czech Republic and two weeks in France.

REASONS FOR STUDYING ABROAD IN NEW ZEALAND AND GENERAL ENGLISH LANGUAGE LEARNING MOTIVES

Chiara’s main goals for studying abroad were having new experiences, getting better at a foreign language, becoming more self-confident, and making new friends. Chiara was not primarily interested in learning English when she decided that she wanted to study abroad. She was open to spending time abroad in an English or a French-speaking country. She decided against the latter because she believed that French teachers talked a lot and so made classes boring. English language learning was important mainly for educational achievements after the sojourn.

Chiara chose New Zealand rather than Canada because she had already visited Canada with her family and because she saw pictures of New Zealand, which she really liked. Chiara chose to be away for one New Zealand academic semester because this enabled her to stay away from her German schooling without missing anything. She also felt that one year away would make her really miss her family while one term (i.e. 3 months) would be too short to make Kiwi friends.

5.2 Social settings

5.2.1 The host family

Chiara had two host parents (Sara and Mike), two host sisters (Becky, 16, and Christina, 19), and a host dog. The family was experienced in hosting students. Both parents worked full-time and were often unavailable when Chiara returned home from school. Mike worked night shifts and either slept or was busy working in his office in the afternoons. Sara usually arrived home at five o’clock and often desired to be alone before she made dinner. Becky attended a boarding
school and was only home at weekends. Christina studied at a university and worked full-time during the semester breaks. Although she lived at home, Chiara reported that she was rarely at home.

Social interactions between Chiara and the host family members were limited and lacked depth. Meal times were the only regular social get-togethers but due to the family’s focus on watching TV during dinner, opportunities for conversation were severely limited. During her stay, Chiara was unable to foster close relationships with any of the host family members and she habitually retreated to her bedroom before and after dinner.

Chiara mentioned that the host family tended to avoid communication beyond basic exchanges with her and broke off conversations that she initiated. Chiara believed that this demonstrated a disinterest in and dislike of her personally. For example, Sara did not seem to show an interest in hearing about Chiara’s daily school adventures and prematurely cut off her anecdotes:

> I started to talking about school how it was normal things like you would also tell your mother at home if someone was picking you up from school … she don't want to hear it she really don't want I talked and as soon as I just had a little break of I don't know twenty seconds she started to turn on on the radio and make it louder so that I don't talk. (Int. 3)

At times, the perceived indifference that the host family expressed towards Chiara discouraged her completely from initiating communication with them. For example, Chiara reported that she stopped talking to Sara during dinner preparations:

> I don't know she doesn't like if I talk to her so also when we were in the kitchen before dinner I didn I don't talk because they're not interested in anything so yeah it's a little bit bad. (Int. 3)

At the same time, Chiara showed an unwillingness to participate in host family conversations and other daily activities because she did not consider them interesting. According to her, the family was mainly interested in “gossip things” (int. 3). These types of conversations did not interest Chiara and she passed up opportunities to engage in them. Moreover, the family typically spent the evenings watching TV. According to Chiara, this was neither interesting nor conducive to social interaction and she did not understand the purpose of just sitting there with the family. Consequently, she shunned the company of the host family during those times.

Chiara typically spent the evenings alone in her bedroom doing homework, listening to music, and watching German TV shows. To maintain a sense of community she had daily two-hour conversations with her German parents over skype. Although Chiara was aware that the
exposure to so much German did not assist with L2 learning, they served as a welcome distraction from her negative feelings about her host family.

Chiara’s choice not to spend evenings with her host family appeared to breed resentment in the host parents, and resulted in what Chiara considered to be ‘punishments’. For example, Chiara reported that her extensive internet use in her room resulted in her being banned without warning from the internet after 10 pm. This made it difficult for her to do her homework (which she usually did after skyping), and exacerbated Chiara’s feeling of being an outsider. Chiara’s comments also indicate her host parents were unhappy that she went to her room early in the evening:

Mike always say in a very ironic way ‘yeah goodbye you're leaving us again’ it's like ‘why don’t you stay?’ and like ‘ah why don't I stay? because you don't care if I stay’. . . I just want to go in my room because they don't talk to me and I don't want to sit there and just doing nothing. (Int. 6)

It appears that Chiara was not prepared to interpret these clashes as stemming from different, yet equally valid, interpretations of the same events. Instead of “confronting misunderstandings and attempting to elucidate misunderstandings in dialogue” (Tan & Kinginger, 2013, p. 162) with her host family, Chiara interpreted them through her own lens, which she appears to have considered ‘truthful’.

The only positive experiences Chiara recounted about her host family concerned outings at weekends. The family once arranged a visit to their holiday house. Christina and Becky once invited her and Alia to go sightseeing with them. And the host family’s relatives once took her to a fun park. Chiara enjoyed these three events, not only because they were an opportunity to discover the region, but particular because they gave her the opportunity to speak and being listened to:

it was really nice because [the homestay grandparents] are so ah they are so nice and they're like talking they are like talking and like we sit in the car and they talk to me it was like that's so new someone is talking someone is interested in me and like I think I never talk that much to someone in the car. (Int. 3)

Although living in this family was mostly unpleasant for Chiara and involved feelings of rejection and loneliness, she ultimately accepted the circumstances and was grateful for the few positive aspects of living with the family, such as having a good internet connection and a comfortable bed.
5.2.2 The high school

Before school started in February 2014, Chiara had already established friendships with two German girls attending the same New Zealand high school. She met Alia on the journey to New Zealand and Alexandra during the orientation days organized by the school (see Chapter 3). The two girls later became her best friends and she spent the majority of her breaks and leisure time with them.

5.2.2.1 Inside the classrooms

Chiara studied six subjects: Mathematics (Maths), Media Studies, Mainstream English, Science, French and Textile Technology. She also attended a ‘form class’ each day, which consisted of a mixture of students from different classes. During these classes, the ‘form teachers’ dealt with administrative school-related issues. Unlike most other international students, Chiara was placed in Year 11 because of her age (15). There were not many international students in her classes. Only one German girl attended her French classes in the second school term, and two South American students attended her English class (Table 25).

Table 25: Student constellations in the different subjects

<table>
<thead>
<tr>
<th>Subject</th>
<th>Year</th>
<th>Local students</th>
<th>German speaking internationals</th>
<th>Non-German speaking international students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maths</td>
<td>11</td>
<td>√</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Media Studies</td>
<td>11</td>
<td>√</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Mainstream English</td>
<td>11</td>
<td>√</td>
<td>x</td>
<td>√</td>
</tr>
<tr>
<td>Science</td>
<td>11</td>
<td>√</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>French</td>
<td>11</td>
<td>√</td>
<td>√</td>
<td>x</td>
</tr>
<tr>
<td>Textile Technology</td>
<td>11</td>
<td>√</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Tutor Class</td>
<td>11</td>
<td>√</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

Chiara did not consider being a Year 11 student advantageous. In fact, she believed that being one of the only international students deprived her of teacher attention because they focused on the local students’ education. She also complained that she missed out on Year 12 options such as PE, which she believed would have been ‘fun’ with all the international students. Chiara also thought that studying in Year 11 without other European or South American international students resulted in her being treated as a ‘normal’ (i.e. local) student, and demanded greater effort to keep up with the class requirements:
the excuse inte of inte being international student for me don't really work because they know my English is good and then I'm like normal for them so I think that's better than Year 12 where mehr [more] more internationals and you can say ‘I'm international student I don't understand English’ so yeah so yeah I'm more like normal. (Int. 6)

Chiara’s data shows conflicting evidence about her levels of engagement in school subjects. Explicit comments suggest that she decided to put in minimal effort. She said she wanted to simply ‘pass’ exams as that would be sufficient for reintegration into her German school:

I know when I would learn I I would have excellence in all the subjects because it's really easy but I don't want to learn because yeah why. (Int. 2)

However, Chiara’s actions did not reflect a nonchalant attitude towards her academic performance. She reported spending considerable time each day completing homework (see 5.2.1). She also expressed pride about successful school performances and disappointment about unsuccessful ones.

Chiara’s classes were predominantly attended by local students. The small number of international students did not restrict opportunities for social interaction with Kiwis. Possibilities for interaction for Chiara seem to have been mainly determined by 1) seating arrangements (i.e. sitting next to people interested in interaction or not), and 2) the type of interaction the subject demanded of the students (i.e. activities requiring teamwork vs. individual student tasks). Chiara furthermore placed high importance on good rapport with teachers.

During her sojourn, Chiara made a number of good Kiwi friends and had a variety of opportunities to converse with them during class time. However, she rarely spent the school breaks with these Kiwis and only once met a Kiwi friend outside the classroom.

Chiara expressed a strong preference for classes that involved group work as this created possibilities to communicate with her classmates. All subjects, except Maths, afforded opportunities of this kind. For example in English, Chiara had to prepare a speech together with three classmates which allowed for much talk. She greatly enjoyed this opportunity for social interaction and referred to the presentation of the speech as “another moment of happiness” (report 4). The activity also gave her confidence in her English skills:

First I was a little bit scared because I was afraid that the class could laugh about me because I don’t know most of them but from the second on I started speaking I enjoyed it. It was so much fun . . . I think that my pronunciation was quite good, that I had eye
contact and I think the content of my speech (why you should become an international student overseas) also wasn’t that bad. (Report 4)

Similarly, the dynamics of Chiara’s Textile Technology classes appeared to be facilitative in the establishment of relationships. Chiara felt particularly drawn to Caroline who showed a real interest in her life. For a short period, Chiara even spent some of her break times with Caroline and her friends (see 0).

The form class that Chiara attended after each first period of the day offered additional opportunities to establish relationships with local students. Chiara very much enjoyed spending time in her tutor group and developed friendships with Kiwis through this class, such as with Abigale in French (see below).

Seating orders also greatly affected levels of interaction between Chiara and her classmates and thus language-learning affordances. Some teachers changed the first term seating orders in the second term. In French classes, Chiara first sat next to a girl who was “just quiet and didn’t talk” (int. 3). When the teacher changed the seating arrangement in Term 2, Chiara was placed next to Abigail. This new composition appeared to encourage conversation and the development of multiple friendships. Although Chiara never met these friends outside the classroom, she enjoyed their company and ultimately called them “the most amazing friends” (report 5).

A similar process happened for Chiara in Maths classes. Initial seating arrangements threatened Chiara’s position as a member of the class because, as she put it, she sat “on a table where everyone sits who has no friends” (int. 1). During the first three months, Chiara expressed a great aversion to Maths classes. In term 2, seating orders changed and Chiara sat next to Kelly with whom she became friendly and held lengthy conversations in class. Chiara and Kelly both went to the same English and Media Studies class, so the two girls also had conversations between classes. The girls became good friends and met outside of school during Chiara’s last week in New Zealand (see 0).

Seating arrangements in Chiara’s Science class appeared facilitative to social interactions in both terms. In the first term, Chiara shared a table with two girls who she liked. In the second term, seating orders changed and the contact between Chiara and the two girls no longer went beyond brief greetings. Chiara then sat next to a boy and a girl. This group of three had to conduct science experiments, which appears to have stimulated much interaction.
In the English classroom, seating orders changed weekly based on a shuffle system used by the teacher. This system seems to have inhibited the creation of new social networks since Chiara reported to have not had the chance to get to know her classmates well like this. However, Chiara was already familiar with Alison from her French class and a girl she played tennis with who she talked to. While chats with Alison were socially rewarding, they did not appear to Chiara to constitute opportunities for explicit L2 learning. Chiara mentioned her wish for people to explain English words or concepts to her in class. To her surprise, only one girl was able to clarify her questions while the remaining students, including Alison, rarely gave satisfying answers:

I noticed that many New Zealanders can't talk their language because like when I ask Alison for ‘ah can you explain what word it is?’ she's like ‘no I can't’ it's like ‘sorry you're speaking English why can't you explain it to me? (Int. 3)

Different subjects also appear to have created different opportunities for non-social language learning. Particularly the tasks Chiara performed in English and French classes stimulated formal engagement with the L2 (e.g. English essay writing, French translation tasks). For example, Chiara believed that translating French words into English enlarged both her French and her English vocabulary. She also mentioned that these exercises were cognitively challenging because they required the simultaneous use of three languages and that successful performance required much practice:

It is very hard when I have French because then my brain is full with three languages. The classes and teacher is speaking English and wants English answers. The texts and tasks are in French and my brain is in german. I have really often the problem that I can’t remember easy words in English as well as in French. But I noticed it is getting better with some practice. (Report 2)

Chiara also took advantage of every opportunity to participate in class discussions, particularly when they involved interactions with teachers. She explained that the latter afforded ideal opportunities for her to prove her close attention to classroom events and to contribute to them. Apparently, her classmates typically did not engage in these interactions and so discussions often only involved herself and the teacher. As can be seen in 5.2.5, Chiara considered interactions with teachers ideal L2 learning opportunities.
5.2.2.2 Outside the classrooms

With very few exceptions, Chiara spent all break times with her German girlfriends Alia and Alexandra. In term 2, Alexandra’s sojourn ended, and three new German girls joined Chiara and Alia during the breaks.

Chiara very much disliked using English with other Germans and insisted on speaking German with them. She mentioned that it sounds ‘stupid’ and ‘ugly’ when Germans speak English. Conversations with South American international students sometimes occurred on the way home but Chiara did not particularly enjoy speaking to them given their ‘poor’ language performance.

For a short period at the end of May, Chiara spent one break per week with the Kiwi girl Caroline from Textile Technology classes and her friends, because Chiara’s German friends were not available. Chiara mentioned that participation in conversations was difficult given her ignorance of the group’s topics, and because she herself did not have things to talk about. She was thus sceptical about the value of these conversations for her English language improvement. Chiara also expressed a feeling of discomfort in the group setting:

I don’t feel very comfortable because they're always hanging out in groups and her friends are very special and interesting and . . . they're like just eh just like komisch [strange] I don't know just like I really can't describe them they're just like also a little bit exhausting to speak with them because they're laughing about everything and yeah it's like not that they're not nice but it's just like yeah because they're also talking a lot about things I don't know about so yeah…. I'm always sitting there and thinking like ‘ok wow yeah I really don't talk that much’ yeah I also don't have the feeling that my English is getting better from one break. (Int. 5)

Given these conditions, which Chiara interpreted as unfavourable, she soon left the group and spent the breaks with her German friends again.

5.2.3 Outside of the high school

5.2.3.1 Social activities and trips

In her leisure time, Chiara pursued a hobby on some days (see 5.2.3.2) and met with her German friends on most other days. During the entire sojourn, Chiara spent only one afternoon with a Kiwi friend. She preferred to meet with Alia but also regularly saw Alexandra and the German girl Karin. In Germany, Karin attended the same Gymnasium as Chiara and had been Chiara’s best friend for many years.
Chiara and the German girls undertook a lot of activities both in the afternoons and at weekends. They went away on a number of weekend trips to typical New Zealand tourist destinations that were specifically organized for international students. The German girls usually spent all their traveling time together, from sitting in the bus to seeing the attractions, sharing meals, and in their room at night. Together with Karin, Alia, and Alexandra, Chiara also took part in a ten-day South Island trip that the school recommended to international students during the Easter holiday. With the exception of a few South Americans, the travel group consisted of German-speakers only, including the tour guides. Consequently, Chiara spoke, as she termed it, “too much German” (report 3) during the holiday.

Outside school Chiara only met a Kiwi friend once. On one of her last Saturdays in New Zealand, Chiara joined Kelly and her friend at the zoo and the girls subsequently had dinner and watched movies at Kelly’s place. Spending time with Kelly was particularly enjoyable to Chiara not only because Kelly was “very übermotiviert [extremely motivated] and aufgedreht [hyperactive]” (int. 6) but also because the girls did not run out of topics to discuss. In fact, Chiara said that the scarcity with which she interacted with Kiwis outside the classrooms could be explained by her insecurity regarding topics of discussion. She mentioned that she was not afraid to ask friends out, but that she did not share stories with them beyond classroom talk:

I’m not scared to ask them to do something I just don’t’ know what to do that’s a problem because yeah what should we do sitting there and like a fact because we’re just always talking about lessons and what stuff that’s the only problem. (Int. 4)

5.2.3.2 Hobbies

Chiara participated in a number of extra-curricular activities organised by her college. In the first school term, she joined the dragon-boating team, the tennis club, and she participated in a choir. In the second term, Chiara did rock climbing. Chiara explained that she decided to participate in these hobbies to prevent the boredom of being home alone and to make Kiwi friends.

After joining these clubs, Chiara no longer believed that they were facilitative of the establishments of Kiwi friendships. The tennis club was only joined by international students, particularly Asians and a few Germans whom she already knew. Chiara explained that she did not generally object to making friends with Asians, but that their low English proficiency levels made interaction problematic. The dragon boating club consisted mainly of Year 12 and Year 13 students, who Chiara considered too old to make friends with. Joining the choir was also
not conducive to the establishment of friendships for Chiara because the activity itself did not lend itself to interaction. Chiara did not make friends with local adolescents during rock climbing because she chose to spend the climbing sessions with her German friend Alexandra (who she spoke German with).

In the second term, Chiara attended a two-day choir camp to prepare for a school concert. All attendees were domestic students and the camp facilitated a number of social interactions that required the use of English (e.g. chatting to her dorm mates). The concert that took place the following week was a chance for families to meet. Chiara’s host parents did not attend the event but her host grandmother did. The weekend away was generally very enjoyable for Chiara and gave her opportunities not only to speak English but also to grow in confidence in unfamiliar surroundings.

This Weekend I went to the choir camp. I was a little bit nervous because I didn’t knew anybody (but I already noticed that I am not nervous any more before something new happens). I just take everything as it comes…All in all I really enjoyed the Weekend and I am happy that I went there although that I didn’knew anybody at the beginning. (Blog 25/05/2014)

5.2.4 Summary

The opportunities Chiara had for English language interactions varied considerably between the three social contexts. Yet, they did not significantly change within the contexts over time. The description above showed that Chiara predominantly spoke German with her L1 peers in the school breaks, during social activities outside of school (including weekend trips), in the home (i.e. via skype and text messaging) and, when possible, during club activities (i.e. rock climbing). English language exchanges occurred during the lessons with classmates and teachers, sporadically in her host family and during hobbies.

5.2.5 Perception of L2 progress

Below, I will summarise Chiara’s language learning goals and perceptions of her L2 development during her sojourn.

Chiara had two main goals of learning English, relating to different stages of her sojourn: accent-related goals pertained to the sojourn itself while grammar-related goals were important for her academic studies after the sojourn:
I just want to improve like the accent because I think like not at the moment but normally my grammar isn't that bad and it's getting really good so it's mainly the accent which is really bad so I really don't like it . . . it's not really a goal before I leave it's when I come back that when I write a text for English at school because I have eh choose English [English intensive course] so that when I write a text the teacher read it and don't see that much mistakes and the teachers he like that I'm quite good in English that would be really important for me. (Int. 4)

Hence Chiara’s orientation to language learning during her SA appeared principally communicative (i.e. she focused on meaning). She believed that she could improve her language proficiency through extensions of social networks with Kiwis, and simply by increasing her speaking time. When she concluded that it was difficult to establish friendships with local students, she accepted that her English proficiency level might not increase significantly over time but that it was going to be “good and better than now” (int. 2). Chiara also emphasized in interview 6 that establishing social networks did not require English at a high level of proficiency as long as she could communicate what she wanted to say:

so I don’t think language is so important because I also know [a German boy] he his English is quite interesting ehm and he has many Kiwi friends so I don’t think the English matters as long as you can show what you want or what you don’t want. (Int. 6)

Conversely, for the EFL classroom context in Germany, Chiara felt a need to produce error-free essays. Besides having a clearly instrumental goal (i.e. the desire to pass exams), it seems that Chiara’s grammar focus was also a result of peer-pressure. She admitted feeling anxious with native speaker peers in her EFL classroom in Germany because she felt their English was better than hers and that this would make her performance less successful in comparison.

Chiara’s perceptions of her own L2 development were mixed. In interview 1, she noted that she became better at switching between English and German, that understanding Kiwis had become easier, but that her speaking skills remained on the same level as when she left Germany.

In interview 2, Chiara believed that understanding English was now ‘really really good’, including understanding her English teacher who spoke very fast and everything on TV including the news. Chiara additionally noticed that she could write faster in English than at the beginning and that she would exclusively think in English when it came to ‘bus numbers’. However, she explained that there were also times when she got lost listening to classmates when they spoke very fast because they had their “own language” (int. 2) (she presumably refers to teenage slang here):
someone says something and I really don’t get it because there are a view [i.e. a few] words which Kiwis pronounce slightly different how I learned and that’s sometimes a little embarrassing because that are mostly very very easy things but I don’t get them. (Report 1)

In interview 3, Chiara reported that she often let people know that she had understood them when really she did not. Chiara still found it difficult to understand jokes, swear words, and English text messages. In interview 4, Chiara noticed a rather dramatic deterioration in her English, French, and German as a result of her trip to the South Island:

so at the moment all my languages are really horrible my German is horrible I always can’t find any words or just have really bad grammar which I normally don’t have ehm my English is a little bit badder [i.e. worse] than when I left to South Island but I think it’s not that bad like at the beginning and yeah you know also my French is getting bad but that’s another thing. (Int. 4)

From May (interview 4/report 3) onwards, Chiara started to notice oscillations in her English performance and started to draw a distinction between good L2 performance inside and poor L2 performance outside the classroom. She generally felt that listening to the teacher was the best English practice that she could get and that even when her English worsened as a result of significant exposure to German, it soon improved again as a consequence of being attentive during school lessons:

After the trip my English got so bad again but after a couple of hours in school again just by hearing the teachers talking my English was at the same level as when I left. (Report 3)

On many occasions, Chiara felt that her frequent involvement with German speakers temporarily took a toll on her L2 performance. She mentioned that break times and trips with Germans resulted in poorer performances for a while. Despite this awareness, she did not want to change her circumstances. Chiara seemed to generally prioritize participation in social activities with friends who were trustworthy and able to grasp her situation as an international student in New Zealand rather than to search meaningful L2 interactions. For Chiara, these friends were Germans, particularly Alia.

Chiara also turned down opportunities to speak English to Germans, which could have afforded her L2 practice despite the scarcity of native speaker friends. For example, Chiara refused her friend Karin’s expressed desire to speak English with her due to her aversion to Karin’s English accent. At the beginning, Chiara also mentioned attempts to speak English with Alia. However,
the girls’ attempts at interacting in English were not successful due to eventual communication breakdowns:

    with Alia I'm always speaking German because like we've always say ‘ok now we're speaking English’ and we're keep it for ten minutes and then there are so many words and somebody say ‘ok it's hard to explain in English and then we fall in German. (Int. 2)

Finally, in interview 6, Chiara felt that she could have a normal conversation with her classmates, unless it got “a little bit zweideutig [ambiguous] or unappropriate” (int. 6). Overall, she felt that her self-confidence increased most during her sojourn but she generally concluded her sojourn feeling “really happy” about her English progress (int. 6).

5.3 L2 development

Below I will describe how each CAF measure developed over time for Chiara.

5.3.1 Grammatical complexity

5.3.1.1 The subordination index

Chiara’s subordination index (SI) increased overall from 1.49 clauses per unit (cpu) in interview 1 to 1.66 cpu in interview 6 (+11%) (Figure 20). In interviews 2 and 4, Chiara’s SI registered below her score in interview 1. Peak performances were reached in interviews 3 and 6. A correlation coefficient of .658 reveals that the linearity of the development was moderately positive.
With an average of usaha 1.53 cpu, Chiara scored below those measured on average for the NNSs and NSs in Mora and Valls-Ferrer’s (2012) study (NNSs: usha 1.60/1.59/1.62, NSs: usaha 1.83), in Llanes et al.’s (2012) study (NNSs usha 1.70/1.87) and in Polat and Kim’s (2014) study (NSs usha 1.97; range: 1.82-2.22). However, her SI levels were within the ranges obtained by Polat and Kim’s NNS (usaha 1.15-2.00) and Nippold et al.’s (2005) 25-year old NS participants (usaha 1.08-1.82). As mentioned in Jana’s chapter, these differences are probably more reflective of data analysis procedures than (just) proficiency levels. The variation in Chiara’s SI development tends to mainly reflect the use of direct speech reports as will be shown below.

**Range and Frequency of Subordinate Clauses**

Figure 21: Subordinate clauses for Chiara

With the exception of non-finite adverbial clauses, Chiara made use of all types of subordinate clauses right from the beginning (Figure 21). Finite complement clauses formed the majority of subordinate clauses in her speech (usaha 16-35 per interview), followed by finite adverbial clauses (usaha 5-17). Chiara infrequently used finite relative clauses (usaha 1-6) and non-finite adverbial clauses (usaha 0-2). Non-finite complement clauses occurred nine times in interview 1 but very infrequently in the remaining interviews (usaha 0-3 clauses). Interviews 3 and 6, which displayed the highest subordination index (see 5.3.1.1), contained the largest numbers of finite complement clauses. I will show below, that they primarily took the form of direct speech complements.

As observed for Jana, and supported by Nippold et al. (2005), Chiara’s distribution of finite subordinate clauses can be considered ‘normal’, or, at least, appropriate to conversational discourse. Infrequent use of demanding non-finite verb structures is considered appropriate to
the speech activity. In terms of frequency, Chiara seems to have underused finite relative clauses. Her use of finite complement and adverbial clauses resembled those found for the 17-year old NSs in Nippold et al.’s (2005) study.

It was beyond the scope of this study to examine the nature of Chiara’s clause types and their relationship to L2 progress.

**DIRECT SPEECH COMPLEMENTS**

Like for Jana, direct speech played a significant role in Chiara’s output. I observed that interviews 3 and 6 stood out for their comparatively high SI scores and atypically frequent finite complement clauses. Jana’s analysis showed that high SI levels can be a reflection of extensive use of direct speech reports. A closer analysis of Chiara’s finite complement clauses reveals the same results: her highest SI levels (int. 3 and 6) co-occurred with the highest numbers of direct speech complements, both in terms of raw numbers and percentages per total number of clauses (Table 26). In interviews 3 and 6, direct speech complements occurred in a total of 36 and 35 utterances (as dependents of matrix clauses and independently in a following unit), respectively and accounted for 26% (int. 3) and 29% (int. 6) of all clauses (including independent clause)\(^{iv}\).

<table>
<thead>
<tr>
<th>Interview</th>
<th>Raw number of matrix clauses per 100 AS-units (and number of dependent direct speech clauses)</th>
<th>Raw number of following AS-units with direct speech clauses (and number of clauses involved)</th>
<th>Total number of utterances containing direct speech per 100 AS units</th>
<th>Ratio of direct speech complements / total number of finite complement clauses</th>
<th>Total number of direct speech clauses (ratio direct speech clauses / total number of clauses)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9 (10)</td>
<td>2 (2)</td>
<td>11</td>
<td>9/16 (56%)</td>
<td>12/149 (8%)</td>
</tr>
<tr>
<td>2</td>
<td>11 (11)</td>
<td>3 (4)</td>
<td>14</td>
<td>11/17 (64%)</td>
<td>15/138 (11%)</td>
</tr>
<tr>
<td>3</td>
<td>23 (26)</td>
<td>13 (16)</td>
<td>36</td>
<td>25/32 (78%)</td>
<td>42/160 (26%)</td>
</tr>
<tr>
<td>4</td>
<td>16 (21)</td>
<td>4 (4)</td>
<td>20</td>
<td>20/20 (100%)</td>
<td>25/160 (16%)</td>
</tr>
<tr>
<td>5</td>
<td>13 (17)</td>
<td>2 (2)</td>
<td>15</td>
<td>16/24 (66%)</td>
<td>19/148 (13%)</td>
</tr>
<tr>
<td>6</td>
<td>28 (33)</td>
<td>7 (11)</td>
<td>35</td>
<td>33/35 (94%)</td>
<td>44/154 (29%)</td>
</tr>
</tbody>
</table>

As with Jana, it will be seen in the following sections that Chiara’s use of direct speech reports had a significant impact not only on the scores she achieved for subordination but also on mean length of clause and the fluency indices speech rate and mean length of run. Peaks in
performances in interviews 3 and 6 for these measures must thus be treated with caution, as they are hardly indicative of Chiara’s general abilities to produce more complex or faster speech, but reflect processes related to the communicative choices she made in these interviews.

5.3.1.2 Mean length of clause

Figure 22: Mean length of clause for Chiara

Figure 22 illustrates that Chiara’s mean length of clause (MLC) changed from 5.95 words per clause (wpc) in interview 1 to 5.30 wpc in interview 6 (-11%). Two dips marked the development (int. 3 and 6). The remaining interviews contained very similar MLC scores (5.86-5.95 wpc). A correlation coefficient of -.536 implies that the development was moderately linear.

Chiara’s average performance (≈5.71 wpc) was very similar to the average values and the ranges calculated for the NS and NNSs in Polat and Kim (2014) (NS: ≈5.68 wpc; range: 5.05-6.76 wpc; NNS: 4-6.67wpc). Although Polat and Kim seemed to employ a different method of measuring non-finite structures (i.e. they counted item-based non-finite structures as clauses), Chiara’s performance regarding this measure can most likely be considered native-like.

I have suggested above that the drops in interviews 3 and 6 are not reflective of L2 regression but variations associated with Chiara’s use of speech modes.

Jana’s analysis showed that narrations with much direct speech resulted in less grammatically complex utterances than descriptions. In Chiara’s case, utterances with direct speech reports typically contained formulaic matrix clauses of the structure (conjunction) + subject + ‘be like’/’think’/’say’/zero-quotative’ with clause lengths ranging between 3 to 4 words (see
Example 40 to Example 45). The direct speech complements, on the other hand, varied greatly in length, both in interview 3 and 6. Some complements consisted of one-word exclamations (Example 40 and Example 41), some were of medium length (Example 42 and Example 43) and some contained ten words and more (Example 44 and Example 45).

Example 40 (int. 3): I was like :: ‘ok’.
Example 41 (int. 6): and was like :: (.59) ‘ok’.
Example 42 (int. 3): because first I thought :: ‘maybe it's just me?’
Example 43 (int. 6): and it was ::’ok then we also can talk German’.
Example 44 (int. 3): I was like :: ‘no he's just a very very good friend in Germany I'm sorry no’.
Example 45 (int. 6): and then I went into kitchen :: and say :: ‘ah I'm sorry I really don't like it that much’.

As with Jana, utterances containing descriptions tended to be longer. They included complexifying devices such as prepositional phrases (‘in Germany’), adjectives (‘only’), and adverbs (‘normally’ ‘really’) (Example 46 and Example 47).

Example 46 (int. 1): because normally when I come home in Germany :: (.70) my mom is there.
Example 47 (int. 4): (.70) and last time we had them :: (.47) my homestay family really they freaked out.

Example 46 and Example 47 show that longer clauses of this kind are not necessarily indicative of complexity in the sense of more advanced language use. They are simply the consequence of a speaker’s choice to structure the dialogue in a comprehensible manner. Brown and Yule (1983) stated that when communication goes beyond informal chats and speakers are describing or justifying events or people, they take “responsibility for creating a structured sequence of utterances which must help the listener to create a *coherent* mental representation of what he is trying to say” (p. 17, emphasis in original). This means that a speaker must “make it clear who or what he is talking about, and specify any relevant properties” of the narrative (p. 17). Hence depending on the purpose of Chiara’s account, longer clauses (e.g. for descriptions) or shorter clauses (for stories) seemed more appropriate.
5.3.2 Lexical diversity

Figure 23: Lexical diversity for Chiara

Chiara began her sojourn with a lexical diversity (LD) score of 8.23 (as denoted by ‘R’) and finished it with R=7.52 (Figure 23), making for an overall change of -9%. The figure illustrates that all scores before interview 6 increased in comparison to the score in interview 1. The linearity of the graph is negligible as indicated by a very low correlation coefficient (r= -.285).

Compared to Guiraud’s Index results measured in previous studies, Chiara’s values were higher at all times (Levkina & Gilabert, 2012: R= 5.84-6.24; Daller & Xu, 2007: R= 5.03-6.18; Lu, 2012: R= 3.3-6.9; Mora & Valls-Ferrer, 2012: R= 6.76-6.97). With an average of R=8.46, Chiara’s values were also above the average scores measured for the NSs in Mora and Valls-Ferrer’s study (2012) (≥7.86).

While Chiara’s increase in LD was not enormous, it is still remarkable that they were higher than those measured for NSs in Mora and Valls-Ferrer’s (2012) study. The authors explained that their LD scores were affected by the “relatively structured and cognitively simple” task (p. 616). It is possible that the freedom Chiara was given regarding the language she produced in response to the interview questions of this study encouraged her to formulate more complex messages than Mora and Valls-Ferrer’s NSs, which pushed a need for lexically more diverse language.

Chiara’s extensive use of direct speech reports did not appear to impact LD scores dramatically, at least not in interview 3 (i.e. LD scores did not decrease). I rather suspect that during the first five interviews, LD levels increased as a result of Chiara’s interactions in the naturalistic context (i.e. she broadened her linguistic repertoire). Clearly, differences in topic or the number of elements Chiara wove into her narration may have also affected her LD scores.
5.3.3 Fluency

5.3.3.1 Breakdown fluency

**Empty pauses**

Figure 24: Empty pauses for Chiara

Figure 24 shows that Chiara reduced the number of empty pauses (EPs) from 14.91 per hundred words (phw) in interview 1 to 10.77 phw in interview 6 (-28%). Development did not progress in a linear fashion but in a moderately negative (r= -.656) manner, marked by two temporary increases in EPs in interviews 4 and 5.

With a mean score of 12.65 EPs phw, Chiara used fewer EPs than the highly advanced NNSs in Götz’s (2013) study (≏15.13 EPs phw), but performed within their range (8-30 EPs phw). Similarly, Chiara performed within the range measured for Riggenbach’s (1991) ‘fluent’ NNS group (10-22 EPs phw) and well above that of the ‘nonfluent’ NNS group (33-73 EPs phw). Compared to the NS group in Götz (2013) (≏3.99 phw) Chiara significantly overused EPs. However, she approximated the highest score Götz reported for one NS (9 EPs phw) in interviews 3 and 6. Thus when individual scores are considered, it could be argued that Chiara paused in a near-native-like way in these two interviews.
Chiara started her sojourn with 3.37 filled pauses (FPs) per hundred words (phw) and finished it with 1.76 FPs phw (Figure 25). Apart from small increases in interviews 4 and 6, she gradually reduced the number of FPs by a total of 48%. A correlation coefficient of -.788 shows that the negative development was highly linear.

Compared to Götz’s (2013) NNSs (ӹ5.12) and Hasselgren (2002)’s more fluent adolescent NNS group (ӹ5.98), Chiara used significantly fewer FPs on average (ӹ2.15). She also used slightly fewer FPs on average than Götz’s (2013) adult NSs (ӹ2.27) and Hasselgren (2002)’s adolescent NSs (ӹ2.81). When considering individual speakers’ performances, Chiara scored within the range both of Götz’s (2013) NNSs (1-14 FPs phw) and NSs (0.5-8 FPs phw). 16% of Götz’s German participants ‘underused’ FPs like Chiara, which Götz explained with individual speaker pausing preference. Possibly, Chiara preferred using pausing phenomena other than FPs to plan her speech (e.g. smallwords).
5.3.3.2 Speed fluency

SPEECH RATE

Figure 26: Speech rate for Chiara

Chiara’s Speech Rate (SR) developed from 149.31 words per minute (wpm) in interview 1 to 191.83 wpm in interview 6, accounting for a total change of 28% (Figure 26). The SR development is marked by peaks in interviews 3 and 6. The graph reveals that the scores in interviews 2, 4, and 5 were relatively similar. The correlation coefficient (r=.705) indicates a highly linear development.

With an average SR of 174.34 wpm, Chiara scored well above Lennon’s (1990) intermediate level NNSs (76-162 wpm), as well as Götz’s (2013) advanced NNSs (≥160 wpm, 117-190 wpm). Chiara’s speech was slower compared to the average SR produced by Götz’s NSs (≥213 wpm) but within their range (approximately 174-280 wpm) in four out of six interviews (int. 2, 3, 5 and 6). According to Tauroza and Allison’s (1990) categories of British English SR (i.e. NS scores), Chiara would have initially been classified as a ‘moderately slow’ (120-160 wpm) speaker (interview 1) but a speaker of ‘average’ speed (160-210 wpm) for the remainder of her sojourn.

All comparisons suggest that Chiara not only started out as a fast NNS, but improved to an extent that she could no longer be distinguished from the average NS in terms of this measure. Faster SR is believed to reflect higher proceduralized knowledge and/or the use of more formulaic sequences (Wood, 2012). The peaks in interviews 3 and 6, which coincide with high numbers of direct speech reports (see 5.3.1.1) suggest that high SR scores were not just a matter of generally more automatized speech processes but of processes related to Chiara’s use of a particular type of narration – one which has shown to be partly formulaic in nature.
Chiara began her sojourn with 6.57 words between empty pauses (of at least 0.3 seconds) and ended it with 9.11 words per run (wpr) (Figure 27), accounting for an overall increase of 39%. The MLR development mirrors that of speech rate (SR), showing peaks in interviews 3 and 6, and three very similar scores in interviews 2, 4, and 5. The correlation coefficient shows that the development was moderately positive ($r=.636$).

Chiara produced average runs (≈7.70 wpr) that were significantly longer than those obtained for upper-intermediate/advanced proficiency levels (Mora & Valls-Ferrer, 2012: ≈6.17/5.93/7.50 wpr; Götz, 2013: ≈5.88 wpr). Her overall rate of change (+2.54 wpr) was much higher than the average found in Mora and Valls-Ferrer (1.33 wpr over 9 months). Compared to NS scores (Pawley & Syder, 1983: ≈8–10 wpr; Götz, 2013: ≈10.3 wpr; Mora & Valls-Ferrer, 2012: ≈12.49 words), Chiara performed less native-like, yet approximated average scores in interviews 4 and 6. In the following section, the changes in Chiara’s MLR scores will be interpreted.

As in Jana’s case, it appears that the differences in MLR scores are to a large degree attributable to the use of direct speech reports. I discussed in Jana’s chapter that the production of direct speech appears cognitively less demanding because complements do not require syntactical embedding and are mostly uttered in the present tense. Primarily, though, it seems that the cognitive load is lessened for direct speech reports because the reporting clauses function as “formulaic introducer[s]” (Tannen, 1986, cited in Romaine & Lange, 1991, p. 236) and serve as planning phases. As mentioned in 5.3.1.2, the reporting clauses in Chiara’s speech exhibited the same formulas repetitively ((conjunction) + subject + ‘be like’/’think’/’say’/’zero-quotative’).
It is interesting to note that Chiara rarely paused between reporting clauses and direct speech complements, or within direct speech complements. She sometimes produced very long runs (e.g. 28 words) which transitioned from one utterance to the next (utterance termination indicated by ‘|’) (Example 48):

Example 48 (int. 6): (.30) so yeah I don't know | and then he said like :: (.57) ‘yeah bye Chiara’| and was like (.57) and then he expect :: like that I'm :: ‘oh I'm sorry :: that I'm leaving’ | and I was like :: ‘ok yeah bye Mike’ | eh because I'm sorry | it's like (1.02) I don't care | so (.30) and also when I told him today :: (.35) ‘yeah I'm not there for dinner | I'm (.38) going out | but I probably b want to have something out :: when I came back’ | He was like :: (2.52) ‘ok’ | and was like :: ‘yeah I told Sara this morning’ | (1.15) but again was like :: (.58) ‘sorry I'm going’.

It appears from this example that there was a reduced need for Chiara to pause before reporting clauses and direct speech complements, suggesting that direct speech production was cognitively undemanding for Chiara. I have mentioned in Jana’s chapter that direct speech reports are unlikely to account for all changes in MLR. Other factors might be the dynamics inherent in L2 processing abilities, changes of discourse topic, communicative intentions, rhetorical effects or more complex language use. It is also possible that Chiara’s very long runs contained hesitation pauses of less than 0.3 seconds which were disregarded in this study.

5.3.3.3 Repair fluency

REPETITIONS

Figure 28: Repetitions for Chiara

Figure 28 illustrates that Chiara decreased the number of repetitions from 1.23 per hundred words (phw) in interview 1 to 0.83 phw in interview 6 (-33%). The range with which she used repetitions during the six interviews was small and her development fluctuated considerably (r=-.219).
On average, Chiara (⌀=0.88 phw) used 1.3 times more repetitions than Götz’s (2013) German speakers of English (⌀=0.69 phw) and a similar number of repetitions as Götz’s NSs (⌀=0.82 phw). Götz argued that repetitions provide a means for speakers to plan ahead utterances. In interviews with very low numbers of repetitions (int. 2 and 4), it is possible that Chiara used performance phenomena other than repetitions when she required additional planning time (e.g. filled pauses or smallwords).

**REFORMULATIONS**

Chiara increased reformulations from 2.15 per hundred words (phw) in interview 1 to 3 reformulations phw in interview 6 (+ 40%) (Figure 29). Despite this overall positive trend, most interviews (int. 2, 3 and 5) reported lower scores than those measured in interview 1. The sharpest dip occurred in interview 2. A correlation coefficient of .627 indicates that the variable increased with a moderately positive linearity.

When it comes to L2 progress, it is important to distinguish between reformulations of content and reformulations of form. Reformulations of content have been considered a natural phenomenon of spontaneous speech because the pressures of online production cause speakers to ‘erase’ what they just said and start anew with a different string of words (Götz, 2013; Riggenbach, 1991). Reformulations of form (retracings) constitute a feature that is more prevalent in L2 than L1 speech and, depending on their frequency, can be considered markers of disfluency (Götz, 2013). However, reformulations of form may also reflect a speaker’s awareness of problematic structures/words and the ability to monitor output, and have thus also been considered a signal of proficiency (Riggenbach, 1991).
Like Jana, Chiara much more frequently changed the content/structure of an utterance, than its form. Often, false starts reflected a concern for precision (Example 49). Other reformulations appear to have been made to ensure maximum comprehensibility or to show awareness of communal knowledge (i.e. to cater for the needs of the listener). In Example 50, Chiara reformulated what appears to have been intended ‘and then my host father’ to the sequence ‘then Mike . . .’. I suspect that she realised before finishing the phrase that I was familiar with her host father’s name, and, to signal awareness that knowledge was shared, she made changes accordingly. These types of reformulations have been considered native-like (Götz, 2013; Riggenbach, 1991; Wood, 2012).

Example 49 (int. 2): then (.44) <mostly I take my eh> so first I take my shower.

Example 50 (int. 6): and <then my ho> then Mike always say in a very ironic way :: (.30) ‘yeah goodbye you're leaving us again’.

Chiara’s data contained only a small number of unambiguously linguistic modifications. Some involve changes of grammatical forms (Example 51) and others corrections of lexical items (Example 52). At times, she changed an originally correct word form into an erroneous form (Example 53). Furthermore, Chiara sometimes ‘stumbled’ and delivered slurred speech that she subsequently corrected (Example 54).

Example 51 (int. 1): (.40) because ehm (.38) she were like (.68) eh not in lessons and ehm (.60) eh (.91) <don't> didn't go in lessons.

Example 52 (int. 2): because (.45) the only eh goal I have is like the <mai> (.55) eh head teacher in Germany says (.34) ok you can go in oberstufe [senior high school] now but yeah.

Example 53 (int. 1): one fun (.33) fun subject <is is> ehm (.87) were ok.

Example 54 (int. 6): <of oc> of course I (1.59) sometimes don't eat everything.

These examples show that Chiara paid attention to form. However, they also illustrate that she did not always possess the necessary competence to successfully repair her output. Moreover, Chiara left a good number of errors unmodified (see 5.3.4). This suggests that she either did not notice all her errors during online speech (possibly due to attentional deficits or insufficient declarative knowledge) or that she was aware of them but chose not to repair them (e.g. due to a focus on meaning) (or a combination of the two).

The benefits of speech with low numbers of grammatical modifications is that it may resemble that of more fluent speakers in terms of frequency of error-correction (Riggenbach, 1991), and
that it increases speed fluency due to decreased processing burdens. However, failure to actively engage with syntax, can also come at the expense of interlanguage development (Skehan, 1998).

5.3.3.4 Smallwords

Figure 30: Frequency of smallwords for Chiara

In terms of the frequency of smallwords Chiara began her sojourn with 9.09 smallwords per hundred words (phw) and finished it with 7.66 (-16%) (Figure 30). Her development reached a low in interview 3 and a peak in interview 5. A very low correlation coefficient of -.023 reflects a lack of linearity.

Compared to Hasselgren’s (2002) participants, Chiara (SimpleName) used almost twice as many smallwords phw as Hasselgren’s NS group (SimpleName) and over three times as many as her two NNS groups (SimpleName). This difference is particularly striking because Hasselgren measured a greater range of smallwords than I did in this study (see Chapter 3).

In terms of range, Chiara used between six (int. 1, 6), seven (int. 3) and eight (int. 2, 4 and 5) smallwords at any one time (Table 27). She used a total of ten different smallwords from the twelve that were measured in this study. The vast majority of Chiara’s smallwords were the single-word smallwords ‘like’, ‘yeah’, and ‘so’. She also used ‘just’ and ‘oh/ah’ in almost every interview, yet with a lower frequency. In contrast, Chiara hardly used multi-word smallwords apart from ‘I think’ and ‘I don’t know’. All other multi-word smallwords were used two times or less:
With regard to Hasselgren’s (2002) three stages of acquisition of smallwords, Chiara made frequent use of the stage 1 smallwords ‘I think’ and ‘just’, yet very infrequent use of ‘or something’ and ‘a bit’. Compared to Jana, she never used the stage 1 smallword ‘you know’. Out of the stage 3 smallwords (NS level), Chiara frequently used ‘ah/oh’ and ‘like’, but scarcely used ‘and everything’. Like Jana, she did not implement the stage 3 smallwords ‘and stuff’, ‘and things’ into her speech.

Chiara’s development approximated target language norms in some respect, but deviated from these norms in other respects. Progress was marked by increased uses of ‘like’ and ‘just’. Conversely, Chiara underused ‘a bit’ and ‘I think’ and ‘or something’ which is untypical of NSs. Like Jana, Chiara frequently used the smallwords ‘so’ and ‘yeah’ which remain unmentioned in Hasselgren. Especially the latter is assumed to be rare in NS speech (Grieve, 2010). Overall, the very frequent use of ‘like’, ‘yeah’, ‘so’ and ‘just’ implies that Chiara, clung to ‘teddy bear smallwords’ that provided her with ‘islands of safety’ during online speech performance.

Table 27 shows that some smallwords only came into use from interview 2 (‘I don’t know’ and ‘ah/oh’), interview 4 (‘and everything’) or interview 5 (‘a bit’) onward. Given that they all constitute the latest acquired types of smallwords (i.e. stage 3 in Hasselgren)\textsuperscript{xvi}, it is possible that spending a few months abroad was necessary for Chiara to notice these words in NS input, and make use of them herself. Ultimately, it is unclear whether Chiara acquired them at later
stages in the sojourn, or whether she had known and used them before going abroad but did not consider them appropriate prior to their first appearance.

Interestingly, Chiara’s data reveals a relationship between the use of direct speech and smallwords. Interviews with most direct speech (3 and 6) contained the smallest numbers of smallwords. Since smallwords can function as fillers (Carter & McCarthy, 2006; Götz, 2013; Gut, 2009; Riggenbach, 1991; Wong, 2000), it could be concluded that smallwords served the function of fillers in interviews with descriptions (and possibly higher cognitive loads) while the formulaic reporting clauses performed this function in interviews with much direct speech.

Overall, Chiara’s data suggest that she tended to cling to the same ‘teddy bear’ smallwords that she had already acquired before going abroad. Studying abroad did not seem to assist her greatly with the acquisition of multi-word native-like smallwords. However, smallwords may have contributed to a sense of ‘smoothness’ in her speech. It is also possible that she sometimes replaced empty pauses by smallwords, which may have contributed to increased temporal fluency.

5.3.4 Accuracy

5.3.4.1 Error-free clauses

Chiara increased the ratio of error-free clauses (EFC) from 77% in interview 1 80% in interview 6 (+4%) (Figure 31). Besides a peak performance in interview 5, her scores ranged within a very small scale of 77% and 80%. Development was moderately positive which is reflected in a correlation coefficient of $r=.566$. 

162
With an average of 81% EFC, Chiara scored above the mean values obtained by the intermediate-level students in Skehan and Foster (1999) and Kormos and Dénes (2004) (≥ 60% and ≥ 53%, respectively) and the intermediate/advanced-level groups in Mora and Valls-Ferrer (2012) (≥70.50-80.91). However, she had lower mean scores than the advanced-level groups in Kormos and Dénes’s (≥86%). Compared to the NS group in Mora and Valls-Ferrer’s (2012) (≥99.60), Chiara performed less well.

Overall, Chiara’s EFC levels only moderately increased. It is unclear to what extent her interlanguage grammar underwent development and/or remained unchanged. Given that Chiara was a particularly fast speaker, it is possible that the pressures of online production prevented her from planning ahead the forms of her utterances which led to inaccurate, incomplete or simplified language use (Wood, 2012). Maintaining the speech flow rather than delaying articulation for the sake of more accurate language may also have been a conscious choice on Chiara’s part. According to Lennon (2000) this is rather normal as “[s]peakers are quite naturally concerned with maintaining perceived fluency, since failure to do this may result in both loss of listener attention and speaker loss of face” (p. 27).

I have alluded in Jana’s chapter to the danger of confusing language accuracy with language progress. Learners can produce perfectly accurate output with poorly developed grammatical forms, or output containing erroneous forms with more advanced language structures (Pallotti, 2009). Unfortunately, the scope of this study did not allow an analysis of language structures and their relationship to L2 progress.

5.3.4.2 Correct finite verb phrases

Figure 32: Correct finite verb phrases for Chiara
Chiara’s correct finite verb phrase (fvph) scores changed from 84% in interview 1 to 88% in interview 6 (+5%) (Figure 32). The development was marked by a peak performance in interview 4 and moderate linearity (r=.532).

With an average of 90% correct fvph, Chiara scored above the mean obtained by the NNS high-proficiency group in Wigglesworth (1997) (max. 85.51%).

A more in-depth analysis was carried out to examine the nature of the development that occurred for fvphs. Chiara made more verb form errors (a total of 53) than tense-choice errors (34).

During the sojourn, she made the same four types of morphological errors and the same type of tense error. Morphological errors concerned 1) omission of the auxiliary ‘be’ in compound verb tenses (Example 55). 2) ‘Was’/’were’ confusion (Example 56). 3) Agreement errors on the 3rd person (Example 57). 4) Agreement errors on lexical verbs of compound verb tenses (double past-tense marking) (Example 58). Tense errors (often in combination with morphological errors) involved failure to mark past tense (Example 59). Errors of other kinds were rare.

Example 55 (int. 1): (.48) or today (.32) no tomorrow she gonna (.54) no on Fri (.30) on Friday she gonna come back. (<is gonna)

Example 56 (int. 1): so (.71) first thing were to learn how to eat on your legs. (<was)

Example 57 (int. 3): (.39) and also I think she don’t want to hear it. (<doesn’t)

Example 58 (int. 2): I didn't found out about other choirs.<<find)

Example 59 (int. 1): (.48) and but eh she was really nice because Christina visit her back in Argentina (.73). (<visited)

Apart from error types 2 (was/were) and 4 (double past-tense marking), Chiara’s inaccuracies all concern structural simplifications. According to Ellis (2008), omitting “grammatical functors such as auxiliary verbs, articles and bound morphemes like plural –s and past tense-ed” (p. 80) indicates earlier stages of interlanguage. He explained that structural simplifications “may occur either because learners have not yet acquired the necessary linguistic forms or because they are unable to access them in the production of specific utterances” (p. 81).

As mentioned above, Chiara may have also consciously prioritized fluency over accuracy based on the belief that accurate output is less important than erroneous output and that the latter would still satisfy her communicative needs. This would provide evidence for Faerch and
Kasper’s (1983) argument that learners sometimes perform utterances which are not correct but appropriate in the given situation in order to produce fluent speech.

One particular error deserves more attention. Without exception, Chiara used informal ‘gonna’ for ‘be going to’ throughout all interviews. As shown in Example 55, she produced this form without the auxiliary be, which was considered ‘absolutely wrong’ in this study xvii. Given the systematicity of this error, it is doubtful whether it can be attributed to a temporary cognitive overload (de Bot & Larsen-Freeman, 2011). Rather, it seems that Chiara stored the wrong form of the grammatical feature in her long-term memory (maybe due to its phonological resemblance to ‘wanna’ which does not require the auxiliary be). From a DST perspective it could be argued that the erroneous ‘pronoun + gonna’ form represents an attractor state in her developing language which was sufficiently deep that even the naturalistic context (i.e. changing resources) failed to facilitate re-organisation of the system. Maybe Chiara’s exposure to informal NS input was too limited for her to notice this feature and thus made internalization and production of it impossible.

5.3.5 The relationship between CAF

Figure 33: CAF relationships for Chiara

Figure 33 depicts the relationships between the global measures of fluency (SR and MLR), accuracy (EFC), lexical diversity (LD) and grammatical complexity (MLC). To make comparisons possible, I plotted the measures on the graph using Z-scores (see Chapter 3).
According to Figure 33, three supportive and one competitive relationship(s) between the variables can be detected (with greater and lesser strength). The strongest supportive relationship existed between the two fluency measures SR and MLR. There was also a partially supportive relationship between EFC and LD, and a relatively weak partially supportive relationship between MLC and LD (two movements differ). Moreover, a strong competitive relationships existed between the two fluency measures and LD. Below I will turn to a more detailed description and interpretation of each of these relationships.

5.3.5.1 Supporting growers

**Speech rate and mean length of run**

![Figure 34: Relationship between speech rate and mean length of run for Chiara](image)

As can be seen in Figure 34, the two growers SR and MLR follow exactly the same pattern. This relationship implies that faster SR went hand in hand with longer runs (int. 2, 3, 6) while slower SR was accompanied by shorter runs (int.1, 4, 5). This correlation has also been observed in Gut (2009) who showed that speakers with a faster SR produce longer runs. The discussions above about SR and MLR prompted two explanations about the measures’ development. On the one hand, I concluded that the developments mirror to a large extent Chiara’s choice of speech mode (i.e. the use of direct speech), rather than her general ability to speak faster or make longer runs. On the other hand, I suggested that the growth of both measures between interviews 1 and 6 can also be related to higher proceduralization skills.
A partially supportive relationship existed between EFC and LD (Figure 35). With the exception of the move between interviews 2 and 3, the two growers moved up and down in an identical way: both variables increased between interviews 1 and 5 (at different rates) and dropped significantly in interview 6.

The relationship between accuracy and grammatical complexity has been much discussed in the area of task-based research (e.g. Robinson, 2007; Skehan, 1998, 2009; Skehan & Foster, 2005), yet the relationship between accuracy and lexical diversity is hardly mentioned (lexical diversity has generally received little attention).

In analogy with Skehan’s (2009) interpretation of the relationship between accuracy and grammatical complexity, I assume that the supportive relationship between EFC and LD in Chiara’s speech is explained by the co-existence of different factors inherent in speech production processes that were facilitative of or debilitating to the development of accuracy and lexical diversity at the same time. For accuracy, the discussion of frequently discussed topics (e.g. host family) with similar microstructures may have eased conceptualizing (i.e. planning) demands and allowed Chiara to pay more attention to syntactical encoding processes, which led to more automatic retrieval of the same structures over time (i.e. higher accuracy levels). Maybe, automatization of syntactical encoding processes then made room for lexical encoding processes, leading to more lexically diverse output. Another possibility is that the supportive relationship between EFC and LD is coincidental, and that EFC may not be well suited to comparisons of this kind.
GRAMMATICAL COMPLEXITY AND LEXICAL DIVERSITY

Figure 36: Relationship between mean length of clause and lexical diversity for Chiara

According to Figure 36, a partially supportive relationship existed between the MLC and LD. With two exceptions (movements between int. 1-2 and 4-5), the two growers exhibited the same developmental patterns (albeit with rather different rates of change). The graph illustrates that the more lexically diverse Chiara’s speech was, the more grammatically complex it was (int. 4) and vice versa (int. 3 and 6).

As discussed in Jana’s chapter, evidence for this relationship can be found in Kormos’ (2005) speech production model which links the production of lexis with that of syntax.

In Chiara’s case, two key differences between the two variables must be noted. The first one relates to the overall development of the two variables. While none of the MLC values reached a score higher than in interview 1, all LD scores but the last one increased over time. The second difference concerns the dips in interviews 3 and 6. MLC scores reached an overall low in both interviews 3 and 6, while LD only reached an overall low in interview 6 (lows were associated with direct speech reports). The relationship between these two growers did not seem particularly strong.
5.3.5.2 Competitive growers

**FLUENCY AND GRAMMATICAL COMPLEXITY**

Figure 37: Relationship between fluency and grammatical complexity

![Figure 37: Relationship between fluency and grammatical complexity](image)

Figure 37 reveals that there was a perfectly asymmetrical relationship between fluency (SR and MLR) and MLC. In interviews 2, 3, 5 and 6, higher levels of fluency went hand in hand with lower levels of complexity. In interview 4, lower levels of fluency were accompanied by higher levels of complexity.

As discussed above, Chiara’s use of direct speech reports appears to have significantly impacted the scores of these measures in interviews 3 and 6. The cognitive ‘undemandingness’ of direct speech reports appears to have allowed Chiara to produce longer runs and more fluent speech than otherwise possible. The repetitive structure of the reporting clauses appeared to lead to language of lower than usual complexity (i.e. length). This trade-off between fluency and complexity has been repeatedly noted in task-based research (e.g. Foster & Skehan, 1999).

Chiara’s overall increase in SR and her overall decrease in MLC points to language production with a strong focus on meaning. Skehan and Foster (1999) assumed that NNSs who prioritize meaning over form might draw on a more lexicalized system, which seems to explain the rise in Chiara’s lexical diversity and the repetitiveness of her direct speech quotatives. Conversely, Chiara’s prioritization of speed over grammatical complexity (MLC) seems to have led to performance with stabilised complexity levels as well as ‘fossilized’ language (e.g. ‘gonna’). While this communicative orientation may have been beneficial for Chiara during chats with NS peers, it may have presented a disadvantage for her long-term interlanguage development (Skehan, 1998).
Lastly, it is interesting to note that the measures that were sensitive to Chiara’s use of direct speech formed close supportive (SR and MLR) and close competitive (SR/MLR versus MLC) relationships. Conversely, the measures that were less affected by speech mode (accuracy and lexical diversity) showed weaker supportive relationships with other measures (e.g. MLC) and no clearly competitive relationships with any of the measures. Maybe strong supportive or competitive relationships between the CAF measures are indicative of difficulties pertaining to the communicative activities. Conversely, a weak relationship between CAF measures may suggest that language progress took place, and that language production for these variables occurred (more) independently of situation-specific demands.
Table 28: Results of all CAF measures for Chiara

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>average</th>
<th>Overall development</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subordination Index</td>
<td>1.49</td>
<td>1.38</td>
<td>1.6</td>
<td>1.48</td>
<td>1.54</td>
<td>1.66</td>
<td>1.53</td>
<td>+11%</td>
<td>.658</td>
</tr>
<tr>
<td>Mean length of clause (wpc)</td>
<td>5.95</td>
<td>5.87</td>
<td>5.41</td>
<td>5.88</td>
<td>5.86</td>
<td>5.30</td>
<td>5.71</td>
<td>-11%</td>
<td>-.536</td>
</tr>
<tr>
<td>Lexical diversity (Guiraud’s-Index)</td>
<td>8.23</td>
<td>8.75</td>
<td>8.50</td>
<td>8.85</td>
<td>8.89</td>
<td>7.52</td>
<td>8.46</td>
<td>-9%</td>
<td>-.285</td>
</tr>
<tr>
<td>Empty Pauses (phw)</td>
<td>14.91</td>
<td>12.70</td>
<td>11.36</td>
<td>13.17</td>
<td>13.01</td>
<td>10.77</td>
<td>12.65</td>
<td>-28%</td>
<td>-.656</td>
</tr>
<tr>
<td>Filled Pauses (phw)</td>
<td>3.37</td>
<td>2.31</td>
<td>1.71</td>
<td>2.21</td>
<td>1.54</td>
<td>1.76</td>
<td>2.15</td>
<td>-48%</td>
<td>-.788</td>
</tr>
<tr>
<td>Speech Rate (wpm)</td>
<td>149.31</td>
<td>174.29</td>
<td>186.39</td>
<td>170.48</td>
<td>173.73</td>
<td>191.83</td>
<td>174.34</td>
<td>+28%</td>
<td>.705</td>
</tr>
<tr>
<td>Mean Length of Run (wpr)</td>
<td>6.57</td>
<td>7.40</td>
<td>8.56</td>
<td>7.24</td>
<td>7.34</td>
<td>9.11</td>
<td>7.70</td>
<td>+39%</td>
<td>.636</td>
</tr>
<tr>
<td>Repetitions (phw)</td>
<td>1.23</td>
<td>0.58</td>
<td>0.96</td>
<td>0.63</td>
<td>1.02</td>
<td>0.83</td>
<td>0.88</td>
<td>-33%</td>
<td>-.219</td>
</tr>
<tr>
<td>Reformulations (phw)</td>
<td>2.15</td>
<td>1.39</td>
<td>2.04</td>
<td>2.21</td>
<td>1.95</td>
<td>3.00</td>
<td>2.12</td>
<td>+40%</td>
<td>.627</td>
</tr>
<tr>
<td>Smallwords (phw)</td>
<td>9.09</td>
<td>9.84</td>
<td>6.22</td>
<td>8.64</td>
<td>11.17</td>
<td>7.66</td>
<td>8.77</td>
<td>-16%</td>
<td>-.023</td>
</tr>
<tr>
<td>Error-free clauses</td>
<td>77%</td>
<td>80%</td>
<td>80%</td>
<td>80%</td>
<td>88%</td>
<td>80%</td>
<td>81%</td>
<td>+4%</td>
<td>.566</td>
</tr>
<tr>
<td>Correct finite verb phrases</td>
<td>84%</td>
<td>89%</td>
<td>90%</td>
<td>94%</td>
<td>93%</td>
<td>88%</td>
<td>90%</td>
<td>+5%</td>
<td>.532</td>
</tr>
</tbody>
</table>

wpc: words per clause; phw: per hundred words, wpr: words per run

The development of lexical diversity (LD) was devoid of a clear trend while the grammatical complexity measures, subordination index and mean length of clause (MLC) developed in a
moderately linear positive and negative fashion, respectively (i.e. between \( r=+/-0.50 \) and 0.70). All MLC levels dropped below the first one (Table 28).

Chiara’s fluency variables underwent the most considerable changes, both in terms of overall growth, and linearity of development. The biggest rate of change occurred with filled pauses (FPs) (-48%), followed by reformulations (+40%), mean length of run (MLR) (+39%), repetitions (-33%), empty pauses (Eps) (-28%), speech rate (SR) (+28%), and smallwords (-16%). In terms of linearity of development, SR and FPs developed in a highly linear way (i.e. between \( r=+/-0.70 \) and .90). MLR and EPs developed in a moderately linear way (i.e. between \( r=+/-0.50 \) and 0.70). Smallwords showed no linearity in their development.

Both accuracy measures (ratio of error-free clauses and ratio of correct finite verb phrases) increased in a moderately positive fashion by a total of +4% and 5% respectively.

### 5.4 Learning opportunities and L2 development

The discussion of Chiara’s second language (L2) development above showed that her speech became more fluent, and more accurate but overall less grammatically complex and lexically diverse. The lexical diversity levels only decreased in the last interview, but increased during all other times. The discussion of these results revealed that many measures ‘waxed and waned’ in particular as a result of the choices she made in each interview regarding the use of narration techniques. Speech modes appeared to affect her speed fluency and grammatical complexity measures. Other measures, such as lexical diversity and accuracy, seemed less affected by her choice of speech mode and changes in these measures are likely to be more reflective of L2 progress. I will discuss below potential reasons for changes in Chiara’s grammatical complexity, accuracy, lexical diversity and fluency. I will finish this section with a short summary.

Chiara’s L2 activities involved a mixture of informal and more formal conversations, as well as analytical language learning activities. Informal conversations mainly occurred in the host family setting and inside classrooms, and to a small degree in her free time (while she was doing her hobbies). The discussions Chiara had with her teachers about subject-relevant topics are assumed to have been of a more formal nature. Analytical language learning appears to have taken place in the English and the French classrooms as a result of activities such as essay writing, speech preparations and translation tasks. Chiara’s serious engagement with
homework, which included looking up words in the dictionary, provided further opportunities for analytical language learning.

Informal language learning activities changed with regard to their frequency in the classroom setting and during activities outside of school, but did not appear to vary much within the host family setting. Informal chats during classes occurred more often with the onset of Term 2 than during Term 1, due to new seating orders. Chiara had more opportunities for informal conversations in the clubs she joined in Term 1 (dragon boating, tennis, choir) than in Term 2 (rock climbing), particularly because she pursued the latter with a German friend. The frequency with which more formal language exchanges occurred is less clear. Similarly, it is unclear to what extent (both regarding frequency and quality) analytical language learning activities took place, as Chiara infrequently referred to activities of this kind. However, Chiara made reference to homework (which can be considered an analytical language learning activity) right from the start. The extent to which Chiara used German appeared relatively stable throughout her whole sojourn. She used her L1 during almost all break times at school and for all social activities outside of school, except for her hobbies. She also consistently skyped with her German family for roughly two hours every day.

As discussed in Jana’s chapter, conversations with native speakers (NSs) have the potential to improve learner’s L2 comprehension and production skills in a number of ways. Comprehensible input can prompt active listeners to notice ‘gaps’ between their interlocutor’s use of grammatical structures and lexical items, and their own. Producing output might force “syntactic processing, testing hypotheses, and developing automaticity” (Skehan, 1998). Language progress has been in particular linked to communication that involves the negotiation of meaning, corrective feedback, and the use of communication strategies.

Chiara arrived in New Zealand with native-like clausal complexity levels (MLC). During the sojourn, her MLC levels decreased when she made frequent use of direct speech reports, but remained relatively stable during all other interviews. Chiara’s subordination index (SI) scores also fluctuated as a consequence of speech mode choices, but increased overall, approaching NS norms in interview 6.

As observed for Jana, her failure to increase her MLC levels cannot be considered a lack of L2 progress but reflective of appropriate language behaviour given the social contexts she participated in. Both adolescent and adult native speakers are said to produce output of low complexity in informal spontaneous communication - characterized by colloquial, syntactically
non-complex, paratactic, incomplete, lexically non-specific language features (Brown and Yule, 1983). In the SA context, then, Chiara may simply not have been required to formulate grammatically more complex output to sound ‘native-like’. Likewise, the dips in Chiara’s MLC development can hardly be interpreted as a sign of deteriorating language skills. They indicate convergence to the SA context in that they reflected heavy use of direct speech reports. This speech mode is typical for female adolescents’ speech (Lampropoulou, 2013; Renouf & Kehoe, 2011; Romaine & Lange, 1991). Although the data showed that Chiara was able to produce this speech mode right from the beginning, increased use in interviews 3 and 6 may indicate that using direct speech became more important over time to enable her to converge with adolescent NS standards (Giles & Coupland, 1991).

Chiara’s higher accuracy levels are believed to be mainly a function of using English of relatively low complexity (discussing similar topics) in the naturalistic context. A potential advantage of producing output with “limited syntax” (Brown & Yule, 1983, p. 7) as typical (or ‘demanded’) in informal conversations is that it allows speakers to pay more attention to encoding and monitoring processes. The more resources a speaker has available for these processes, the more accurate and fluent the output can become (Kormos, 2011; Skehan & Foster, 1999). Chiara’s specific accuracy results (significantly more correct finite verb phrases) suggest that she was able to apply her knowledge of morphological and syntactical rules increasingly more automatically. Possibly, these forms were particularly salient or occurred particularly frequently both in the NS input she received (i.e. she may have noticed gaps) and in her own output (she may have tested L2 hypotheses). It is also possible that Chiara’s accuracy levels increased as a result of her more formal analytical language learning activities in the New Zealand classrooms.

Despite the overall progress, Chiara continued making errors. Of particular interest was the inaccurate use of the construction be going to (e.g. ‘I gonna’) during her sojourn (it appeared ‘stabilized’). Two factors might explain this. Chiara repeatedly reported having paid close attention to what her teachers said during formal classroom talks because she believed that they provided her with the best language practice she could get. Assuming that their speech differed from informal language use in a number of ways during these classroom interactions (e.g. more specific vocabulary, fewer colloquial structures, slightly more complex language), it is possible that they did not provide the input required for Chiara to correct more colloquially-inspired errors in her own speech. Secondly, Chiara never mentioned receiving corrective feedback on
grammatical forms. It appears, then, that Chiara’s focus on the input received by teachers, and the lack of an explicit focus on form, may not have been sufficient for this error to change.

Chiara’s output became marginally more lexically diverse during her sojourn, but the lexical diversity scores plummeted during interview 6. Different explanations may account for this development. It is possible that Chiara’s increase in lexical diversity suggests that her overriding concern while speaking was communicative effectiveness (i.e. prioritization of meaning, not form). She may not have restricted herself to “item-and-rule approaches” to language learning but was more “lexical in [her] mode of communication, and correspondingly in the repertoire of language knowledge that [she] possess[ed]” (Skehan, 1998, p. 38). The disadvantage of this approach to language learning is that it does not generate enough pressure for “syntacticization” (p. 90), and thus prevents greater interlanguage advancement (which might also explain the stabilised gonna structures). The changes in Chiara’s lexical diversity may also have been impacted by the more formal language learning activities of the New Zealand classrooms. It is possible that as a result of Chiara’s intense attention to her teachers’ speech, she noticed new and less frequent words which she subsequently introduced into her own output (making it lexically more diverse). Chiara’s data also revealed that she initiated orientations to linguistic and grammatical forms in informal conversations with NS peers during classroom interactions. For example, she reported having actively asked her NS classmates for explanations of word meanings.

Chiara’s high initial levels of and significant increase in fluency is rather surprising given the relatively little contact she had with NS friends abroad. Possibly, her English immersion classes in Germany had already provided a good basis for her to use “fluency-enhancing strategies” (Götz, 2013), such as the use of repairs, smallwords, and other formulaic language. Her fluency increase during the sojourn can be assumed to be a consequence of ‘more’ opportunities for chatting to NS peers and host family members involving low structural complexity and similar topics, which seem to have facilitated increasingly automatized lexical and syntactical encoding processes (i.e. higher fluency). However, a closer analysis showed that more automatized language performance only seems to explain higher fluency scores in the interval between the first and second interviews. The subsequent speed fluency scores reflect speech mode choices. Apart from the peak performances in interviews 3 and 6 (which were associated with heavy use of direct speech reports), Chiara’s speed fluency levels did not increase.
In sum, Chiara’s focus on classroom interaction for her language interactions are likely to be responsible for the increase in accuracy, as well as the increase in lexical diversity (though marginal) in her speech. Regular opportunities for informal interactions in comparison may have resulted in more automatized speech performances (i.e. as reflected in higher fluency scores). However, the increase in fluency is perhaps more likely to reflect the use of speech patterns typical of female adolescents – direct speech mode. This choice affected both complexity and fluency scores, which fluctuated depending on her choice of speech mode. I believe that informal speech interactions also provided opportunities for Chiara to use the same linguistic structures, leading to more accurate output. Her lack of propensity to take risks with speech probably also lead to her using similar speech structures and modes in the interviews.
6 Case Study 3: Alia

6.1 Background information

Alia and Her Family in Germany

Alia grew up in a small town in Germany with her two parents and younger brother. Her mother worked as an assistant in a phone company and her father changed jobs mid-way through Alia’s sojourn from being a manager to being self-employed. Her father’s former job involved a lot of travelling in English-speaking countries, which according to Alia, made him a fluent English speaker. At home, they spoke German although Alia’s father occasionally (and unsuccessfully) prompted her to speak English. Alia’s family had never lived outside of Germany but had travelled to a variety of European countries on holidays, as well as Africa, North America, and New Zealand.

Alia had been a youth elite athlete in Germany and spent the majority of her free time at training (up to six days a week, once or twice a day). The sport significantly affected most aspects of her life. It challenged her work commitments at school, put significant pressure on friendships, and took a toll on her physical health.

Alia turned 16 before she went abroad, the average age of international students.

German School and Language Learning

Alia was half-way through Year 10 of a public German Gymnasium before she left for her sojourn in New Zealand. The curriculum devoted four periods per week to English language classes. Besides grammar learning, the students had to read and discuss English texts about political, historical and cultural events of English-speaking countries. Topics included Mexican immigration into the USA, Apartheid in South Africa and globalisation. EFL classroom practices involved interpreting English prose and poetry and studying the use of rhetorical devices in English literature.

Alia had studied English for seven years before she went to New Zealand. She reported having been a quiet student in Germany and expressed a desire to change this situation after her return. Alia had also studied Spanish for three years and French for seven years but disliked the latter as she felt she could not speak it very well.
Reasons for studying abroad in New Zealand and general English language learning motives

Originally, Alia was not interested in studying abroad. However, she changed her mind about the overseas experience in response to a friend’s positive accounts of New Zealand a year prior to her sojourn. Besides being ‘far away’, New Zealand felt like the right choice for Alia because she wanted to study in an English speaking country and believed that New Zealanders were friendlier than Americans (her other option). Moreover, Alia could imagine using English in a future job (as yet unspecified). She also admitted that her father exerted an influence on her choice of language.

Alia chose to go abroad for one academic semester because she believed it was wrong to spend more of her parents’ money and because she would miss her mother and sports competitions in Europe in summer. She was given the option to extend her stay to September 2014. The longer she stayed in New Zealand, the more difficult this choice became but she finally decided to leave New Zealand after one semester because she felt she had made full use of the sojourn.

6.2 Social settings

6.2.1 The host family

Alia’s host family consisted of four members: host parents Hanna and Sam, host sister Chloe (17) and host brother Mike (19). Mike attended university in another city and spent little time at home. Chloe attended the same college as Alia and worked part-time at weekends. The family had accommodated six German girls before they hosted Alia. They also hosted two adolescent Thai girls during a period of four weeks while Alia was living with them.

Jana was provided with many opportunities for social interactions with the host family. While not all activities involved in-depth language engagement (i.e. they often watched TV), Alia interpreted all get-togethers as opportunities to mingle. She preferred spending time with the family over being alone in her room.

Unless Alia was at training, she enjoyed dinner with the family in the evenings. Meals were enjoyed away from the TV and provided opportunities for conversations. Alia mentioned sharing her daily experiences with her host family during these times. The family also seemed to have made an effort at sharing cultural knowledge with Alia, such as explaining local table manners. On her training days, Alia joined the family later in the evening and either ate alone
while the family was watching TV in the living room or with her host father who returned late from work. After dinner, the family normally watched movies together or read in the living room. The former provided limited opportunities for interaction and conversations were mainly restricted to discussions about the movie. However, from Alia’s accounts, it appears that it was not the interaction per se, but simply being in the presence of the host family in the evenings and adopting their habits that made her happy.

During weekends at home the family shared the physical space of the house and interactions seem to have occurred during meal times or sporadically during the day. Apart from spending time at home, the family went away on a few weekend trips. Alia had the chance to meet the extended family at their holiday house. The family and grandparents also organised a picnic on mother’s day, followed by a movie and dinner. Just before Alia left New Zealand, her host family organised to see family friends with whom they went for a walk, had dinner and played games. This trip not only gave Alia the opportunity to gain deeper insight into the New Zealand culture, it also strengthened her sense of belonging to her host family:

> on Sunday we were sitting in the car I was sitting in the middle Chloe on my right and Hanna on my left and we we had a lot of snacks and ehm Mike played music and the eh weather was so good and I felt so that it it is my family and I belong to that family. (Int. 6)

Except for Mike, who lived away from the family, Alia enjoyed a friendly and close relationship with all host family members. Hanna seemed to act ‘in loco parentis’, keeping a close watch over Alia’s whereabouts and attended carefully to her needs. She provided emotional and motivational support when Alia was feeling down and explicitly expressed her affection for Alia. Hanna also performed other caretaker duties such as taking Alia to doctor’s appointments or picking her up from school when necessary. Alia described the times together as enjoyable and interpreted Hanna’s behaviour as ‘caring’, which made her feel at ease in the host family. Hanna also assisted with Alia’s language learning. The following excerpt demonstrates that Hanna was eager to explain words that she believed were difficult for Alia:

> yesterday she said ehm ‘the shop is in in the opposite of this shop’ and she thought that ehm I don't understand 'opposite' but I I ehm shake my head and so that I understand but then she ehm start started to ehm to explain to eh what opposite means and I I know . . . but then ehm in the morning I said ‘I make photos’ and so s she said ehm 'take photos' that's ok that's good and now I I remember. (Int. 1)

Mostly, Alia appreciated her support and often also requested it, particularly towards the end of her stay:
I said to my host mom that she has to help me ehm in my just three more weeks ehm with my English … she has to help me like hundred percent cent and with a lot of power ehm to correct me … and she was she's correcting me a correcting me always all the time and I'm I'm le learning new words and when I have a new word I always repeat it. (Int. 5)

For Alia, Hanna was one of her favourite native speakers to converse with not only because she corrected her output and taught her new words, but also because her English was easy to understand (she was British). She also appeared to genuinely like and care for Hanna. Alia tried to help with cooking and cleaning both to express her gratitude for living with them and because she was used to doing so in Germany. To Alia’s disappointment, Hanna however kept on declining her help for reasons of practicality and efficiency.

Alia’s relationship with her host father Sam was not as close, but was friendly and affectionate. According to Alia, Sam treated her like “his own daughter” (int. 3). Alia said Sam often made jokes and was a source of cultural knowledge. Sam sometimes gave Alia lifts, and they spent one evening watching a movie together which Alia greatly enjoyed. In contrast to Hanna, Alia could not understand Sam’s English very well because of his Kiwi accent.

Alia enjoyed a close relationship with her host sister Chloe. Their friendship was based on similar interests and hobbies. At home, the girls often spent time together in their bedrooms talking about topics of mutual interest and giving each other ‘dating’ advice. Alia and Chloe habitually went for runs together both in the evenings and at weekends. They sometimes also went shopping together or went to the beach. Alia generally evaluated the experiences with her host sister very positively. Chloe, like Hanna, corrected Alia’s output. However, she did so much more seriously, “like a teacher” (int. 6), which made Alia insecure at times. Even though Alia believed that Chloe gave her feedback in order to assist her language learning progress, the conduct raised doubts whether or not Chloe really liked her. According to Alia’s interpretations, their communication difficulties sometimes even resulted in Chloe’s withdrawal from the communicative situation. This left Alia feel like a ‘stranger’ in the home: she's not talking with me and I'm not talking to her because ehm always what I'm saying it's a little bit umgangssprachlich [colloquial] what I want to ehm to explain or say so she's always saying ‘what?’ and then I'm explaining and she said ‘ah ok I don't understand’ and is walking away … so if I see something in the television and I'm I want to explain something what I'm seeing or thinking or what is my opinion then I try to explain something what every or which what every person in German Ger in Germany would understand but Chloe doesn't get … it's just I feel that I'm a fellow is it fellow fellow like fremder [stranger]. (Int. 5)
Alia enjoyed a friendly relationship with her host grandparents. They grandparents regularly got together with the family on Sundays and Alia played games with them, ate local food, and had friendly conversations. Alia also visited her host grandmother at the hospital when she injured her leg. This event provided her with the welcome chance to get an experience that was different from her ‘hospital experiences’ in Germany.

When the host family hosted the two Thai girls, Alia reported enjoying their company and their efforts at communicating with her. Having the girls around boosted Alia’s confidence about her English proficiency level because in comparison to the girls, Alia felt her communication was much more effective.

While Alia enjoyed the company of her host family, her accounts indicate that she was very conscious of her English language use with them. She felt that her communication in the family was limited and that it lacked quality. According to her own interpretations, these difficulties stemmed from an insecurity regarding her English language knowledge as well as an uncertainty regarding her ability to initiate conversations:

my English is bad in the home . . . in the host family it's bad but outside it's good . . . because ehm I don't speak a lot eh in the family and when I speak I'm a little bit not so un unsure or not sure sure what I say and then I'm a little bit ‘is my grammar right?’ or yeah. (Int. 1)

Alia also reported not having communicated much during outings with the family, for which she later felt ashamed.

6.2.2 The high school

Alia attended the same college as Chiara. Like Chiara, Alia had established friendships with Germans before starting at the school. She met Chiara prior to her sojourn and the girls spent the week before school began together. Alia then met Alexandra and other international students during the college’s two-day ‘international student orientation’. She first met the local students in the classrooms at the beginning of the first semester. Apart from becoming friendly with a few classmates, no significant relationships with local students developed from these contacts.
6.2.2.1 Inside the classrooms

Due to her age (16), Alia was placed in Year 12 like most other international students from Germany. She chose six subjects: Physical Education (PE), Biology, Photography, Mainstream English, Mathematics, and Health. Both PE and Photography classes were specifically designed for international students, and not attended by domestic students. Student compositions in the different classes were varied and ranged from classes exclusively attended by internationals (PE and Photography), to classes without other German-speakers (Biology and Maths), a class that only contained local students (Health), or a class that contained a mix of all types of students (English) (see Table 29).

<table>
<thead>
<tr>
<th>Local students</th>
<th>German speaking internationals</th>
<th>Non-German speaking internationals</th>
</tr>
</thead>
<tbody>
<tr>
<td>PE</td>
<td>x</td>
<td>√</td>
</tr>
<tr>
<td>Biology</td>
<td>√</td>
<td>x</td>
</tr>
<tr>
<td>Photography</td>
<td>x</td>
<td>√</td>
</tr>
<tr>
<td>Mainstream English</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Mathematics</td>
<td>√</td>
<td>x</td>
</tr>
<tr>
<td>Health</td>
<td>√</td>
<td>x</td>
</tr>
</tbody>
</table>

School generally did not constitute an important part in Alia’s life. Her attitude toward homework or assignments was nonchalant and Alia preferred to pursue her hobby:

sports is more is for me more important than school and I said it to my mom in Germany that I ehm ‘sport sport geht vor [sport comes first] for me’ and it's ehm normally should this not be ehm it should be that the school is ehm more ehm important but for me just sport. (Int. 1)

Alia’s data reveals that she engaged in classroom activities and completed tasks but did not put in extra work. Alia was more interested in enjoying social relationships with classmates (L1 or L2 speakers) than engaging in formal English learning opportunities. For example, Alia was passionate about her PE class precisely because it was attended by international students (German and non-German speakers):

The school last week was normal and made fun. I love my P.E. class because there are a lot of international students and it’s fun to do sports with them. (Blog 02/03/14)

Alia generally shared few accounts about classroom experiences. She referred to the social experiences she had in English, Health, and PE, as these subjects appear to have brought her
the most enjoyment. Apart from English and Health classes, Alia did not interact with local students in the classrooms, but she often spoke English with non-German-speaking international students.

The opportunities for English interactions in the different classrooms appear to have been determined by 1) the composition of the class, 2) seating arrangements, and 3) the type of L2 engagement the subjects demanded of the students. Some classes also encouraged non-social engagement with the English language. Alia did not report having interacted with teachers beyond brief chats about her work.

Seating arrangements in English and Health appear to have encouraged the development of friendships with local students. In Health, Alia shared a table with two Kiwi girls and a Russian boy. Alia liked one of the Kiwi girls, Alice, and became friendly with her. The relationship between Alia and Alice did not extend beyond the classroom and Alia was unwilling to invest more in the relationship given the little time she had in New Zealand. In English, Alia shared a table with a German boy, Lorenz, and a Kiwi girl, Christina. According to Alia’s comments, the trio had a lot of fun. During informal chats in the classroom, Alia and Christina grew close. Christina occupied an important role in Alia’s language learning. For example, during conversations with her, Alia was able to pick up colloquial expressions and use them herself:

> when I'm speaking with her I feel like that I can't speak German and so the the normal slang so like 'hang on' or 'I don't know' so and she's say speaking really slow and clear so that I can understand everything and yeah. (int. 6)

Christina also appeared to assist Alia’s L2 development by giving explanations and feedback on her pronunciation. She also took away some of her fear of speaking English in public, which Alia appreciated.

A number of classroom tasks in PE, English and Health classes afforded opportunities for the use of English. In activities during PE classes, Alia mentioned often conversing in English with South American boys – an activity which caused L2 anxiety but, at the same time, could be performed in a zone of relative safety:

> I have to talk a lot of En eh in English with the boys and now I mean the internationals and from ehm a lot of South American and ehm it's like ehm ueberwin ueberwindung eh eine ueberwindung [a risk] to speak with them because you sometimes un don't understand … and when when I when I have to ueberwinden myself [overcome my inhibitions] it is like the ice is broken . . . yeah and with the Kiwis you don't know ehm if they are like if they like you and if they are interested and want ehm to ehm if they
Some PE classroom tasks also required Alia to use her English. For example, she had to describe to blindfolded classmates how to erect a tent. This task challenged Alia’s L2 abilities to the extent that the teacher had to alter it for the purpose of completion.

A classroom task in Health required Alia to conduct an interview with somebody outside the classroom. The activity not only demanded the use of English with different native speakers, but it also appeared to extend Alia’s vocabulary:

Health ehm I had to do an interview with my ehm with one of my parents about ehm do you ah what is resilience do you know ‘resilience’? it’s like eh if something worse happen and then you are sad and then you have to bounce back like you have to think pos positive and so … I want want to ask Hanna but she said ‘oh I don't know what to say maybe thi this with my cat’ but I said ‘no not that with the cat’ because so I asked Chloe although she's just ehm seventeen and has to be an adult but I didn't care. (Int. 4)

English demanded a lot of in-depth engagement with the English language. Activities included writing academic and creative essays, studying literature and films, and taking the school internal exams. Writing essays was especially challenging for Alia, and the results not satisfactory. Alia received feedback from her teacher regarding her written English assignments and often had to re-write her essays. While she appreciated the teacher’s assistance, she remained unhappy about her own performance, particularly in comparison to that of other German students. However, Alia admitted that she did not attempt to become better at writing in English despite her awareness that the skill would be useful for her education in Germany. She claimed not to care about her writing during her sojourn.

**6.2.2.2 Outside the classrooms**

With very few exceptions, Alia spent all school breaks with German speakers. During the breaks at school, the international students sat physically apart from the domestic students. In term 1, Alia predominantly interacted with her two close German friends Chiara and Alexandra (see Chapter 5) but she also established friendships with other Germans at the college such as Lorenz, and Viktor (see 0). At the end of term 1, Alia’s friend Alexandra left the school, which reduced the already secluded triplet (Alia, Chiara, and Alexandra) to a pair (Alia and Chiara) and temporarily caused anxiety in Alia regarding her future place in the group of internationals. The arrival of three new German girls (Inge, Karin, and Fabia) in term 2, who Alia and Chiara befriended, increased the group size and removed Alia’s worry of being alone with Chiara.
Alia reported that being with Germans was exciting and she did not desire establishing close friendships with internationals from other cultural backgrounds. Alia also did not show a strong motivation to interact with Kiwis beyond the classrooms (with one exception, see 6.2.3.2):

I don't have ehm Kiwi friends and I say ‘ok ehm I take it easy’ if I don't have Kiwi friends when I'm coming back it's ok for me because I'm back and have my friends there . . . I have the Germans and I like the Germans and the other internationals ehm I like them too but ehm they are Braz from Brazil and Ch Chile and they are different and ehm I'm not sure because ehm they are so m different and …I don't want to ehm meet them because it's not so my imagine or picture. (Int. 2)

Alia identified different reasons why she did not spend time with Kiwis during the breaks. She believed that timetables inhibited interaction between classes; that she was unable to distinguish potential friends from foes among students wearing uniforms; and that the Kiwis would not take an interest in establishing friendships with international students:

However, Alia admitted that she herself did not devote much effort to changing her circumstances. She did not actively ask her New Zealand classmates to spend breaks with her and seemed reluctant to take risks when opportunities for interactions emerged.

Partly, Alia attributed this behaviour to her personality. She considered herself shy and expressed anxiety regarding other students’ perception of her. Alia also shied away from interactions with Kiwis that were challenging, preferring to enjoy her time in New Zealand in the safety of her German friends. For example, in mid-March, when Alia became closer to Christina (the Kiwi girl from English classes), she was invited to spend the breaks with her and her friends. Alia joined them but only for a little while. She felt that she could not contribute much to their conversations, and that being with internationals was more enjoyable, and so she soon returned to her German group.

Alia reported having always spoken German during break times. She also did not attempt to speak English to her German friends because she considered this hard. Although she was aware that her exclusive engagement with Germans during breaks deprived her of English learning opportunities, she ultimately just accepted this:

**Luzia:** what are your goals for your English for the next three weeks?

**Alia:** ehm yeah just speak more English than German but then I always say ‘ok you're not going to lunch with the Germans’ but then I ‘ah no I want to go with the Germans’ so I'm going with the Germans. (Int. 5)
6.2.3 Outside of the high school

6.2.3.1 Social activities and trips

Alia spent the majority of her free time with her close friend Chiara (see Chapter 5) and her other German friends. Alia and Chiara met regularly, and for extended periods, at weekends and some school nights. Throughout her accounts, she emphasised how much joy Chiara’s company brought her and how keen she was to continue spending time with her. She also mentioned that during times of Chiara’s absence, she felt ‘bored’. Her other close German school friend, Alexandra, sometimes joined their activities in the afternoon, and the three took cultural trips to some tourist attractions. In the second term, Alia sometimes also met the other German girls, Karin, Michelle, and Inge to go shopping and to the movies. The German girls visited famous tourist sites, tried out adventure sports, and ate German food.

Alia also took part in the South Island trip during the Easter holiday. With the exception of a few South Americans, the travellers were all German-speaking international students. Alia mentioned that most activities during the fortnight were carried out in German, even when they were in the company of South American students. In the last week of her sojourn, Alia went to the school ball, which was the last major occasion for her to go out with her German friends.

Alia very rarely engaged in social activities with friends that required the use of English outside of school (except for her hobby). Occasionally, she joined a group of internationals of mixed cultural backgrounds that regularly met by the beach on Friday evenings. It appears, however, that even during these evenings she mainly spoke German, and did not engage a lot with South Americans in English. Alia once went on a romantic date with a Venezuelan boy, which required the use of English, but she broke off contact with him after the evening.

Alia never met the Kiwi friends she made at school, outside school. With one exception, Alia did not seem to mind this. During the interviews, she often expressed a desire to meet the Kiwi girl, Christina (from her English classes). As some comments suggested, a get-together did not occur because Alia was not assertive or confident enough to organise a meeting (albeit being confident enough to ask for Christina’s phone number). She also stated that she was too busy or too lazy to do so.

While Alia generally seemed content with her social networks in New Zealand, her parents were not. This clash of expectations caused her some worry:
and ehm I have Kiwi friends but I don't do something with them outside of school it worries me a little bit so that it's like a little bit pressure and I told my parents about that because my parents say say ehm yeah 'you should have Kiwi friends’ and stuff like that. (Int. 4)

6.2.3.2 Hobbies

Alia was a successful athlete and sports training and competitions occupied a large part of her free time, both in Germany and New Zealand (see section 6.1). In New Zealand, Alia joined a premium athletics club and she assigned involvement in that community first priority. Alia’s training involved three double sessions a week. On Mondays, Tuesdays, and Wednesdays, she trained after school with two different coaches in succession, and typically returned home at eight o’clock. Before competitions, Alia spent extra days at training. She also attended training sessions when she was injured and unable to actively participate.

Alia’s team consisted of local athletes as well as the German international student Viktor (who left New Zealand after the first school term). Alia established close friendships with most athletes and one of the coaches (James). During Term 1, she was particularly close to the German boy Viktor and their mutual friend Lorenz (who also attended Alia’s PE and English classes) and the trio regularly met before training sessions for a snack. In the company of these two boys, Alia spoke German. With all other members, Alia exclusively communicated in English.

Being in the athletics team resulted in a range of opportunities for English conversations and social interactions, not only during regular training sessions but also outside of them. Outside training sessions, Alia regularly had appointments with doctors and the club's physiotherapist, which required her to use specific words related to the human anatomy and sports activities. Interactions during training also involved the use of vocabulary specific to athletics, such as terms for training equipment, events and place names for competitions, the human anatomy and so on. Interactions with her coach James in particular appear to have provided opportunities for Alia to broaden her lexical repertoire:

_**Luzia:** is there one person in particular that you can think of where speaking with him or her in English is particularly helpful?

_**Alia:** ehm yeah I think James because he's he's ehm explaining me ehm the technique and it's so he's doing that with his hands or with his whole body but with ehm with vocabulary like ‘angles’ I've never kno I don't know what ‘angles’ is is ehm winkel [angle] yeah. (Int. 2)
Speaking English with male athletes offered opportunities for Alia to practice her English in a relaxed manner. She reported having experienced lower L2 anxiety and more confidence when talking to the boys, rather than the girls, despite the fact that she believed they sometimes thought she was ‘stupid’ because she did not always understand what they were saying.

Alia, in line with her male friends, believed that New Zealand girls were ‘touchy’ and more difficult to get along with than New Zealand boys:

I notice that I can be or that I can speak better with boys here in Ge eh in New Zealand instead of girls and in Germany it's the other way … I feel more confident with the boys because I think they are nicer … a lot of boys told me that they [the girls] are so bitchy here and because they're all coloured her hair their hair and a are a little bit zickig [bitchy] . . . or maybe maybe they are not not easier to get along here maybe it's just myself that I that can ehm be better with the boys. (Int. 2)

Over time, however, Alia also became close friends with the female athletes.

Competitions required Alia to use English at many different events and for different purposes. For example, she spoke English with team members and coaches during the long drives to the competitions, while waiting for her race, or during dinners and breakfasts with the team. During one competition, Alia reported having spoken in English for extended periods of time over four days:

we are driving I think ten hours … and talked about everything like about boyfriend girlfriend and all this stuff … because James's the most relaxed person I’ve ever met . . . and then we had like . . . introduction in what we have to do with James so he explained things and yeah then we had to go into a call room and have to warm up . . . and just speaking English four days just speaking English. (Int. 3)

Overall, being a member of the sports team brought Alia professional success and some of her happiest times in New Zealand. Saying goodbye to her athletics team at the end of her sojourn was so difficult for Alia that she compared it to a couple’s breakup.

6.2.4 Summary

The different social contexts Alia engaged in varied greatly in terms of opportunities for English language interactions. Alia regularly spoke English in her host family and the athletics club, as well as during classroom interactions (yet to a much smaller degree it seemed). Conversely, Alia predominantly communicated in German during the school breaks and in her leisure time, when she was not at training.
6.2.5 Perception of L2 progress

In interview 1, Alia reported that she wished to become more fluent in English, to broaden her vocabulary, and to become better at grammar in general and using tenses in particular. Her main motivation for language progress was to obtain good grades to attend an intensive English course (‘Leistungskurs’) in Germany after her sojourn. Alia believed that frequent speaking opportunities would assist that goal. Alia did not believe that studying abroad for an academic semester leads to much linguistic improvement. She believed that progress could only be achieved after an SA period of ten months.

In interview 2, Alia furthermore expressed a wish to improve her English pronunciation, to produce output without having ‘to think’, to dream in English more often, and to start thinking in English about complex matters. In interview 4, she was still unhappy about her use of tenses and wished to acquire words like ‘as well’. In subsequent interviews, Alia felt content about the way she spoke and simply desired to speak English more frequently.

Alia’s evaluation of her own language learning progress changed from mainly positive in the first half to mainly negative in the second half of her sojourn. In the early stages (int. 1 to 2), she was positively surprised at her communication skills and noticed improvements in the areas of vocabulary, the use of different sentence structures, and the ability to engage in conversations during service encounters. During this time, she was also aware of language-related areas that remained challenging for her. They mainly involved forming English questions, the use of tenses, and subject-verb agreement.

In mid-April, Alia felt that her L2 development stagnated somewhat, except for vocabulary learning. She believed that despite sufficiently noticing (or being alerted to) communication problems, she was unable to successfully learn from them for future situations:

I think so the last time I thought it [her English] becomes better and now I'm not sure I think it's the same like the last and ehm I knew I I learn new words but then when when sometimes somebody says to me ehm that I have made a mistake and I repeat it and the next time I say it wrong again because I'm too lazy to think about. (Int. 3)

After the South Island tour and the weekend trips where she spoke only German, Alia experienced a real decrease in her ability to speak English:

[After the South Island trip] I couldn't say one sentence so sentence ehm it was so bad really it was so bad it was like at the beginning ehm they had always to correct me in everything and then after a while I was just laughing and they were laughing because it
was so bad … and then yesterday ehm I was sitting ehm in the [training hall] and then Eva was coming to my table and we did homework together and then Hillary and then James so it was like really confused with everything I felt so not confident with myself couldn't speak English. (Int. 4)

Alia reported being embarrassed and sad about her temporary ‘inabilities’ to speak English after ‘German-only’ activities, particularly when chatting to her host family. She sometimes attempted to change the circumstances by speaking English to Chiara in her free time but these attempts failed and were not pursued any more. Alia ultimately accepted the high level of German interactions during school breaks and in her free time although this conflicted with her mother’s wish for her SA:

and ehm in school in the break we always speak German and then we when I'm with German it's always like twelve hours eh I speak German and then I then I always say ‘I don't care’ but then my my mo my mom says ehm you have to speak English and then I say ‘yes I know’. (Int. 4)

At the end of her sojourn (interview 6), Alia’s perceptions about her progress were mixed. She believed that her English was generally ‘good’ (albeit not ‘perfect’) and that she had developed the ability to easily switch between German and English, to count in English, to have a short conversation in English, and to notice intonation patterns. She also believed being abroad increased her vocabulary and her listening ability. However, Alia stated that the use of tenses had not improved, that she did not dream in English as desired, and that writing short notes automatically occurred in German, not English.

Besides changing L2 perceptions over time, Alia generally felt that her English was better when she felt confident and after extended periods of L2 conversations. Conversely, Alia noticed that her English worsened after physical exhaustion, and after activities during which she only spoke German (see above). She also complained about insecurities regarding her L2 proficiency in the company of native speakers when they did not understand her or paid excessive attention to her English (see 6.2.1 for her L2 insecurities in the host family). Alia also felt that speaking English to native speakers in a group of Germans resembled a language contest, which made her insecure and upset. It appears that Alia experienced the least L2 anxiety in interactions with native speakers with whom she felt comfortable, and who did not make ‘harsh’ L2 corrections (e.g. her host mother, her coach, and her Kiwi classroom mate Christina).

At the end of her sojourn, she stated that social integration had little to do with language abilities but only with attitude and behaviour:
I think it's hard to integrate ... you in a group if you're like acting a little bit scared and shy and ehm I think if you're friendly and if your language isn't that good so at the beginning my English wasn't eh wasn't that good so but I always said that I'm like from Germany and then I'm can't really say what I want to say and they helped me and yeah so I think like how you act and eh showing respect and yeah. (Int. 6)

She also believed that making mistakes is appropriate and even funny (albeit slightly embarrassing) at times.

6.3 L2 development

6.3.1 Grammatical complexity

6.3.1.1 The subordination index

Alia started her sojourn using 1.44 clauses per AS-unit (cpu) and ended it with 1.29 cpu (-10%) (Figure 38). The development is marked by a general decrease but also a peak performance in interview 5. A low correlation coefficient of -.395 implies considerable fluctuation in the development.

On average, Alia (ƺ1.36) performed well below the scores measured in Mora and Valls-Ferrer’s (2012) study (NNSs: ƺ1.60/1.59/1.62, NSs: ƺ1.83), Polat and Kim’s (2014) study (NSs ƺ1.97; range: 1.82-2.22), and Llanes et al.’s (2012) study (NNSs ƺ1.70/1.87) study. However, she performed within the broad range observed for Polat and Kim’s (2014) NNS (1.15-2 clauses per AS-unit) as well as the 17-year old NS participants in Nippold et al.’s (2005) study (ƺ1.30).
The comparisons suggest that Alia’s utterances were appropriate to her age and the communicative situation. It was also noted that decreased SI levels are not necessarily reflective of decreased L2 levels but indicative of spontaneous speech (Ferrari, 2012).

**RANGE AND FREQUENCY OF SUBORDINATE CLAUSES**

Figure 39: Subordinate clauses for Alia

<table>
<thead>
<tr>
<th>Type</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-finite adverbial</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>non-finite complement</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>finite relative</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>finite adverbial</td>
<td>13</td>
<td>14</td>
<td>10</td>
<td>7</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td>finite complement</td>
<td>19</td>
<td>22</td>
<td>8</td>
<td>13</td>
<td>22</td>
<td>13</td>
</tr>
</tbody>
</table>

Alia used all types of subordinate clauses in interview 1, implying that the constructions were acquired before going abroad (Figure 39). She predominantly used finite complement clauses (8–22 clauses per 100 AS-units) and finite adverbial clauses (7–18 clauses) albeit with great variation. Finite relative clauses played a minor part in Alia’s utterance complexification (1-3 clauses) and she only sporadically produced non-finite adverbial and complement clauses (0-4 clauses). Alia used significantly more finite complement and adverbial clauses in interviews 1, 2 and 5 than in interviews 3, 4, and 6.

The distribution of her finite subordinate clause types is appropriate to conversational discourse (Nippold et al., 2005). She used a similar frequency of finite complement clauses as Nippold et al.’s (2005) 17-year old NS group, but fewer finite adverbial and relative clauses than they did. As in Jana’s and Chiara’s chapter, I paid detailed attention to the particular case of direct speech complements.

**DIRECT SPEECH COMPLEMENTS**
Jana’s and Chiara’s L2 analyses revealed that they both used large numbers of direct speech reports during certain interviews, and that the use of this speech mode considerably affected their L2 performance, particularly speech complexity (Jana and Chiara) but also fluency (Chiara). A closer examination of Alia’s finite complement clauses revealed that she also made use of direct speech but only to a small degree. As can be seen from Table 30, 17 utterances (out of 100) contained direct speech complements in interview 1, while in the remaining interviews, only 2 to 6 utterances contained direct speech complements. The ratio of direct speech complements per total number of clauses was small at all times (2-14%). With the exception of interview 1, the majority of finite complement clauses were not direct speech complements.

Table 30: Finite complement clauses and direct speech for Alia

<table>
<thead>
<tr>
<th>Interview</th>
<th>Raw number of matrix clauses per 100 AS-units (and number of dependent direct speech clauses)</th>
<th>Raw number of following AS-units with direct speech clauses (and number of clauses involved)</th>
<th>Total number of utterances containing direct speech per 100 AS units</th>
<th>Ratio direct speech complements / total number of finite complement clauses</th>
<th>Total number direct speech clauses (ratio direct speech clauses / total number of clauses)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9 (11)</td>
<td>8 (9)</td>
<td>17</td>
<td>10/19 (53%)</td>
<td>20/144 (14%)</td>
</tr>
<tr>
<td>2</td>
<td>5 (5)</td>
<td>1 (1)</td>
<td>6</td>
<td>5/22 (23%)</td>
<td>6/142 (4%)</td>
</tr>
<tr>
<td>3</td>
<td>2 (2)</td>
<td>0</td>
<td>2</td>
<td>2/8 (25%)</td>
<td>2/129 (2%)</td>
</tr>
<tr>
<td>4</td>
<td>5 (6)</td>
<td>0</td>
<td>5</td>
<td>5/13 (38%)</td>
<td>6/126 (5%)</td>
</tr>
<tr>
<td>5</td>
<td>5 (8)</td>
<td>1 (1)</td>
<td>6</td>
<td>6/22 (27%)</td>
<td>9/146 (6%)</td>
</tr>
<tr>
<td>6</td>
<td>3 (2)</td>
<td>0</td>
<td>3</td>
<td>3/13 (23%)</td>
<td>3/129 (2%)</td>
</tr>
</tbody>
</table>

Given that Alia was able to use this speech mode right from the beginning, L2 proficiency does not appear to be the main reason for its limited use. That is, Alia did not ‘fail’ to notice this speech mode in NS input and integrate it into her own speech – which could have been indicative of lower proficiency levels. However, it is possible that the direct speech mode was not as well proceduralized or target-like for Alia as it was for Jana and Chiara. Example 60 shows that her use of direct speech report was less dramatic (e.g. the use of ‘say’ is unmarked and lacks a pragmatic effect, Romaine and Lange, 1991) and more complex than Jana’s and Chiara’s.

Example 60 (int. 1): and she pushed me | because she said :: ‘ah you don't ehm (.56) have time’ | (.56) or ‘you don't like me :: because you are in in your group | and then you ehm (.47) don't come to me’ or something like that | and (.56) ehm she's ehm (.30)
very motivated | and I'm motivated too | but ehm she's so motivated :: that she ehm (.43) pushes me down (.30) with her motivation | (.32) because she said :: 'ah I'm good in here | I'm good in here' |

Another reasons for Alia’s scarce use of the direct speech mode may be related to her performance of her identity. This aspect will be discussed in section 6.4.

**6.3.1.2 Mean length of clause**

![Figure 40: Mean length of clause for Alia](image)

Alia’s mean length of clause (MLC) increased from 5.83 words per clause (wpc) in interview 1 to 7.22 wpc in interview 6 (+24%) (Figure 40). The development was gradual with the exception of a slight increase in interview 3. A very high correlation coefficient (r= .951) confirms the linearity of the development.

With an average of 6.48 wpc, Alia’s clauses were longer than those measured for the three NSs in Polat and Kim’s (2014) study (≤5.68 wpc). During the first four interviews, however, her scores were within NS range (5.05-6.76 wpc) as well as within the range measured for their NNS participant (4.-6.67 wpc).

Jana’s and Chiara’s L2 analyses showed that high numbers of direct speech complements correlated with lower numbers of MLC. Conversely, narrations that involved descriptions and opinions went hand in hand with longer clauses. In section 6.3.1.1, I showed that Alia’s speech contained few direct speech reports and it is highly unlikely that they affected MLC levels in any way. Most narrations in Alia’s data concerned descriptions of events, people and things.

The question is how exactly Alia made her clauses more complex (i.e. longer), and what clause lengthening (i.e. complexification) tells us about L2 progress. According to Norris and Ortega (2009), clause complexification (i.e. longer clauses) is the result of adding single words to
phrases (e.g. adjectives) or adding multiword sequences to clauses (e.g. prepositional phrases). Example 61 demonstrates that Alia was capable of using clause-lengthening devices such as prepositional phrases (‘after the period’), adverbs (‘fast’), and adjectives (‘next’ and ‘another’) already in interview 1. The example also demonstrates that right from the beginning, she was able to produce clauses of considerably length (‘because (.48) after the period…’ contains 13 pruned words).

Example 61 (int. 1): (.41) and because ehm it's bad :: because (.48) after the period we d we have to go fast to our next lesson | and so you (.35) can't ehm (.92) ehm (.36) talk with them :: or go going to the next room | because ehm everybody has another timetable.

It is possible that longer clauses in Alia’s case not ‘only’ mirrored an increased ability to complexify (i.e. L2 progress) but may also have been determined by situation-specific choices Alia made regarding what she wanted to say and how. Brown and Yule (1983) explained that the primary purpose of ‘transactional’ communication is message comprehension (as opposed to ‘interactional’ communication that is concerned with the maintenance of social relationships). To achieve this goal, successful transactional speech tends to involve more specific vocabulary and language that is clearer than primarily interactional language. In other words, for a speaker to make him/herself understandable, s/he must provide a full account of “where and when the events happened and who the main participant was, before he recounts the series of events” (p. 17). Example 62, taken from interview 6 (i.e. highest MLC levels), lends itself particularly well to demonstrating the clarity with which Alia described when and where some of the activities related to athletics competitions took place, and who participated in the activities. The excerpt contains clause-lengthening devices such as prepositional phrases (‘into like this air paper’, ‘with these bubbles’), adverbs (‘yesterday’, ‘probably’, ‘as well’), adjectives (‘next’, ‘another’, ‘big’, ‘a lot of’) and a number of smallwords (‘I think’, ‘like’).

Example 62 (int. 6): (.31) and yesterday they packed all the poles in (.41) into like this air paper like with these bubbles (.96) and ehm yeah they're flying first to Canada :: to have a competition next Sunday :: when I’m flying | (.51) and then they have another competition probably next week in (.39) Vancouver or Seattle | (1.32) and then (.53) I think on the twenty seventh of July (.46) they have the (.41) big competition | (.49) and I know a lot of Germans :: who are there as well |.

Given Alia’s sharp increase in MLC levels, it is possible that supplying details “to create a coherent mental representation” (Brown & Yule, 1983, p. 17) when describing experiences became both easier and more important for her. Thus complexification on the phrasal/clausal level (i.e. MLC) may not only be indicative of proficiency, or speech mode but also of the
individual’s motivation, confidence and willingness to share information. Isabelli-Garcia (2004) made a similar finding regarding her examination of communicative functions.

While these elaborate narrations helped Alia to draw clear pictures of her experiences in New Zealand, they did not conform to norms of spoken language with regard to clause length. Possibly, Alia was not quite aware of the pragmatics of informal conversations (i.e. that clauses are short) and instead produced utterances more characteristic of formal and written narratives (a possible consequence of practicing English in formal EFL classrooms in Germany). Alternatively, it is possible that Alia made a conscious choice to specify relevant properties in her speech to ensure comprehension. The interview contexts lent themselves well for elaborate narrations – possibly better than informal speech situations with Kiwi friends.

### 6.3.2 Lexical diversity

[Figure 41: Lexical diversity for Alia]

Alia improved the lexical diversity (LD) of her output from R= 8.01 in interview 1 to R= 9.60 in interview 6 (+ 20%) (Figure 41). There is a slight dip in her development in interview 4. The overall improvement is linear and shows little variation, which is supported by a very high correlation coefficient (r=.916).

Compared to previous studies which measured LD of NNSs’ spontaneous oral speech by means of the Guiraud’s Index, Alia’s values were higher at all times (Levkina & Gilabert, 2012: R= 5.84-6.24; Daller & Xu, 2007: R= 5.03-6.18; Lu, 2012: R= 3.3-6.9). With an average of R=8.94, Alia’s values were also above the average scores measured for the NSs in Mora & Valls-Ferrer’s (2012) study (≈7.86).
Although studying abroad has typically been shown to positively affect vocabulary growth (Badstübner & Ecke, 2009; Dewaele & Pavlenko, 2003), the increase in Alia’s already high initial lexical diversity (LD) scores are rather remarkable. Milton (2009) explained that increased LD levels can go hand in hand with increases in speech rate development. He stated that as a rule of thumb, “as learners improve in level and fluency, they are likely to increase the variety of words they use in production” (p. 127).

It appears from Alia’s data that she was willing to take risks when it came to language production. A number of instances show that ‘deficits’ in her linguistic storage (Lennon, 2000c) triggered the use of communication strategies which exceeded her linguistic repertoire. When her vocabulary searches for specific words were unsuccessful, she did not avoid the problem by changing the communicative goal, but “tackle[d] the problem directly by developing an alternative plan” (Faerch & Kasper, 1983). For instance, in Example 63 Alia made use of a communication strategy (i.e. circumlocution) to deliver the intended message.

Example 63 (int. 6): (.31) and yesterday they packed all the poles in (.41) into like this air paper like with these bubbles.

In terms of L2 proficiency, communication strategies like this may have boosted Alia’s vocabulary development.

Lastly, Alia’s range of vocabulary is believed to have increased most significantly as a result of her membership in the sports club. Dewaele and Pavlenko (2003) argued that an L2 learners’ “familiarity with the domain” (p. 125) (i.e. specialised knowledge) can affect lexical diversity levels. They explained, for example, that when asked to talk about trees, a “gardener is more likely to give more detailed information because his/her expertise allows him to distinguish different types of trees” (p. 124). In this study, the participants were free to elaborate on topics (within constraints) of interest during the interviews. With the exception of interviews 1 and 4 (note that these interviews show the lowest LD scores), Alia predominantly shared the experiences of her athletics group. As described in section 6.2.3.2, Alia’s membership in the group led to her using increasingly more low-frequency words specific to athletics, such as terms for training equipment, physiotherapy, competitions, the human anatomy and so on. Example 64 demonstrates the range of training apparel vocabulary Alia possessed at the end of her sojourn (i.e. ‘pants’, ‘jumper’, ‘singlet’, ‘training suit’, ‘vest’, and ‘crop top’).

Example 64 (int. 6): (.97) and ehm (.59) and then James gave me ehm because I said on Friday that that I really want like a singlet with New Zealand (.88) because it it wou
it would make me so proud (.78) to wear it in Germany (.57) and then he gave me a 
plastic bag and was so much New Zealand stuff in there (.54) ehm it w was (.70) like 
(1.28) how is it called (.41) pants and (.63) like a jumper like a full (.69) training suit 
(.63) ehm with (.51) NZL on the back (.98) it's really cool then a (.42) vest (1.02) vest 
with New Zealand ehm a crop top with New Zealand (.32) I think it's a li little bit older 
from London two thousand twelve when it was in England (.90) and then other stuff 
great black and white (.57) New Zealand (.66) things. 

Thus talking about ‘athletics’ seems to have elicited uncommon words in general and Alia’s 
growing familiarity with technical terms of the field (as an expert) in particular, are assumed 
to be largely responsible for her significantly increased lexical repertoire.

6.3.3 Fluency

6.3.3.1 Breakdown fluency

EMPTY PAUSES

Figure 42: Empty pauses for Alia

Figure 42 shows that Alia reduced the number of empty pauses (EPs) from 19.55 phw in 
interview 1 to 14.68 in interview 6 (-25%). Despite the temporary increase in interview 5, a 
correlation coefficient of -.760 for EPs indicates a highly linear development.

Compared to the NNSs in Götz’s (2013) study, Alia’s mean value (≈16.53 EPs phw) was 
slightly above their average (≈15.13 EPs phw), but she finished her sojourn with a score below 
this number (14.68 phw). Like these NNSs, Alia showed a clear overuse of EPs compared to 
NSs (Götz, 2013), and performed outside their range of variation (1.5-9 EPs phw). However, 
Alia’s results are comparable to those found for the ‘fluent’ participants in Riggenbach’s (1991) 
study (10-22 EPs phw) and well below the range within which the ‘nonfluent’ group performed 
(33-73 EPs phw).
An overuse of EPs (i.e. extensive searching for ideas, lexical items or grammatical forms) has been classified as one of the key problem areas even for very advanced L2 English speakers (Götz, 2013). High frequencies of EPs have also been shown to be a strong indicator of dysfluent speech (Chambers, 1997; Riggenbach, 1991) – a characteristic which probably pertained to Alia’s output, at least at the beginning of her sojourn.

**Filled pauses**

![Figure 43: Filled pauses for Alia](image)

Alia reduced the number of FPs by 59% from 8.36 per hundred words (phw) in interview 1 to 3.39 phw in interview 6 (Figure 43). Despite a temporary increase of FPs in interview 3, the development was highly linear (R = -.871).

With a mean value of 4.98 FPs phw, Alia used fewer FPs than Götz’s (2013) highly advanced adult NNSs (±5.12) but more than twice as many FPs as Götz’s NSs (±2.27). However, Alia always performed within the range measured for Götz’s NSs (0.5-8 phw). In comparison to Hasselgren’s (2002) results, Alia’s mean value can also be situated somewhere between the more fluent adolescent NNS group (±5.98) and the NS group (±2.81). Overall, Alia’s FPs development indicates increasingly more native-like behaviour.
6.3.3.2 Speed fluency

**SPEECH RATE**

As can be seen in Figure 44, Alia improved her unpruned speech rate (SR) from 129.96 words per minute (wpm) in interview 1 to 158.60 wpm in interview 6 (+22%). There was a slight dip in Alia’s SR level in interview 3. A correlation coefficient of .857 implies that the development was highly linear.

With an average SR of 150.03 wpm, Alia scored above three of Lennon’s (1990) four German speakers of English (76-99 wpm). Her SR was very similar to that of Lennon’s fastest speaker (138-162) and within the range of Götz’s (2013) upper-intermediate/advanced learners (≥160 wpm, 117-190 wpm). However, Alia’s values never reached those scored by Götz’s NSs (≥213 wpm; 174-280 wpm). According to Tauroza and Allison’s (1990) categories of British English SR (i.e. NS scores), Alia would be classified as ‘moderately slow’ at all times (120-160 wpm) with the score of the last interview reaching the upper limit of this classification.

Given the scarcity of direct speech reports in Alia’s speech, her SR development appears to generally reflect better automatized syntactic and lexical knowledge which probably developed through repetition and practice in the naturalistic context.
Alia gradually increased the number of words she used per run (wpr) between EPs of at least 0.3 seconds (Figure 45). In interview 1, she used 5 wpr compared to 6.56 wpr in interview 6 (+31%). A correlation coefficient of .764 implies that the development was highly positive.

Compared to previous results, Alia’s average runs (≈5.91 wpr) were relatively similar to or below the results obtained for NNSs of upper-intermediate/advanced proficiency levels (Mora & Valls-Ferrer, 2012: ≈5.93-7.50 wpr; Götz, 2013: ≈5.88 wpr). Alia’s overall rate of change (+1.56 wpr) was higher than the averages found in Mora and Valls-Ferrer (1.33 wpr). Compared to NS values (e.g. Götz, 2013: ≈10.3 wpr; Mora & Valls-Ferrer, 2012: ≈12.49 wpr) Alia’s average performance was less native-like. Since measures differ with regard to cut-off points for EPs, these comparisons must be viewed with caution.

Higher MLR levels means that a speaker “is able to process more language within a single time span” (Wood, 2012, p. 121). Typically, improvements in this measure are explained by greater proceduralization skills, the use of more formulaic language, or cognitively less demanding tasks (Götz, 2013; Towell, 2002). Towell (2002) also found that syntactic complexification of the run (e.g. the addition of adverbials) can contribute to longer runs. In Jana’s and Chiara’s case, speech modes (i.e. especially direct speech reporting clauses) impacted on speech production processes, and hence changes in MLR.

As illustrated in 6.3.1.1, Alia’s use of direct speech reports was marginal and an effect of this speech mode on MLR values can be discounted as a reason for her increased MLR. Presumably, a combination of adding words to clauses (i.e. complexifying clauses) and higher proceduralization skills led to longer runs in Alia’s speech. The former is confirmed by
generally longer clause productions (i.e. higher MLC scores, see 6.3.1.2). Alia may have also replaced EPs with non-lexicalised fillers over time, such as smallwords. When used instead of EPs, smallwords would forestall the premature end of a run (as indicated by an EP) and increase productive fluency (as measured by MLR) (Götz, 2013). This would be considered a planning strategy typical of advanced L2 speakers (Raupach, 1980).

6.3.3.3 Repair fluency

REPETITIONS

![Figure 46: Repetitions for Alia](image)

Alia reduced the frequency of repetitions from 4.14 per hundred words (phw) in interview to 2.35 phw in interview 6 (-43%) (Figure 46). Despite the fluctuations, a correlation coefficient of -0.806 implies that the development was highly linear.

Compared to Götz’s (2013) participants, Alia significantly overused repetitions, both with regard their averages, and their ranges. On average, Alia (≈2.87 phw) used 3.5 times as many repetitions as Götz’s NSs (≈0.82 phw) and 4 times as many repetitions as the advanced NNSs (≈0.69 phw). With the exception of interview 5, Alia never performed within their range (0.05 - 2.07 repetitions phw for both groups). This result is interesting in so far as Götz’s NNS group significantly underused repetitions. Their performance was explained by a lack of internalization of repetitions as a speech planning strategy. Although Götz also suspected that given the broad range the use of repetitions in native and non-native speech is idiosyncratic, Alia’s overuse of the variable is probably more indicative of speech processing difficulties, rather than personal preference. However, the decrease in the measure suggests an approximation to these norms towards the end of her sojourn.
From Figure 47 it can be seen that Alia’s reformulations dropped from 4.14 per hundred words (phw) in interview 1 to 2.63 phw in interview 6 (-36%). The development followed a smooth downward path with only a slight increase towards the end. A correlation coefficient of -.823 affirms a highly linear negative decrease.

Lennon (1990) reported that the majority of his participants increased ‘self-corrections’ during their 6-month sojourn (involving reformulation of discourse and form). I will discuss Alia’s reformulations in more depth to see whether the decrease in the development relates to L2 progress. In the following examples, reformulations were placed in angle brackets.

Throughout all interviews, the majority of Alia’s reformulations concerned changes of content. They concerned both false starts (Example 65), and ‘listener sensitive’ changes. The latter included additions to the message for precision (Example 66) or hedging purposes (Example 67). Both are reflective of restructuring processes typical of spontaneous speech.

Example 65 (int. 1): <and then ehm (.91) today> and there's one girl and I like her very much.

Example 66 (int. 1): (.84) ehm (.31) <m my friends> (1.12) two years ago my friends ehm didn't understand that I want to do so (.75) much sport.

Example 67 (int. 1): (.53) and ehm (2.09) <I I> (.67) in my opinion I can't speak the eh the language English fluent in five months.

Unlike Jana and Chiara, Alia also frequently changed grammatical structures. She mainly modified finite verb forms (Example 68), prepositions (Example 69), and nouns (Example 70). In interview 1, most retracings concerned finite verb forms, while in interview 6, most
retracings concerned prepositions. Changes in form decreased over time (from 17 in interview 1 to 10 in interview 6).

Example 68 (int. 1): and <I (.91) don’t li> <want to> (.34) <I (.37) don't> (.71) <I (.61) would not like eh huh> (1.95) <I want not> (.76) won’t ehm (1.30) meet them because (.62) it's not so (1.32) my (1.63) imagine or (.62) picture.

Example 69(int. 6): and (1.09) maybe the reason that I'm not going anymore (.31) because I could go <of> on Friday.

Example 70 (int. 6): (.47) because they have holiday friends there two ehm (.59) think nineteen or twenty years old <girl> (.33) girls.

Alia’s changes of form signal that she was aware of (some of) the shortcomings in her speech during online production and drew upon her analytical knowledge system to correct them. In terms of L2 progress, this could be considered the first step to interlanguage change (Skehan, 1998). Self-correcting output might also indicate that Alia did not shy away from experimenting with language and taking risks when it came to L2 production. As has been mentioned in 6.3.2, it is unclear if Alia displayed this behaviour only during the interviews given her understanding of my role as an L2 expert, or in the naturalistic context in general. According to Lennon (2000), a learner who seeks feedback from his/her interlocutor on features s/he is uncertain about or who “deliberately tr[ies] out a locution about which she is unsure” is an “ideal learner” (p. 31). However, Skehan (1998) argued that extensive ‘retracing’ is not representative of natural discourse and may thus be “irritating” for listeners (Skehan, 1998).

This short analysis seems to suggest that Alia’s reduction of reformulations reflects advanced L2 levels because she reduced features that are characterised as dysfluent L2 speech (i.e. changes of form). An increase in reformulations can only ever be understood as more advanced when changes of content are measured independently of changes of form.
6.3.3.4 Smallwords

Figure 48: Frequency of smallwords for Alia

Figure 48 shows that Alia increased the number of smallwords from 2.44 per hundred words (phw) in interview 1 to 5.64 phw in interview 6 (+131%). Between interviews 3 and 5, her speech contained a similar frequency of roughly 5.60 smallwords phw. A correlation coefficient of .838 reflects highly linear positive growth.

With an average of 4.63 smallwords phw Alia performed in a very similar manner to Hasselgren’s (2002) NS participants (although this comparison must be made with caution since Hasselgren measured a greater range of smallwords than I did). Her adolescent NS group used an average of 4.46 smallwords phw, the more fluent NNS adolescents used 2.79 smallwords phw, and the less fluent speakers used 2.31 smallwords phw.

In a more qualitative analysis, I investigated the types of smallwords that Alia used in her speech. Table 31 shows that, the range of smallwords Alia used varied between 6 (int. 1, 2 and 4), 9 (int. 3 and 6) and 10 (int. 5) at any one time. Overall, she used 11 different smallwords (out of a total of 12). Alia predominantly used the single-word smallwords ‘yeah’ and ‘so’, as well as the multi-word smallword ‘I think’. From interview 3 onwards, she also frequently used ‘like’. Her speech contained isolated occurrences of ‘just’, ‘and things’, ‘and everything/and (all this) stuff/and things’, ‘and or something like’, ‘a (little) bit’ and ‘ah oh’. Alia never used the smallwords ‘you know’ and ‘I mean’.
Compared to the three stages of smallword acquisition identified by Hasselgren (2002), Alia’s findings are mixed. She used smallwords both of stage 1 (i.e. ‘I think’, ‘just’, ‘and/or something like’, ‘a (little bit)’ and stage 3 (i.e. ‘ah/oh’, ‘like’, ‘and things’, ‘and (all this) stuff’), yet she did not make use of the stage 1 smallword ‘you know’. With the exception of ‘ah/oh’, Alia acquired all stage 3 smallwords later than the stage 1 smallwords, as hypothesized by Hasselgren. ‘Like’, ‘and things’ and ‘and (all this) stuff’ appeared for the first time in interview 3. The smallword ‘I don’t know’ first occurred in interview 5. It can thus be concluded that at least three months of residing abroad were necessary for Alia to produce these later-acquired smallwords. From the third month onward, Alia used smallwords at a considerably higher frequency than she previously had (Figure 48).

However, her extensive use of the same four smallwords (‘yeah’, ‘so’, ‘I think’, and increasingly more ‘like’) shows that she clung to the same lexical teddy bears (or ‘safety zones’) throughout her 5.5 months abroad. As discussed in Jana’s and Chiara’s chapter, ‘so’ is untypical of English native speech and may be explained by L1 interference (Grieve, 2010). The smallword ‘I think’ is often overused by NNSs. Only the increase in ‘like’ reflects more native-like behaviour. However, acquiring smallwords has generally been shown to be hard for NNSs, as they are used in an unpredictable manner by NSs and hardly noticed by NNSs (Götz, 2013). Götz also explains that a lack of systematic implementation of smallwords in German syllabi leads to the underuse of smallwords in Germans’ English. Alia’s sluggish growth
regarding multi-word smallwords in particular suggests that acquiring them was difficult for her.

I suspect that as with Jana and Chiara, smallwords often functioned as ‘fillers’ in Alia’s speech (see Carter & McCarthy, 2006; Götz, 2013; Gut, 2009; Riggenbach, 1991; Wong, 2000). In this function, smallwords can assist speakers in managing the processing burden of online communication without making it sound less ‘smooth’ or native-like (Götz, 2013). For example, it seems that Alia used the smallword ‘like’ in places where she needed to pause in searching for a lexical item (Example 71).

Example 71 (int. 6): (.59) because they had like a house (.62) with a big ehm (.52) like (1.30) farm for f kiwi fruit (.92) and ehm (.58) yeah.

**6.3.4 Accuracy**

**6.3.4.1 Error-free clauses**

Alia’s error-free clauses (EFC) changed from 74% in interview 1 to 79% in interview 6 (+7%) (Figure 49). She reached her peak performance in interview 4 (82%). The development is devoid of a clear trend, as demonstrated by a negligibly low correlation coefficient of .199.

With a mean score of 78%, Alia performed above the mean values reported for the intermediate-level students in Skehan and Foster (1999) (approximately =60%) and Kormos and Dénes (2004) (=53%). Alia was equally as accurate as the intermediate/advanced-level group in Mora and Valls-Ferrer (2012) (=70.50-78.91). Compared to the intermediate/advanced-level group in Kormos and Dénes (2004) (=86%), and the NSs in Mora and Valls-Ferrer (2012) (=99.60), Alia performed less well.
I have discussed the danger of confusing language accuracy with language progress in Jana’s and Chiara’s chapters. I also mentioned that I refrained from conducting a detailed analysis of the participants’ single errors, as such an investigation was considered unreliable.

6.3.4.2 Correct finite verb phrases

Figure 50: Correct finite verb phrases for Alia

Alia’s use of correct finite verb phrases (fvphs) increased from 89% in interview 1 to 95% in interview 6 (+7%) (Figure 50). Unlike her global accuracy measure, the variable proceeded in a highly linear fashion following an upward trend (r=.785).

On average, Alia (≈92%) performed well above the scores obtained from the NNS high-proficiency group in Wigglesworth (1997) (max. 85.51% correct fvphs).

I applied a more detailed analysis of the types of finite verb phrase errors Alia made to potentially account for language progress. With minor exceptions (8 morphological errors), fvph errors in Alia’s speech were tense choice errors (total of 57 errors). Mostly, she overgeneralized the use of a present (simple or continuous) tense to linguistic contexts that unambiguously demanded the use of a past tense or present perfect tense (Example 72). She also failed to backshift verbs in indirect speech complements (Example 73). Especially between interviews 1 and 3, Alia sometimes used present tense for future events (Example 74) or erroneously used a past tense for contexts that required a present tense, like generalizations (Example 75).

Example 72 (int. 2); and (.97) he is living (.57) since five years (.34) here in New Zealand. (<has been living for)

Example 73 (int. 6); (.97) and ehm (1.59) and then James gave me ehm because I said on Friday that hat I really want like a singlet with New Zealand. (< wanted)
Example 74 (int. 1): (.53) and ehm (2.09) I I (.67) in my opinion I can’t speak the eh the language English fluent in five months. (< won’t be able to speak)

Example 75 (int. 2): (.78) because he's with every girl on facebook (.58) he's with ehm (.33) he's friend with every girl (.58) and so she eh she he ehm always said ehm (.54) ‘this girl’s pretty and that's my my chick’ and (.51) and I hate that. (< says)

The few form-related errors in Alia’s data involve missing auxiliaries in compound verb tenses (Example 76), errors on the lexical verb of compound verb phrases (Example 77), or subject-verb agreement failure (Example 78).

Example 76 (int. 1): (.33) and I hope I be fit there. (< will be)

Example 77 (int. 3) (.33) but then it starts (1.07) wen to went well. (< to go)

Example 78 (int. 5): (.37) and she have been in a relationship with (.33) from my friends the father. (<has)

This brief error analysis suggests that with regard to fvphs, Alia mainly made tense choice errors, which is typical of more advanced learners (Götz, 2013; Lennon, 1991).

It therefore seems that for Alia, being abroad was particularly beneficial in terms of producing correct finite verb forms but less effective in correcting other language features, such as prepositions. This observation would be in line with Ellis and Barkhuizen (2005) who claim that “learners do not acquire grammatical features concurrently. Rather, some features are acquired early and others late” (p. 151).
6.3.5 The relationships between CAF

Figure 51: CAF relationships for Alia

Figure 51 displays the nature of the interconnectedness of the six CAF measures speech rate (SR), mean length of run (MLR), error-free clauses (EFC), lexical diversity (LD) and mean length of clause (MLC) (using Z-scores).

The figure shows a very interesting pattern: between the first and the last interval (int. 1-2 and 5-6), all growers increased. It seems that these times provided a powerful spur for Alia’s L2 development. The period in between (interviews 3-5) is marked by more fluctuation. Competitive relationships only ever occurred during these times. Like for Jana’s and Chiara’s CAF discussion, growers were only considered supportive or competitive when at least 3 out of 5 movements were identical or distinct, respectively.

Alia’s CAF growers displayed 1 competitive and 2 supportive relationships: a fully supportive relationship existed between MLC and LD and a partially supportive relationship existed between SR and EFC. Conversely, there was a competitive relationship between the complexity measures (LD and MLC) and EFC. In the following, I will consider these relationships in more detail.
6.3.5.1 Supportive growers

**GRAMMATICAL COMPLEXITY AND LEXICAL DIVERSITY**

Figure 52: Relationship between mean length of clause and lexical diversity for Alia

Figure 52 shows that lexical diversity (LD) (measured with Guiraud’s Index) and grammatical complexity (MLC) developed in unison at all times (i.e. they were supportive growers). That is, longer clauses went hand in hand with lexically more diverse output and vice versa. A supportive relationship implies that attention can be equally distributed among both dimensions during online production. Developing both lexical and grammatical complexity together seems logical: the more word types Alia had at her disposal the greater her capacity was to form more complex (i.e. longer) clauses (Verspoor & van Dijk, 2011). Kormos’ (2011) explained that syntactic processes are dependent on the choice of lexis.

Although the growers fluctuated to some extent, both measures exhibited a highly linear upward development. The temporary dip that both measures underwent in interview 4 could indicate that the choices Alia made in that particular interview (e.g. regarding the macrostructure, topic, level of details of the narrative etc.) were less demanding on the conceptualizer than those in interview 3, for example. However, Alia’s changes in MLC and LD seem to primarily reflect L2 progress, not ‘task influences’.
SPEECH RATE AND MEAN LENGTH OF RUN

Figure 53: Relationship between speech rate and mean length of run for Alia

Figure 53 shows that a partially supportive relationship between SR and MLR existed. During three intervals (int. 1-2; 3-4 and 5-6) the growers developed in unison. Between interviews 2-3 and 4-5, there was an inverse relationship among the variables. These two fluency measures have been shown to correlate because they both reflect proceduralization processes (Gut, 2009). The results for the specific fluency measures further showed that the dip in SR in interview 3 correlated with higher levels of FPs, while the dip in MLR in interview 5 correlated with higher levels of EPs. An explanation of the variables’ divergences during these two times could therefore be that Alia chose to use different ‘fillers’ when she encountered planning problems.

ACCURACY AND FLUENCY

Figure 54: Relationship between fluency and accuracy for Alia

Figure 54 shows that SR and EFC were supportive growers except for the interval between interviews 4-5. MLR and EFC were supportive growers except for the interval between
interviews 2-3. These relationships suggest that for most of the time, Alia could distribute her attention among the fluency and accuracy dimension without a trade-off effect in either. According to Skehan (1998), this finding is not surprising since both accuracy and fluency are affected by formulator-linked attention. It seems that during most times, Alia gave sustained attention to fluency given that both SR and MLR increased in a highly linear fashion and EFC did not.

6.3.5.2 Competitive growers

**Complexity versus accuracy**

Figure 55: Relationship between complexity and accuracy for Alia

![Graph showing the relationship between complexity and accuracy](image)

Figure 55 shows that the fully supportive growers MLC and LD formed a competitive relationship with EFC between interviews 2 and 5. Lower EFC scores (int. 3 and 5) co-occurred with higher complexity scores and vice versa (int. 4). In other words, longer and lexically more diverse clauses were produced, resulting in more overall errors per clause (and vice versa).

This is in line with Skehan’s (1992) trade-off hypothesis, which argues that limited attentional resources allow learners to pay more attention to either accurate or complex output but not both at the same time. In other words, a speaker’s need for tight organization of demanding information exerts pressure on the planning phase (i.e. the conceptualizer), making less attention available for processes related to syntactical and lexical encoding (Kormos, 2011). Applied to the growers in Figure 55, this could mean that interviews 1, 3 and 5 and were cognitively more demanding (i.e. decrease in accuracy) than interviews 2, 4, and 6 (i.e. increase in accuracy). It was noted above that fluency developed in a similar way to accuracy.
Overall, the highly linear upward development of LD and MLC suggests that Alia generally allocated fewer resources to accuracy and more attentional resources to complexity. I showed above that increasing complexity levels indicated that Alia was willing to use language in an uncontrolled way and to experiment with challenging structures. While this appears to have come at the expense of making fewer overall errors (6.3.4.1), it did not seem to come at the expense of fluency. Skehan and Foster (1999) showed that the benefits of this approach (i.e. willingness to experiment with language) can foster overall L2 progress due to the restructuring processes that occur in the rule-based system.
### 6.3.6 Summary

Table 32: Results of all CAF measures for Alia

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Average</th>
<th>Overall development</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subordination Index</strong></td>
<td>1.44</td>
<td>1.42</td>
<td>1.29</td>
<td>1.26</td>
<td>1.46</td>
<td>1.29</td>
<td>1.36</td>
<td>-10%</td>
<td>-0.395</td>
</tr>
<tr>
<td><strong>Mean length of clause (wpc)</strong></td>
<td>5.83</td>
<td>5.86</td>
<td>6.62</td>
<td>6.49</td>
<td>6.84</td>
<td>7.22</td>
<td>6.48</td>
<td>+24%</td>
<td>0.951</td>
</tr>
<tr>
<td><strong>Lexical diversity (Guiraud’s-Index)</strong></td>
<td>8.01</td>
<td>8.70</td>
<td>9.03</td>
<td>8.74</td>
<td>9.53</td>
<td>9.60</td>
<td>8.94</td>
<td>+20%</td>
<td>0.916</td>
</tr>
<tr>
<td><strong>Empty Pauses (phw)</strong></td>
<td>19.55</td>
<td>17.40</td>
<td>15.27</td>
<td>15.36</td>
<td>16.91</td>
<td>14.68</td>
<td>16.53</td>
<td>-25%</td>
<td>-0.760</td>
</tr>
<tr>
<td><strong>Filled Pauses (phw)</strong></td>
<td>8.36</td>
<td>4.91</td>
<td>5.91</td>
<td>3.87</td>
<td>3.45</td>
<td>3.39</td>
<td>4.98</td>
<td>-59%</td>
<td>-0.871</td>
</tr>
<tr>
<td><strong>Speech Rate (wpm)</strong></td>
<td>129.96</td>
<td>150.21</td>
<td>148.91</td>
<td>156.15</td>
<td>156.35</td>
<td>158.60</td>
<td>150.03</td>
<td>+22%</td>
<td>0.857</td>
</tr>
<tr>
<td><strong>Mean Length of Run (wpr)</strong></td>
<td>5.00</td>
<td>5.44</td>
<td>6.25</td>
<td>6.47</td>
<td>5.73</td>
<td>6.56</td>
<td>5.91</td>
<td>+31%</td>
<td>0.764</td>
</tr>
<tr>
<td><strong>Repetitions (phw)</strong></td>
<td>4.14</td>
<td>4.33</td>
<td>2.24</td>
<td>2.69</td>
<td>1.45</td>
<td>2.35</td>
<td>2.87</td>
<td>-43%</td>
<td>-0.806</td>
</tr>
<tr>
<td><strong>Reformulations (phw)</strong></td>
<td>4.14</td>
<td>4.04</td>
<td>2.75</td>
<td>2.15</td>
<td>2.18</td>
<td>2.63</td>
<td>2.98</td>
<td>-36%</td>
<td>-0.823</td>
</tr>
<tr>
<td><strong>Smallwords (phw)</strong></td>
<td>2.44</td>
<td>3.17</td>
<td>5.60</td>
<td>5.59</td>
<td>5.36</td>
<td>5.64</td>
<td>4.63</td>
<td>+131%</td>
<td>0.838</td>
</tr>
<tr>
<td><strong>Error-free clauses</strong></td>
<td>74%</td>
<td>81%</td>
<td>77%</td>
<td>82%</td>
<td>75%</td>
<td>79%</td>
<td>78%</td>
<td>+7%</td>
<td>0.199</td>
</tr>
<tr>
<td><strong>Correct finite verb phrases</strong></td>
<td>89%</td>
<td>91%</td>
<td>92%</td>
<td>91%</td>
<td>91%</td>
<td>95%</td>
<td>92%</td>
<td>+7%</td>
<td>0.785</td>
</tr>
</tbody>
</table>

wpc: words per clause; phw: per hundred words, wpr: words per run

Alia’s results showed that grammatical complexity and lexical diversity increased in a very highly linear fashion (i.e. between r= .90 and 1) and to a substantial degree (+24% and +20%,
respectively). The development of clause subordination (SI) showed considerable fluctuation and was devoid of a clear trend (Table 32).

Regarding fluency, the biggest overall rate of change occurred with smallwords (+131%), followed by filled pauses (-59%), repetitions (-43%), reformulations (-36%), mean length of run (+31%), empty pauses (-25%), and speech rate (+22%). In terms of linearity of development, all variables developed in a highly linear way (i.e. between r = .70 and .90).

Alia’s global accuracy levels fluctuated considerably during the sojourn without a clear trend (.199). Conversely, the specific accuracy measure of correct finite verb phrases developed in a highly positive manner over time (.785).

6.4 Learning opportunities and L2 development

In this section, I will examine possible relationships between the types of language learning activities Alia engaged in during her sojourn and her second language (L2) development. The discussion of Alia’s L2 development revealed that her oral English performance became more fluent, more accurate (only with regard to the ratio of correct finite verb phrases, not error-free clauses) and more complex during the period of her sojourn. Unlike Jana and Chiara, Alia did not make extensive use of direct speech reports, and her complexity and fluency measures were therefore unaffected by this speech mode. Her development was often highly linear and appeared to be representative of L2 progress. This is not to say, however, that other factors such as the nature of the interview data, data analysis processes, or learner affective factors (e.g. ‘task’ motivation), did not also contribute to changes in CAF.

Alia regularly engaged in informal, ‘uncontrolled’ conversations with native speakers (NSs) in her host family, in different classes, and outside school in her athletics club and during doctor’s appointments. In all contexts, she was also involved in more analytical, ‘controlled’ English language learning activities. At school, Alia performed tasks such as essay writing, literature and film analysis, and one-to-one interviews. In the host family, Alia received corrective feedback and was given explanations of English words. During athletics training, Alia was taught technical terms specific to athletics.

The nature of Alia’s informal language learning activities in her host family setting and at school did not seem to change much during her sojourn (these findings only pertain to Alia’s self-reports). From the beginning, Alia enjoyed friendly and close relationships with all of her
host family members. At school, Alia became friendly with some South American international students and some Kiwi girls early in the semester. Except for spending a few school breaks with a group of Kiwi girls, interactions with these local students only occurred inside the classrooms. English conversations at the sports club seemed ample from the beginning. During training, she mainly interacted with the male athletes, not the female athletes, and her male coach (in English). Alia used German in a relatively stable manner throughout her sojourn during (almost) all break times at school and for all social activities outside of school (except for her hobby and activities related to it).

Below, I will first discuss potential reasons for Alia’s significantly increased levels in grammatical complexity (mean length of clause) and lexical diversity (as measured by Guiraud’s Index). I will then examine what types of language activities may have impacted on the fluctuations of her global accuracy levels (ratio of error-free clauses) and the increased specific accuracy measure (ratio of correct finite verb phrases). Finally, I will offer some suggestions regarding the activities that may have affected Alia’s increasing fluency levels.

Like Jana and Chiara, Alia arrived in New Zealand with native-like levels of grammatical complexity and lexical diversity. I assumed that this was the result of their formal analytical language learning activities in Germany which are likely to encourage students to produce output that is relatively complex (i.e. long) and lexically diverse. Unlike Jana and Chiara, however, Alia’s speech variables increased notably during her sojourn. In fact, they rose to a level that has not typically been measured for NSs in informal naturalistic contexts. I suspected that the growth in these measures can be explained by a focus on form during interactions with NSs in all three social settings (i.e. the host family, the classrooms and particularly the sports club), a convergence to the language variations used by the NSs in her communities, as well as her orientation to language learning.

During conversations with NSs (in both informal and formal settings) Alia was exposed to authentic NS input. Processing input for meaning and/or for form can drive interlanguage development (i.e. accuracy and complexity) forward because input “provides the crucial evidence from which learners can form linguistic hypotheses” (Gass & Mackey, 2007, p. 177). Alia’s accounts of her conversations with NSs provide evidence that some of her interlocutors modified their language to make themselves more comprehensible. Alias’s host mother regularly explained words that she considered difficult for Alia. Her coach modified his output (technical vocabulary related to athletics) to facilitate Alia’s participation in the athletics
community. A Kiwi friend explained relevant concepts to Alia to help her follow the English classroom instructions. By modifying their language, these NSs not only enabled Alia’s participation in the activities of the different communities, but they also created opportunities for her to pay attention to particular words and expressions that she otherwise may not have noticed; modified input is hypothesized to assist learners in constructing their L2 grammars because it provides ‘positive evidence’ of what is correct (Gass & Mackey, 2007). I suspect that these ‘collaborative’ processes enabled Alia to internalize new words and structures, and contributed, to a large degree, to the higher levels of lexical diversity and grammatical complexity.

Possibly, some more formal classroom activities also assisted Alia in extending the range of her vocabulary, although Alia did not put much effort into classroom work. For example, she had to do research about somebody’s experience of ‘resilience’ (i.e. conduct an interview), which appears to have led to the acquisition not only of the word ‘resilience’ but also concepts associated with it (e.g. ‘bounce back’, see section 0).

The age and gender of Alia’s NS interlocutors may have influenced Alia’s speaking style, and, in turn, the levels of grammatical complexity and lexical diversity of her output. The analysis of Alia’s L2 development showed that her speech was almost devoid of direct speech reports. Most of the NS input that Alia received came from adults (her host mother and father, teachers, her coach) and males (the athletes). Romaine and Lange (1991) stated that direct speech reports primarily occur in the speech of female high school students. It is possible that the L2 speech communities in which Alia participated used speech modes other than direct speech for communication (e.g. descriptions). In this sense, using direct speech reports would have shown communicative behaviour divergent from her community and would indicate a departure from its norms (Giles & Coupland, 1991). It is also possible that Alia consciously chose not to use more direct speech reports because she did not want to be associated with speech norms used by female adolescent Kiwis because her male friends considered them ‘bitchy’.

I also suspected that Alia’s risk-taking speech behaviour played an important role in her development of grammatical complexity (Faerch & Kasper, 1983; Gass & Mackey, 2007; Skehan, 1998). When learners encounter communication problems and resort to achievement strategies (e.g. paraphrase, word coinage) to resolve the temporary breakdown, they are given the opportunity to test L2-related hypothesis. Faerch and Kasper (1983) argued that “depending on the feedback obtained, hypothetical rules either get rejected or incorporated into
the IL system as fixed rules” (p. 53). A basic condition for interlanguage change to occur is for a learner to “extend his resources without abandoning the IL system completely” (p. 54). Alia’s data implies that during the interviews, at least, she did not shy away from trying out (successfully or not) more advanced language structures. Engaging in more “cutting-edge language” (Skehan, 1998, p. 97) was accompanied by difficulties, which Alia, tried to overcome by means of communication strategies (often cooperative strategies, appealing to me for help). This behaviour may have broadened Alia’s range of L2 structures and vocabulary.

Conversations with NSs can also encourage an L2 learner to use more accurate language. When learners produce erroneous output, they sometimes receive information about its incorrectness (e.g. by means of negative evidence, such as overt correction or through negotiation) (Gass & Mackey, 2007). Alia reported having received overt corrective feedback from her host mother and her host sister. Possibly, more improvements took place in the specific accuracy measure (finite verb phrases) than the global accuracy measure (error-free clause) because when corrected, the finite verb phrase errors granted Alia more generalization power than errors on forms or structures that occurred less frequently in her output, such as prepositions. I also assume that Alia’s ‘nonchalant’ attitude toward school work (e.g. essay correction), which kept her from paying close attention to her teachers’ corrections, deprived her (to some extent) of increasing her overall accuracy levels. Having said that, error-free clauses may simply not be a sensitive enough measure to capture L2 progress (Lambert & Kormos, 2014).

Lastly, the output during conversations in both informal and more formal settings is assumed to have contributed to Alia’s higher fluency levels. Speech is assumed to become faster and more automatic when learners are able to “retrieve the appropriate lemma matching the activated concept and to perform the ensuing syntactic and morphological encoding procedures” in a relatively unconscious manner (Kormos, 2011, p. 52). Alia regularly engaged in chats with classmates, host family members, and sports club members during her sojourn. Her accounts imply that these conversations tended to revolve around relatively similar topics (e.g. ‘boys’, ‘relationships’, ‘food’ etc.), allowing for “routinization of language use”, which, in turn, may have promoted automaticity (Gass & Mackey, 2007, p. 180) (as mirrored in faster speech and fewer hesitations) – and, probably, hypothesis testing, which also led to increased accuracy levels (ratio of finite verb phrases). I also assume that noticing frequent expressions during informal conversations with NSs helped Alia automatize existing L2 knowledge (Kormos, 2011).
In sum, regular informal and formal language learning activities seemed to positively affect Alia’s speech fluency. The language learning assistance that Alia received through more formal, controlled activities in the host family, the school, and the athletics club appeared to stimulate the growth in the grammatical complexity and lexical diversity measure, as well as the increase in correct finite verb phrases.
7 Synthesis: Key Issues in Case Studies

7.1 Introduction

This thesis examined the inter-relationships between the three study abroad participants’ motivations for learning English, the opportunities they experienced to use and learn English, and their actual L2 development over a 5.5-month period. I begin this chapter by examining the participants’ motivations and how these changed over time. The three central aspects of motivation (i.e. motivational orientation, behavioural motivation and attributional motivation) are discussed in separate sections and without taking into account contextual influences. It is acknowledged that such a strict separation is somewhat arbitrary as all three aspects of motivation and the learning contexts are understood to form part of the same dynamic system. However, I believed that examining each central aspect separately is essential for an understanding of their respective influence on motivation. I will then focus on the nature of the language learning opportunities that existed for the participants by considering the contextual constraints they experienced and their agency. Finally, I will examine how their motivation and learning opportunities may be said to have impacted on their L2 proficiency. In this chapter, I attempt to identify the general factors that shaped the participants’ study abroad experiences, pointing out commonalities and differences among them.

The discussion of the learners’ motivation and language learning opportunities, and the discussion of their L2 development are followed by concluding sections that tie the findings in with those from previous studies.

7.2 Motivation

7.2.1 Motivational orientation

When the students were asked to state their reasons for studying abroad, they reported that learning English was not a primary motivation. Jana felt a need to ‘spice’ up her life with new experiences; Chiara wanted new experiences and personality changes; and Alia wanted ‘time-out’ from the pressures of her life in Germany, and to build confidence.
However, all three participants believed that by way of studying abroad in an English-speaking country, they would develop their L2 proficiency, which would help them reach medium and long-term goals. Their medium-term goals were linked to German education requirements (i.e. a focus on passing exams, or achieving entry into an English ‘Leistungskurs’ at the Gymnasium). Their long-term goals were unspecified future studies (Jana), jobs (Jana, Chiara and Alia), travel (Alia), or pleasing parents (Alia). Education, job and travel motivation can be seen as instrumental in orientation (i.e. they had economic and practical advantages), providing both a promotion-focus (i.e. for personal satisfaction) and a prevention-focus (i.e. to satisfy externally imposed educational goals) (Dörnyei, 2009b). Given the dominance of English as an international language, these goals appear to reflect the prevailing discourse in non-English speaking countries, which associates English language skills with higher social status, and, hence, more life opportunities (Dörnyei, 2009a; Kim, 2006).

Of particular interest were the context-specific reasons the learners had for using English during their SA. Education, job, and travel motivation may seem relevant for future success, yet they may not provide a sufficiently powerful stimulus to use English in the L2 environment, particularly in naturalistic L2 interactions. Pellegrino Aveni (2005) distinguished different types of communicative goals that motivated her participants to “attempt communication” (p. 28) during SA: 1) information exchange goals (i.e. the L2 is used for the purpose of providing or receiving information to satisfy a physical, affective or intellectual need, or to share opinion/knowledge with others), 2) social networking goals (i.e. communication is performed to develop relationships and maintain etiquette) and 3) second language practice goals (communication is initiated or sustained for the purpose of L2 progress).

Given the learners’ education, travel and job goals, which were tightly integrated into their desire to study in an English-speaking country, I believe that second language practice was an integral motivating factor in most of the communicative interactions they experienced during their sojourn, although it may not have always been their primary goal. They also did not often make this goal explicit.

Establishing and developing friendships with locals was an important reason for the use of the L2, as explicitly reported by all three participants. Social networking motivations reflected the participants’ desire to become members of L2 communities, indicating an integrative orientation to language learning. The participants also expressed a desire to maintain ties with
their compatriot community, which appeared to be based on their need for psychological wellbeing and ‘having fun’.

Finally, the participants had information exchange goals, especially in the social settings in which they pursued goals related to non-linguistic, personal motives. Chiara often initiated or sustained L2 conversations in the classroom with fellow students and teachers to exchange information or opinions. This was motivated by her need and desire to perform well academically (an educational goal). Jana and Alia were generally less interested in performing well in their New Zealand high schools, and exchanging information to fulfil an intellectual need did not seem compelling for them.

Alia used English for information exchange in her athletics club. One of her main goals in New Zealand was to improve her sports performance (goals related to hobbies) and achieving this goal rested on her ability to successfully communicate sports-related information to her coach and other athletes. For Jana, information exchange appeared to be a compelling reason to use the L2 with her Kiwi friends – not primarily to provide or receive information but to discuss “a favorite topic of interest” (Pellegrino Aveni, 2005, p. 29). Her information exchange goals were incidental to her primary goal of establishing social networks.

In summary, all participants used English during their SA for three communicative purposes. An important goal for all participants was social networking. They were also motivated to practice and improve their English to achieve medium-term educational and long-term career and travel related goals. Finally, in order to function in English-speaking settings (school, home-stay or hobby) and to satisfy their personal non-linguistic goals, they were motivated to use English for information exchange purposes. Like Pellegrino Aveni (2005), I assume that SA students experience multiple communication goals in any one interaction and that the significance of the goals differs with each situation.

7.2.2 Behavioural motivation

In this section I discuss the effort that the participants were willing to expend to achieve progress in their English. Section 7.2.1 showed that the three participants were motivated to improve their English for goals related to post-SA situations. During the SA, the students were motivated to use their English for functional and integrative purposes. The former implies a primary focus on form (particularly for educational goals), the latter on meaning (I acknowledge that drawing a rigid distinction is to some extent arbitrary but useful for
conceptualisation purposes). Hence, during SA, English learning was both the objective, and the instrument the participants used to fulfil their non-linguistic goals.

The analysis showed that the extent to which the participants exerted effort to use/improve their English was neither high nor low in general but crucially dependent on what non-linguistic goal they chose to focus on during SA. Their effort also depended on their contextual affordances (see 7.3 for a discussion of affordances). In this sense, the degree to which the participants ‘put in effort’ to reach their preferred goals (and the extent to which this was possible given contextual constraints), could be said to have equated with the effort they put in to learning English. In line with Kim (2006), I equated motivational behaviour with the participants’ “conscious realization of personal significance of an L2-related activity” (p. 51). In this section, I discuss what areas were of most personal significance to the participants as a way of understanding their persistence with using/learning English.

Jana consistently spoke about the importance of establishing and strengthening relationships with host nationals and practicing her English. Every social setting was characterised as an opportunity to achieve this goal. She expended considerable and sustained effort to establish, and later develop, her Kiwi friendships. She used English even when she believed that remaining in her “expatriate bubble” (Ward & Kennedy, 1993) and speaking German would have been easier. She also made considerable effort to adhere to “etiquette and social propriety during interactions with others” (Pellegrino Aveni, 2005, p. 28). For example, she spoke English with a German friend in front of her Kiwi friends to include everyone in the conversation. Jana did not consider challenges to achieving her goals as ‘too costly’ or outweighing the benefits of pursuing them. However, when Kiwi friends were unavailable (e.g. in her free time and during some school breaks) Jana sometimes resorted to spending time with her German friends, prioritizing a desire for fun and comfort over absolute persistence with the L2 learning goal.

Chiara’s efforts for language learning/use were heavily focused on the classroom environment. In this setting, she put in constant effort to achieve her L2 goals. Although the supportive classroom learning environment (i.e. Chiara indicated that she was welcomed in her classes) and the fact that the use of English was a requirement for communication (i.e. there were, with one exception, no German-speaking students in her classes) were likely conducive to learning in general, I believe that Chiara’s personal desire to perform well in her New Zealand high
school explains the persistence with which she pursued her L2 goals in this social setting. In all other settings, Chiara did not demonstrate perseverance with using/improving the L2.

Alia devoted considerable and sustained effort in communicating with members of her host family and the athletics club to forge social bonds. In these social settings, she persisted in pursuit of her L2 goals even when confronted with threats (i.e. her fear of not understanding others fully, or of corrective feedback), and when in the presence of German speakers. Outside these settings, Alia’s language learning effort decreased. In the classrooms, she appreciated the contacts she had with Kiwis and non-German international students but did not seem to make an extra effort to develop these social networks or to practice English. During break times and in her free time, she prioritised having ‘fun’ with Germans over the challenges that came with social networking or L2 practice with Kiwis.

The discussion shows that despite professing very similar goals for SA (i.e. social networking goals, L2 practice goals), the intensity of effort expended to achieve these goals varied considerably between the participants and was heavily influenced by their motivation to achieve non-linguistic goals. This supports Gardner’s (1985) claim that individual differences in motivation are reflected not in goal setting but in the intensity with which these goals are pursued. He stated that “given the same goal, two individuals could differ in their effort to achieve this goal” (p. 52). It must be noted that for these advanced learners of English, whose communicative L2 competence enabled successful participation in L2 conversations right from the beginning of their SA, improving their L2 did not constitute a fundamental need but only a desire to better participate in L2 interactions.

### 7.2.3 Attributional motivation

L2 learners’ subjective attributions of their L2 performance can affect future learning motivation and behaviour. Studies have shown that when language learners believe that 1) they perform poorly, and that their poor performance is the result of their own disability or lack of effort rather than external factors, 2) but that the outcome of their performance can be changed, and 3) that they themselves are in control of the factors that lead to change, then they are more likely to increase their effort to learn the language in future (Ellis, 2015; Ushioda, 2001).

Clearly, attributional motivation rests on two assumptions. Firstly, a desire for a change in learning behaviour can only come about when learners perceive L2 ‘gaps’ between their current and their desired L2 proficiency. Secondly, even when learners perceive L2 gaps and view
change as within their control, if the change is not deemed necessary it will not likely take place. I believe that for the three participants, attributional motivation increased only when their experiences of L2 deficiencies were related to their personal goals for language use (i.e. related to education, friendships, hobbies).

Jana regarded herself as a perfectionist, who was never satisfied with her English performance. This belief indicates that she located blame for imperfection internally. Jana mentioned struggling most with ‘spontaneous’ speech (mainly speech production processes, but also aspects related to grammar, sentence structure, and appropriate lexis choices) She attached considerable significance to getting better at English and considered that the best method of improving was to speak with NSs more frequently – a strategy that overlapped with her social networking goal. She saw this as impossible while staying with German speakers. To achieve greater speaking time with L2 speakers, she exerted agency and left her L1 compatriots. Possibly, regular interactions with NSs made Jana sensitive to recognizing increasingly more L2 deficiencies, which further stimulated her to learn more. It appears that the gap between her actual abilities and her desired abilities, at any one time, combined with the significance she attached to her L2 learning and networking goals, provided a powerful stimulus for her to enhance subsequent learning-specific actions. At the end of her sojourn, Jana was disappointed that English had not become her default language of thinking and that speaking English did not always feel ‘normal’. She attributed this ‘failure’ both to ‘uncontrollable’ external constraints (i.e. the limited duration of her sojourn) and her own lack of effort (i.e. the fact that she still spent a considerable amount of her free time with German internationals).

Chiara generally did not perceive a large gap between her existing level of English and her ideal level and was mostly satisfied with the improvements she considered had occurred. Only a few L2 aspects caused difficulty during her sojourn: understanding Kiwi friends’ jokes and taboo words, speaking without a German-sounding accent, and speaking English after prolonged periods of speaking German. Chiara acknowledged that she could choose to speak less German to improve her L2, but she consciously chose not to, because the satisfaction of being with her Germans friends outweighed any desire for change. Perhaps, her classmates’ compliments of her level of English also contributed to her lack of motivation to change her learning behaviour.

Additionally, Chiara regularly expressed the belief that improving her English accent and increasing her speaking ability (when it reached a temporary low) was possible simply by
listening to her teachers. In fact, she attributed all of her positive L2 outcomes to listening to her teachers’ English during classes (see 5.2.5). Hence, although Chiara evaluated her L2 proficiency as deficient in some areas, this recognition did not induce in her a desire to change her learning behaviour. The external validation of her current level of English, combined with her belief in listening as a method of improving, appears to have caused her to determine that she did not need to change her learning behaviour. It is also possible that this, in turn, impacted her willingness to evaluate subsequent L2 communication as anything but evidence of her success. Chiara seems to have considered that she had agency (i.e. that the L2 deficiencies were internally located) but had little motivation to take action to actively develop her L2 skills as she believed they were good enough. In fact, Chiara stated in interview 2 that she no longer had language learning goals. She also finished her sojourn being “really happy” (int. 6) about her L2 progress.

Alia’s evaluations of her L2 progress were mixed. She identified aspects she felt had improved (speaking abilities, L2 confidence, a wider vocabulary and better listening abilities) and aspects she had continuous difficulties with (grammar and writing abilities). She also experienced times of L2 lows. Alia typically attributed negative L2 outcomes to her own disabilities or lack of effort (e.g. spending extended time speaking German and not making an effort at correcting essays). At school, she did not change her learning behaviour despite perceiving these gaps and lows. She admitted that speaking ‘more English than German’ or paying more attention to classroom activities would help but that she preferred being lazy over making an effort. Perhaps, this lack of effort had to do with the fact that against her expectations, she was surprised how easily she could communicate with and understand New Zealanders – a perception that might have eliminated a need for further improvements. Conversely, poor L2 performance in her host family and her athletics club proved a hindrance to the achievement of her goals. In these settings, Alia, who believed that she was to blame for her deficiencies, persisted in asking for corrections and explanations in an aim to close L2 gaps (i.e. she took control of her learning). Alia believed that listening to members of these two groups (in particular her coach and her host mother) would assist her learning best and associated all positive L2 outcomes with extended time spent speaking English with these interlocutors.

7.3 Language learning opportunities

The participants’ L2 learning motivation appears to have influenced the extent to which they sought L2 learning opportunities for themselves and their preparedness to address the
constraints (real or perceived) in obtaining such opportunities. The focus of this section is on the actual opportunities that the three participants experienced in the different social settings (in the host family, at school, and outside of school). Below I will describe how the participants acted (i.e. exercised learner agency) in response to their social environment, as they perceived it, and the affordances that arose in the process. In particular, I set out to determine 1) whether or not they took advantage of the language learning opportunities that were available and what the nature of these opportunities were, and 2) whether or not the individuals actively created language learning opportunities and what the nature of these opportunities were.

7.3.1 The host families

The host family setting is commonly believed to be a favourable environment when it comes to language learning in a study abroad (SA) context (Rivers, 1998). Benefits of living with a local family can also include cultural learning and psychological wellbeing (Schmidt-Rinehart & Knight, 2004). However, studies have shown that homestay experiences can vary considerably for SA students and that homestays can be challenging. The outcome of a homestay experience depends “not only on the intentions of the host families, but also on the dispositions and stances adopted by students in relation to their experiences” (Tan & Kinginger, 2013, p. 157). The findings of this study provide further evidence that the “homestay is a mixed blessing” (Rivers, 1998, p. 495).

Jana’s first host family and Chiara’s host family were similar in terms of the opportunities they provided for social interaction and the extent to which they were perceived by them as providing affordances for language learning and social networking. Both families ate dinner and spent evenings in front of the TV, which limited the number and quality of opportunities for talking. Other daily communication was limited to ‘small talk’ with lengthy conversations occurring rarely. The girls were only occasionally invited on outings with their families. Both girls interpreted their host families’ behaviour as indicative of a lack of care and interest in them and were generally dissatisfied with their homestay experiences.

Jana and Chiara managed these perceived limitations differently. Jana embraced the few opportunities for social networking and L2 practice available in her first host family (e.g. cooking dinner with the host mother). Chiara, on the other hand, dismissed most (if not all) potential opportunities for language learning and social networking at home. She expressed disinterest in the family’s topics of conversations and evening routines and did not interpret
either activity as conducive to meaningful interaction or social networking. Chiara also felt that her own attempts at initiating conversation about her daily experiences were met with indifference by the host family members. As a result, she stopped initiating conversations.

Jana changed host families in the hope of finding a linguistically and socially more rewarding environment. Chiara did not change host families despite her feelings of isolation and alienation. Instead of spending time with the host family using English, she retreated to her bedroom where she sought emotional support through Skype conversations with her German family, did homework, or watched German TV.

Jana’s second host family was similar to Alia’s host family. Both girls felt that the environment offered affordances for language learning and social networking and they engaged in daily informal interactions. Both families enjoyed dinner at the dining table away from the TV. This allowed for daily meaningful conversations. Evening activities in both families involved watching TV and both Jana and Alia actively took part. Both families also invited the girls to a number of social activities outside the home at weekends. Jana believed that her second host family supported her language learning goals through in-depth conversations, which provided high levels of engagement and required her to use more advanced English. Alia was grateful for the corrective feedback and word explanations she received. Given these enriching environments, Jana and Alia often preferred spending time with their host families to being alone in their bedroom or out with their friends. The girls also participated actively in their homestays by helping with household chores.

In sum, the extent to which the three participants viewed their host families as providing affordances for language learning/social networking, and the extent to which they actively sought language-learning opportunities, varied. External opportunities arose as a consequence of exchange programme requirements (e.g. sharing dinner) and/or were offered ‘voluntarily’ as a result of the host family’s positive attitudes and behaviour towards the participants (e.g. outings). The extent to which the learners created conditions for language learning was crucially dependent on their sense of validation in the homestay family and the meaning they attributed to the different types of social interactions.

7.3.2 Inside the classrooms

Research on the foreign high school classroom has demonstrated that participation in classroom activities is influenced by the host community’s treatment of the exchange students (e.g. their
attitudes towards them and the efforts they make to integrate them) and the exchange student’s stance towards the foreign practices and the community members (students and teachers) (Churchill, 2006; Kinginger, 2009; Llanes et al., 2012).

While the findings of this study generally support the findings noted above, an additional factor was pivotal when it came to the student’s learning in the high school classrooms. The presence of other L1 speakers, not only in ESL but also mainstream classrooms, proved to significantly detract from the networking opportunities with host nationals in classrooms and during SA in general. This finding has previously only been made for university-level learners. Rather than being the ‘best environment’ for total immersion and language learning, the classroom environment at high schools when there are numerous L1 students limits language learning opportunities often presumed to be integral to SA (Freed, 1995; Spenader, 2005).

Chiara, who studied with domestic students, had significantly different experiences from Jana and Alia, who both studied alongside many other students sharing their L1. The Year 11 classrooms provided Chiara with many opportunities for English language learning. With one exception, Chiara’s classes were populated with local students. She felt that her presence was acknowledged and welcomed by most students and teachers. Sitting next to chatty classmates or interacting with them during group work provided plenty of English speaking opportunities. Chiara’s accounts showed that she also engaged in analytical language learning tasks both at school (through individual classroom tasks), and at home (through homework). Besides taking advantage of these externally afforded opportunities, Chiara actively created additional opportunities for interaction/language learning. In particular, she sought involvement in classroom discussions with her teachers (which she saw as particularly stimulating L2 affordances), and took on an active role in classroom activities with local students (e.g. when doing experiments in science classes). Her increasingly close relationships with Kiwi girls in the different classrooms resulted in frequent opportunities in class for informal chats.

Alia and Jana were placed in Year 12 (Alia and Jana) and Year 13 (Jana) classrooms, and studied alongside other L1 students in a number of subjects. Although opportunities to interact with Kiwis in classes with L1 speakers were still possible, they had to be actively sought. Jana was much more persistent than Alia in building friendships with Kiwis and more actively sought out opportunities for language learning and social networking (e.g. sitting away from an international student and next to a Kiwi student in PE). Alia seems to have been more inclined to interact with L1 or other international students, possibly because this was easier and
did not detract from her goal of having friends at the athletics club. When L1 students attended their classes, seating arrangements and team activities afforded Jana and Alia opportunities for L2 practice, of which they took advantage. Classes populated with L2 students exclusively were more beneficial for the establishment of relationships with NSs. Both Jana and Alia established social networks in their classrooms, although only Jana’s friendships became stronger and extended beyond the classroom environment.

Unlike Chiara, Jana and Alia did not have educational goals in New Zealand, which was reflected in their low levels of active engagement in classroom activities. Jana and Alia aimed to meet their teachers’ expectations but only to fulfil the minimal school requirements. For Alia, who was placed in mainstream English classes, formal language learning activities (e.g. essay writing) seemed less open to negotiation than they did for Jana.

7.3.3 Outside the classrooms and outside of the high school

Building social networks and engaging in regular interaction with NSs has been shown to be difficult (Churchill & Dufon, 2006; Hernandez, 2010; Ife, 2000). This study shows that social interactions between the participants and host nationals outside the classrooms, host family settings, and clubs were particularly challenging with this regard. The three participants’ experiences during the school breaks differed greatly. Jana spent increasingly more time with local students while Chiara and Alia almost exclusively remained in their “expatriate bubble” – a place that has been associated with psychological wellbeing but not socio-cultural adaption (Ward & Kennedy, 1993, p. 143).

Jana’s and Mona’s early decision to physically separate themselves from their international groups during breaks in order to create conditions that would facilitate the establishment of Kiwi friendships opened up a number of opportunities for social interaction and L2 practice. Once they separated, the girls were approached by Kiwis and spent a number of break times in different L2 groups. After roughly two months, some of Jana’s classroom friendships became stronger and eventually extended to encounters in the breaks where she made more L2 friends. As a result of the positive response of the Kiwi students and Jana’s own pro-active behaviour, the school environment provided her with increasingly more affordances for informal language learning and social networking.

Jana took advantage of many opportunities for social get-togethers with her Kiwi friends and her own efforts to suggest meetings contributed to these opportunities. In the second school
term, her decision to join the football team with some of her Kiwi friends allowed for additional L2 interactions. However, Jana did not fully abandon her friendships with Mona and other German friends and occasionally met with them in her free time.

In contrast to Jana, Alia and Chiara spent most break times and most of their free time with Germans (unless they were at their hobbies). Neither participant reported seeking to create L2 learning opportunities with their Kiwi classroom friends during the breaks. Neither took full advantage of the opportunities that were offered to them. They accepted invitations from classmates to spend a small number of break times together, but they later refused their invitations because they felt excluded from the ensuing conversations. They felt that spending time with L1 students was more comfortable or fun. Both participants mentioned having proposed meetings outside of school to some of their Kiwi friends. This was taken up on one occasion by one of Chiara’s friends but never by Alia’s classroom friends.

### 7.3.4 Hobbies

All three participants joined a sports club (or several) at some stage during their sojourn. Chiara joined a number of extra-curricular activities to find opportunities for social networking and L2 use in places other than the school context (e.g. dragon boating). Although she appears to have been granted “legitimate peripheral participation” (Lave & Wenger, 1991) in these clubs, Chiara did not consider the ‘environmental properties’ of these contexts (predominantly Asian international students or ‘older’ students) as affordances and no friendships developed. Similarly, when Chiara attended rock-climbing classes in Term 2, she preferred to perform all activities with another German friend instead of mixing with the local students. During these activities, it can be assumed that Chiara used English only for purposes of information exchange.

Joining the local athletics club automatically provided Alia with plenty of opportunities for social interaction and L2 use. She used English with team members and coaches in training sessions for friendly chats and information exchange, which often involved the use of more technical vocabulary. Furthermore, regular visits to doctors and physiotherapists added opportunities to hear and use these technical terms. Sometimes, Alia communicated with her team members on an extended basis, such as during multi-day sports competition at weekends or during team dinners. Alia devoted considerable effort to attaining both sports and social-
network related goals in this community. Jana’s football experience was briefly mentioned in section 2.3 (it involved the same L2 community that she was involved in at school).

7.3.5 Concluding comments

This study provides evidence that SA as an L2-rich context does not in and of itself offer language learning opportunities, but that learning dynamically emerges in the “mutually constitutive relationship between persons and the contexts in which they act” (Ushioda, 2009, p. 218). The findings reported above demonstrate that the learners’ participation in L2 communities (or a lack of it) was influenced by their motives, goals and behaviour and changed dynamically in relation to the perceived behaviour and stances of their L2 communities and the level of access they were given.

Isabelli-Garcia (2004; 2006) and Hernandez (2010) found that SA students who were highly motivated to learn the target language (TL), had an integrative orientation to learning, and positive attitudes towards the host culture in general were better able to establish social networks during their SA than the learners with low motivation. Isabelli-Garcia (2004; 2006) also observed that integrative motivation and positive attitudes were only maintained during the SA for learners who participated in strong social networks. This thesis has found that there was no clear linear relationship between the participants’ language learning orientation and attitudes toward New Zealanders, and their abilities to establish social networks in general. All three participants continuously expressed an instrumental and an integrative orientation to language learning. The degree to which the participants were able to establish strong social networks depended on their behavioural motivation (i.e. the actual effort that they made to build relationships), the perceptions they had of specific social groups, and the degree to which they believed these social groups were positively disposed towards them. In other words, the participants’ motivation differed for each social group and setting – it was neither generally low or high, nor integrative or instrumental.

Willis Allen (2010) and Kim (2011) found that unless participants understand that their L2 learning activities are “efficacious and meaningful to their L2 development” (Kim, 2011, p. 115) L2 learning motivation and L2 learning are not facilitated. They both pointed to the importance of goal setting and goal internalization – arguing that learning in the L2 environment can be maximised when clear goals are followed. The findings of this study indicate that the three participants subordinated their L2 goals to their non-linguistic personal
goals most of the time, making L2 goal-setting arguably redundant. The activities they performed primarily to satisfy their personal goals (i.e. being studious, doing sports, being friendly) were ‘automatically’ considered meaningful to their L2 development, while activities motivated by other purposes were considered less meaningful. The findings showed that all three participants experienced beneficial language learning opportunities in one or multiple social settings during SA.

Motivation also relates to the participants’ abilities to balance contact with L1 and L2 communities without this seriously affecting their language learning. Maintaining close relationships with compatriots has generally been considered detrimental to SA students’ language learning (Isabelli-Garcia, 2004; Kinginger, 2008) – often because it is related to the students’ inability to form close relationships with host nationals and to acculturate to the new environment (Isabelli-Garcia, 2006). Spenader (2005) argued that these relationships are more complex and that only when interacting with compatriots totally replaces contact with L2 members is learning negatively affected. This study provides evidence in support of Spenader’s (2005) findings. A crucial factor that seems to have made L1 friendships possibly less detrimental to language learning was the students’ high initial proficiency level. Their ability to communicate did not depend on consistent L2 practice. This gave them the freedom to ‘relax’ with their compatriots without detrimentally affecting their L2. This contrasts with many previous findings, particularly those concerning Americans abroad, which showed that only through consistent effort at language learning were the students able to successfully participate in L2 interactions (e.g. Kinginger, 2008).

The findings on motivation and language learning confirm those made in recent motivation research, namely that motivation is a dynamic construct that “changes over time as a result of multi-level interactions with environmental factors and other IDs” (Dörnyei & Ushioda, 2011a, p. 198). Motives are “inherently unstable, gaining or losing power depending on the conditions, content, and course of activity” (Willis Allen, 2010, p. 31). It is the actual L2 learning experiences that influence retention, modification, or rejection of learners’ initial motives (Kim, 2011; Churchill and duFon, 2006).

The findings also illustrated that participation in L2 communities is a “site of contestation and renegotiation” of identities (Lantolf & Pavlenko, 2001, p. 149) and that learning is a ‘situated’ activity where agency is co-constructed. Unless the learners positioned themselves and were received as “persons of consequence” (Tan & Kinginger, 2013, p. 156) in and by their host
communities, they “remain[ed] on the margins of a particular community of practice” (Lantolf & Pavlenko, 2001, p. 152). Hence, a crucial factor that seemed to impact the learners’ efforts at initiating and maintaining L2 conversations and building social networks was their identity.

Kinginger (2008) and Spenader (2005; 2011) indicated that national identity can have a strong influence on students’ practices in their host communities. Kinginger (2008) showed that if the host culture has negative perceptions of the learners’ nationality then that can create unfavourable conditions for interactions during SA. She found that those who felt personally challenged when host nationals criticised the United States tended to adopt a sense of “national superiority” and sometimes resisted engaging in further L2 interactions. Spenader (2005; 2011) found that the American learners who desired to adopt a Swedish identity during their high school SA were more integrated into the host culture and showed higher L2 gains than those who rejected being ‘Swedish’.

In this study, the students’ social identities seemed to have been the significant factor that determined success or failure of their integration into L2 communities. Chiara seems to have had a particular motivation toward investing in her identity as a good student. Alia was motivated toward investing in her identity as an athlete. Jana seems to have been motivated toward investing mostly in her identity as a popular friend. As with motivation, different social settings were either more or less conducive to the portrayal of these preferred identities, which created different affordances for learning and social networking. For Chiara, the high school classrooms provided ideal affordances to negotiate her position as a competent classroom member. For Alia, the athletics club afforded ideal opportunities to participate as a successful athlete. Jana’s investment in her identity as a friend was less confined to specific social settings and more confined to a specific social group (her Kiwi friends) where she used her social skills. Possibly, being ‘German’ in New Zealand, compared to being American in France or Sweden, was viewed as unproblematic or even desired both by Kiwis and the exchange students, and therefore did not become the locus of challenge.

In line with Pellegrino Aveni’s (2005) research, these findings suggest that the participants strove to “foster a sense of validation through social interaction, that is, a sense that their presence is welcome, even sought after” (p. 19). When the contexts allowed them to engage as a successful student, a successful athlete, or a popular friend, respectively, the participants were able to minimize the possibility of experiencing a sense of inferiority to their L1 self.
Having different investments in different groups or members of the TL community also confirms Norton and Toohey’s (2005) research. They argued that “people in whom learners have the greatest investment may be the very people who represent or provide access” (p. 119). Through the use of popular social identities, the three participants were successful at gaining access to L2 communities, in which they gained the right to speak. Making themselves “socially salient” (Churchill & DuFon, 2006, p.20) provided them with symbolic power and, hence, language learning opportunities. The findings also concord with Lantolf and Pavlenko’s (2001) observation that learners “actively engage in constructing the terms and conditions of their own learning” (p. 145). They mentioned that learner agency can be derived from the identities they adopt, and that “being positioned as an athlete and a popular friend” (p. 148), for example, can be more satisfactory than purely focusing on developing L2 skills.

It seems that the social settings or groups where the participants could draw on their preferred social identities was where they increased their motivation to use the L2 for practice and social networking. These were the places where they felt most validated, but they were also the places that were “familiar to the learners with common rituals and routines” which put them in control and gave them a sense of accomplishment (Pellegrino Aveni, 2005, p. 140). Sticking to familiar beliefs and identity patterns seems to have made participation more successful.

This, in turn, seems to explain why Chiara and Alia abandoned their efforts to use English and preferred the company of L1 communities in settings which did not allow for the portrayal of their preferred identities (i.e. outside the classroom and the athletics club). Conversely, it is possible that Jana’s persistence at establishing social interactions were continuously high, even when German speakers were around, because her investment in her social identity dependent on continuous efforts in a variety of social settings.

7.4 L2 development

In the following section I will discuss the participants’ oral language development as it occurred during their study abroad in relation to the social factors that influenced their development. I will start this section by describing the influence the interviews had on the students’ L2 performance. I will then describe the three students’ L2 development in terms of their similarities and differences, and suggest explanations for the observed changes. Finally, I will discuss the ways in which the CAF constructs interacted and suggest what this shows about the growth and decline of the variables.
7.4.1 Interview factors affecting CAF

Previous research has shown that speech production (assessed in terms of complexity, accuracy, and fluency measures) is affected by several factors related to the cognitive and interactional demands of the elicitation tasks (Kormos, 2011; Pallotti, 2009; Robinson, 2007; Skehan & Foster, 1999). In this study, semi-structured interviews were used to elicit spontaneous oral language data. The interviews resembled friendly informal conversations in style (Labov, 1984; Nippold et al., 2005). That is, they were dialogic and revolved around the same few topics: experiences in the host family, at the college, during their free time with their friends and in sports clubs. The interview questions were generally open. Overall the discourse demands on the participants were relatively low since they were only asked to handle information that was familiar to them. The participants were not given any pre-task planning time, but, as is customary for conversations, “various degrees of planning” (van Lier, 1989, p. 495) marked the actual interaction. The participants’ speaking time was not restricted in any way.

However, although the discourse demands on the participants were constant, the participants were, within the very broad topic boundaries of the interviews, free too chose exactly what to speak about (i.e. the specific events, sub-topics, people), when to speak (they could interrupt me or not, hold very long turns or not) and how to speak (i.e. decisions related to linguistic, pragmatic and discourse choices). The interviews resulted in conversations that went in many directions, depending on the participants’ interests, needs, and abilities, as well as the needs the participants believed the interlocutor had.

One of the discourse-related choices the three participants made during the interviews concerned the speech mode. When narrating past events, the participants were free to give more ‘factual’ or more ‘dramatic’ reports of what had happened. These choices were affected by the events the students narrated (e.g. referring to a personal story vs. an event that regularly happened at the school) (Lampropoulou, 2013). One feature was particularly prominent in two of Jana’s and Chiara’s interviews: direct speech reports (Alia made infrequent use of this speech mode). When analysing their oral speech, it became clear that the use of this device had a profound effect on some CAF measures. The nature of direct speech reports resulted in more complex utterances (higher subordination index), less ‘complex’ (i.e. shorter) clauses (for grammatical complexity), lexically less diverse output (for lexical diversity) and fewer empty pauses (for fluency). It also resulted in longer runs and higher speech rate. Direct speech reports
did not appear to affect accuracy measures. No attempt was made to control the speech modes the students opted for.

Another choice concerned the level of detail the participants chose to provide during each interview. Some descriptions were characterised by provisions of “discrete elements” (Isabelli-Garcia, 2004, p. 52) without elaborations. Other narrations were detailed and explicit, involving explanations and supporting arguments. Given the learners’ high initial L2 proficiency levels, it is unclear to what extent the latter is reflective of more developed communicative abilities or choices related to personality and their idiosyncratic style (as observed by Isabelli-Garcia, 2004).

7.4.2 CAF results

This study used data captured over six data collection periods to measure L2 development. In line with Dynamic Systems Theory (Verspoor, de Bot, & Wander, 2011), which perceives language behaviour as unstable and dynamic, the language features (as measured by various CAF measures) did not develop in clear linear ways. That is, none of the scores on the measures in this study consistently went up or down in relation to previous scores. Rather, the participants’ linguistic profiles were characterised by variation (i.e. peaks and dips in the graphs).

This study used correlation coefficients to determine the extent to which development was statistically linear (remember that the term development in this study does not refer to more advanced language use but to language change over time). These ranged from very weak to very strong correlations. It is important to emphasize, however, that development that was determined as ‘highly linear’ by a correlation coefficient does not denote development devoid of dips and peaks but development lacking unusually sharp peaks and dips.

7.4.2.1 Similarities

The participants L2 development showed commonalities with regard to some fluency-related aspects, as well as the types of clauses they used (grammatical complexity).

**Fluency**

The results for the speed fluency measures, the breakdown fluency measures, and repetitions were very similar for the three participants (Table 33).
Table 33: Similarities in development of fluency

<table>
<thead>
<tr>
<th>Measures</th>
<th>Development</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Speed fluency</strong></td>
<td>• SR and MLR increased</td>
</tr>
<tr>
<td></td>
<td>• The increase occurred in a moderately positive (Jana) or highly linear fashion (Chiara and Alia)</td>
</tr>
<tr>
<td></td>
<td>• SR and MLR resembled NS norms</td>
</tr>
<tr>
<td><strong>Breakdown fluency</strong></td>
<td>• EPs and FPs decreased</td>
</tr>
<tr>
<td></td>
<td>• The decrease in EPs occurred in a positively linear fashion</td>
</tr>
<tr>
<td></td>
<td>• The frequency of EPs remained high and non-native like</td>
</tr>
<tr>
<td><strong>Repair fluency</strong></td>
<td>• The frequency of repetitions decreased</td>
</tr>
</tbody>
</table>

SR: speech rate, MLR: mean length of run, EPs: empty pauses, FPs: filled pauses, NS: native speaker

The participants’ speed fluency variables demonstrated an overall increase for speech rate (SR) and mean length of run (MLR). Although each participant’s development contained idiosyncratic dips and peaks typical of nonlinear systems, improvements for all three participants occurred in a moderately positive or highly linear fashion. They performed within a range that was considered native-like from the start of their sojourn. As discussed above, Chiara’s speed fluency measures also reflected her use of the direct speech mode.

The breakdown fluency of all three participants decreased in terms of empty pauses (EPs) and filled pauses (FPs), as well as repetitions during their sojourn. EPs decreased in a positively linear fashion for all three students while the linearity of the development for FPs and repetitions varied between the participants. Despite the overall decrease in EPs, all three participants significantly overused filled pauses at all times compared to native speaker values. This was not observed for FPs and repetitions.

The overall changes in these fluency measures were interpreted as reflecting higher proceduralization skills. That is, for the three participants, word retrieval became quicker and grammatical rules could be applied more automatically over time. Higher speed fluency (i.e. automatized L2 knowledge) has been shown to result from substantial exposure to informal L2 input and the learners’ own productions in social interactions (Ellis, 2008; Mora & Valls-Ferrer, 2012; Skehan, 1998; Wood, 2012). Spontaneous, informal L2 use with NSs can be assumed to involve relatively low structural complexity and similar topics (Collentine, 2004) – and this was confirmed by some of the students’ comments. Kormos (2011) claimed that attending to “frequently-heard constructions can raise the activation level of the learner's representation of linguistic items and can help accessing and automatization of existent
knowledge of the L2” (p. 50). Frequent opportunities to speak are considered crucial for students to “go beyond carefully constructed utterances and achieve some level of natural speed and rhythm” as “only by frequent use is the fluency side of speech likely to be improved” (Skehan, 1998, p. 18). In section 7.3, I demonstrated that all three students had substantial exposure to L2 input and sustained opportunities for L2 production in informal and/or formal interactions during their study abroad (although the quantity of these types of interactions differed for the three students). These opportunities provided them with good opportunities to use syntactical structures and lexical items with more ease and speed.

**Grammatical Complexity**

Table 34 shows a very similar pattern in the three participants’ use of clause types.

<table>
<thead>
<tr>
<th>Measures</th>
<th>Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clause types</td>
<td>Majority of finite complement clauses</td>
</tr>
<tr>
<td></td>
<td>Followed by adverbial clauses</td>
</tr>
<tr>
<td></td>
<td>Very few (or no) finite relative clauses, non-finite complement clauses, non-</td>
</tr>
<tr>
<td></td>
<td>finite adverbial clauses</td>
</tr>
</tbody>
</table>

The participants’ use of subordinate clause types was very similar. They predominantly used finite complement clauses and finite adverbial clauses and only rarely finite relative and (complex) non-finite clauses. These findings can be explained by the nature of spontaneous speech, which, as I have shown, the learners were exposed to and produced frequently during their SA. Non-hierarchical utterance structures are not a sign of lower-level language abilities but a reflection of informal speech (Brown & Yule, 1983; Carter & McCarthy, 2006; Ferrari, 2012). Carter and McCarthy (2006), for example, argued that “[t]he needs of real-time communication do not allow the speaker time to construct over-elaborate patterns of main and subordinate clauses” (87a). Likewise, the use of mainly finite complement and adverbial clauses is less indicative of ability than of communicative appropriateness in spontaneous speech situations (e.g. Nippold et al., 2005). However, relative and complex non-finite clauses are later-acquired structures (Diessel, 2004) and perhaps constitute less frequently used and more demanding linguistic structures.

It must be borne in mind, however, that the results may reflect the way utterances were segmented into AS-units in this study, particularly with regard to the treatment of non-finite verb structures, direct speech reports, and which and because-clauses (see Chapter 3).
7.4.2.2 Differences

While the three participants’ L2 development was marked by similar paths for speed and breakdown fluency measures (and repetitions) and similar frequencies for subordinate clauses, most of the CAF measures did not develop uniformly. These include the remaining complexity measures, the accuracy measures, as well as the fluency measures reformulations and smallwords.

FLUENCY

Table 35: Differences in development of fluency

<table>
<thead>
<tr>
<th>Measures</th>
<th>Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reformulations</td>
<td>• Jana and Chiara: overall increase with clear and positive linear development</td>
</tr>
<tr>
<td></td>
<td>• Alia: overall decrease, clear linear development</td>
</tr>
<tr>
<td>Smallwords</td>
<td>• Chiara: overall decrease, great fluctuation</td>
</tr>
<tr>
<td></td>
<td>• Jana and Alia: overall increase in a moderately (Jana) and highly (Alia) positive way</td>
</tr>
</tbody>
</table>

Alia’s behaviour regarding the use of reformulations was different from Jana’s and Chiara’s behaviour, whose development was almost identical, in so far as her reformulation scores decreased over time, while Jana’s and Chiara’s scores increased (Table 35). However, from interview 4 onwards, the three participants’ behaviour converged regarding the frequency of reformulations, suggesting that their involvement in L2 interactions during their study abroad period resulted in the attainment of a ‘norm’ for advanced German speakers of English. The three participants’ developments were highly linear (Jana and Alia) or moderately linear (Chiara).

Reformulations can occur with either content- or grammar-related elements. Both NNSs and NSs correct content during unplanned speech while correcting grammar (‘retracing’) more frequently occurs in NNS speech. Two patterns stood out for retracings. Jana and Chiara hardly modified grammatical forms while Alia did so frequently. Changes of form were interpreted as signs that Alia was aware of (some of) her shortcomings during online production and drew upon her analytical knowledge to correct them. Jana’s highly accurate output did not provide many opportunities for linguistic modification. Chiara, whose output was less accurate, often ignored her errors. It is unclear whether her lack of monitoring indicated a lack of awareness of her errors (which could be interpreted as evidence of a lower level of proficiency), or simply a lack of concern with correctness. The analysis of content modifications, which concerned the
majority of the participants’ reformulations, reflected a concern for precision and highly developed conversational skills. They were made in response to the listener’s needs (e.g. by providing additional information or clarification) or for the purpose of self-positioning (e.g. hedging).

I found different patterns in the participants’ development of smallwords: Jana and Alia increased the number of smallwords in a moderately positive (Jana) and highly positive (Alia) way over time while Chiara’s smallword development was marked by fluctuation and an overall negative trend. However, as for reformulations, they used an equal number of smallwords towards the end of the sojourn, suggesting that their L2 practice opportunities abroad brought about a ‘normal’ pattern for advanced German speakers of English.

The qualitative analysis of smallword types revealed that the participants used a similar range of smallwords in similar ways. They all overused one-word smallwords which they preferred to multi-word smallwords; they all clung to a small number of mostly non-native-like ‘teddy-bear’ smallwords (e.g. ‘so’), and they all underused some more native-like smallwords (e.g. ‘just’, ‘I don’t know’). In line with Götz (2013), I suspected that acquiring smallwords appears difficult for NNSs, even when exposed to regular L2 input, as they are hardly noticed in speech and seem to be used in unpredictable ways (Götz, 2013).

**ACCURACY**

<table>
<thead>
<tr>
<th>Measures</th>
<th>Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratio of error-free clauses</td>
<td>• Jana and Alia: no trend detectable</td>
</tr>
<tr>
<td></td>
<td>• Chiara: moderate increase</td>
</tr>
<tr>
<td>Ratio of correct finite verb phrases</td>
<td>• Jana: overall decrease</td>
</tr>
<tr>
<td></td>
<td>• Alia and Chiara: overall increase in a moderately (Chiara) and highly positive way (Alia)</td>
</tr>
<tr>
<td></td>
<td>• Jana and Alia made errors typical of advanced NNSs</td>
</tr>
<tr>
<td></td>
<td>• Chiara made a variety of errors, often typical of lower-level NNSs.</td>
</tr>
</tbody>
</table>

As can be seen in Table 36, one of the differences in the participants’ global accuracy measure (i.e. error-free clauses) was that Jana’s and Alia’s development did not manifest linearity whereas Chiara’s development did to some extent although her improvements were small, with one exception.
The results for correct finite verb phrases (fvph) showed different trends (Table 36). Jana’s development decreased over time in a moderately negative way (although marginally). Chiara’s development increased in a moderately positive (Chiara) way and Alia’s in a highly positive (Alia) manner. The more fine-grained analysis of the different error types revealed that Jana and Alia made predominantly tense-choice errors (typical of more advanced learners), while Chiara also made a number of morphological errors (typical of lower-level learners) (Lennon, 1991), including what appeared to be ‘stabilised’ language errors (‘I gonna’).

Overall, Chiara’s and Alia’s accuracy results suggest that spending 23 weeks abroad was beneficial for finite verb phrases. Chiara’s focus on more formal classroom talk and analytical language learning activities may have helped her accuracy. Alia, besides being engaged in some formal English learning at school, additionally received corrective feedback from her host family, which may have helped her accuracy. Jana’s accuracy levels were highly advanced right from the beginning and the fluctuations interpreted as a reflection of the normal ebbs and flows of spontaneous speech. She had little room for improvement – a finding that has been made before (Freed, 1995; Isabelli-Garcia, 2004). However, it is possible that the nature of her language learning activities (predominantly informal conversations), did not help her eliminate the errors that she did make. It was acknowledged that more accurate output, as measured by the global accuracy measure (error-free clauses) does not automatically reflect more developed language use. It could also reflect the use of less advanced, and thus more accurate language structures (Lambert & Kormos, 2014; Norris & Ortega, 2009; Pallotti, 2009).

**Grammatical Complexity**

Table 37: Differences in the development of grammatical complexity

<table>
<thead>
<tr>
<th>Measures</th>
<th>Development</th>
</tr>
</thead>
</table>
| Mean length of clause| • Jana and Chiara: overall decrease, great fluctuations  
                       • Alia: overall increase, highly linear development |
| Subordination Index  | • Chiara: overall increase, moderately linear development  
                       • Jana and Alia: overall decrease, great fluctuations |

The three participants’ first and second mean length of clause (MLC) scores were almost identical. Subsequent to this, their development followed two different routes. Jana’s and Chiara’s developments fluctuated and decreased overall while Alia’s development increased in a highly linear way (Table 37). It was assumed that the initial two scores were reflective of the participants’ formal schooling in Germany, which had required them to produce
syntactically complex utterances. From interview 3 onwards, Jana’s and Chiara’s non-linear MLC development primarily reflected changes in speech modes (see section 1.2). Alia’s MLC development seemed to demonstrate increased L2 proficiency and her willingness or ability to talk in greater detail about her experiences.

These changes were likely to have reflected the speech styles of the L2 communities they chose to participate in. Jana and Chiara mainly interacted with adolescent females who are believed to frequently communicate by means of dramatic and emotive techniques, such as direct speech reports (Lampropoulou, 2013; Renouf & Kehoe, 2011; Romaine & Lange, 1991). It is possible that Jana’s and Chiara’s desire to adapt to the norms of their Kiwi friends, including their linguistic norms, encouraged them to use this structure more frequently over time when relationships became stronger. Alia, on the other hand, often interacted with same-age and older male athletes who may have preferred narrative forms other than direct speech reports for communication (e.g. descriptions).

Two different developmental trends existed for the subordination index (SI). Chiara’s SI scores showed a moderately linear overall increase while Jana’s and Alia’s SI scores decreased over time, albeit with much fluctuation. As for MLC, Jana’s and Chiara’s SI development mirrored to a great extent the choices they made with regard to the use of speech modes. Jana’s and Alia’s decreasing SI levels can probably be explained by the nature of spontaneous speech, which demands output of relatively low complexity levels (Brown & Yule, 1983; Ferrari, 2012). Ferrari (2012), for example, found that grammatical complexity (SI) is lower in interactional speech and that little subordination is more appropriate than much subordination in tasks like semi-structured interviews. Conversely, the higher SI scores that Chiara reached in interviews with few direct speech reports could reflect the impact of the more formal and grammatically complex nature of classroom communication.

**LEXICAL DIVERSITY**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical diversity</td>
<td>Jana and Chiara: overall decrease, great fluctuations</td>
</tr>
<tr>
<td></td>
<td>Alia: overall increase, highly linear development</td>
</tr>
</tbody>
</table>

The findings for lexical diversity (LD) showed that the three participants started their sojourn with similar and native-like LD scores but finished it with quite different scores. Alia’s output
became increasingly more lexically diverse while Jana’s and Chiara’s output fluctuated and was marked by an overall decrease (Table 38) although this is deceptive for Chiara’s development as her LD scores increased marginally between interviews 1 and 5 and only decreased in interview 6.

It is possible that the first score reflects the participants’ language learning background in EFL classrooms in the German Gymnasium, which was relatively similar. The remaining LD scores seem to reflect, once again, the students’ different socialization processes. That is, the overall decrease in Jana’s LD development was interpreted not as a sign of language ‘regression’ (her performances remained within NS norms) but as an indication of a shift in speech patterns towards the norms of her L2 communities. Adolescent NS speech is marked by “general, nonspecific vocabulary” (Brown & Yule, 1983, p. 9). Thus, producing lexically less diverse output might have been contextually appropriate and indicative of L2 progress. Chiara’s engagement with more formal language in the classroom may have been the source of the increase in lexical diversity in her speech although the increase was marginal. For Alia, it was suggested that in addition to the formal L2 activities in the classroom, her membership in the athletics group most significantly impacted her LD development. Alia greatly increased her technical vocabulary related to athletics, which seemed reflected in her LD development.

7.4.3 CAF relationships

This study also sought to examine the relationships among CAF components that could be meaningfully compared in order to detect whether they worked in unison or competed against each other.

7.4.3.1 Supportive relationships

<table>
<thead>
<tr>
<th>Measures</th>
<th>Participants affected by these relationships</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical diversity and mean length of clause</td>
<td>Jana, Chiara and Alia</td>
</tr>
<tr>
<td>Speech rate and mean length of run</td>
<td>Jana, Chiara and Alia</td>
</tr>
<tr>
<td>Accuracy and fluency</td>
<td>Jana and Alia</td>
</tr>
<tr>
<td>Accuracy and lexical diversity</td>
<td>Chiara</td>
</tr>
</tbody>
</table>
For all three students, the relationship between mean length of clause (MLC) and lexical diversity (LD) was supportive (Table 39), suggesting native-like ways of speech processing. More sophisticated (or here diverse) choice of lexis may have driven the participants’ syntax “because these [more complex] lexical choices are more likely to require more complex syntactic clothing” (Skehan, 2009, p. 516). Conversely, the use of more highly frequent words (as indicated by lower MLC and LD levels) may have triggered less ‘sophisticated’ syntax.

All three participants’ development was furthermore marked by a correlation between MLR and SR. This relationship has also been observed before, for example by Gut (2009) who showed that speakers with a faster speech rate also produce longer runs. It could thus be argued that fluency variables which reflect similar speech processes (e.g. levels of automatization) grow together.

Jana’s and Alia’s speech was characterized by a supportive relationship between speech accuracy (error-free clauses) and fluency (both MLR and SR). This relationship reflects previous findings in task-based research. For example, tasks that require learners to handle familiar information with few elements and a clear macrostructure have been shown to conjointly raise accuracy and fluency. Chiara’s speech showed a partially supportive relationship between error-free clauses and lexical diversity. This relationship has been observed before by Robinson (2007) who claims that both form-based measures (i.e. accuracy and complexity) occur in tandem in complex tasks. I further suspected that in Chiara’s case, automatization of syntactical encoding processes made room for lexical encoding processes, which resulted in more accurate and more lexically diverse output.

However, as can be seen in the individual case study chapters, the relationships between these sets of growers were not supportive all the time (i.e. during some intervals they competed against each other). The different ‘strengths’ of these relationships show that “relations […] change throughout the course of development” (Larsen-Freeman, 2009, p. 583) as a result of their dynamic interaction with each other and with the environment.
7.4.3.2 Competing relationships

Table 40: Competing CAF relationships

<table>
<thead>
<tr>
<th>Measures</th>
<th>Participants affected by these relationships</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy and grammatical complexity</td>
<td>Jana and Alia</td>
</tr>
<tr>
<td>Fluency and grammatical complexity</td>
<td>Jana and Chiara</td>
</tr>
</tbody>
</table>

Table 40 shows the competing CAF relationships that existed. These relationships did not exist for all three participants. A competitive relationship existed between accuracy (error-free clauses) and grammatical complexity (MLC) for Jana and Alia. Jana’s and Chiara’s speech was marked by a competitive relationship between fluency (SR and MLR) and grammatical complexity (MLC). In task-based research, these findings are explained in terms of attentional deficits that occur during speech processing (Skehan, 2009). Accuracy and fluency, which are believed to be affected by encoder-linked attention, compete for attentional resources with complexity, which is affected by conceptualizer-linked attention. Thus, depending on where a speaker allocates attention either accuracy and fluency are raised, or complexity.

7.4.4 Concluding comments

The findings regarding the students’ L2 development partially confirm what the literature on students’ language development during SA has found, and partially contradict it. The participants’ overall increase in speed fluency and decrease in breakdown fluency and repetitions confirms previous research, which has shown that SA leads to greater oral fluency – and that this can be interpreted as a consequence of their informal language learning activities (Freed, 1995; Lennon, 1990; Llanes et al., 2012; Mora & Valls-Ferrer, 2012; Tonkyn, 2012; Towell et al., 1996; Wood, 2012). However, two problem areas were observed: empty pauses were significantly overused – suggesting that the participants’ proceduralization skills did not reach a native-like level and multi-word smallwords were underused – suggesting that their acquisition caused difficulty. Reformulations of content appeared to occupy two functions in speech. One was associated with repair as a result of the cognitive demands of spontaneous speech and one related to repair made to cater for the listener’s needs, reflecting discourse competence. It is questionable to what extent the latter can be considered a sub-construct of fluency – a “pure performance phenomenon” (Housen et al., 2012a). The discussion of fluency
confirms that the construct is multi-dimensional (Housen et al., 2012a) and that the subcategories reflect different cognitive underpinnings.

The findings for grammatical and lexical complexity were mixed. Only Alia’s speech became more complex over time and did not vary as a result of ‘task’ effects. This finding challenges most previous research which has shown the SA context to be beneficial for the acquisition of vocabulary and structures (Daller & Xu, 2007; Kinginger, 2009; Wood, 2012). However, it confirms findings made by researchers who have shown that different discourse types require different structures and different lexis (Ferrari, 2012; Mora & Valls-Ferrer, 2012; Pallotti, 2009). These researchers suggest that advanced learner skills are not simply reflected in higher scores but also in the ability to perform in a variety of different discourse modes – and that lower complexity scores can be reflective of more advanced language use.

Unlike what has been typically reported, two participants in this study increased their accuracy levels during SA while the third participant (Jana) already performed at a near-native like level. However, these findings were only made for the specific accuracy measure (i.e. the ratio of correct finite verb phrases), not the global measure (i.e. the ratio of error-free clauses). In this regard, the findings do replicate earlier findings that show that SA benefits specific grammatical features, particularly for students who had much room for improvement because they entered the SA programme with lower levels of accuracy (Isabelli-Garcia, 2006; Llanes, 2011; Mora & Valls-Ferrer, 2012). It appears that as proposed by Lambert and Kormos (2014) and Pallotti (2009), global measures of accuracy are too elusive to make judgements about progress. However, specific measures of accuracy, like the one used in this study, seem more sensitive to the assessment of progress, particularly for advanced level learners, as also shown in Tonkyn (2012).

The findings also support Dynamic Systems Theory (DST) regarding the variation inherent in language use. Language is not ‘acquired’ but constantly develops, and variability reveals a great deal about actual language use. Correlation coefficients were used to provide objective assessments of the levels of variability of each variable. The results demonstrated that some patterns were relatively smooth and linear, while other patterns were ever-changing and “jumpy” (van Dijk et al., 2011, p. 84). The measures used to capture oral fluency displayed a tendency for steady linear patterns for all individuals. The global accuracy measures revealed highly non-linear patterns for all individuals. The variability patterns for specific accuracy and complexity (grammatical and lexical) measures differed for the individuals and seemed to vary
with proficiency levels (the higher the more variation) and speech mode preferences (the use of different speech modes resulted in variation). Smoother trajectories and relatively clear patterns tend to reflect more advanced language learning processes while irregular patterns tend to reflect more experimental learning processes (i.e. making errors and experimenting with linguistic structures for accuracy or complexity) (van Dijk et al., 2011, p. 59).
8 Conclusions and Implications

8.1 Introduction

As mentioned in the introduction, a longitudinal study that investigates individual learners is best equipped to shed light on the idiosyncrasies and complexities of language learning motivation, L2 opportunities, and language development. This is a longitudinal study, combining multiple quantitative language assessments with in-depth qualitative analyses of three female German adolescents’ high school experiences in New Zealand. It addressed three main questions: what the learners’ motivation was to use and learn English during their sojourn; what type of language learning opportunities existed for the learners in different social settings; and what the nature of the learners’ second language (L2) development was during their SA period.

8.2 Motivation and language learning opportunities

The results of this thesis provide evidence that language learning motivation is dynamic and complex. It symbiotically develops in interaction with the L2 environment, and is affected by the identities the learners adopt in different social settings.

All learners were oriented towards language learning in both integrative and instrumental ways, but neither type proved sufficiently powerful as a determiner of motivation per se. The data showed that the level of persistence with learning the L2 was inseparable from the effort they put in reaching goals un-related to language. The three participants were not ‘just’ L2 learners in the study abroad (SA) context; they were also host daughters, athletes, friends, and high school students. Their language learning motivation was not generally low or high but gained or lost momentum contextually. Only settings and communities in which the participants’ desired to invest – as owners of particular social identities - and which in turn, allowed them to invest, generated high L2 motivation.

This study showed that language learning affordances emerged dynamically as a result of environmental contingencies, agency and motivation, and were influenced by the learners’ investments in different social identities. Opportunities in the environment had to be noticed and interpreted as affordances to facilitate L2 use.
These findings support those made by researchers of social motivation theories (Dörnyei & Ushioda, 2009, 2011a; Lantolf & Pavlenko, 2001; Norton & Toohey, 2005) who claim that motivation is not quantifiable, and that its relation to social and psychological elements are “unpredictable, nonlinear and always unique, since every person and context are unique” (Dörnyei & Ushioda, 2011a, p. 77). In particular, the findings demonstrate that only a “person-in-context relational view of motivation” (Ushioda, 2009) that takes into account the learners’ goals, motives, identities and contexts can adequately capture the nature of learning motivation.

8.2.1 Limitations and future research

Examining motivation from a socio-dynamic perspective is costly and challenging. One of the difficulties is the problem of limiting the focal point of interest. This thesis is limited in so far as it could not attend to all factors related to facets of motivation or language learning opportunities – although it is unclear where these examinations begin and end and doubtful that rigid boundaries exist. Its scope moreover did not allow for a sufficiently detailed description and analysis of the aspects that were considered relevant. Particularly aspects related to identity, which demonstrated great potential to account for “the challenges that students face in study abroad” (Kinginger, 2013) warrant more detailed investigation. Evidently, researchers must prioritise elements among the relevant influences and accept that “not every single project can cover all relevant factors” (Dörnyei & Ushioda, 2011a, p. 199). Despite the challenges, this study demonstrates that examining learners in context can generate insightful “stories of motivation” which help identify “specific learner variables that form optimal patterns with environmental factors” (Ushioda & Dörnyei, 2012, p. 406) Future research should consider motivation from this perspective and conduct research that is longitudinal and qualitative in nature to adequately capture these dynamics. Concepts of agency, affordances, and identity were suitable for investigating the learners’ motivation and language learning opportunities during SA but future studies could complement these aspects by introducing other concepts relevant to learning motivation such as L2 learning anxiety and willingness to communicate.

Another limitation of this study concerns the nature of learners’ self-reports. Although self-reports elicited rich data, introspection is limited insofar as this type of data only represents the learners’ own perceptions of the reality. An investigation of other relevant persons’ perspectives on the students’ sojourns (e.g. their host parents, their teachers) could give a more well-rounded picture of SA experiences.
8.3 L2 Development

Complexity, accuracy and fluency (CAF) measures were used to capture the individuals’ L2 development over several data collection points. I adopted an “organic approach” (Norris & Ortega, 2009, p. 556) by measuring and describing a range of CAF variables based on previous studies (e.g. Bulté & Housen, 2012; Housen et al., 2012a; Norris & Ortega, 2009). The results showed that the three individuals’ L2 development was non-linear and non-uniform, supporting the arguments made by dynamic systems theory (DST) (Verspoor et al., 2011; Larsen-Freeman, 2009).

The findings showed that the students’ fluency developed in a smoother and more linear way than accuracy and complexity (grammatical and lexical), pointing to more advanced proceduralization skills. This improvement was captured by a relatively steady increase in speed-related aspects of speech production (speech rate and mean length of run), and steady decrease in breakdown indices and repetitions. Accuracy and complexity showed considerable inter-individual differences – although the specific accuracy measure (i.e. the ratio of correct finite verb phrases) showed a linear increase for two participants. This study also examined the interactions of the global CAF constructs. The results demonstrate that trade-offs occur, supporting the claims made by Skehan (1992) and DST researchers (Verspoor et al., 2011). They also show that attentional resources are not allocated to the same domains for all individuals, supporting the importance of individual case study research.

When taking into account the overall trends in the learners’ L2 development, the results for fluency and accuracy confirm previous findings for SA learners (i.e. fluency improves, accuracy results are mixed) (Freed, 1995; Lennon, 1990; Llanes et al., 2012; Mora & Valls-Ferrer, 2012; Tonkyn, 2012; Towell et al., 1996; Wood, 2012). The results produced in this study for lexical diversity in particular, but also grammatical complexity, extend our current understanding of L2 learning during SA. The complexity constructs did not just remain ‘unmodified’ let alone increase, as typically observed (Daller & Xu, 2007; Foster & Tavakoli, 2009; Mora & Valls-Ferrer, 2012), but decreased for two of the three participants as a result of being abroad. Rather than reflecting language ‘regression’, these findings were interpreted in terms of changes in the nature of the learners’ communicative competence.

This study also suggests that there is a connection between the learners’ experiences abroad and their language use. Increased fluency levels were believed to be a consequence of the
learners’ informal language learning activities. Lexical diversity seemed to increase as a result of exposure to input that was highly context specific. Grammatical complexity seemed to reflect speech norms of the learner’s L2 communities. Accuracy appeared to increase as a result of controlled language activities, and only for those students with ‘room’ for improvement – as has been observed before (Freed, 1995; Isabelli-Garcia, 2004). However, arriving at definite conclusions regarding these relationships would be to ignore their complexity and it is questionable whether such conclusions are possible.

8.3.1 Limitations and future research

This study showed that not all CAF measures chosen for analysis proved equally sensitive for capturing the underlying cognitive correlates they purported to capture, nor were they equally valid for describing language progress. Problematic measures were ‘reformulations’ for fluency, ‘error-free clauses’ for global accuracy, and the grammatical complexity measures ‘subordination index’ and ‘mean length of clause’. These measures could not unambiguously distinguish between language progress, interview ‘task’ effects, and learner variables – which is considered a limitation of this study. Specific accuracy and complexity measures were more promising than multi-dimensional, global constructs. Future research must establish whether or not it is sensible to use complexity measures for advanced speakers’ oral performance given the non-complex nature of informal speech.

A major difficulty concerned data analysis processes for oral language elicited through semi-structured interviews. Numerous, often difficult decisions had to be made and not all decisions could be based on previously established criteria. Difficulties mainly concerned the choice of the nature of the oral language samples (i.e. regarding its length, position in the interview, and issues regarding topics); the establishment and application of valid and reliable AS-unit segmentation criteria, and several other choices regarding the treatment of specific language features, such as non-finite verb phrases, discourse markers, and direct speech reports. Interviews appear to be a valid way of eliciting spontaneous oral language, but future research must provide more detailed guidelines for the segmentation of spoken data into units, and the treatment of non-finite verb structures and other problematic features. This study provides a starting point for a principled way of segmenting data.

Future research should acknowledge that measuring proficiency levels on the basis of one or two data collection points is arbitrary and cannot provide a true picture of learners’ L2
development. Multiple data collection points and relatively uncontrolled elicitation methods can open up avenues for important insights into language use.

An investigation of advanced learner language must also include the “variation in language in general and particularly with non-pedagogic language use” (Kinginger, 2009, p. 82). A more richly developed baseline is required for comparisons of this kind. Data interpretations must be based on the learners’ communicative competence and their ability to use a variety of discourse modes. Improved L2 performance cannot simply be defined by more fluent, accurate or complex speech. The use of direct speech was critical in these advanced learners’ speech analysis and the effects of direct speech on speaker performance warrant a careful investigation in future studies.

8.4 Pedagogical implications

On a more practical level, my goal was to produce results that could afford interesting insights into the quality of high school exchange programmes both for educational institutions and for exchange students, parents and teachers.

This study shows that ‘agency based’ high school exchange programmes in New Zealand do not ‘guarantee’ students access to the use of the L2 as interactions inside and outside the classroom can be and often are carried out in the L1. The presence of other L1 members renders this context similar to an “’island’ student group context” and hence linguistically and culturally less challenging and stimulating than immersion programmes (Engle & Engle, 2003).

Although this study provides evidence that language learning opportunities still exist in this context, sustained language interactions with host nationals is only possible with significant learner investment. This study questions the linguistic and cultural value of high school exchange programmes that place students of the same L1 into the same high school. In particular, it questions some of the policies adopted by high schools – such as separating international students from domestic students in classes.

This study shows the need for high schools in New Zealand to make sojourns linguistically and culturally more meaningful to international students by challenging them to leave their comfort zone. Grouping international students into specially designed courses should be avoided for advanced level learners who require the linguistic challenges of the mainstream curriculum to
further their language development. The high school classrooms also provide a superior context for international students to make initial contacts with host nationals. It also seems pivotal for teachers to treat the international students like local students, so as not to deprive them of formal language learning activities which will be advantageous for their EFL classroom in Germany.

This study showed that the homestay, despite its potential advantages, can constitute a rather unattractive environment with few language and cultural learning opportunities. SA programmes need to ensure that homestays make an effort to take an active interest in the students, to treat them as “person[s] of consequence” (Tan & Kinginger, 2013, p. 156), and to provide them with opportunities for interaction.

The study also appeals to exchange programmes to make transparent to prospective international students whether or not they will study alongside other L1 speakers and what the consequences of such contexts could be. Given the ‘osmosis myth’, students often hold unrealistic expectations about potential linguistic gains during their SA and are, accordingly, easily disappointed with their L2 development.

Lastly, this research provides strong support for an approach that advises prospective international students to embrace challenges and treat them as potential opportunities for language or cultural learning. SA students should be “mindful” of their newcomer status and be “positively disposed toward opportunities to learn” (Tan & Kinginger, 2013, p. 156). To really progress their learning, SA students must make substantial and sustained efforts.
The term ‘study abroad’ in this study is used as an umbrella term for ‘student exchanges’ and sojourns experienced by international fee-paying students.

ii The early longitudinal case studies of L2 learners in the naturalistic context examined grammatical language features (Schumann, 1978; Schmidt, 1983; Ellis, 1984). All studies showed that the learners’ overall communication skills increased, particularly due to an increased repertoire of formulaics, but that their linguistic abilities developed slowly, or even stabilised. However, since these studies were concerned with naturalistic learners, not SA students, they will not be considered in this thesis.

iii From Verspoor and van Dijk’s (2011) interpretations of the relationship between variables, the rate of change between measurement points is not of key relevance when it comes to determining the type of relationship.

iv Every effort has been made to anonymise the collected data. In the examples from the students comments, people mentioned frequently in the students’ data were given a pseudonym (i.e. host family members, friends etc.). People who were mentioned infrequently were referred to as [name] or, alternatively, by their role in brackets [my trainer] (e.g. ‘my friend [name] didn’t show up’/ ‘[my trainer] is going to Europe for a competition’). Place names were referred to as [place] (e.g. ‘we went to [place] on Friday night’).

v These processes could not start earlier due to the international student manager’s heavy workload with the new arrivals.

vi In fact, I only every conversed with these three participants in English; before, during and after the interviews, as well as in emails and text messages.

vii No attempt was made to create a systematic distinction between discourse markers and pragmatic markers. This would have extended the scope of this thesis, and clear boundaries do not always exist.

viii The fluency and complexity measures were not double coded (i.e. no interraters were used). Data analysis processes, which required me to continuously re-check the coding, were considered reliable for these more objective CAF measures.

 ix Percentage changes were calculated as follows: (int. 6 value – int. 1 value)/ int. 1 value x 100.

x Götz measured pauses of any ‘silence’ and Riggenbach’s cut-off points were set at 0.5 seconds, while this study used a cut-off point of 0.3 seconds. The differences in measures means any comparison is not a one-to-one comparison of the data sets and for that reason any conclusions drawn from the comparisons may be weakened.

xi New Zealand English adult speakers have been shown to speak faster, overall, than other English native speakers (Robb & Gillon, 2007). It is likely, therefore, that Jana would be considered slow speakers of New Zealand English.

xii Not all smallwords that Hasselgren (2002) identified feature in the list of smallwords that she divided into the three hypothetical stages of acquisition.

xiii These findings only apply for the selected AS-units that were analysed in this study. It is possible that Jana made extensive use of direct speech reports during earlier interviews.

xiv When direct speech spanned over several clauses and included fresh starts with independent clauses, they were counted as independent clauses, not dependent clauses (see Methodology Chapter).

 xv These comparisons must be regarded with caution since measures differ with regard to cut-off points for empty pauses.

 xvi I don’t know’ does not feature in that list.

 xvii I am aware that ‘lshe gonna’ can be dialectal and an appropriate alternative to ‘I’m gonna’ in some communities (usually in the USA). The adolescent New Zealand NS who made grammar judgments in this study, however, considered the form erroneous for the local context.

 xviii These findings only apply for the AS-units that were analysed in this study. It is possible that the participants made more extensive use of direct speech reports during earlier interviews.
References


Badstübner, T., & Ecke, P. (2009). Student expectations, motivations, target language use, and perceived learning progress in a summer study abroad program in Germany. *Die Unterrichtspraxis / Teaching German, 42*(1), 41–49.


260


QSR International PTy L’Td. (2012). NVivo qualitative data analysis Software.


Appendix 1 – Ethics Approval

Office of the Vice-Chancellor
Research Integrity Unit

UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE

02-Oct-2013

MEMORANDUM TO:
Dist Prof Rederick Ellis
App Lang Studies & Linguistics

Re: Application for Ethics Approval (Our Ref, 010455)

The Committee considered your application for ethics approval for your project entitled The dynamic nature of the L2 development and the language learning motivation of German speaking exchange students in New Zealand: A case study.

Ethics approval was given for a period of three years.

The expiry date for this approval is 02-Oct-2016.

If the project changes significantly, you are required to submit a new application to UAHPEC for further consideration

In order that an up-to-date record can be maintained, you are requested to notify UAHPEC once your project is completed.

The Chair and the members of UAHPEC would be happy to discuss general matters relating to ethics approvals if you wish to do so. Contact should be made through the UAHPEC Ethics Administrators at: humanethics@auckland.ac.nz or the first instance.

All communication with the UAHPEC regarding this application should include this reference number: 010455.

(This is a computer generated letter. No signature required.)

UAHPEC Administrators
University of Auckland Human Participants Ethics Committee

cc, Head of Department / School, App Lang Studies & Linguistics
Ms Luzia Sauer
Mr Keith Montgomery

Additional Information:
1. Do not forget to fill in the ‘approval wording’ on the Participant Information Sheets and Consent Forms, giving the dates of approval and the reference number, before you send them out to your participants.

2. Should you need to make any changes to the project, write to the UAHPEC Administrators by email (humanethics@auckland.ac.nz) giving full details of the proposed changes including revised documentation.

267
3. At the end of three years, or if the project is completed before the expiry, please advise UAHPEC of its completion.

4. Should you require an extension, write to UAHPEC by email before the expiry date, giving full details along with revised documentation. An extension can be granted for up to three years, after which a new application must be submitted.

5. If you have obtained funding other than from UniServices, send a copy of this approval letter to the Manager - Funding Processes, UniA Research Office. For UniServices contracts, send a copy of the approval letter to the Contract Manager, UniServices.

6. Please note that UAHPEC may from time to time conduct audits of approved projects to ensure that the research has been carried out according to the approval that was given.
Appendix 2 – Interview Template

BACKGROUND QUESTIONS (only used in first interview)
- How old are you?
- What is your address in New Zealand?
- What is your place of origin?
- How long have you studied English for in Germany before you came to NZ?
- When did you come to NZ and when will you leave?
- What languages have you studied besides English?
- Have you lived outside of Germany before?
- Have you been on a language exchange before?
- Have you travelled outside of Germany?
- Does your family use a language other than German at home?

GENERAL GUIDING QUESTIONS

Warm-up
- How are you?
- Has anything special happened in the past month?

New Zealand culture
- Have you had a particularly interesting cultural experience?
- Is there anything in particular that you liked or didn’t like?

Social networks
- How is it going with your host family?
  - E.g. dinner, weekends, people
- How is it going at the high school?
- How is it going with your friends and family from Germany?
  - How often are you in touch with them?
- What are your expectations regarding these social networks in NZ?

Hobbies
- What do you normally do after school?

English
- How is it going with your English?
- Is learning English still important to you?...as important as before? Why? Why not?
  - Goals at the moment? … long-term goals?
    - Do you think your English will get better in the next month?
- Does spending time with anyone in particular help increase your English?

Identity
- How do you feel about yourself here in NZ?

Stories
- What was your best and your worst moment in the past month?

BLOG/MONTHLY REPORT FOLLOW-UP QUESTIONS
- What happened to the awkward situation with the boy that invited you to the prefect party?
Appendix 3 – Diary Entry Guidelines

Please write about...

- Any event that is related to your English learning and that you consider important (for example chatting with host parents and friends, talking to teachers, reading the news, doing homework).
  - What are your feelings and opinions about these events?
  - What is your perception of your language progress?
- Any other positive or negative events that have recently occurred and your feelings and opinions about them.
- The differences and similarities between Germany and New Zealand (for example traditions, foods, transportation systems, education systems and friendships).

Example of a diary entry – taken from Alia (Post #2)

First school days

Hello diary,

This week was my first week in school. On Monday the internationals, who are arrived a few days ago had orientation days. I'm an international student from Germany, so I had to go too. It was interesting but a little bit confusing at the first view. I was glad that I had five days with me although I didn't want to speak a lot German. But I was very good to speak the mother language for a little bit.

After school I went to my first track and field training at the stadium behind the college. It was very hot and I could not train a lot because my body was confused with everything!

On Wednesday we had our first real school day. We had to find our rooms, running in the right direction and a lot more.

It was luck that we had on Thursday a day of because Waitangi day. Jhuhh!! Realled! The weather was bad and it rained the whole day. So my friends and I went to the mall in order to watch the scary movie paranormal activity. Before the movie I felt very sick and thought I have to cont over. This was too much. But when I saw the movie (I just saw 10 minutes because then I hit me behind my jacket) I felt much better!
Appendix 4 – Monthly Report Guidelines

Please reflect on your linguistic and cultural experiences of the past month and describe with a minimum of 500 words in English (approximately one page of typed text):

1. …your thoughts about your own language-related performance so far
   - What was your most pleasant/unpleasant memory as a language learner in the past month, how do you feel about your linguistic progress, and why?
2. …your perceptions of the practices of the New Zealand culture
   - What is your attitude towards Kiwis, how different do you think they are compared to people from Germany? What are your experiences in adapting to the New Zealand culture?
Appendix 5 – Transcription Excerpt

(Taken from Chiara, interview 1)

*CHI:  (.54) &ehm (1.68) I think it would [% fv] be (.41) great [% indcl] to stay in contact with some of the Germans [% scl:co:nf].

*CHI:  (.76) but (.66) &ehm also to get (.41) some international students like (.68) from Brazil or [% indscl].

*CHI:  (.30) like they're [% fv] all from all over the world here [% indcl].

*CHI:  (1.65) &ehm (.77) of course find some Kiwis [% indscl].

*CHI:  but I think that's [% fv] hard [% indcl] to stay with them in contact [% scl:co:nf].

*CHI:  because it's [% fv] like the other end of the world [% indcl].

*CHI:  (.80) so I would [% fv] like to have Kiwi friends [% indcl] because &ehm (1.57) yeah I want [% fv] to (.99) get to know [% scl:adv] how they live [% fv] [% scl:co].

*CHI:  (.43) because my homestay sisters (.72) are [% fv] away all the time [% indcl].

*CHI:  and I think they don't [% fv] like me [% indcl].

*CHI:  (.63) and &ehm (3.08) yeah the Germans also don't [% fv] know it [% indcl].

*CHI:  and (1.80) yeah I would [% fv] like to (4.82) just stay with them [% indcl].

*CHI:  and then I think they're [% fv] (.68) <much relax (.40) so> [/\] (1.34) more relaxed than (.51) German students [% indcl].
## Appendix 6 – Error Coding Principles

Principles applied for the global accuracy measure error-free clauses:

<table>
<thead>
<tr>
<th>Error type</th>
<th>Procedures and examples from my data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-collocational lexical choices</td>
<td>Were treated as errors. (e.g. ‘or my mom (.39) my mom’s always afraid or (.44) like get (.27) gets panic as well when she hears that and (1.26) yeah’ &lt; feels/gets into).</td>
</tr>
<tr>
<td>Lexical choice is inaccurate in the linguistic context</td>
<td>Were treated as errors. (e.g. ‘but ehm she’s so motivated that she ehm (.43) pushes me down (.30) with her motivation’ &lt; gets me down)</td>
</tr>
<tr>
<td>Lexical choice definitely expresses a meaning other than attempted</td>
<td>Were treated as errors. (e.g. ‘I can go in the next class’ &lt; year)</td>
</tr>
<tr>
<td>Lexical choice constitutes a non-TL word coinage or substitution</td>
<td>Were treated as errors. (e.g. ‘and then they drove like into a (2.20) ending street’ &lt; dead end)</td>
</tr>
<tr>
<td>Subject omission</td>
<td>Acceptable with conventional expressions (e.g. ‘gotta go’), dummy subjects (e.g. ‘was really hot there’), and anaphoric deletion.</td>
</tr>
<tr>
<td>Something/anything</td>
<td>Both accepted in questions and offers; only ‘something’ considered correct in independent positive statements (e.g. ‘something is wrong’). Only ‘anything’ considered accurate in negative statements (e.g. ‘I don’t know anything about it’).</td>
</tr>
<tr>
<td>Preposition following a German word</td>
<td>When German words were followed by a preposition that cannot be used in combination with the translated English word, this preposition was coded as erroneous (e.g. ‘but then I ehm (0.88) zweifeln [doubt] (0.84) ehm (0.36) on my (0.88) myself’ &lt; -)</td>
</tr>
<tr>
<td>Direct speech vs. indirect speech reports</td>
<td>Any tense use in direct speech reports was considered accurate. If the speech was a mixture between reported and indirect speech, I accepted the unchanged tense form (e.g. ‘and she was like :: oh no she has to go to school.’)</td>
</tr>
<tr>
<td>Conversational historical present</td>
<td>Considerable investigations would have needed to be pursued to guarantee that present simple was used appropriately when describing past events (i.e. the use of conversational historical present) and not a learner error. This lay beyond the scope of this study. I thus refrained from considering present tense use for a recounting of past events as accurate in this study.</td>
</tr>
<tr>
<td>Quotative verbs of direct speech reports</td>
<td>Formal reporting verbs (e.g. ‘say’ or ‘ask’) and colloquial ones (e.g. ‘be like’, ‘go’, ‘be’, ‘zero verb’) were all considered accurate.</td>
</tr>
<tr>
<td>Verb tense backshift in indirect speech reports</td>
<td>In indirect speech with past tense quotatives, failure to backshift verbs and change deictic elements was considered inaccurate with the following exceptions: 1) General statements (e.g. ‘the sun rises in the east’) and statements that are still true at the time of reporting (e.g. ‘my brother studies at Uni Leipzig’). 2) Failure to change spatiotemporal elements and pronouns to the linguistic context of the speech complement was not considered incorrect</td>
</tr>
</tbody>
</table>
when the learners did so deliberately to adjust to the needs of the interviewer.
(e.g. ’and then ehm yesterday I wa I was (.69) ehm (1.29) I was I or I I was afraid that he’s angry with me when I say I want to participate in (.53) today in one week (.63) at the so Nation Day from [College] (.87) and yeah’).

:: clause separator

Principles applied for the specific accuracy measure (correct finite verb phrases)

<table>
<thead>
<tr>
<th>Intention of speaker in the context</th>
<th>Present simple &amp; Present continuous</th>
<th>Present Perfect &amp; Past Perfect</th>
<th>Past Simple &amp; Past Continuous</th>
<th>Future tenses: Going to + infinitive, present-continuous, shall/will, be to + infinitive, present simple (for scheduled events outside our control), be about to + infinitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>To refer to a present event</td>
<td>√</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>To refer to a future event</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>√</td>
</tr>
<tr>
<td>To refer to a past event or an event that began in the past but continuous to the now.</td>
<td>x</td>
<td>√</td>
<td>√</td>
<td>x</td>
</tr>
</tbody>
</table>

√: considered accurate, x: considered inaccurate