A CORPUS-BASED APPROACH TO WRITING IN GERMAN AS A FOREIGN LANGUAGE IN THE SOUTH AFRICAN TERTIARY CONTEXT

A thesis submitted in fulfilment of the requirements for the degree of

MASTER OF ARTS

OF

RHODES UNIVERSITY

BY

GWYNDOLEN JEANIE ORTNER

16 NOVEMBER 2015

PREPARED FOR:
German Studies Section, School of Languages
Rhodes University
Grahamstown

SUPERVISOR:
DR UNDINE S. WEBER

CO-SUPERVISOR:
DR SYLVIA JAWORSKA

ABSTRACT

German Studies students at Rhodes University have normally never studied the German language before enrolling for the first-year course and face the challenge of a fairly rapid linguistic advancement in order to engage meaningfully with the literatures and cultures of German-speaking countries. This thesis investigates the process of teaching and learning to write in a more academic way in German as a foreign language at Rhodes University, using corpus linguistic tools for both analysis and instruction. The past 20 years have shown a shift from traditional teaching methods resting on notions of an underlying prescriptive grammar, to teaching based on insights from real-life language data (Gabrielatos, 2005; Krummes & Ensslin, 2012; Sinclair, 1997) and applications of corpora to teaching and learning have shown to be highly successful in many European contexts (Aijmer, 2010; Johns, 1991; Granger, *et al.*, 2002; Varley, 2009). In the South African context however, this is a relatively new concept with few publications on the application of corpus linguistics to language teaching (Van Rooy, 2008), and one which does not seem to have reached its full potential. A writing course was instituted whose aim was two-fold: 1. to teach learners "every-day academic" German words (TAG words) and phrases (collocations) based on German mother-tongue corpus evidence; 2. to have learners write short assignments in German at regular intervals (Homstad & Thorson, 1996; Estes, *et al.*, 1998); both aims with the overarching objective to improve the students' academic register in German. After the writing course, 80% of the participants perceived that their writing had improved and specifically attributed this to the corpus-based instruction received during the writing course, and regular writing in German. Quantitative data (from the learner corpus created) shows a marked improvement in the use of the collocations taught. Moreover, participants (weaker students in particular) also found that their writing in English had improved as a result of the various exercises they had to complete as part of our German writing course.

Keywords: Academic literacy, everyday academic writing, corpus linguistics, learner corpora, German studies, case study, writing-to-learn, transferrable skills/procedure

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CHERTL	Centre for Higher Education Research Teaching and Learning
DaF	*Deutsch als Fremdsprache* (German as a Foreign Language)
FALKO L1	*Fehlerannotiertes Lernerkorpus* (First language control corpus)
FLT	Foreign Language Teaching
GFL	German as a Foreign Language
LL	Log-Likelihood
NF	Normalised Frequency
RUDaF	Rhodes University Deutsch als Fremdsprache
SLA	Second Language Acquisition
TAG word	Typical (everyday) Academic German Word
WHiG	What's Hard in German

PREFACE

Acknowledgements

I am indebted to a number of people who have been instrumental in seeing this master's thesis through to fruition. I have greatly appreciated the interest shown, the conversations had, and the support given.

My especial thanks go to:

Dr Undine S. Weber, my super supervisor, whose enthusiasm for the project and weekly conversation have been a constant support and driving force behind this undertaking; my co-supervisor, Dr Sylvia Jaworska whose knowledge and expertise in the field are invaluable; the German Studies Section of the School of Languages, a home away from home; the participants, for their valuable contributions and willingness to explore new methods; Prof Mark de Vos and Kristin van der Merwe for their keen interest and 'idea sparking' conversation; The Department of English Language and Linguistics at Rhodes University for always having an open door to their tea-room; Dr Carol Thomson, co-ordinator of the Writing Intensive Project for her kind words of encouragement; Prof Dr Anke Lüdeling who allowed access to the Falko-L1 corpus of mother-tongue German writing; and lastly to my family and friends who have spent valuable time trying to understand and contribute meaningfully this research. You are all simply the best.

The financial assistance from the Rhodes University Postgraduate Prestigious Scholarship, the Ada and Bertie Levenstein Bursary and the DAAD In-Country Scholarship, towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to Rhodes University, or the donors.

INTRODUCTION

Every year thousands of students entering university for the first time make the choice to learn a new language. There are many reasons why one may choose to do so: linguistic or cultural interest, prospective economic benefit, a family connection or a simply practical need, such as reading for scientific or sociological research (Gethin & Gunnermark, 1997; Masgoret & Gardner, 2003). Learning a new language presents a challenge, as learners must make the leap in three years from re-learning the pronunciation of the alphabet to producing short academic texts. Learning German has its own specific challenges, including grammatical gender and case, both of which are unfamiliar concepts to an English mother-tongue[1] speaker.

German Studies at Rhodes University is a multi-faceted field, encompassing not only the teaching of grammar and vocabulary, but also literature, cultural studies (including the histories and some of the current political debates of the German-speaking countries), and translation (Rhodes University, 2013; German Studies, 2015). Traditionally, due to the fact that German Studies at Rhodes University mostly has *ab initio* students, there has been little focus placed on what would be deemed 'academic literacy' skills in German, and a specific focus on formal academic writing skills has not always been possible due to time and curricular constraints.

This mirrors a general trend in foreign language classrooms where "Foreign language instructors, believing that writing in a foreign language may prove too frustrating and difficult for students, have often been reluctant to incorporate writing-to-learn activities into their pedagogy" (Homstad & Thorson, 1996, p. i). Studies which have introduced regular writing activities in the foreign language classroom have provided much qualitative evidence of an improvement in writing fluency and confidence (Homstad & Thorson, 1996; Estes, *et al.*, 1998). However, such studies often struggle to provide quantitative evidence to support their qualitative perceptions, due to methodological barriers (Homstad & Thorson, 1996).

Related to the assessment of writing improvement is the teaching of writing skills. Corpus linguistics, a relatively new branch of language enquiry and methodology, has argued for language pedagogy which is based on actual language usage, rather than on mother-tongue speaker intuitions

[1] This thesis makes use of the terms "mother-tongue" or "first-language" as opposed to "native" speakers when referring to one's first language, and "non-mother-tongue" or "second language" speakers as opposed to "non-native speakers" when referring to second language speakers. This is as I recognise that in the South African context, and the larger ex-colonial context, the word "native" has particular pejorative connotations, and can be perceived as insensitive.

(Gabrielatos, 2005, p.5; Sinclair, 1997, p. 32-34). This has led to various applications for language research, which impacts on language pedagogy. For example, in 2011, research conducted in the UK for the project "What's Hard in German?" (WHiG) employed corpus linguistic methods in order to identify "linguistic structures that pose a specific difficulty for the acquisition of German as a foreign language" (Krummes & Ensslin, 2012; Krummes & Ensslin, 2014; Jaworska, *et al.*, 2015). The researchers posed that foreign language students avoid difficult structures, and thus difficult structures will be underrepresented in their writing. Not only did the authors identify how UK foreign language learners' writing differed from mother-tongue writers of German in Germany, they also created useful materials for teaching academic collocations for German writing based on real-life language data (Krummes & Ensslin, 2012).

Thus, it can be said that "Electronic language corpora and their attendant computer software are proving increasingly influential in language teaching as sources of language descriptions and pedagogical materials" (Gabrielatos, 2005, p. 1). This trend can be observed in the increasing number of publications concerning corpora and language teaching in the past 20 years. In fact, by 2010 there had been "at least twenty-five authored or edited volumes" published on the topic of teaching and language corpora, and numerous articles (see McEnery & Xiao, 2010, for a comprehensive reading list). The various uses of corpora and their relation to the many aspects of teaching and learning include using corpora to develop materials such as dictionaries and word lists to be used indirectly in the language classroom, as well as the direct use of corpora in the classroom to provide real life data for learning (either through computer based, or paper-based exercises) (Leech, 1997).

There is specifically a growing interest in learner corpus research as learner corpora are both practically and theoretically valuable to the fields of second language acquisition and foreign language teaching (Granger, *et al.*, 2002; Granger, 2004). Two main approaches to the analysis of learner corpus research have arisen, namely: Error Analysis (EA) – which seeks to identify, quantify and analyse learner errors, so that areas where learners need specific help and attention can be identified through the analysis of their writing, in different contexts and at different levels of competence (Corder, 1981; Zeldes, *et al.*, 2003, López, 2009) – and Contrastive Interlanguage Analysis (CIA) – which seeks to compare learner language with language produced by mother-tongue speakers, to identify overuse, underuse or misuse of formulaic phrases so that insights into the process of language learning, and learner interlanguage, can be formed (Granger, *et al.*, 2002; Krummes & Ensslin, 2014). Both of these approaches have contributed to the description of learner language, and have been used to improve teaching and learning materials, with very recent applications to be found relating to the teaching and learning of German which has traditionally relied on "anecdotal evidence" and has not been as corpus-informed as English teaching and learning materials (Krummes & Ensslin, 2014).

Thus inspired by recent research and developments in the field of corpus linguistics as it relates to foreign language learning, this thesis explores the process of teaching and learning to write in a more academic way in German as a foreign language[2] at Rhodes University, using corpus linguistic tools. The materials used are based on recent research findings for German (Krummes & Ensslin, 2014). Coupled with this new approach to language instruction based on real life data, an intensive writing programme was instituted in order to give participants an opportunity to "write to learn". The submissions from this intensive writing programme form the base of a longitudinal learner corpus which spans three years of data collection both before and during a structured corpus-based everyday academic writing course in German Studies. Where other researchers have struggled to quantitatively define the impact of their intensive writing courses on student writing (Homstad & Thorson, 1996; Estes, *et al.*, 1998), this research methodology seeks to harness the computational power afforded by the creation of learner corpora in order to compare learner writing before and after the writing course. This thesis therefore draws on the expanding field of applied corpus linguistic methods for both the instruction and analysis of an everyday academic German writing course.

This research came about as a consequence of a pilot study that was undertaken at the German Studies Section of the School of Languages at Rhodes University in 2013 (Ortner, 2013), seeking to introduce writing tasks which aimed to asses our students grasp of *alltägliche Wissenschaftssprache* (everyday academic language) (Ehlich, 1995). This study was undertaken with the third-year German Studies students in 2013, where a focus was placed on producing short 300 word descriptive and argumentative texts. No extra lessons were given to the participants during the pilot study, and the voluntary writing tasks fell outside of their normal academic schedule for the year. Data collected was used to form the beginnings of the learner corpus, RUDaF (Rhodes University *Deutsch als Fremdsprache*) and used comparatively with data drawn from a corpus of mother-tongue academic German writing, WissDe (Jaworska, 2011), in order to examine the use of specific lexical items in context.

This 2013 study found that the participants were not yet fully able to produce everyday academic German texts (Ortner, 2013), showing an ability to write only "simple connected text on topics that are familiar or of personal interest" (Council of Europe, 2012). In many cases, as was shown through a corpus-based examination of both morpho-syntactical errors, as well as lexical errors, students often produced work which resembled "English dressed as German" (Jaworska, 2011, p.3). The many lexical

[2] As German is not an official language of South Africa, this thesis will refer to the teaching and learning of German in the South African tertiary context as teaching and learning German as a foreign language, rather than a second or additional language.

errors produced by participants within their short texts were attributed to a reliance on English-based writing procedures employed by participants in writing their assignments, as shown in the results of the questionnaire which formed a part of the methodology for the pilot study (Ortner, 2013).

Participants showed a propensity to rely on unreliable translation software, and displayed a tendency to produce many errors based on ill-informed word choices from a range of dictionary options (Ortner, 2013). Furthermore, the participants did not appear to be able to use every-day German words and collocations metaphorically as is characteristic of academic writing. It was further revealed that very little engagement in reading of any German text outside of the classroom was taking place amongst the participants (Ortner, 2013), and this was seen as a particular obstacle to learning, as individual reading engagement plays a highly important role in developing vocabulary, and determining academic success (Cunningham & Stanovich, 1998).

While participants found the tasks challenging, through the process of writing short texts, the participants' writing became progressively better in terms of the types of errors made, and the length of the texts submitted, as the weeks progressed (Ortner, 2013, p. 26). Improvement was seen in every case (Ortner, 2013, p. 26). However, whilst there was improvement overall, specific mistakes were repeated by participants over the 7 weeks, and written feedback given in the assignments did not appear to be internalised (Ortner, 2013).

The results from this study highlighted the need for better resources which students could use when searching for lexical items to use in writing pieces, as well as specific instruction in everyday academic German writing at Rhodes University, with a focus on collocational knowledge, as well as continuous opportunities for students to practise writing. The 2013 study informed the way in which the methodology for this thesis was developed, and the research goals outlined below find their roots in the insights into student writing developed in 2013.

This thesis expands on the findings of the 2013 pilot study outlined above (Ortner, 2013), by creating and instituting a corpus-based writing course in 2014 and 2015, which makes use of the findings presented by authors of the WHiG and FALKO projects (Jaworska, *et al.*, 2015; Krummes & Ensslin, 2012; Krummes & Ensslin, 2014), with the second, third and Honours (i.e. fourth) year level German Studies students in order to improve their everyday academic German writing skills. The writing pieces produced both before and during the writing course collectively contribute to the ongoing German learner corpus, RUDaF, which uniquely portrays the specific learner writing of South African students at Rhodes University who have only begun learning German at the tertiary level. When compared to research undertaken for the study of English, there are few German learner corpora which have been created, and these differ in terms of the learners' first language, the skills level, mode of production,

text genre, compilation purpose, retrieval method, and depth of annotation (Goossens & Granger, 2013; Krummes & Ensslin, 2014). There are very few learner corpora created in South Africa, and none for German as a foreign language in South Africa[3]. There is thus a wide research gap pertaining not only to German, but all languages taught and learnt in South Africa. This thesis aims to show how learner corpora may aid the teaching and learning of languages – both foreign and indigenous – in the South African tertiary context, using German as a foreign language as a specific case study. This thesis aims to achieve a number of specific goals relating to the creation and analysis of a South African learner corpus of German, and the creation, institution and analysis of an everyday academic writing course for German Studies.

Research Goals

Through a qualitative analysis of questionnaires and interviews conducted both before and after the institution of the writing course, I identify the writing strategies and methods employed by our participants when writing in German. For example, whether they write directly in German (that is to say that they attempt to think in German) or whether they write out the assignment in English first and then translate it. In addition to this, I explore participants attitudes towards writing (in the general tertiary context), and their attitudes towards writing in German Studies specifically. Insights from the pre-writing course questionnaire informed the creation of the writing course. Insights from the post-writing course questionnaire inform the qualitative evaluation of the writing course and the evaluation of the successfulness of a corpus-based approach to teaching writing in our unique context.

Currently, our students in German Studies 2 should be writing at A2 level, German Studies 3 students at B1 level, and German Studies Honours students at B2 level (levels according to the Common European Framework of Reference for Languages (Council of Europe, 2012)). Together with establishing an ongoing learner corpus, I specifically assess the writing submitted for RUDaF in order to provide metadata for the corpus and confirm the level of writing which are learners are at. These assessments are made according to rubric specifications set out by the Goethe Institute (Goethe Institut, 2015).

In order to provide a deeper quantitative analysis of the effect of the corpus-based writing instruction on learner writing in German Studies, I further analyse our learner writing using insights from contrastive interlanguage analysis (CIA). Corpus tools allow for an analysis of learner writing in German

[3] To date, in all internet searches and through conversations with colleagues, the only learner corpus to which I have found reference in South Africa is the Tswana Learner English Corpus (Van Rooy & Schäfer, 2002; Goossens & Granger, 2013).

at Rhodes University, by examining the use of key everyday academic words and phrases as used by participants before the writing course and during the writing course. This analysis is furthered by comparing learner writing to mother-tongue German writing in the corpus FALKO L1. Insights from this analysis allow for recommendations for the creation of GFL teaching materials in our South African context.

As the introduction of any new teaching and learning tool affects both the teacher and the learner, I critically self-evaluate the process of instituting corpus methods in the classroom from my perspective as a teacher; and I assess the attitudes of the students towards the use of corpus tools to learn new academic German vocabulary and phrases. Lastly, I evaluate the writing course as a whole in terms of its usefulness and effectiveness, in order to make recommendations for future writing courses in the German Studies Section of the School of Languages at Rhodes University.

This thesis is organised into seven chapters. Chapter 1 outlines the context of the research as it relates to the fields of literacy, academic literacy, German Studies and the South African context, placing this research within the broader societal context. The historical evolution of corpus linguistics, and its convergence with the fields of teaching and learning is divulged in Chapter 2 which provides the theoretical context of the research. Following on from this, Chapter 3 provides a review of the literature covering the topics of everyday academic German, formulaic language and learner corpus research as they relate to the description and teaching of everyday academic German. Chapter 3 also provides evidence of current methods of teaching using language corpora, the challenges for teachers in implementing these tools, and the attitudes of students towards using corpus tools. The review of the literature reveals gaps in current knowledge pertaining to corpus-based teaching of German and sets the scene for the methodology which is expanded on in Chapter 4.

Chapter 4 outlines the methodological approach taken in addressing the above research goals. This includes an overview of the participants and their language backgrounds as well as a review of the procedure for designing a corpus-based writing course for our particular setting, creating and assessing the writing tasks, creating the RUDaF learner corpus and designing the questionnaires and interviews. Quantitative results are presented and discussed in Chapter 5, which explores the use of everyday academic German words within the RUDaF learner corpus (before and during the writing course) and draws further comparisons with the FALKO L1 corpus of mother-tongue German writing. Following on from the quantitative analysis, Chapter 6 reviews the topics addressed in the questionnaires and interviews conducted, and addresses participant motivation, and attitudes towards writing and corpus tools, providing qualitative insights into learner writing in German Studies at Rhodes University and the growth of participants during the writing course. Finally, in Chapter 7 a summary of the essential tenets of this thesis is made. Challenges and successes of the project are

addressed in order to illuminate the way for further researchers. The main findings of the research are addressed, and the qualitative and quantitative results drawn together to form a holistic view of writing in German Studies at Rhodes University, and how this was positively affected by the institution of a corpus-based everyday academic writing course. The long term effects of the research, as well as the implications for further research are discussed.

In summary, data-driven teaching and learning methods, with a focus on collocational knowledge, have shown to be highly successful in many European contexts (Johns, 1991; Varley, 2009; Gabrielatos, 2005; Granger, *et al.*, 2002; Aijmer, 2010). The past 20 years have shown a shift from traditional teaching methods resting on notions of an underlying prescriptive grammar, to teaching using real-life language data, as will be exemplified in Chapter 2. In the South African context however, this is a relatively new concept with few publications on the application of corpus-linguistics to teaching and learning to be found (Van Rooy, 2008), and one which does not seem to have been utilized or reached its full potential. This case study of the implementation of a corpus-based approach to teaching a course in everyday academic literacy in one's foreign language (in this case German), coupled with an intensive writing programme and therefrom the creation of a learner corpus, provides valuable insights which are transferrable to all second and foreign languages taught in South Africa. This research may provide valuable insights to any person who also wishes to implement a corpus-based writing course, and thus may encounter many of the same obstacles and challenges.

CHAPTER 1. SOCIETAL CONTEXT: ACADEMIC LITERACY AND GERMAN STUDIES IN SOUTH AFRICA

This chapter seeks to place the research undertaken for this thesis within the South African context. Beginning with an overview of academic literacy, I outline the challenges involved in developing academic literacy, particularly for a second-language learner. As German is the specific second language which this thesis explores, writing in German as a foreign language and the semantic category of everyday academic German is focussed upon. The corpus-based everyday academic German writing course which is central to this thesis is placed in the context of discipline-specific writing intensive courses which are currently being instituted in South Africa to improve academic literacy (Thomson, 2013). As corpus linguistics informs the way in which better teaching and learning materials can be created, the role of corpus linguistics and the creation of learner corpora is briefly considered in light of the dearth of learner data in South Africa. South Africa is a multilingual society with 11 official languages all striving to seek recognition and status. Given this context, the teaching and promotion of German as a foreign language in South Africa is considered in terms of educational policy. However, German as a Foreign Language (GFL) has seen a worldwide decline in student numbers (Reinbothe, 2006), thus the role of GFL in the South African tertiary context is outlined, and an argument for the promotion of German as a foreign language in the South African context is presented.

Reading and writing are skills which are often taken for granted. Once the basic alphabetic code has been "cracked" (Wolf, 2008), we continue to hone our skills with every act of 'literacy' in which we engage. By the time a student reaches the tertiary education sector, they have progressed over 12 years of education from interaction with basic texts such as 'the cat sat on the mat' to more challenging academic constructions with complex syntactical structures and discipline-specific vocabulary (with the need to understand and decode implicit and nuanced meaning) such as: 'According to observations by Panthera *et al.* (2001, p. 62) the common domestic housecat, *Felis catus* or *Felis Sylvestris Catus*, has been recorded to sit on many coarse woven or plaited fibrous materials used for covering floors or furniture'[4]. Producing academic texts requires a certain skills set on the part of the reader or writer. This skills set is termed academic literacy.

Torgesen, *et al.* (2007, p. 3) define academic literacy in terms of a type of reading proficiency which is required in order to "construct the meaning of content-area texts and literature [...] and [...] the ability

[4] The example and study have been fabricated.

to make inferences from text, to learn new vocabulary from context, to link ideas across texts, and to identify and summarize the most important ideas or content within a text". However, understanding an academic text is not sufficient; one must also be able to reproduce work within the given academic discourse (Trupe, 2011). In order to reproduce work within the academic discourse effectively, one must have an understanding of 'academic culture', that is to say an understanding of the "values, etiquettes and assumptions" of the academic world (Revisionspiral, 2011; Ortner, 2012, p. 4). This idea is reflected in Boughey's writing on the topic of academic literacy (2000, p. 283): "The concept called literacy involves knowing how to speak and act in a discourse. Academic literacy involves knowing how to speak and act in academic discourses".

The learner who chooses to begin learning a foreign language (L2) at the tertiary education level must not only make great advances in grammar learning, but also learn how to 'speak and act' on paper, within the academic discourses which are specific to the target language (Swales & Feak, 1994; Campbell, 1998; Jaworska, 2011). Whilst in German Studies an effort is made to teach students about German culture, a study of German academic writing culture is not included in the syllabus. Thus, becoming academically literate, or acquiring academic 'cultural norms' is one of the most important facets of one's university education; however, it is often taken for granted that students will acquire this skill simply by virtue of being in attendance at a university, and formal instruction into the values, norms and etiquettes surrounding academic discourse and the forms it may take is seldom given in any subject (Boughey, 2005; Boughey, 2012).

The challenge of developing writing skills, and in our case specifically academic writing skills for foreign language students of German, is not unique to the German department at Rhodes University South Africa. Writing across all disciplines in the tertiary sector is an important aspect of academic success, and is often not explicitly taught. Writing in German has its own particular set of obstacles and conventions[5]. As Laurillard states, "It is a universal feature of academic discourse, no matter what the discipline, that it represents ideas using a particular formalism. Most subjects use a special language, often using everyday words but with specific connotations and precise definitions that are importantly different from the normal meanings" (Laurillard, 1997, p. 172). German is no exception, and for foreign language learners of German, these phrases are often not directly translatable from the student's

[5] For a comprehensive introduction to the teaching of Deutsch als Fremdsprache, see Helbig/Gotze/Henrici/Krumm (2010), Rosch (2011), Huneke & Steinig (2009), Bausch/Christ/Hüllen/Krumm (2003).

mother-tongue (in our case often, but not exclusively English), resulting in peculiar non-idiomatic phrases which may be characteristic of learner texts (Krummes & Ensslin, 2012).

As with every language, academic German writing has its own particular set of norms, with specific nuances and turns of phrase, allowing the reader to note that one is communicating within the set discourse of academic writing (Fairclough, 2001)[6]. This empowers the writer, as they draw on the power held within the discourses of formalized knowledge, and conversely a writer is disempowered when they fail to use these forms correctly (Fairclough, 2001). The pursuit of most language learners is to achieve as natural and mother-tongue-like language use as possible. Acquiring academic literacy in German involves becoming familiar with many set phrases and items of vocabulary which may appear in everyday language, but which have specific set uses in formal writing.

We may call this 'everyday academic language', or as coined by Ehlich (1995) "*alltägliche Wissenschaftssprache". Alltägliche Wissenschaftssprache* is a semantic category used specifically to describe, and thereby analyse, the use of everyday, ordinary German vocabulary for the purpose of academic writing (Ehlich, 1995; Steinhof, 2007; Skrandies, 2011). The vocabulary which is typical of this semantic category is characteristic of both oral and written communication across most academic disciplines, as asserted by Ehlich (1995) Steinhof (2007) and Skrandies (2011). This entails an often metaphorical use of vocabulary items (Fandrych, *et al*., 2012). That is to say that the vocabulary has undergone a "functional change" (Skrandies, 2011). Examples of this include noun-verb collocations such as *einer Frage nachgehen* (to pursue a question) or *ein Problem beleuchten* (to 'illuminate' a problem, or rather to gain a deeper understanding of a problem).

Academic writing is usually one of the last skills that one acquires in learning a language, if at all, and as such, can incorporate some of the most challenging higher order tasks in any language (Van Schalkwyk, *et al.*, 2009), which may include aspects such as the reconceptualisation of writing and the negotiation of voices within a text (Boughey, 2000, p. 6). It is particularly difficult for non-mother-tongue speakers of German to use *alltägliche Wissenschaftssprache* correctly, due to its frequent metaphorical usage (Jaworska, 2011). This means that students who have a good competency in German may experience problems in writing where lexical choices are not effective, and writing may then appear as an exact translation of the mother tongue: or "English dressed as German" (Jaworska, 2011, p. 3).

[6] The following section is largely based on research undertaken in 2013 (Ortner, 2013).

There is currently a country-wide effort to institute writing-intensive courses within disciplines to foster and improve discipline specific academic literacy (Thomson, 2013). This research, which seeks to introduce a course in everyday academic German where students are tasked to write fortnightly paragraphs, is thus in line with the current drive at Rhodes University (and all tertiary institutions in South Africa; at Rhodes University instituted through CHERTL[7]), to design 'writing intensive' courses where students are tasked to write more, more often, within their disciplines (Thomson, 2013). This approach is one which seeks to give direct instruction of the norms of academic writing, and ensure that plentiful amounts of constructive formative feedback is given to students throughout the various stages of the writing process. Writing intensive courses in other foreign-language German classrooms have produced good, qualitatively-informed results which indicate an increase in learner writing fluency and confidence (Homstad & Thorson, 1996; Estes, *et al.*, 1998).

The use of corpus linguistic tools has proved to be a valuable asset to researchers in the field of second language acquisition (SLA) and foreign language teaching (FLT) (see Granger, *et al.*, 2002, Tribble, 2002; Granger, 2004; Gabrielatos, 2005; Yoon, 2008; Jaworska, 2011). Corpus linguistics is a field of language study which seeks to examine large bodies of real-life language data which are typically stored in an electronic corpus (McEnery & Hardie, 2012(a), p. 2; Baker, 2006, 2010). Corpus linguistics can be defined as a set of procedures, tools and concepts employed to study different areas of language, rather than as a study of any one particular aspect of language in isolation (McEnery & Hardie, 2012(a), p. 1), yielding a large amount of data which is authentic, and which can be rapidly and reliably processed by computer software designed specifically for this purpose (Baker, 2010; McEnery & Hardie, 2012(a)). Corpora have been used not only to develop curricula based on the examination of mother tongue corpora (Gabrielatos, 2005; Mindt, 1997), but also to examine learner corpora and study the intricacies of learner language so as to identify areas where learners need specific help (Granger, *et al.*, 2002, Granger, 2002; Meunier, 2002). Corpora have become importantly linked with language teaching and learning, and therefore with the language learner. As Krummes and Ensslin state, there are three different kinds of relationships which language learners can have with corpora: "they can be the donors of corpus data towards the creation of learner corpora, they can use corpora themselves, and they can use corpus-attested language learning materials" (2012, p. 112).

This research into academic literacy, and the bond between language teaching and corpus linguistic tools is particularly poignant as currently, South Africa finds itself in the midst of a "literacy crisis" which affects every level of education and learning (Fleisch, 2008; Howie, *et al.*, 2006; Howie, *et al.*,

[7] Centre for Higher Education Research Teaching and Learning (CHERTL).

2011). Academic literacy (as discussed above), and the specific challenges faced by students trying to acquire this skill at the tertiary education level, is a topic which has received much attention in the past 20 years (Boughey, 2000; Boughey 2005; Lillis, 2001; Lotz-Sisitka, 2009; Morrow, 1993; Weideman, 2003; van Schalkwyk, 2009). The government has recognised this and made large budget provisions for all aspects of education, from base literacy to tertiary education (Gordhan, 2013; National Treasury, 2014).

The creation of South African mother-tongue language corpora and learner corpora would provide valuable insights for creation of better learning and teaching materials and methods (Granger, 2004). However, there is very little learner data which has been collected and published in South Africa for German, other European, or indigenous languages (Goossens & Granger, 2013). Where corpora of learner writing are found in South Africa, they tend to typically be focused on learner English writing (cf. Van Rooy & Schäfer, 2002, Van Rooy, 2005, Van Rooy, 2008), or the compilation of learners' corpora[8] to assist creation of word lists or glossaries which are subject-specific for second language learners of English (Van der Walt & Fourie, 2005).

This focus on learner English in South Africa is not surprising, as English is seen as the language of 'prestige' in South Africa, and is also the most widely used language in the tertiary education sector (Oliver, 2009). According to the Language-in-Education Policy, the main aims of the Ministry of Education policy for language in education are (paragraph 5):

> 1. to promote full participation in society and the economy through equitable and meaningful access to education;
>
> 2. to pursue the language policy most supportive of general conceptual growth amongst learners, and hence to establish additive multilingualism as an approach to language in education;
>
> 3. to promote and develop all the official languages;
>
> 4. to support the teaching and learning of all other languages required by learners or used by communities in South Africa, including languages used for religious purposes, languages which are important for international trade and communication, and South African Sign Language, as well as Alternative and Augmentative Communication;
>
> 5. to counter disadvantages resulting from different kinds of mismatches between home languages and languages of learning and teaching;
>
> 6. to develop programmes for the redress of previously disadvantaged languages.
>
> (Department of Basic Education, 1997)

[8] See Chapter 3 for an explanation of a learners' corpus.

Of particular importance to myself as a researcher of German is point 4, which highlights the aim to "to support the teaching and learning of […] languages which are important for international trade and communication" (Department of Basic Education, 1997). In addition to the above, provision is made in the constitution under point 6.5, that it is the responsibility of a Pan South African Language Board, which is established by the national legislation, to "[…] (b) promote and ensure respect for — (i) all languages commonly used by communities in South Africa, including German, Greek, Gujarati, Hindi, Portuguese, Tamil, Telegu and Urdu […]" (Department of Basic Education, 1997). Here we find explicit reference to the promotion of the German language in South Africa.

More recently, in general notice issued by the Government Gazette on the 7th of April 2015 calling for comments on the Language Policy of the Department of Basic Education, it is stated that "The Constitution of the Republic of South Africa recognizes that cultural diversity is a valuable national asset and hence promotes multi-lingualism" (Department of Education, 2015, p. 10). In a further paragraph this notice makes direct reference to the language in Education Policy discussed above:

> The Language in Education Policy is conceived of as an integral and necessary aspect of government's strategy of building a non-racial nation in South Africa. It facilitates communication across barriers of skin colour, language and region; whilst at the same time creating an environment in which respect for languages other than one's own is encouraged. The underlying principle is the maintenance of home language(s) while providing access to and the effective acquisition of additional language(s) by all learners. (Department of Education, 2015, p. 10)

Certainly on paper there is room for the respect and development of the German language in South Africa, especially as it relates to ideas of intercultural competence (Weber & Domingo, 2011; Weber, 2015).

Germany and the study of German as a foreign language (an important language in the areas of science and philosophy – among other disciplines – before 1900) experienced a loss of popularity after the two World Wars in the 20th Century (see Reinbothe, 2006). Nonetheless, Germany is still an influential and important language of communication today, which is learned in 101 countries worldwide (Jaworska, 2009), above all in Central and Eastern Europe (Vasagar, 2013). German is "the most widely spoken mother-tongue in the European Union", spoken as a mother-tongue by "16 per cent of the EU's 500 million-strong population, followed by Italian and English, both spoken by 13 per cent" (Vasagar, 2013, p. Online). French is spoken as a mother tongue by 12 per cent of the EU (Vasagar, 2013). Germany has one of the world's leading economies, currently ranked as the largest national economy in Europe, the fourth largest by nominal GDP in the world and the fifth by GDP (Münkler & Winkler, 2014).

Germany and South Africa have a strong economic relationship and Germany is South Africa's second largest trade partner (German Missions, 2015). Apart from these economic ties, there are many socio-historical ties between South Africa and Germany as many German settlers came to areas such as the Eastern Cape in the 19th century and settled here, forming language communities and settlements with names like Berlin, Stutterheim, Hamburg and Hannover (Zipp, 2013). These place names serve as strong reminders of this shared past (Zipp, 2013), and there are still many direct German descendants living in South Africa today. In the tertiary education sector, the German Government sponsors hundreds of students worldwide, including South Africans, to study in Germany each year through the DAAD (*Deutscher Akademischer Austausch Dienst,* or German Academic Exchange Service). In 2013 alone the DAAD funded 600 German Students to study in South Africa, and 406 South Africans the opportunity to study in Germany (DAAD, 2014). This is in stark contrast to the prejudices still held against Germany for their part in the world wars. Learning a particular language and engaging with a particular culture does not necessarily mean condoning previous heinous behaviour, but serves to build bridges of understanding, on an intellectual, cultural and personal level (Weber, 2015; Harden, 2000). In today's growing multilingual and global society, the study of any foreign language, but particularly German, can only be considered as an asset aiding one's intercultural competence and awareness, respect and understanding for others, one's communication abilities and awareness of one's own language structures (Weber, 2015), and by extension of all of the above: one's prospects in terms of employability.

German as a foreign language (GFL) as an independent academic discipline is a relatively new field, although the German language and literature (*Germanistik*) has been taught and studied for many decades (Jaworska, 2009, p. 6). GFL is not focussed solely on grammar, but rather is concerned with providing students with language skills, as well as cultural, political and social aspects of the German speaking world in order to "grant a more profound access to the culture in question" (Jaworska, 2009, p. 10). GFL abroad places the student between two cultures as well as two languages (or more) (Jaworska, 2009, p. 10). Increasingly, universities in the UK have to accept students with low initial language proficiency, and universities such as Rhodes University in South Africa accept *ab initio* students, which means that there is a great focus placed on language teaching and applied linguistics in GFL. In three years, students must progress from a basic knowledge of the language, to language competency, which is accompanied by introductions to literature, translation and cultural studies of Germany and other German-speaking countries.

The study of German as a foreign language in the South African Tertiary sector has been on the decline since 1994, due to a number of factors pertaining to the politics, policies and perceptions permeating the South African context, and some have gone so far as to say that German is "facing a crisis" (Witte,

2002/2003, p. 173) across the region of Sub-Saharan Africa as a whole (Witte, 2002/2003) (Annas, 2002/2003). Whilst this mirrors a general trend in the decline of the study of GFL in many European countries (Jaworska, 2009, p. 6) (mainly due to the First and Second World Wars and a boycott of 'German Science' (Reinbothe, 2006)), despite the wide-spread of German as a mother tongue in the EU; South Africa's own unique multilingual context – with no less than 11 official languages vying for attention and research – has created a situation in which, over the past two decades, several German departments have closed down, and many more have been called into question, having to fight a constant battle to prove their worth.

In November 2014, during the writing of this thesis, there was a petition which was circulated calling for the executive management structures of the University of KwaZulu Natal to allow the disciplines Afrikaans, French, German, Italian and others to remain open. The author of the petition writes: "Language departments serve a vital role within the infrastructure of any university, UKZN included". The author writes the following motivation for the petition which gives an indication that it is not only German which is under threat at the tertiary level in South Africa, but all foreign languages not indigenous to South Africa and of little economic benefit to the government.

> [T]he students enrolled currently in these programmes are constantly caught in a state of limbo, because they never know if their respective majors will be available in the next semester and on which campus. The students of UKZN should be taken into account and informed, at the very least, well in advance about the future of their courses. UKZN has been eroding the languages for a number of years now, not renewing staff, complaining about low student numbers in the languages, while causing such confusion in students' and lecturers' minds that numbers are in freefall. To be a language student at UKZN is to be very courageous and determined. Parallel to this, UKZN is constantly courting foreign students, from Africa as well as from the other continents, and the foreign students who come to UKZN to follow a Humanities programme invariably look for a foreign language to add to their course. If this is allowed to carry on we might as well say goodbye to our foreign relations and the prospect of creating links with other countries that are not English speaking. The language departments which you wish to destroy and have already starved of resources, teach your students to communicate effectively with countries outside of Africa as well as with our fellow African nations. The university management intends on the termination of these disciplines beginning as of the first semester next year. It must be emphasised that none of this information has been openly discussed with or distributed amongst the students! (S., 2014)[9]

This emotive motivation to sign the position gives a clear indication of the type of feeling surrounding the closure of language departments.

The situation in South Africa used to be much like that of the UK and France, with the teaching of German following the traditional '*Auslandsgermanistik'* model which was considered a replica of the

[9] The author of this petition withheld their surname for the purposes of anonymity.

Germanistik taught in German-speaking countries (Jaworska, 2009, p. 9), with a focus on Literary Studies and Historical Linguistics, and requiring mother-tongue or near mother-tongue competency from the students (Weber, *et al*., 2015, p. In review). *Germanistik* had been criticised from as early as the 1970s "as being irrelevant to modern day society" (Altmayer, 2001 in Jaworska, 2009, p. 9), and thus adapted to the social and academic challenges by reducing the significance of literature and adding "a range of subjects from the fields of cultural and social studies, with greater focus on the events of the 20th century" instead (Jaworska, 2009, p. 10). Thus, the "concept of interdisciplinary German Studies was born" (Jaworska, 2009). As Jaworska points out, the whole field of *Germanistik* underwent great changes, from being one quite homogenous discipline it has "emerged as a regional, autonomous discipline, heavily influenced by the academic profiles of individual departments, teaching and learning traditions, the position of German in school curricula, attitudes towards German language and culture, proximity to German speaking communities, and the intensity of commercial and cultural exchanges, and other factors" (Hinkel, 2001 in Jaworska, 2009, p. 10).

The following excerpt from Weber, *et al.*, (2015) similarly critiques the traditional *Germanistik* practises of the past in South Africa, starting by calling them "a very exclusive approach to our discipline" and adding that "our subject cannot just be a personal enrichment course for a few students coming from either a privileged background or one that by chance enabled them to fulfil the preconditions of registering for *Germanistik"* (Weber, *et al*., 2015). The continuation below serves to illustrate the discipline that German Studies at Rhodes University has become, and its own unique conceptualisation of the field as influenced by all of the factors quoted from Jaworska (2009) above:

> Neither can German Studies be solely a service course to satisfy the needs of the South African political and economic market; a balance between those two needs to be reached, fostering a love of knowledge which "cannot be sold in the marketplace or serve the cause of some apparent material progress" (Horn 1999: 81) without ignoring or engaging with said marketplace since it is, whether we like it or not, strongly influenced in South Africa by the same policy makers that determine directives for higher education. Many of our students enrol for our subject because they still have personal ties to the German-speaking world, because a particular aspect of what they conceive to be German culture fascinates them (e.g. sport, classical composers) or simply because they assume having a European language will enhance their 'marketability'. (Weber, *et al*., 2015, p. In review)

According to Rolf Annas, in South Africa there are "fewer lecturers [of German], fewer students [of German] at postgraduate level, and there are more students who start their study of German without any previous knowledge of the language" (Annas, 2002/2003, p. 181). Statements such as this beg the question of why German as a foreign language has been impeded. I have already mentioned some reasoning for this above, and will make this view clearer here: it seems evident that South Africa's unique multilingual context actually serves to hinder further multilingualism in some contexts, for

example the learning of European foreign languages. South African tertiary education language departments are engaged in a 'battle' for students. Where taking an indigenous language has become compulsory, student numbers in other language departments may drop.

Multilingualism as it is used in the South African context is a misnomer as it appears not to refer to the development 'foreign' languages, especially in terms of language policies; but rather to the elevation and intellectualisation of the various African languages. Rhodes University has recently been awarded a SARChI Research Chair for the period of 2013 to 2018 in the field of intellectualisation of African languages, multilingualism and education as part of the South African Research Chairs Initiative (SARChI)[10] (African Languages, 2015).

Whilst of course this is a necessary and very positive development in the South African context, it is unfortunate that burgeoning research and interest in one area should exclude the other, rather than seeing all language learning as a valuable and irreplaceable asset; as linguistic awareness and sensitivity as well as intercultural competence are transferrable skills upon which a high value should be placed if we wish to educate a generation of scholars who are committed to positively transforming our country and forming internal as well as international bonds. Teaching, learning and researching foreign languages can go hand in hand with the teaching, learning and researching of African languages.

Anne Baker wrote in 2007: "Arguments for the promotion of indigenous languages should not be used as reasons against teaching German in a multilingual South African context" (Baker, 2007, p. 30). Baker put forward the argument that the elevation of African languages can be in harmony with the teaching of German as a foreign language, especially if a learner's first language is taken into consideration when teaching a foreign language. Baker (2007) provided, for example, some valuable insights for teachers in multilingual classrooms, by performing an error analysis of the written work of grade 10 and 12 learners of German as a foreign language. Unfortunately, the University of Johannesburg no longer has a German department/section, and valuable research such as this is no longer being conducted. Similarly, research undertaken in 2013 in the form of a questionnaire involving 18 participants at the German Studies Section at Rhodes University, revealed that the study of German, which includes a module on cultural studies, may indeed further one's intercultural competence (ICC), and foster the type of critical thinking which can "aid in the transformation of post-Apartheid South

[10] One of the main focus areas of the SARChI Research Chair project in the field of intellectualisation of African languages, multilingualism and education is corpus development in African languages (African Languages, 2015). This serves as an indicator of the importance of corpus research in the development of teaching and learning practices in South Africa.

Africa" (Weber, *et al.*, 2015; Weber, 2015). In short, whilst the main thrust of this research will be to explore the implementation and efficacy of a corpus approach to teaching an everyday academic writing course (in this case, for German) in South Africa, one should keep in mind the greater context of foreign language teaching in South Africa. This thesis is not primarily an argument for German Studies, but such is the context of the research that an argument for German is necessary at all.

In conclusion, given the context of tertiary education in South Africa as outlined above, where students struggle to acquire academic literacy in English, developing academic literacy in a new foreign language such as German poses a particularly challenging obstacle to most students. While there are shared values norms and etiquettes across the realms of both English and German everyday academic writing, the formal language employed in academic discourse is often used metaphorically, and is not directly translatable, and therefore needs to be explicitly taught. This can be achieved through a combination of writing more, more often, coupled with language instruction based on insights gained from mother-tongue and learner corpora, as will be further shown in Chapter 2 and Chapter 3. Even as German as a foreign language has seen a decline in student numbers in South Africa, South African educational policies do make room for the promotion of German, and many good arguments for the promotion of German in South Africa can be made, including economic benefit and improved intercultural competence (Weber, 2015).

CHAPTER 2. THE THEORETICAL CONTEXT: A HISTORICAL CONVERGENCE

Granger states the following about corpus linguistics, which I find to be a particularly poignant: “It is neither a new branch of linguistics nor a new theory of language, but the very nature of the evidence it uses makes it a particularly powerful methodology, one which has the potential to change perspectives on language” (Granger, *et al.*, 2002, p. 4). As this thesis aims not only to analyse the writing of our students, but also to aid future modern language instruction at higher education institutions in South Africa, a short overview and description of corpus work in the field is necessary. Thus, this chapter will cover some base tenets of corpus linguistic methodology, as well as how corpus linguistics and language teaching came to be united, providing a base for the discussion of learner corpus research and corpus-based teaching methods in Chapter 3.

‘Corpus’, from the Latin, literally means ‘body’, and is, in a linguistic context, used to refer to a body of text. More precisely, a corpus may be defined as “a principled collection of text”, rather than an arbitrarily complied set of writing. Flowerdew (2012) outlines four criteria for defining a corpus as follows, a corpus “consists of authentic naturally occurring data”, “is assembled according to explicit design criteria”, “is representative of a particular language or genre” and lastly “is designed for a specific linguistic or socio-pragmatic purpose” (Flowerdew, 2012, p. 3). The idea of a “principled collection of text” will be expanded on in Chapter 3 where considerations in building a learner corpus are explored.

In the 13th century, the first biblical concordance was created manually by Hugo of Saint-Cher and a team of scholars. This involved indexing every instance of each word in the bible, and was considered a great feat – a life’s work. A similar task could be performed today with the aid of computer software in a matter of minutes (McCarthy & O'Keeffe, 2010). We can now use computers to “read, search and manipulate” (McEnery & Hardie, 2012(a), p. 2) corpus data, thus the possibility of having extremely large data sets is feasible, rapid and reliable, where the human effort alone is not. The surge of corpus-related studies is directly mirrored in the surge of computer advances which were experienced in the twentieth century, and would not have been possible without these technological advances (Bonelli, 2010, p. 14). There has always been the drive to observe the patterns in language, and to examine language in context; however, creating a concordance for even one relatively small body of text remains virtually impossible for the linguist without the help of technology.

Real instances of language use make a corpus an ideal language learning tool

The first electronic corpora are attributed widely to lexicographers, and to pre-Chomskyian structural linguists whose interests lie in looking at ‘real’ instances of language use as the heart of linguistic

research (McCarthy & O'Keeffe, 2010). This notion of 'real' or 'authentic' language use is an essential driving force behind all corpus linguistic studies, and is perhaps the indispensable factor which makes corpora so ideally suited to language teaching: a corpus provides a huge sample of actual language use, produced in real settings by actual users of a language. Many teachers and linguists have relied on introspection and intuition to provide descriptions of language. Language intuition has proved helpful, however, this is not always accurate or reliable, as "[b]eing a native speaker does not automatically mean that a user has a conscious, clear, and comprehensive picture of the language in all its contexts of use, nor do all native speakers share the exact same intuitions" (Gabrielatos, 2005, p.5). There are often large discrepancies to be found between actual use (observed through corpus analysis) and traditional, introspection-based views on language (Sinclair, 1997, p. 32-34).

By using specialized software called 'concordancers', patterns in language become clear through a combination of frequency lists and collocations. Frequency refers to the number of times a particular word 'type' occurs in a corpus of a total amount of 'tokens' (normally the total word count), for example the most frequent 'type' that occurs in most corpora is the definite article *the* (Baker, 2006, p. 47). A frequency list is usually the first step of analysis for any corpus linguist, as it provides one with an idea of what to focus on, particularly when focussing on lexical items (Baker, 2006, p. 121). This may then be followed up by looking at the collocations and keywords in context (KWIC) in the concordance lines, as well as by comparing the term to other similar lexical items to provide a comparison (Baker, 2006, p. 121).

Collocations refer to the tendency for words to co-occur frequently in regular patterns with other words, for example in English [powerful] + [argument] (Durrant, 2009, p. 158). Collocation "often involves relationships between words which may be separated by other, non-fixed, or semi-fixed words, and which may differ in their position relative to one another" (Durrant, 2009, p. 158). These relationships between words and the patterns that consistently repeat themselves in a language are highly important to consider when teaching and learning a language. This is because much of language use consists of semi-preconstructed phrases, as shown by Sinclair (1991) who undertook pioneering work in corpus-driven lexicography. Collocation was originally conceived of as "a purely textual phenomenon" (Firth, 1968).More recently however, 'neo-Firthian' corpus-linguists have interpreted collocational patterns to show something about the psychological reality of language (Durrant, 2009). For example Hoey (2005) has theorized that words do not exist in isolation, but rather that words exist in psychological relation to other words (Hoey, 2005, p. 5). Semi-preconstructed phrases (or "chunks" of language (Ellis, 2003)) are now considered to constitute single choices for the mother-tongue speaker (Sinclair, 1987, p. 320). These 'chunks' may also be termed 'formulaic sequences' (Jaworska, *et al.*, 2015), and the concept of 'formulaicity' will be expanded upon in Chapter 3.

Notions of a 'common sense' prescriptive grammar of English, where descriptions of language do not match up with corpus evidence, have been highly problematized (Sinclair, 1997, p. 30). Sinclair argues that the first lesson one should learn from corpus study is that "language cannot be invented, only captured" (Sinclair, 1997, p. 31), and that one should present only real examples in the context of language teaching. He argues that it is also more practical to search through a file of "relevant examples" than to think of an example which "sounds natural" (Sinclair, 1997, p. 31), something which teachers and linguists alike have been doing since the origin of their profession.

> When 'sounds natural' is examined closely, it usually transpires that it is almost impossible to invent an adequate example [of natural language use]; attempts by language teachers, lexicographers, and others to represent usage are often embarrassing and never reliable. (Sinclair, 1997, p. 31)

Thus the corpus linguist and language teacher takes a descriptive, rather than a prescriptive approach to grammar, concerning themselves with actual language and becoming interested in finding out how speakers of the language use that language in context, rather than with what is stated to be correct in textbooks.

Corpus linguistics- part of the linguistic toolbox

In the 1980s and 1990s the term 'corpus', which had previously been used to describe a body of text, made its way into the linguists' vocabulary as a tool and method of investigation; by the late 1980s, the term 'corpus linguistics' (CL) was established (McEnery & Hardie, 2012(a), p. 2). The revolution in hardware and software in the 1980s and 1990s made this possible, as computers were no longer a tool only able to be used by few, they could be used by all who had the resources to access them, and with a fair amount of ease (Jones & Fortescue, 1987). Corpus linguistics can now be defined as a set of procedures, tools and concepts employed to study different areas of language, rather than as a study of any one particular aspect of language in isolation (McEnery & Hardie, 2012(a), p. 1), yielding a large amount of data which is reliable, as well as authentic, and which can be rapidly processed by computer software designed specifically for this purpose (Baker, 2010; McEnery & Hardie, 2012(a)).

Parallel to the developments in desktop computers is the development of the internet (O'Keeffe & McCarthy, 2010; McEnery & Hardie, 2012(a)). The ability to upload and download data at speed means firstly an easy transfer of knowledge between scholars which was previously inhibited by factors such as actual geographical distance; and secondly, the easy accessibility of texts already in electronic form, alleviating the necessity for large and clumsy text scanners (McCarthy & O'Keeffe, 2010). Consequently, the size of one's corpus became almost unlimited, as texts became available electronically. Thus, from the 1990s onwards linguistics ceased to be constrained by a scarcity of data, and there is more data available now than with which current methodologies can cope (Bonelli, 2010,

p. 17). Small, specialised corpora began to be of interest during this time, and the value in creating a corpus of specific language use, such as that produced by learners of a language, was realised. While the potential for creating huge corpora remains, the purpose of one's investigative needs now determines the size of the data set (Ragan, 1996).

As an aside, it is necessary to note that the majority of the initial and subsequent corpus compilations were undertaken for the study of the English language, and various varieties thereof. As asserted by Granger (1998, p.3), "English is undoubtedly the language which has been analysed most from a corpus linguistics perspective". Corpus research in English has dominated in every sector, from large multi-million word corpora (e.g. the British National Corpus and the Bank of English), to small specialised corpora of learner language (e.g. the International Corpus of Learner English). Models based on English corpus analysis have paved the way for research in other languages, as will be shown in the Chapters below.

Methodology defines domain

A highly significant point in the development of corpus linguistics was the realisation that the 'methodological enhancement' afforded by corpus software – which allowed for accurate computational power at the command of the researcher – coupled with an 'explosion' in the quantity of data available to be accessed via computer, caused a theoretical and qualitative revolution in the field of linguistics and language research. As Bonelli states: "The methodology has ended up defining the domain of the discipline" (Bonelli, 2010, p. 18), as new issues were raised regarding language 'competence' vs. language 'performance', and patterns of language which were not visible before emerged in the vast amounts of data at the fingertips of researchers. Corpus linguistics as a field therefore focuses on performance rather than competence, description rather than universals, and quantitative as well as qualitative analysis, which can be viewed in sharp contrast to a Chomskyan approach to language (Granger, 1998, p. 3). The two approaches are not mutually exclusive however, and corpus linguistics can provide evidence for or against claims made by traditional "armchair linguists" (Granger, 1998, p. 3).

For example, based on corpus evidence, there has been much debate on what constitutes a 'word' for counting and analysis purposes, which has had ramifications on pedagogical theories and practises (Gardner, 2007). Corpus methodology has also shown its effects in the field of second language (L2) acquisition, where there has been a shift in focus from examining learner texts from the perspective of manually-encoded error analysis, to an in depth corpus-based analysis of the typical lexico-grammatical patterns of learner language (Jaworska, *et al.*, 2015). This will be further expanded upon in Chapter 3.

As mentioned above, before the 1980s, the term corpus was virtually absent from publications concerning teaching and learning, which begs the question of how this "convergence" began. Leech (1997) asserts that many researchers in the field of corpus building, design, and research in the late 1980s and early 1990s were also teachers in many different capacities, and there was an increasing tendency for corpora to be used informally for language teaching as well as research purposes. This change to corpus-based methods in teaching had, according to Wichmann (1997, p. xvi), three prerequisites. The first was mainly practical: the technology had to be fairly cheap, easily accessible, and with a greater storage capacity for teachers and learners alike, as has been described above. Leech asserts that "The corpus, purely as a resource, is rather like a shelf in a university library: it is there to be exploited, and the same resources are equally usable for research and learning" (Leech, 1997, p. 2).[11] The second was a focus on a descriptive, rather than a prescriptive approach to grammar, where the focus lay in actual language usage; as seen in Sinclair's argument for the use of real examples only. The last prerequisite required a changed view of the role of teacher and learner, where agency for learning lies with the learners themselves (Wichmann, 1997, p. xvi). This shift in teaching philosophy in the 1990s has been described by Laurillard (1993, p. 13-15) as 'teaching as imparting knowledge' to 'teaching as mediated learning'. This in itself brought about challenges within the classroom, subverting the traditional order of learning and challenging teachers and learners alike (McCarthy & O'Keeffe, 2010, p. 7).

A 'trickle-down' from research to teaching

Where these prerequisites were fulfilled, Leech (1997, p. 2) describes how many researchers experienced a "trickle down" effect from their corpus research into their everyday teaching, even though there was very little published on the educational use of corpora at that time. Johns (1986, 1988, 1991), was one such researcher, and the one of the first to advocate for data-driven learning methods which focus on corpus data as a tool for foreign language learners. Leech's own experience of this phenomenon began as early as 1976 while he was using the LOB (Lancaster-Oslo/ Bergen) Corpus for postgraduate teaching, and he writes in 1997 that "the original 'trickle down' from research to teaching is now becoming a torrent!" (Leech, 1997, p. 2).

With the publication of the Collins COBUILD English Dictionary, the "first corpus-based dictionary for learners", many more teachers became aware of what a corpus is, and began to use corpus based

[11]In 'developing countries', the question of access again becomes important, as even though the technology has been developed, it may not be accessible to many teachers and students who lack even the most basic of resources, never mind personal computers and internet connections.

materials (Gabrielatos, 2005, p. 1, McEnery & Xiao, 2010, p. 5). Leech acknowledges that this metaphor of a 'trickle down' may be interpreted as unhelpful, or even offensive, implying that research is more important than teaching, an assumption held by many academics (Leech, 1994, p. 4). Leech (1997, p. 4) notes that this is in no way a one-way dependence, and that of course the "'trickle up' from teaching to research can be just as important". However, while Leech may have regarded the increasing awareness for the use of corpus research at the time as a "torrent" washing into the realm of teaching, many teachers and academics remained (and still remain) unaware of corpus research and its applications in teaching and learning (Breyer, 2009). Whilst many teachers may make indirect use of corpus-based products such as dictionaries, many remain unaware of what a 'corpus' is and how this may potentially be used in the classroom.

The convergence between teaching and language corpora

According to Leech (1997) the convergence between teaching and language corpora has three foci, with the first being the 'central focus area' of enquiry: (1) the direct use of corpora in teaching (teaching about corpora, teaching to exploit corpora, exploiting to teach corpora), (2) the indirect use of corpora applied to teaching (reference publishing, materials development, language testing), and (3) further teaching-orientated corpus development (Language for Specific Purposes (LSP) corpora, L1 and L2 developmental corpora, bilingual/multilingual corpora). All three areas of convergence are important in attempting to find solutions to the questions raised in the introduction chapter, namely, how do we measure writing success and levels of academic literacy? How does one evaluate whether a student's writing is 'mother-tongue-like' and what type of help should be given to students to aid in the acquisition of writing skills? Chapter 3 will expand on these areas of convergence, reviewing specific research in these areas, in order to provide answers to some of the above questions.

With this historical development of corpus linguistics as it relates to teaching and learning in mind, the following Chapter provides a review of the literature surrounding everyday academic German and the use of corpus linguistic tools to facilitate the description and teaching thereof.

CHAPTER 3. LITERATURE REVIEW: USING CORPORA FOR INSTRUCTION AND ASSESMENT

This chapter builds on the historical foundations of the convergence between corpora and language learning as set out in the chapter above, expanding on how researchers have used corpora for language teaching and learner language description – keeping in mind the questions relating to the assessment of academic literacy and student writing, and what help can be given to improve writing in the foreign language, as set out in the introduction to this thesis. As academic writing in German and other languages involves a specific metaphorical use of language, and as correct lexical choices are an important facet of the evaluation of academic texts, lexical choices and formulaicity will be explored as an important dimension of learner writing. Following this, the use and creation of learner corpora and what they can tell us about the nature of learner writing as compared to mother-tongue writing are discussed. With an idea in mind of the assessment and description of learner writing in place, the discussion orientates towards the types and use of corpus-based materials, with reference to relevant research undertaken by leading scholars. As there are two main approaches to teaching using corpora, namely computer- and paper-based corpus approaches, the advantages and disadvantages of both are addressed. Finally, as teaching involves both teacher and learners, the role of the teacher in implementing a corpus-based approach to language learning is considered, and the overview concludes with a discussion on students' attitudes towards using corpus tools.

Lexical choices and academic literacy

In Chapter 1, everyday academic German was discussed with a focus on the metaphorical use of everyday language in an academic context. It was asserted that many of the phrases that we use in academic language are not directly translatable, and that language learners may make use of grammatical, but non-idiomatic phrases (Krummes & Ensslin, 2014). In Chapter 2, collocational patterns in corpora were said to be highly important as researchers such as Sinclair (1987) have demonstrated that collocations are indicative of psychologically related, semi-preconstructed phrases (or 'chunks' of language (Ellis, 2003)) which are now considered to constitute single choices for the mother-tongue speaker (Sinclair, 1987, p. 320). This observation has serious implications for language teaching and learning, and this Chapter will provide evidence for why this is so, and how this knowledge can aid language researchers and teachers, with a specific focus on everyday academic German.

Correct lexical choices form a key component of what it means to be academically literate. In the definition of academic literacy developed by Weideman (2003, p. 7) at the University of Pretoria,

(upon which the TALL[12] test of academic literacy is based), the first ability for which researchers look in a text upon which an assessment of academic literacy is being undertaken, is the student's ability to "understand a range of academic vocabulary in context" and secondly to "interpret and use metaphor and idiom, and perceive connotation, word play and ambiguity". Students should furthermore "understand relations between different parts of a text, be aware of the logical development of (an academic) text, via introductions to conclusions, and know how to use language that creates cohesion between the different parts of a text" (Weideman, 2003, p. 7). Both this point as well as the next, which calls for students to be able to "interpret different kinds of text types (genres), and show sensitivity for the meaning that they convey, and the audience at which they are aimed" (Weideman, 2003, p. 7), are simply variations of an ability to make the correct lexical choices, and recognize collocational patterns.

Understanding a range of academic vocabulary in context is easily addressed with the use of corpus tools. Questions raised in English such as "is there an academic vocabulary?" have led to research such as that performed by Coxhead (2000), and Gardner and Davies (2014). Coxhead (2000) compiled a corpus of 3.5 million running words of English academic and non-fiction texts to create an Academic Word List (AWL) for English. Coxhead (2000) used corpus methods to identify the words which are most often encountered by university students when engaging with academic texts, thereby providing students who entertain 'academic goals' with a list of words which would be most worth their while to study (Coxhead, 2000, p. 213). What is particularly useful in the case of the AWL is that the word list has been divided into smaller sub-lists which are based on frequency (Coxhead, 2000, p. 213). This allows for a sequenced approach to teaching academic vocabulary (and creating materials) based on the idea that the words which occur more often will be more important, or indeed necessary, to learn.

Whilst the AWL was limited, as it only contained single words and not "multi-word collocational patterns which are known to be prevalent in language" (Durrant, 2009, p. 158), it provided a good basis for further research. Gardener & Davies (2014) for example, improved on the research performed by Coxhead (2000) and created a 'new academic vocabulary list' (AVL) based on over 120 million words. Gardner and Davies have created a web-based interface where AVL words can be accessed, identified and interacted with (including collocational patterns and n-grams for each word), creating an easily accessible tool for all teachers and students (Gardner & Davies, 2014, p. 305). These types of 'academic vocabulary' lists are very helpful to learners of English, as well as mother-tongue

[12] Test of Academic Literacy Levels (TALL)

English speakers attempting to acquire or improve their academic literacy in English. Unfortunately, such comprehensive resources do not yet exist for many languages other than English.

Learners' corpora

The task of creating a full academic word list in any language is a massive feat, and not one to be undertaken lightly by a solitary researcher or language teacher. However, those wishing to provide a resource of the type of academic vocabulary found in a specific discipline may do so by creating a relatively small corpus of discipline-specific texts in order to highlight for example, the most frequent lexical items and their collocations, as well as characteristic grammatical structures found in a certain genre (Leech, 1997, p. 18). These types of corpora which are created for the use of the learner as an aid are termed learners' corpora. Learners' corpora are not to be confused with learner corpora, which are collections of language produced by learners (Granger, 2004), as will be elaborated upon below. Learners' corpora are often termed 'Language for Specific Purposes' (LSP) corpora (Leech, 1997), as they provide evidence of language used in specific contexts, specialist subjects or genres; for example, academic writing in the field of history, science, law or medicine.

In South Africa there is a growing trend to create learners' corpora of syllabus materials in order to provide subject specific glossaries which include a translation and explanation of the term in the students' home language, e.g. isiXhosa (Van der Walt & Fourie, 2005). Van der Walt & Fourie (2005) for example sought to create a learners' corpus for a first-year course in Mathematics at Stellenbosch University. They further aimed to compare their specialized word list for mathematics, to Coxhead's Academic Word List in order to ascertain whether a general academic word list such as that created by Coxhead may or may not "include enough terms to render tailor-made learners' corpora unnecessary" (Van der Walt & Fourie, 2005, p. 242). Van der Walt and Fourie (2005) found that the most frequent terms in the corpus created from the mathematics course content did not appear in Coxhead's Academic Word List, or that the words in the module had such a "specialised meaning that general academic support on the basis of the word list would probably not be very useful for students" (Van der Walt & Fourie, 2005, p. 242). This type of research demonstrates the value of learners' corpora as a tool for understanding vocabulary in specific contexts or genres.

Research gap: few German corpus resources

There are few corpus resources available for the study of academic German, whether spoken or written, or comprised of language collected from mother-tongue or second-language German speakers. Those who have tried to address this issue for German include Fandrych *et al.* (2012) and Jaworska (2011). Fandrych, Slavcheva & Meißner (2012), created a corpus of spoken academic German (Fandrych, *et al*., 2012). While this was mainly comprised of spoken mother-tongue German,

researchers also aimed to include therein the first instances of a corpus of spoken German learner language (Fandrych, *et al.*, 2012, p. 319).

Jaworska (2011) created a relatively small corpus of written academic German writing in order to help the final-year students in her course to write their dissertations in German. The corpus created consists of 132 samples of published work from the University of Regensburg, from the realms of sociology, history and linguistics, resulting in the million-word corpus titled "WissDe" (Jaworska, 2011, p. 5). This corpus was intended as a resource to be used by students in order to examine words in context, and gain access to the metaphorical, and functionally academic usages of everyday words, such as those described in the section above.

Even as these two studies show a recent increase in corpus resources for German, the research gap pertaining to written academic German remains unfulfilled as there is to date no comprehensive corpus of academic German available for research (or other) purposes. There are also few subject or discipline specific corpora such as WissDe (Jaworska, 2011) which are publically available. This raises many avenues for further research into German, as well as for all languages for which no such list currently exists (such as the indigenous languages in South Africa).

Thus, whereas the topic of "English for Academic Purposes" has been well documented using corpus linguistic methodology (Coxhead, 2000; Durrant, 2009, Gardner & Davies, 2014), such is not the case for *Deutsch als Wissenschaftssprache*, or for *Deutsch als alltägliche Wissenschaftssprache* (Krishnamurthy & Kosem, 2007; Jaworska, 2011). *Deutsch als Fremdsprache* (DaF) has thus been influenced by corpus linguistics methodology following the work undertaken in English, resulting in a plea for new DaF didactics.

Learner corpora, a bond with SLA and FLT

Whilst much information can be gained from mother-tongue corpora, and used effectively to describe language as it is used by those born speaking and writing it, valuable insights can also be gained from the analysis of learner writing, and how it reflects or digresses from how mother-tongue speakers of a language use the language (an approach termed contrastive interlanguage analysis CLA (Granger, 2002)). This is particularly valuable for the fields of second language acquisition (SLA) research which "tries to understand the mechanisms of foreign/second language acquisition and foreign language teaching (FLT) research which aims "to improve the learning and teaching of foreign/second languages" (Granger, 2002, p. 5). Due to a wide range of applications, learner corpora have been welcomed as one of the most exciting recent developments in corpus-based language studies (Gabrielatos, 2005). Learner corpus research has created an important bond between corpus linguistics and foreign/second language research (Granger, 2002, p. 4), as both SLA and FLT can, and

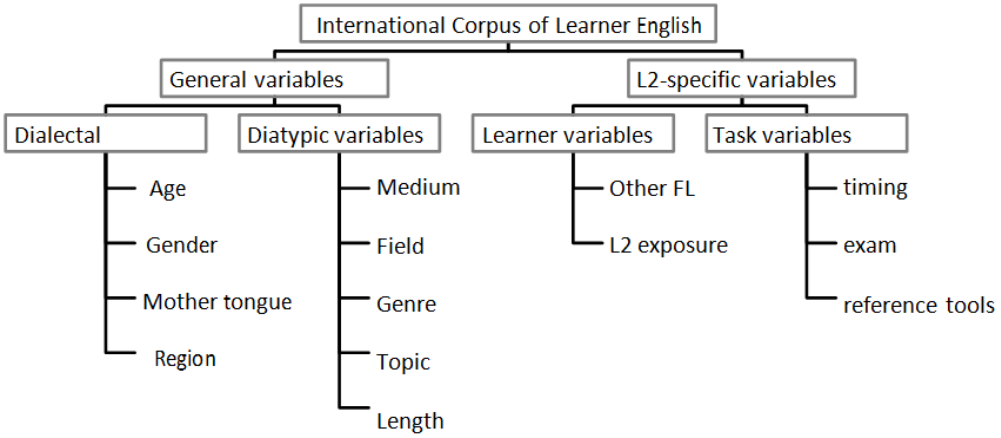
have been informed by learner corpus research, which describes learner language and better helps us to understand it, thus informing better teaching practice (Granger, 2002, p. 4).

Teachers have always collected (unconsciously, and sometimes consciously) learner data. There is—as indeed many authors who write on the topic of corpus linguistics and language teaching point out—"nothing new" about collecting learner data (Granger, 2004, p. 123). What makes computer-assisted learner corpora different, and highly advantageous to natural and elicited data types, is size, variability, and automation (Granger, 2004; Flowerdew, 2009). The significance of this with regard to corpus application will be exemplified below.

Learner corpora may be defined as collections of texts (spoken or written) produced by foreign or second-language learners (Granger, 2002). Learner corpus data is not simply a collection of errors, a learner corpus must consist of "continuous stretches of discourse which contain both erroneous and correct use of the language" (Granger, 2002, p. 9). Furthermore, learner corpora are always collected with explicit design criteria in mind, as there is so much variation in second/foreign language learner production influenced by linguistic, situational and psycholinguistic factors pertaining to the learner and their individual background. As Granger states: "A random collection of heterogeneous learner data does not qualify as a learner corpus" (Granger, *et al.*, 2002, p. 9). A learner corpus is only as useful as the care taken to control and encode the variables (Granger, *et al.*, 2002, p. 9). The variables which are normally controlled for in learner corpus design include those pertaining to the learner (for example learning context, mother-tongue, level of proficiency, and command of other foreign languages), and those relating to the task setting or language situation (for example time limit, use of reference tools, exam, audience/interlocutor) (Granger, *et al.*, 2002, p. 9; Flowerdew, 2008, p. 5).

For example, Granger (2004, p. 126) lists all the variables which are controlled for and recorded for the *International Corpus of Learner English (ICLE)* database. These are represented in the figure below (Granger, 2004, p. 126).

Figure 1, ICLE Corpus Design Criteria



(Granger, 2004, p126)

The considerations taken in creating the ICLE allow one to draw conclusions about learner language within very specific parameters, depending on the sub-corpus selected for analysis. Take for example Aijmer's (2002) study of the range and frequency of modal verbs of L2 English writers. Given the well-constructed nature of the International Corpus of Learner English (ICLE) which takes into account all of the above considerations and has several sub-corpora representing different learner varieties, Aijmer (2002) can draw conclusions about the use of modal verbs of L2 English writers, who have Swedish, German or French as a mother-tongue and are advanced-level university students, as compared to mother-tongue English writers. Aijmer's preliminary investigation of the modals in L1 and L2 English varieties revealed "a generalised overuse of all the formal categories of modality" in advanced learner writing (Aijmer, 2002, p. 72). Whilst Aijmer notes that the study in question was non-exhaustive, as it did not account for every possible expression of modality (Aijmer, 2002, p. 72), these findings are nonetheless valuable as they point to areas of interest which can improve the teaching of a second language and the experience of writing for second language learners.

One particular aspect of learner corpora research which exemplifies the bond between corpus research and teaching and learning, is formulaic language. Whereas in the past, the term 'formulaic' was traditionally used to describe fixed phrases such as idioms, proverbs or sayings that may be rare in language use (e.g. 'a little bird told me'), recent corpus-driven evidence has demonstrated that "formulaicity is a much more salient and ubiquitous linguistic phenomenon" (Jaworska, *et al.*, 2015). Formulaic language as it is recognised by corpus linguists today does not only include fixed strings of words, but takes a wide range of flexible structures with morphological and syntactical variability into account, some examples of which include collocational patterns e.g. *exert/wield influence*, or phrase

frames e.g. *if you look at* *, *I don't know what* *, *well I mean* * (O'Donnell, *et al.*, 2013; Jaworska, *et al.*, 2015)

Formulaic language may be presented under many different labels, depending on the methodology followed in the research (Jaworska, *et al.*, 2015). Formulaic sequences are the "basic building blocks of discourse" (Jaworska, *et al.*, 2015). This does not mean that they are necessarily complete or well-defined linguistic structures, but rather lexio-grammatical fragments which function as basic building blocks of discourse, ensuring cohesion and coherence of speech and writing. This is directly related to the concept of collocation (the tendency for words to co-occur with other words in recognisable patterns), as discussed briefly in Chapter 2. Formulaic language has become an important indicator in differentiating between advanced second/foreign (L2) language learners and first language/mother-tongue (L1) speakers, as learner corpus research has shown that formulaic language is often "underused, overused or misused in learner language" (Jaworska, *et al.*, 2015).

Formulaic language in German

German scholars and applied linguists, such as Jaworska, Ensslin & Krummes (Jaworska, *et al.*, 2015) have begun to address the research gap identified earlier in this Chapter, by analysing formulaic sequences in mother-tongue and non-mother-tongue argumentative writing in German. Researchers have found that although there is "substantial scholarship on formulaic language in L1 and L2 English", there is less so for other languages such as German (Jaworska, *et al.*, 2015). Jaworska *et al.* assert that there is little corpus evidence for typical lexico-grammatical patterns of learners, as traditionally L2 German research has been based on manually encoded error analysis which has focussed on the "cognitive mechanisms […] underlying the acquisition process" (Jaworska, *et al.*, 2015).

Jaworska, Krummes & Ensslin (2015) undertook a comparison of a corpus of FL German argumentative essays, written by advanced students of German (WHiG), and a corpus of mother-tongue German argumentative essays (FALKO L1) (Jaworska, *et al.,* 2015; Lüdeling, *et al.*, 2008). This led to the conclusion that German mother-tongue speakers are less likely to use formulaic language in argumentative writing than British advanced learners of German, whose writing showed a larger number of 3-grams under corpus-driven analysis (Jaworska, *et al.*, 2015). Jaworska *et al.* (2015) classified the formulaic sequences found according to function, and drew further conclusions, stating that "native speakers of German prefer discourse structuring devices to stance expressions, whilst British advanced learners display the opposite references". British learners of German were finally showed to prefer "macro-discourse-structuring devices and cautious language" whilst mother-tongue German speakers were shown to prefer "micro-discourse-structuring devices" and a tendency towards more direct language (Jaworska, *et al.*, 2015). This study is an example of the great potential

of learner corpus research. Findings relating to the WHiG and FALKO L1 corpora are not limited to formulaic language, and will be returned to again in this Chapter.

Why not traditional Error Analysis (EA)

It is natural for language teachers to want to have a grasp on the types of errors to which their students are most prone; one possible route to follow when wanting to categorize or examine learner language, is a traditional error analysis. Traditional error analysis focusses on manually identifying and categorising learner errors according to an error typology, and aims to gain better insights into a learner's interlanguage (Corder, 1981; Dagneaux, *et al.*, 1998; Ellis, 1994). However, traditional error analysis has been heavily criticised as being methodologically weak and subjective (Dagneaux, *et al.*, 1998; Vann, *et al.*, 1984). These criticisms are not unfounded, and are briefly expanded upon below, as it is because of these limitations that new methods of studying learner language have been developed.

The first methodological weakness that traditional error analysis presents is a basis on "heterogeneous learner data" (Dagneaux *et al.*, 1998, p. 164), which makes the data "difficult to interpret and almost impossible to replicate" (Ellis, 1994, p. 49), compounded by the fact that traditional error analysis categories are "fuzzy" as they are ill-defined (e.g. grammatical and lexical errors are not explained) and subjective in nature (Dagneaux *et al.*, 1998, p. 164). Furthermore, traditional error analysis, unlike learner corpus enquiry as described above, "cannot cater for phenomena such as avoidance" as the focus placed on the identification of overt errors creates a climate of disregard for both instances of non-errors (correct use), and non-use or underuse of words and structures (Dagneaux *et al.*, 1998, p. 164). According to Dagneaux, Dennes & Granger (1998, p. 164) EA is thus restricted to examining and categorising only "what the learner cannot do", and so presents a "static picture of L2 learning" (Dagneaux *et al.*, 1998, p164).

Like many other fields of linguistic enquiry, error analysis has benefited enormously from corpus-based methodology. Learner corpora collected according to design criteria eliminate criticisms of heterogeneous data, and the ability of corpus software to have multiple layers of coding allow for the implementation of better, well-defined, and more nuanced error-typologies. Thus, corpus-aided error analysis has proved to be more beneficial than traditional error analysis, as large amounts of controlled data can be studied reliably and instantaneously through a corpus tagging system, revealing patterns of specific errors in context, where a researcher is able to add and subtract layers of coding in order to reveal certain types of errors (Krummes & Ensslin, 2014). Computer-aided error analysis has led to such materials developments such as "L2 aware" grammar and style checkers (Dagneaux,

et al., 1998) (which are programmed to anticipate the types of errors which learners of a specific mother-tongue will make in the target language) and corpus-based language learning materials.

Applications of the theory: examining what's hard in German (WHiG)

Given the above insights into learner corpora and error analysis, German linguists have undertaken to identify and address "linguistic structures of German that pose a specific difficulty for the acquisition of German as a foreign language (GFL)". This project titled 'What's Hard in German' (WHiG), undertaken in the UK and Germany, has shown many practical applications, some of which have already been mentioned in the sections above. In line with traditional error analysis, previous researchers such as Borin & Prütz (2004 cited in Krummes & Ensslin, 2012) or Westergren-Axelsson & Hahn (2001 cited in Krummes & Ensslin, 2012) have observed learner German errors. However, as with previous traditional EA studies, they did not take into account the fact that learners may actively or unconsciously avoid difficult linguistic structures. The authors of the WHiG project assert that the relative underrepresentation of structures in learner data implies that these structures are difficult to acquire (Krummes & Ensslin, 2012).

A deep analysis of learner errors has been made using the annotated corpora FALKO L2 (*Ein fehlerannotiertes Lernerkorpus des Deutschen als Fremdsprache*) and WHiG (*What's Hard in German*) and provide good insight into the types of constructions with which European leaners of German as a foreign language struggle (Zeldes, *et al*., 2003; Lüdeling, *et al*., 2008). This was further achieved via a multi-frame error tagging system (Krummes & Ensslin, 2014). Through the comparison of learner corpora with mother-tongue corpora, further limitations of EA are stripped away, as researchers can now describe exactly what language forms learners are using (correctly) and which forms they are avoiding, through the use of over- and under-use statistics, as described in the sections above.

Thus, by using error analysis in conjunction with corpus methods, researchers have been able to identify the types of errors made by learners of German (and their frequency), as well as the types of structures which learners avoid, or conversely, overuse in their essays. These insights have been instrumental in the creation of useful materials for teaching German writing. The research findings published as a consequence of the WHiG project are an example of the direct practical applications of corpus linguistic tools, which can be transferred to many different contexts, as will be shown in Chapter 8 where I expand on the WHiG resources, and how they were adapted to be used successfully to teach academic German collocations in the South African tertiary context.

In conclusion, learner corpora provide a good opportunity for the analysis and assessment of student writing, as seen in the section above. Furthermore, learner corpora have assisted researchers in the creation of practical learning and teaching materials which can be transferred to many different

contexts. However, in South Africa there are few actual learner corpora for languages other than English from which to draw insights, which is why the creation of a corpus of learner German writing (as is one of the aims of this thesis) is an important and valuable asset to the field of foreign language teaching in South Africa, as insights from this endeavour could be transferred to other South African, German teaching contexts, and to other foreign language teaching contexts where educators may wish to create their own similar learner corpora following the methodology exemplified in this thesis.

When learner writing has been assessed and specific problems identified, the next logical step is to provide some form of aid in order to improve learner writing, especially so that learner writing may achieve as natural an appearance as possible. Once again, this can be addressed through the use of corpus tools, and learning materials can be created based on the insights gleaned from corpus-driven analysis of learner and mother-tongue corpora as shown above with the WHiG project. There are further ways in which to design corpus-based teaching materials and these will be discussed below, along with the considerations taken in designing and implementing a corpus-based academic writing course in German.

Helping students through the use of a corpus: creating corpus-based exercises

Researchers are generally optimistic about what corpus tools can offer language learners (Aijmer, 2010; Wichmann, *et al*., 1997). As formulaicity has proved to be such an important facet of the psychological reality of language (Sinclair, 1991; Hoey, 2005), many researchers have argued for the direct use of corpus tools in the language learning classroom, in order to expose learners to real-life data. One of the first researchers to suggest using corpora to aid foreign language learners was Johns (1986). Johns defined the approach as data-driven learning (DDL), a method which sought to encourage students to explore the "regularities of patterning in the target language" via the use of computer generated concordances (Johns & King, 1991, p. iii). This approach was viewed by Johns as one of the most powerful tools which a language learner can possess (Johns, 1988), and forms a key component of what is described by Leech (1997) as being at the core of the convergence between teaching and language corpora.

Yoon (2008) asserts that one of the great benefits of students manipulating a corpus themselves is that through learning to interact with corpora, students learn about the language itself, which is essentially a study of the language; thus the student who is involved with corpus linguistics is required to take on the role of questioner, and essentially, to become a linguist in their own right (Yoon, 2008, p. 44). As Leech summarizes "[…] the process of getting to grips with the software invariably shades getting to grips with the techniques of linguistic analysis" (Leech, 1997, p. 9). As learners who undertake these methods are supposed to "discover" grammar based on "evidence from authentic

language use" (Stevens, 1995, p. 3), Stevens states that the learner is in these cases is described as a "language detective".

There are, therefore, high expectations on the parts of researchers and teachers of what corpus tools can offer language learners, and higher expectations of the type of inductive learning which is supposed to occur as a result of the inclusion of corpus tools in the language learning environment. The following section will provide a review of how corpus tools may be integrated into the language classroom. Following this, the role of the teacher in implementing corpus tools will be taken into account, and some of the difficulties associated with integrating corpus tools into the language learning classroom will be outlined. The discussion will conclude with an overview of what students themselves have said about the use of corpus tools in the classroom, and what benefits they have perceived as being directly related to corpus resources.

Teaching to exploit corpora, and exploiting corpora to teach

There are two ways in which corpora can be used directly in the language classroom, both of which were incorporated into the academic writing course for German Studies at Rhodes University. The first way in which corpora can be used directly in the classroom can be called "teaching to exploit" (Leech, 1997). Teaching to exploit involves teaching students to use corpus software and having them perform searches of specific grammatical or lexical items in order to deduce patterns of language for themselves. The second way, 'exploiting to teach' (Leech, 1997) is performed by providing students with pre-made printed exercises based on a teacher's corpus exploration, and preparation, of concordance lists. The rationale behind these two approaches is the same, and stems from the main concern of computer-based data driven learning (DDL) as presented in Chapter 2: "Present real examples only" (Sinclair, 1997, p. 31). Examining collocations and keywords in context (KWiC) form the heart of this enquiry, as students taught via data-driven learning methods must look for the patterns in language in order to enhance their "awareness of the lexico-grammatical patterning of texts" (Yoon, 2008, p. 31).

The aim of the language learner is to develop "a feel for the language" or language "intuition" (Gabrielatos, 2005, p. 6), something which mother-tongue speakers develop through exposure to language (Gabrielatos, 2005, p. 6). Through exposure to language use, mother-tongue speakers develop a mental corpus which is informally examined in the mind to produce these intuitions (Gabrielatos, 2005, p. 6). In countries where the target language is not widely spoken, (as is the case of German in South Africa) learners may lack sufficient exposure to be able to form and recognise patterns. Reading facilitates language learning, however this is often expected to be undertaken outside of the classroom, given the limited amount of contact hours in a language classroom. As

discussed in Chapter 5 the participants for this study engaged with very little reading outside of the language classroom. Corpora which are representative of a mother-tongue variety of a language can offer "condensed exposure to language patterns" (Gabrielatos, 2005, p. 10). As Gabrielatos states: "It does not seem worthwhile to spend the time it takes learners to read more than 2,000 words to teach only four collocations" (Gabrielatos, 2005, p.12).

Thus, according to Gabrielatos, "[…] by working on representative examples from language corpora, learners will be helped to recognise recurring patterns of structure and meaning" (Gabrielatos, 2005, p. 6). Furthermore, a corpus can provide such a wealth of instances of use of a specific item, that there is sufficient evidence which is required by learners to "refine their perception of it" (Gabrielatos, 2005, p. 6). It should be noted that I am in full agreement with Gabrielatos on the point that corpora should not replace out-of-class reading, or be the "sole vehicle for the development of reading skills and strategies" (Gabrielatos, 2005, p. 10), but rather that corpus use should be an addition in the classroom which allows learners to engage in a form of "condensed reading" to focus on specific language features (Gabrielatos, 2005, p. 10). With this in mind, the considerations taken in creating corpus based exercises to ensure this type of condensed language exposure can now be discussed.

Three important considerations for creating DDL exercises

When attempting to integrate corpus materials into the language learning classroom (as is one of the main goals of this thesis, see Chapter 1 and 4), there are many factors to take into consideration. These considerations are guided by the aim or purpose of what the teacher wants to teach, for example, beginner level or academic vocabulary, and the collocations associated with these. The first consideration lies in choosing a corpus (Oghigian & Chujo, 2010). The second in controlling the tasks set, and the third in bridging the gap between language research and language production (Oghigian & Chujo, 2010). I will present an overview of the types of considerations here as they are pertinent to the examples of work undertaken by other researchers in the field as presented below. This discussion is returned to in Chapter 4 where I present an overview of the decisions made in implementing an everyday academic writing course for German Studies at Rhodes University.

Choosing a corpus: size, relevance, accessibility

Whether teaching to exploit corpora or exploiting corpora to teach, one needs a corpus as the base for learning and materials creation. There are three factors to consider when choosing an existing corpus, and these are size, relevance and accessibility. Of course, as already stated, there are many resources of this kind available for English and not as many, or none, available for others. Where this is the case, a corpus will need to be created by the teacher or researcher, specifically for this purpose

(Tribble, 1997). In this scenario the following considerations remain important, but the task itself becomes more difficult to achieve, due to copyright issues and accessibility of relevant texts.

Size is the first important consideration as one does not want to "drown" the learners in data (Tribble, 1997, p.2), especially those who are less proficient. Nor does one want a corpus which is too small to provide learners with enough instances of examples to see a pattern (Oghigian & Chujo, 2010, p. 201). The internet for example has been hailed as a great source for corpus consultation (Oghigian & Chujo, 2010, p. 201), however the internet knows no bounds, and often information may be unsuitable or irrelevant to learners, creating a frustrating learning experience for all. The internet may be put to better use to find already existing web-based corpora and corpus tools created for the target language. This leads to the second consideration in choosing a corpus, which is *relevance*. The corpus chosen should also be relevant to the learners, both in terms of content, as well as register. For example, when creating an academic writing course for German (such as is the focus of this study, see Chapter 4) the corpus chosen should not be one of children's language, or of short stories, but rather of texts which are academic in nature, and written by mother-tongue speakers of German. These texts must also be *accessible* to the learners, either through the internet or as a downloadable file.

Controlling the task: 'a guided dynamic partnership'

When one has a specific learning goal in mind, there is little use in simply 'setting students loose' on a corpus, without any guidance or direction. This is all very well for advanced students of linguistics with a knowledge of corpus tools who wish to research certain language phenomena, and pursue "student-designed" and "interest-driven" discovery learning in their own time (Oghigian & Chujo, 2010, p. 202). However, beginner learners will not be able to cope with this sort of level of inquiry (Oghigian & Chujo, 2010). As stated in Chapter 2, since the 1990s there has been a change in the role of teacher and learner from "teaching as imparting knowledge" to "teaching as mediated learning" (Laurillard, 1993, p. 13-15). The teacher and learner must work together in a "guided dynamic partnership" (Oghigian & Chujo, 2010, p. 202), and the language teacher should guide searches which will be fruitful and informative, revealing patterns which can be explained and fit into the context of the learning aims which teachers have. See the examples relating to the research performed by Dodd (1997) in the sections below.

Language researchers must become language producers: corpus resources alone are not enough

Corpus methods of language learning are not in themselves enough for language learning, and must be used in conjunction with other teaching methods, so that learners may make the move from being language researchers to language producers. Oghigian and Chujo identify that "one of the inherent disadvantages in corpus-based learning is that it tends to underscore the idea of L2 as "the study of"

as opposed to "the use of"" (Oghigian & Chujo, 2010, p. 202). A teacher may find it difficult to encourage students to produce texts when they have been "more comfortable with an arm's length analysis of language" which is typical of a corpus approach (Oghigian & Chujo, 2010, p. 202). It is for this reason that Oghigian & Chujo (2010) state that it is "vitally important to have a communicative aspect in which to apply acquired knowledge, whether it is in a spoken or written context" (Oghigian & Chujo, 2010, p. 202). The state of moving from corpus to discourse has been advocated by Braun (2008) who highlights the importance of using multiple modes of learning, including "comments and exercises, explanations, exploratory tasks, study aids and didactic hints for learners and teachers" (Braun, 2008, p. 55). This aspect is highlighted in Chapter 4, where research relating to the concept of writing more, more often (Estes, *et al*., 1998; Thomson, 2013; Torgesen, *et al*., 2007) in order to improve writing skills, is expanded upon.

Teaching about corpora, the first step in exploiting to teach

Along with access, one needs a certain technical skills set in order to work with any corpora and related software programmes. One requires a certain amount of training and knowledge about what a corpus is, how it can be manipulated, what the various manipulations mean, and how to perform these operations. Teaching students about corpora is invaluable as it is the first step in all direct use of corpora in language teaching. Teaching students about corpora is usually undertaken during a set course at either undergraduate or postgraduate level where the origins, development, uses, and, most importantly, the practical skills of corpus linguistics are explicitly taught. This training is quite easily accomplished, but it is necessary. One cannot simply use, or expect students to use, corpora 'intuitively'. Large amounts of data can be overwhelming, and low proficiency users may struggle to comprehend the corpus data, making a direct corpus based approach far better suited to advanced learners.

Examples of computer-based and paper-based exercises

The following section provides examples of how researchers and teachers have used computer-based and paper-based exercises in German and other foreign language learning classrooms, and some advantages and disadvantages of each approach will be presented.

Computer-based, a case study

Computer-based direct corpus consultation, or teaching (learners) to exploit corpora, is often argued for as the preferred mode of corpus integration into the language classroom with statements such as the following made: "By learning to interact with corpora, students find themselves learning a great deal about language, and how to study language" (Leech, 1997, p.9). Consider a study undertaken by Dodd (1997, p. 131) where he relates some of his own experience of 'teaching to exploit' a corpus of

written German for advanced language learning at undergraduate level at the University of Birmingham. Dodd (1997) asserts that 'raw' corpus data (concordance files in unedited form) provide a powerful and simple tool for language learning, and can be exploited with little or no adaptation for advanced foreign language learners. Dodd (1997) made use of the *Bonner Zeitungskorpus* ('Bonn Newspaper Corpus', BZK) for teaching applications accessed through the concordancer MicroConcord on the computer network at the Birmingham University School of Modern Languages. Dodd and his colleagues were particularly fortunate to have experts such as Tim Johns available to help format this 3 million words of running text (taken from East and West German national and local newspapers, over the period from 1949-74) by removing certain tags inserted by the creators, to help make the data appear more natural for learners. Dodd (1997) explains three ways in which he exploited the corpus for teaching purposes: firstly, as an informal resource for students to browse through, secondly as a resource for forming explicit grammatical knowledge, and lastly as a resource for student-led projects on grammatical or socio-linguistic projects. The following examples relate to the first instance, and illustrate how vocabulary may be learnt in context through guided corpus searches.

Dodd (1997) tasked his students to browse through 'raw' corpus data, with a specific goal in mind e.g. vocabulary exercises such as to look for all instances on words ending in various suffixes (e.g. -*freundlich*, -*feindlich*, -*nahe*, -*fern*) and to report back six words which they had not previously known and explain them to their group. German is particularly well suited to morpheme-based exercises such as this. Another exercise could be to check a particular feature of language such as the syntax of modal verbs, where students perform a simple search of, for example, *h*tte* within three words of *sollen.* As Dodd (1997, p. 133) asserts, this type of exercise provides far more variations in the examples and 'contextual clues' of usage, than a textbook would. This is again what Sinclair was so fervently advocating: *use real examples only*. Whilst Dodd maintains that this sort of file could be easily compiled by students who have access to a corpus, were this type of open access not the case, teachers could distribute printouts of the concordance lines of these lists and perform the same task. Information regarding frequency is also a rather important aspect which can be exploited quite simply. Dodd (1997, p.133) provides the simple example that three translation equivalents of 'complete/ly', namely *total* (24), *absolut* (70), *völlig* (409) occur with very different frequencies in the BKZ, which indicates a difference in usage. These are not interchangeable, and looking at these words in context will provide examples for how each is used. Dodd (1997) states that in these types of exercises "formal grammatical description will usually play a secondary role, the main emphasis being on observing text patterns in order to improve students' accuracy and extend their competence in language" (Dodd, 1997, p. 133).

Thus, Dodd provides examples of how he used a corpus of written German to teach vocabulary in context. However, Dodd does not expand upon what effect this has on student performance and understanding, or indeed on the students' experience of using these methods, and whether they found them to be effective.

Examples of paper-based corpus exercises

The benefit of having pre-made exercises or making exercises oneself and acting as an intermediary is that if one wants to avoid having students actually confront (or be confronted by) a computer, similar results can be achieved by providing students with paper printouts from the corpus (Boulton, 2008). Without hands-on knowledge of corpora and corpus software, self-access exploitation of corpora by students themselves is naturally limited. This option avoids having to sit students down in a computer laboratory and teach them how to use the computer software themselves to obtain their own data. This method has been criticized as an 'easy way out', as some say that it does not produce the same long term results as 'teaching to exploit' learning sessions do (Leech, 1997, p. 10). However, as Leech has noted, "'the easy way' ensures that a maximum number of students are able and willing to participate in this kind of learning experience" (1997, p. 10).

Teachers can still encourage all students to pursue their own self-access mode to corpus research in their own time, and those who are interested and willing can be provided with access to the corpus and the software. Certainly, in the realm of foreign language teaching, paper-based exercises also ensure that the sections of corpus data that students are exposed to are manageable, and at the appropriate level for learning. Furthermore, in a country like South Africa, where not all learners have unlimited access to a computer, the internet or even electricity, paper-based exercises certainly present an advantage. Data presented by simple corpus investigations (such as a search for a frequent item viewed it its sorted concordance lines) can often be pages long, and very overwhelming for a beginner or intermediate second language learner. Oghigian and Chujo (2010, p. 202) assert that it is often far more helpful (and time effective) for students if a teacher provides pre-vetted guided searches.

Here we can take as an example the paper-based exercises developed by Krummes and Ensslin (2014) based on findings from the FALKO L1 and WHiG corpora. Krummes & Ensslin (2014, p. 201) state that: "Although spelling, grammar and lexis presented no urgent problems for WHiG participants, they did not use much formulaic language and collocations". Krummes & Ensslin (2012) developed a worksheet which presents five keywords or TAG words (Typical Academic German words) for academic writing in German, and corpus-based exercises for each of these, aiming to develop not only isolated vocabulary knowledge, but also collocation awareness. Krummes & Ensslin (2012, p. 110) observed

that learners sometimes "rely on overusing 'safe' but repetitive language strings or create grammatically sound strings that native speakers would not, however, use". These inauthentic phrases included the following:

to introduce the essay topic: *In diesem Aufsatz werde ich über* x *schreiben* ['In this essay I will write about *x*']

to give examples: *Nehmen wir als Beispiel* x ['Let us take as an example *x*']

to express a personal stance: *Ich bin der Meinung*, *dass* ['I am of the opinion that']

to verbalise a conclusion: *Zum Schluss kann man sagen*, *dass* ['At the end one can say that']

These inauthentic, often directly translated phrases do have authentic counterparts which are used by mother-tongue speakers. Krummes & Ensslin (2012) formulated a worksheet which is based on previous findings from Webb & Kagimoto (2011 cited in Krummes & Ensslin, 2012) that "collocations are best learnt with fewer node words (i.e., keywords) and more examples" (Krummes & Ensslin, 2014, p. 201). These exercises are therefore based on the following three ideas: 1) teaching formulaic language should be explicit, 2) teaching collocations should be explicit, and 3) providing a good number of examples is important. Krummes & Ensslin (2012) provide a 5-5-5 approach in which a worksheet ideally "presents five node words (named 'keywords' in the handout[13] 'TAG words' hereafter) with five formulaic sentences ('phrases') per node word and five concordances ('examples') per phrase" (Krummes & Ensslin, 2014, p. 201). The five TAG words documented were chosen by the researchers for their "ubiquitous use by German scholars and in UK Higher Education German-language settings" (Krummes & Ensslin, 2014, p. 201) and are listed below (see Appendix 4 for a copy of the handout).

Typical Academic German words, TAG words

· *Zweck* ['aim']: in introductions (essay discourse structure)

· *Beispiel* and *beispielsweise* ['example']: in the main body (essay discourse structure)

· *Erachten* and *Ansicht* ['opinion']: in the main body (personal stance expression)

· *laut* and *zufolge* ['according to']: in the main body introductions (impersonal stance expression)

[13] The keywords from the WHiG worksheet will be referred to as TAG words (Typical Academic German words) in this thesis to avoid confusion with the corpus linguistic concept of a 'keyword' which refers to a word which is used significantly differently across two corpora (Baker, 2010).

· *Fazit* [‘conclusion’]: in conclusions (essay discourse structure)

The advantages and disadvantages of a corpus approach to teaching and learning

The above section has provided a background of corpus linguistics and an argument for the use of corpus tools in the area of language teaching. This has highlighted the many positive aspects of using corpus linguistics, such as: real-life, authentic examples as opposed to invented examples (textbook English) (Sinclair, 1997, p. 31), inductive, discovery-based learning (Laurillard, 1993; Leech, 1997), learner autonomy, teacher empowerment, refinement and improvement of existing teaching materials, empirical data for the development of new courses (Language for Specific Purposes- LSP) as well as providing empirical data/ insights into second language acquisition.

There are also, however, many constraints when considering a corpus-based approach to research and teaching. First and foremost, corpus methods require access to technology, and more specifically, access to corpora and related software programmes. In a context where there is a lack of access to resources such as computers and the internet, this type of approach would be futile. Corpus linguistics by its very nature is only truly successful when a large corpus of text is able to be electronically compiled and manipulated. Where technology is absent, in rural parts of Southern Africa for example, a corpus-based approach would be a life-long endeavour similar to that of the unfortunate monks under Hugh Saint-Cher in the 13th Century. Moreover, even if one has technology, corpora are not readily available for languages other than English. A compilation of a corpus is time consuming, and beyond the scope of many teachers.

Breyer (2009) states, despite the “undiminished enthusiasm in the research community” it has become an issue of growing concern that “the application of corpus tools and resources in the classroom remains limited” (Breyer, 2009, p. 153). Whilst many teachers and students may indirectly consume corpus-based educational products such as dictionaries and grammars, “few teachers are clear about the nature of corpora, or their significance for language teaching, and fewer still have ever made direct use of a corpus” (Gabrielatos, 2005). Similarly, Aijmer (2010, p. 1) notes that when it comes to the realm of teaching, corpus linguists are “generally enthusiastic” about what corpora have to offer, “however, the use of corpora in the EFL classroom is a rare occurrence and teachers are still unwilling or lack the skill to use corpora as an aid to get insights into English” (Aijmer, 2010, p. 1).

One needs to know about corpora in order to use them. Aijmer (2010) states that teaching about corpora is in fact often neglected, and that teaching about corpora is essential if one wants to inform students and teachers, not just other researchers in the field, about the benefits of corpora and what corpora can do. Aijmer (2010, p. 2) makes the claim that: “Although courses in corpus linguistics are sometimes included in the university curriculum the direct exploitation of corpora in the EFL classroom

is unusual and the impact of corpora on syllabus and materials design has been slight". As I have observed, there seems to be only a vague idea of what a corpus is at many language departments/ sections in South Africa and what it can be used for, even if many language teachers have inadvertently been using informal language 'corpora' of their own students written work throughout their teaching careers. What is meant by this is that teachers form a 'mental map' of the typical errors made by learners, perhaps even particularly by those made by students of a particular mother tongue or age group or social class, and the typical style of writing that their students produce (Gabrielatos, 2005). This is done by reading and mentally cataloguing student writing pieces over the years. As Granger asserts, there is "nothing new" (Granger, 2004, p. 123) in the idea of collecting learner data, teachers have been doing this for many years. However, as with the monks of ages past, it would take a teacher a lifetime to index all instances of every word in every student essay filed away over the past 20 years, where a computer could do this for them in a matter of minutes, pulling up a keyword in context and showing every instance of the patterns of learner writing to be had at their fingertips. That is of course, only if the teacher had any idea of what a corpus was, how to construct one, what software to use, and how to interpret the information gleaned through the process, sifting relevant examples of the term from irrelevant ones (Gabrielatos, 2005).

For the above reasons, I would like to express my agreement with Leech's sentiment that the notion of the term "corpus revolution" (Leech, 1997, p. 22) which was often used to describe the flood of new research as a result of corpus based/driven studies, was in fact "a misnomer for a change which is taking place gradually, as suitable materials become available" (Leech, 1997, p. 22). New research based on corpus methods is published each year, each shedding new light on applications to teaching and learning, and aiding better understanding of mother-tongue and second language writing. Furthermore, new technology is constantly being developed, and corpus software is not infallible: they (various computer software) often have glitches, may malfunction, or 'crash' when needed most (see Yoon, 2008), and they will not do what is required of them without explicit instruction. That which was true 20 years ago is still true of our situation today. Computers and corpora are still "comparatively expensive resources" (Leech, 1997, p. 23), and education is still "generally underfunded" (Leech, 1997, p. 23) as it was.

The role of the teacher

Whether one is teaching students about corpora and corpus software, guiding students through DDL exercises on the computer, or creating on-paper exercises based on pre-vetted concordances; these pursuits are at the heart of the convergence between corpora and teaching (see Leech, 1997). The role of the teacher in implementing the use of corpora in classrooms is vital, and this will be briefly

discussed as no corpus-based intervention can be possible without a teacher who has both motivation and a necessary skills set to teach with corpora.

The role of the teacher and the specific challenges which may be faced by teachers, in implementing a corpus approach to learning are not significantly outlined in the literature, which usually places the focus on the role of the learner in using corpora. This is surprising, given that the direct use of corpora in the classroom was observed by Johns (1988, p. 14) as creating "a shift in the traditional division of roles between student and teacher" and has been identified as an issue which teachers may have trouble with coming to terms with (Breyer, 2009). Whilst corpus based/driven 'discovery' learning is understood as "student-designed", and "interest driven" (Hunston, 2002), the teacher's role remains an important one, as the teacher and student must work together in a "guided dynamic partnership" as stated in the sections above (Oghigian & Chujo, 2010, p. 202).

Breyer (2009) found that the key factors that influence the teacher's decision to implement corpora are "motivation, availability of materials and the possession of adequate skills to teach with corpora" (Breyer, 2009, p. 154). It is becoming increasingly clear that this experience needs to be quite substantial; that is, teachers need to have a sound understanding of corpus analysis in order to teach with it. This leads to the question as to when and how teachers should acquire such 'corpus literacy' (Mukherjee, 2006, p. 179). Time constraints both in learning the necessary skills, as well as teaching them in class, pose a serious challenge to popularising corpus consultation, particularly at the tertiary level (Chambers, 2005, p. 113).

One important factor which could greatly alleviate this task (which has been identified, but not sufficiently addressed in the literature) is the creation and availability of pre-ordained corpus-based exercises (Al Saeed & Waly, 2010; Oghigian & Chujo, 2010; Breyer, 2009). The "limited availability of ready-to-use corpus teaching materials" (Breyer, 2009, p 155) contributes significantly to the challenge with which teachers are faced when trying to implement corpus methods in the classroom. This is as "currently, direct corpus applications are neither a standard component of curricula nor textbooks" (Breyer, 2009, p 155), which means that teachers must then face the task of producing their own brand new exercises, a task which those in Breyer's (2009) case study found considerably challenging.

Even when pre-ordained corpus-based materials for the target language of choice do exist, and provided one has access to them, they "still leave the teacher with a number of challenges to overcome" (Breyer, 2009, p156). The teacher has to "assess the materials for their appropriateness in the respective learning context; that is language proficiency level, suitable vocabulary, integration into the curriculum and so forth" (Breyer, 2009, p. 156). Thus, a teacher must face the challenge of

successfully understanding and integrating these materials into the classroom, which requires a certain amount of 'corpus literacy'. Where materials do not exist, a teacher then has the choice and challenge to create their own. Challenges include: "finding or creating an appropriate corpus, acquiring and learning how to work with concordancing software, creating meaningful exercises, and producing worksheets" (Breyer, 2009, p. 156).

Thus, in summary of the above section, a teacher who wants to implement the use of corpus tools in the classroom should ideally have a good knowledge of corpus linguistics and corpus tools themselves, as well as a good understanding of the linguistic structures of the language which they are teaching, and motivation to do so. In addition to this, they should have the necessary motivation to embark on the journey of implementing a new type of methodology.

Potential benefits and students' opinions

As has been expanded upon above, researchers are very optimistic about what corpus tools can offer language learners. One method of assessing the successfulness and value of corpus tools is by asking students what their experience of using them has been like, and what benefits they perceive through their own use of corpus tools. Students' attitudes towards the use of corpus tools in learning to write in a second language provide valuable insights into the usefulness and value of corpus tools in the classroom.

Oghigian & Chujo (2010) performed a longitudinal DDL case study at Nihon University. Oghigian & Chujo (2010) created a vocabulary list by comparing two tests for English. Based on exercises implemented in 2005, where vocabulary was taught using DDL methods (2 classes per week, 90mins each), participants displayed enthusiastic feedback (Oghigian & Chujo, 2010, p. 206). Furthermore, participants sowed gains between pre and post test scores. Following this initial study, in 2008 and 2009 Oghigian & Chujo (2010) taught the same vocabulary to two groups of students using different methods of DDL, namely paper-based and computer-based data-driven learning. Feedback from both groups of learners highlights the advantages of each approach.

Responses to DDL were "generally positive" and the researchers stated that "Most students found the DDL course to be novel, appealing and useful, and found value in the related activities" (Oghigian & Chujo, 2010, p. 208). Students who showed a preference for computer-based work in Oghigian & Chujo's study declared that it was "very powerful" and that the learning "fixed in their memories". (Oghigian & Chujo, 2010, p. 212). These students appreciated the control they had over their own learning. Participants who used paper-based materials reported that this saved a lot of time as more tasks were able to be completed in class (Oghigian & Chujo, 2010, p. 212).

Similarly, Varley's study of the integration of corpus tools in a second-year undergraduate English as an Additional Language classroom shows that "Students generally had a positive response to corpus consultation and were able to identify benefits clearly, particularly in the areas of vocabulary acquisition and increased awareness of syntactic patterns" (Varley, 2009, p. 133). In Varley's (2009) study, most students indicated a likeliness to use corpus resources in the future, particularly if they possessed clear goals for their language learning. On the other hand, some students found the volume of data elicited from a concordancer overwhelming, and raised the issue of not "trusting" the corpus as they would trust the dictionary (Varley, 2009, p. 143).

While Oghigian & Chujo (2010), and Varley (2009) were concerned with teaching vocabulary using concordancers, and had many students in their data pool, Yoon (2008) performed an in-depth investigation of six students' writing experiences, investigating the changes in students' writing processes associated with corpus use over a long period of time. "Once the corpus approach was introduced to the writing process, the students assumed more responsibility for their writing and became more independent writers, and their confidence in writing increased[14]" (Yoon, 2008, p. 31). Yoon's study reveals the very personal aspect of teaching and learning. Students are individuals and will experience corpus use in individual ways. Some may have a genuine interest and enthusiasm for corpus tools, while others may perceive it to be much the same as a dictionary: "a necessary inconvenience" which they use because they have to, and ideally hope to be able to write without it (Yoon, 2008, p. 38). A particular finding of Yoon's (2008) study was that students employed a need-based attitude towards corpus use, i.e. the more writing assignments they had to complete, the more they turned to the corpus for help. This is a particularly poignant finding as even though students may perceive the corpus to be very useful, "lack of meaningful engagement with writing would limit its appeal" (Yoon, 2008, p. 44). This reflects what (Oghigian & Chujo, 2010, p. 202) assert, that language students must become language producers in order to cement knowledge.

These studies (as well as others) thus show a general trend of a positive assessment of corpus tools on the part of the students who use them. Raised levels of confidence in writing, and increased linguistic awareness are noteworthy outcomes for the language learner involved with corpus linguistic tools. In Chapter 5 the experience of the participants' of this study in using corpus tools is explored, and these trends are confirmed within our unique learning context.

Chapter summary and conclusion

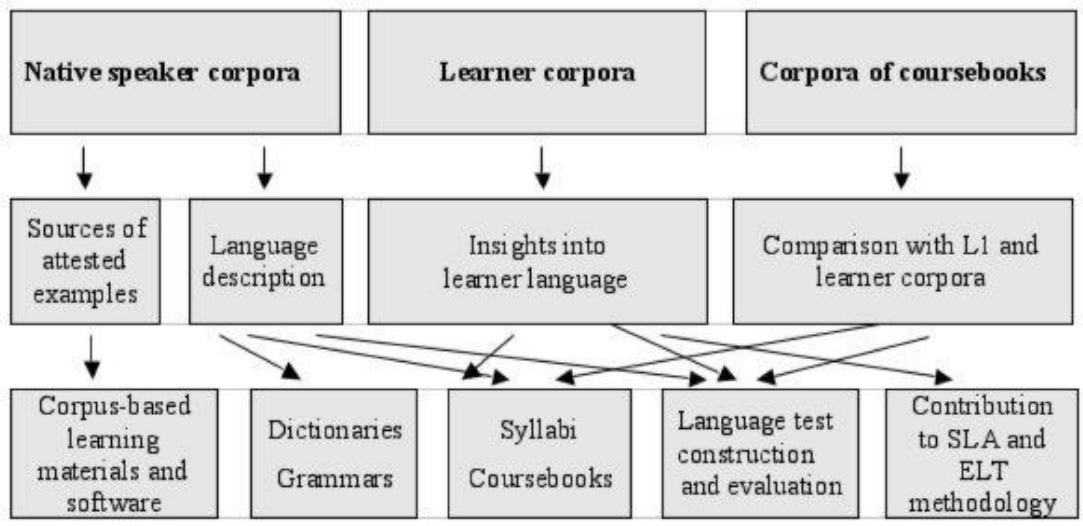
[14] This finding was mirrored in our own participants' responses to using corpus tools for language learning (see Chapter 6: Reflecting on Writing, the Participants' Perspectives).

This chapter has explored what constitutes everyday academic language in German, by exploring the specific metaphorical language used and the importance of correct lexical choices. Formulaicity has been discussed as an important dimension of learner writing, as it has been shown that leaners often underuse, misuse or overuse formulaic sequences in comparison to mother-tongue speakers. Learner corpora have been identified as an important tool for describing and analysing learner language, especially as it compares to mother-tongue use. Where traditional error analysis has been criticised for a number of methodological weaknesses, it has been shown how corpus methodology has adapted and improved upon this field of analysis in order to better describe learner language within specific parameters (in the field of SLA), and provide a base for better teaching methods (in the field of FLT).

Two types of corpus-based approaches to language teaching (computer-based and paper-based) have been discussed, and the respective advantages and disadvantages of each outlined. It has been shown that whilst computer-based language learning was viewed to offer better opportunities for inductive learning in the past, this is not always possible due to both practicalities and time constraints. Exercises in the form of worksheets have proved to be very beneficial in the language classroom (Krummes and Ensslin, 2012). There are various challenges facing the teacher who aims to implement a corpus-based approach, however, with sufficient motivation, knowledge and access to resources, good results can be achieved. This is shown in the positive evaluation of corpus tools by students who have been taught directly using corpus methods.

Gabrielatos (2005, p. 4) provides the following useful diagram which summarizes and illustrates the many interconnecting ways in which corpora are relevant to language teaching, as have been discussed in the sections above.

Figure 2, Corpora and ELT



CHAPTER 4. RESEARCH METHODOLOGY

Qualitative or quantitative approaches

Whilst there have been both qualitative and quantitative studies into the implementation and effect of corpus based methods with language learners (Varley, 2009), qualitative studies are the preferred approach (Chambers, 2007), as they present researchers and practitioners in the field with a good perspective on what has been done, and what does and does not work in specific settings (Chambers, 2007). Following this trend, this research assumes for the most part a qualitative approach; that is to say that there is a focus placed on the process of learning, by trying to understand how proficiency is increased, what the participants think about the improving their proficiency, and how the setting influences them (Heigman & Croker, 2009, p. 8; Terre Blanche, *et al*., 2009). This is achieved within the 'academic literacy' and 'German as a foreign language' acquisition paradigms. Certain sections of this methodology further aim to assume a quantitative approach, seeking to "measure gains in proficiency over a period of time" (Heigman & Croker, 2009, p. 8) in addition to the qualitative analysis of the process.

A mixed methods approach

This implementation of both qualitative and quantitative methods within one study, in the field of applied linguistics, has been termed a "mixed methods" approach (Bryman, 2012; Heigman & Croker, 2009, p. 15; Terre Blanche, *et al*., 2009). Mixed methods research is an emerging field of study, and Ivankova & Cresswell state that "with its focus on the meaningful integration of both quantitative and qualitative data, [mixed methods research] can often provide a breadth and depth that a single approach may lack by itself" (Ivankova & Cresswell, 2009, p. 136).

In this methodology, I expand on research performed for an Honours paper in 2013[15] (Ortner, 2013), which was intended as a pilot study for this dissertation. The aim of both the pilot study, and the main body of the research, was to create a learner corpus of German writing at Rhodes University, and furthermore to institute corpus-based teaching and learning practises in the German Studies section of the School of Languages at Rhodes University, using insights from existing mother-tongue (FALKO L1) (Zeldes, *et al*., 2003) and learner corpora (WHiG, FALKO L2) (Krummes & Ensslin, 2012; Krummes

[15] As is in the nature of an Honours dissertation (max. 10000 words), the initial Honours project which laid the foundations for this thesis was limited in its scope. I sought only to create a corpus of learner writing to identify common patterns in their writing. This was useful as it provided a base for measurement, and highlighted many issues such as availability of data, the technicalities involved in creating a learner corpus, and the need for better didactics in implementing an academic writing course. The resulting methodology as outlined above is largely based on insights gained through the undertaking of previous research (Ortner, 2012; Ortner, 2013).

& Ensslin, 2014) to improve writing. Following this, an assessment of the value of a corpus-based academic writing course for German as a foreign language in South Africa is performed. The methodology is designed to be carried out in three stages. These three stages are not isolated, and many will overlap, as they inform one another to ensure triangulation of the data.

4.1 Research Goals

This research intends to achieve a number of specific goals, as set out in the introduction to this thesis:

1) Identify the writing attitudes, methods and strategies employed by our students when writing in German;

2) Through the implementation of a course that uses and adapts the tools and methods of corpus linguistics improve the writing of students in German at Rhodes University;

3) Assess the attitudes of German Studies learners' towards the use of corpus linguistic tools in developing academic literacy;

4) Evaluate the corpus-based course in terms of its usefulness and effectiveness by comparing the writing of participating students both before and during the writing course with reference to five key everyday academic collocations, as well as to mother-tongue writers of German.[16]

[16] FALKO-L1 is a corpus of mother-tongue German writing in the form of argumentative essays written by mother-tongue German school leavers, and is available under a Creative Commons License which can be accessed via the following link: http://creativecommons.org/licenses/by/3.0/deed.de. Access to FALKO-L1 was granted by Prof Dr Anke Lüdeling (Korpuslinguistik und Morphologie, Institut für deutsche Sprache und Linguistik, Humboldt-Universität zu Berlin)

4.2 Participants

Participant overview

A total of 21 participants were involved in this research over the period of three years, from 2013, to 2015. However, the discussion will focus on a primary group of 13 students, as five of the 21 were involved only in the pilot study, and three were excluded from the study for reasons which will be elaborated upon below.

Of these 13 primary participants, three were male, and 10 were female. Participants had an average of two to three years of prior German exposure. As our students are generally *ab initio* students who (by definition) have no prior knowledge of the language upon arrival in their first year, the primary group of participants are third year students, majoring in German Studies, but the participant group ranges from second year to Honours level (fourth year) German Studies students. As the main study was performed over two years, four members of the second year group of 2014, and one member of the 2014 third year group were able to continue the course in 2015. This continuation has provided particularly good insights, and has allowed for a longitudinal analysis of both the writing ability and the improvement of the writing of these five students.

Excluded participants

Three participants had to be excluded from inclusion within the learner corpus. These participants were P8, P9, and P17. P8 was excluded as she performed exceptionally badly on the writing assessments, and did not meet the level required according to the design criteria developed for the RUDaF learner corpus (Minimum Level A2 according to the CEFRL, see "factors pertaining to the learner" in creating the RUDaF Learner Corpus below). P9 and P17 were excluded as they had many prior years of exposure to German, and almost mother-tongue levels of proficiency, thus they too did not meet the requirements in order to be included in the learner corpus. These three students completed all writing assignments, questionnaires and interviews; however, this data has not been included in the discussion. Participant numbers still range from P1 to P21 in the discussion, with P8, P9, and P17 notably absent. P01, P03, P04, P05 and P07 were involved in the pilot study, but not the main research project.

Language Backgrounds

Of the primary group of 13 participants, 9 identified themselves as English mother-tongue, two as Afrikaans, one as isiXhosa, and one as XiTsonga mother-tongue.

Other languages that the participants identified themselves as being able to speak other than their home language, English or German, included: Afrikaans, French, Chinese, isiXhosa and isiZulu.

The following table illustrates the above information with regard to each individual participant. This information is represented exactly as the students themselves wrote it in Questionnaire 1 (see Appendix 6). It needs to be noted, however, that some of the participants' self-reporting is inaccurate. This can be easily verified by comparing the participants' answers on the questionnaire to other administrative data (which for confidentiality reasons cannot be reproduced here). For example, P15 speaks English (very well) and obviously German as a Foreign Language. P11 who asserts that they had started learning German in 2000, later confirmed that this was only in an informal setting. While P11 was able to orally express some colloquial phrases at the beginning of their first year, their writing remained at the level of their ab initio peers; that is level A2 for second year. The level of German competence displayed by P11 thus conforms to the design criteria set out for the RUDaF learner corpus. For this reason their writing has not been removed from the learner corpus.

Table 1, Participant Overview, Represented as Participants Have Self-Reported

Participant No.	Home language	Other Languages	Began German in
P02	English	Afrikaans	2011
P06	English	German	2010
P10	Afrikaans	English, French	2008
P11	English	German, Afrikaans, French	2000
P12	English	Afrikaans, basic French	2013
P13	English	Afrikaans, Chinese	2013
P14	isiXhosa	English, Afrikaans, isiZulu	2013
P15	Afrikaans		2011
P16	English	Some Swedish, German, some Zulu	2012
P18	English	Afrikaans, German, basic Chinese (Mandarin)	2012
P19	XiTsonga	English, SeTswana, isiZulu, Afrikaans, German, isiXhosa	2012
P20	English	German	2012
P21	English	German, Afrikaans, French	2012

Consent

All participants were asked to fill in an informed consent form at the beginning of the data collection, acknowledging their participation in the study, as well as their anonymity as participants (see Appendix 1). In an introductory lecture, participants were told explicitly that their writing would be used for the formation of a learner corpus which would potentially inform the teaching methods for future years. All participation in questionnaires and interviews was also done with the explicit consent of the participants who acknowledged their role and anonymity in the study.

4.3 Procedure

A short summary of the procedure

In order to address the research goals as set out above, the data for this dissertation was gathered as follows, and can be set out succinctly in 5 points. These will be expanded upon in the sections below.

1) Participants were asked to complete set writing tasks in German alongside their normal assignments. Writing pieces were assessed according to the CEFRL criteria, and feedback was provided to the participants.

2) The writing pieces collected were used to begin creating a learner corpus of German writing at Rhodes University. The learner corpus was used both to compare the writing of our participants to mother-tongue German speakers, as well as to internally compare the nature of our students writing before and during the writing course in order to explore the usefulness of a corpus-based approach to learning new German academic vocabulary in the South African tertiary setting.

3) Participants completed a questionnaire which aimed to assess their attitudes towards, and experiences of, writing in German. The questionnaire was followed by an interview in which specific themes from the questionnaire were elaborated on.

4) A series of lectures was introduced, based on corpus data form the WHiG project and aimed at specifically introducing key everyday academic German vocabulary and phrases which could be integrated into the writing pieces. Participants were given further writing assignments which were also added to the learner corpus.

5) Participants were again asked to complete a questionnaire. This second questionnaire aimed to assess the participants' attitudes towards the writing course, and their experience of using corpus-based materials. This was used to evaluate the successfulness of the course.

Writing more, more often – a rationale based on the literature

This section provides a short rationale for the first section of the methodology, which was to have our participants write weekly paragraphs in German. The aim of this was threefold: firstly, to elicit data from the participants in order to gauge at what level they were writing, and secondly, to improve their writing skills in German through the process of writing more, more often (Estes, *et al*., 1998; Homstad & Thorson, 1996; Thomson, 2013). Lastly, to provide students with a platform to engage with the vocabulary and phrases learnt in the writing course, to encourage the move from language research, to language production (Oghigian & Chujo, 2010).

Various scholars and teachers have described their approach to confronting the challenge of learning to write in German (Homstad & Thorson, 1996; Estes, *et al*., 1998). Estes (1998) maintains that creative writing is an important aspect of all her German classes, right through from first to third year level, and suggests that this constant experience of writing reinforces the learning of vocabulary in context. Estes begins this process with her students by setting 25 to 30 word descriptive *Aufsätze* (paragraphs or essays) for her first-year students, and then expands these *Aufsätze* in the second-year classes to include opinion pieces where students must not only describe things but also present and defend an opinion. This is a "crucial time in their mastery of German" (Estes*, et al*., 1998, p. 69) as they are supposed to write in German, but their thought processes are still in English, and many students may produce lexical and syntactic oddities such as those described by Jaworska (2011), resulting from a direct superimposition of English onto the German language. Estes maintains: "It is apparent immediately when students think in English and translate because their German comes out awkward and strange" (Estes, *et al*., 1998, p. 69).

Estes employs a method of writing and rewriting, tasking her students to hand in a first draft where she highlights mistakes which the students much correct, then hand in again, and then have marked with corrections and some comments by the teacher, followed by a third and final re-submission which is fully marked and corrected by the teacher (Estes, *et al*., 1998). Lopez-Mayhew employs a similar approach, and is particularly concerned with the idea of students writing out a paragraph in their mother-tongue first, as this leads to complicated translation tasks, rather than a stimulation of thought processes in the target language (Estes, *et al*., 1998, p. 71). Frequent writing was viewed by all three authors to significantly contribute to their student's written and oral language competency. Whilst this claim may be subjective as no supporting data was presented, I find it nonetheless valuable as an indicator of success, as teachers constantly have to form subjective judgements of student performance and tailor their classes based on these judgements.

Similarly, in another study which looked at the effects of introducing writing-to-learn activities in the German foreign language classroom, researchers were unable to provide conclusive quantitative results of improvement as they "found it impossible to do a systematic analysis of the graduation proficiency exams, as there were too many variables to isolate the effects of intensive writing" (Homstad & Thorson, 1996, p. 6). Although Homstad & Thorson were unable to quantify their results, their report states that, similarly to Estes *et al.* (1998) findings, the general impression formed by the researchers was that there was an increase in fluency and confidence as compared to previous years (Homstad & Thorson, 1996, p. 6).

In later sections of this methodology chapter, and again in Chapter 5, I expand on the formal assessment of our student writing with reference to specific typical everyday academic German words and collocations taught during the writing course, so as to both quantitatively and qualitatively provide specific evidence of where and how our students improved as a result of writing more, more often. In this way, using corpus linguistic methods, I hope to provide a more objective and measurable approach than that of Estes *et al.* (1998) and Homstad & Thorson (1996) when ascertaining how writing more, more often has affected our students.

This idea of staggered writing pieces, and of writing more, more often, is one which has been recognised as a significant contributor to the academic literacy project. The writing intensive courses introduced in Higher Education institutions across South Africa seek to do exactly this, to task students to write more frequently, within their disciplines, and to receive feedback on this writing (Thomson, 2013). Whilst writing courses may provide examples of how to construct an argument, and what types of phrases to use, it is important for students to have an opportunity to put this type of learning into practice. Through practice, and the process of writing their own texts and constructing their own arguments, students are able to make the move from studying language to producing language (Oghigian & Chujo, 2010). Multilingual writers are also made more aware of the writing conventions across languages, and become able to move between languages with more ease (Canagarajah, 2011).

4.3.1) Writing Tasks

In order to gauge the level at which our students were writing, and if they were capable of producing short everyday texts in everyday academic German, the first data which was gathered were writing pieces on descriptive topics, which thereafter were of a more argumentative nature. This was undertaken during the pilot study in the second semester of 2013 (Ortner, 2013). In the first semester of 2014, writing tasks were continued. These writing tasks were not only designed to gauge student levels of writing, but also to allow them to engage with writing in German on a more regular basis (see

Estes, *et al.*, 1998), as longer, more academic writing pieces were not traditionally a part of the German Studies 3 course, due to time constraints and levels of competency.

All writing assignments were submitted electronically, and marked both in terms of content and language use, by the language lecturer (Undine Weber) using the "track changes" tool on Microsoft Word. These copies, which included comments on where to improve, what word choices would have been better, as well as what was particularly well written, were returned to the participants. The process of written feedback is an important skill in learning to write academically (Swales & Feak, 1994; Weideman, 2003).

Pilot study writing topics, 2013

During the second semester in 2013, seven writing topics were undertaken by seven students, who were then in third year German Studies. Writing topics had a limit of 300 words. Topics covered in 2013 were simply exploratory, and started at the most simple descriptive level of writing. These were as follows:

Table 2, 2013 Writing Topics, Pilot Study

German Studies (GS3) 2013		
Topic	Thema (Deutsch)	Topic (English)
Topic 1	Der Regen	The rain
Topic 2	Der Löwe	The lion
Topic 3	Platonische Freundschaft zwischen Männern und Frauen: kann es sie geben?	Platonic friendship between men and women: can this exist?
Topic 4	Deutsche Sprache: schwere Sprache?	The German language: a difficult language?
Topic 5	Schuluniformen: sind sie nötig?	School uniforms: are they necessary?
Topic 6	Spionage: Darf der Staat seine Bürger bespitzeln?	Espionage: is the state allowed to spy on its citizens?
Topic 7	Hausmänner	Househusbands

As arbitrary as these topics may seem at first glance, they often had to do with elements of various modules (particularly the Cultural Studies module and the conversation classes). Whilst no formal writing course was given, each week's assignment was accompanied by various stimuli which included vocabulary items (identified by Jaworska 2011 by frequency), and their collocations, as well as extra information on the theme, or links to websites or relevant articles. 45 texts were collected from 7 participants (according to the topics listed above) during the pilot study in 2013, as seen in the following table.

Table 3, Participant Submissions 2013

Participant No	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7
P1	1	0	1	0	1	1	1
P2	1	1	1	1	1	1	1
P3	1	1	1	1	1	1	1
P4	1	1	1	0	1	1	1
P5	1	1	1	1	1	1	1
P6	1	1	1	1	1	1	1
P7	1	1	1	0	1	1	1

Writing tasks first semester 2014

Writing tasks in the 2013 pilot study were voluntary, and it was found that participants were less motivated than they would have been had the writing assignments been assigned a score which could contribute to their overall achievement score for German Studies. This is not unique to our study; other researchers such as Estes (1998, p. 69) observed that student motivation for writing pieces where no marks were allocated was low. This was amended in 2014 where writing assignments were designed to run parallel to the course content of the *Landeskunde* module being taught, and assigned marks which formed a small portion of the modules total score. This ensured that the topics were relevant and current, and meaningful to the students, and that motivation was high. Student motivation will be expanded on in Chapter 5: Reflecting on Writing, the Participants' Perspectives.

Assessing academic literacy and language proficiency

In addition to the linguistic feedback given, a rubric and mark was assigned to each writing piece. A formal assessment of the writing contributes greatly to the understanding of the writing levels in the learner corpus, and allows for better control of the variables (e.g. one is able to separate out the better writers from the poor). In order to assess the level of writing, a rubric for each level of study was obtained online from the Goethe Institute. The Goethe Institute bases its criteria on the guidelines set by the Common European Framework of Reference for Languages to ensure that students meet international standards for their level of study (Council of Europe, 2012). These rubrics may be found in Appendix 5.

CEFRL levels

The Common European Framework of Reference for Language (CEFRL), provides a broad base at which learner levels of competence may be assessed (Council of Europe, 2012). This is according to 3 broad

divisions that form 6 levels. Level A is the basic user, with A1 at "breakthrough or beginner" and A2 the "waystage or elementary" learner (Council of Europe, 2012). Level B is categorised as the "Independent User" where B1 is the "threshold or intermediate" and level B2 is "vantage or upper intermediate" (Council of Europe, 2012). The final level is level C, the "Proficient User", who at level C1 have "effective operational proficiency" or at advanced level C2 have achieved "mastery or proficiency" (Council of Europe, 2012). The CEFRL describes what a learner is supposed to be able to do in terms of reading, listening, speaking and writing at each level, and as such is a useful tool when attempting to categorise the writing of any learner of a European foreign language, such as German.

Our students, in third year, are expected to be at level B1 (threshold or intermediate). According to the CEFR, our students at B1 should be able to understand "the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc". Furthermore, in practice they should be able to deal with "most situations likely to arise while travelling in an area where the language is spoken" (Council of Europe, 2012). And lastly, in terms of writing capabilities, they should be able to produce "simple connected text on topics that are familiar or of personal interest", as well as "describe experiences and events, dreams, hopes and ambitions and briefly give reasons and explanations for opinions and plans" (Council of Europe, 2012).

This hardly calls for a great level of academic literacy as displayed by the definition adapted from Weideman (2003) (see Chapter 3). These writing tasks are thus more well defined at the level of B2 which calls for students to "produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options" (Council of Europe, 2012). The writing assignments did not aim for discipline-specific discourse, but for "everyday academic" discourse – *alltägliche Wissenschaftssprache.*

Given the expected CEFRL levels for our participants' level of study, these writing tasks pushed them to achieve above and beyond the normal expectations. As the students themselves explained in their interviews, while the writing tasks were challenging, they were also extremely helpful, as they forced the participants to think and write in the target language. As can be seen in the corpus, the writing in the initial collection stages was not very good. As students progressed and gained practical experience, particularly during the writing course, their writing became better and better, and students reflected (see Chapter 5) that the process of having to write in German was integral in building their confidence in writing and their German language skills. A table of the writing scores may be found in Appendix 1. As the writing scores measure a range of factors, they do not reflect a direct increase in writing proficiency. This is a common issue faced by assessors of writing courses, as writing scores may not reflect large increases in overall proficiency. This is why it is important to explore the usefulness of a

writing course by approaching the examination of writing from many different angles aside from a simple total score evaluation.

Marking and formative feedback

As with the 2013 exercises, writing submitted was returned electronically and marked by the lecturer (Undine Weber) using the track changes function on Microsoft Word. Formative feedback is an important factor in learning to write academically, as students can track their progress, and identify areas of strength and weakness in their written texts, and hopefully learn from their errors (Swales & Feak, 1994; Weideman, 2003; Ortner, 2013). I assigned marks on the rubric (obtained from the Goethe Institute online, and based on the CEFRL levels discussed above) based on the mark up performed by the lecturer. These were counterchecked by the lecturer, and the students were given copies of the rubric pertaining to their level of study and explicitly shown and explained to what each category entails. 2014 Topics (8 to 11 before the writing course, 12 to 14 during the writing course) were as follows:

Table 4, 2014 Writing Topics

German Studies (GS2/GS3/GHons)		
Topic (date)	Thema (Deutsch)	Topic (English)
Topic 8 (06.05.2014)	*Unsere Mütter, unsere Väter*: Können Sie sich in die Charaktere einfühlen? Warum, warum nicht? (200 words)	*Unsere Mütter, unsere Väter*: Are you able to empathise with characters? Why, or why not? (200 words)
Topic 9 (20.05.2014)	Bitte schreiben Sie Ihre Einführung auf Deutsch (100- 200 Wörter).	Read Chrystal Downing's *Staging Ideology and Love in* Good Bye, Lenin! What does Chrystal Downing mean by the Mise en Scéne of Freedom? Use examples from the film to illustrate your answer. Please write your introduction in German (100 - 200 words), and then continue the assignment in English.
Topic 10 (01.08.2014)	Formulieren Sie Ihre eigenen Ideen, wie es zu grammatischen Geschlechtern im Deutschen gekommen ist. Sie können Ihrer Phantasie freien Lauf lassen, aber achten Sie bitte darauf, dass Sie Ihre Sätze sinnvoll gestalten und verbinden.	Formulate your own ideas as to how grammatical gender was formed in German. You can let your imagination run free, but please make sure that your sentences make sense and are connected.
Topic 11 (08.08.2014)	Was halten Sie persönlich von Martin Luther, soweit Sie das beurteilen können?	What is your personal opinion of Martin Luther, insofar as you can judge?

Topic 12 (22.09.2014)	Was halten Sie persönlich von Sprachgesellschaften? Sind sie heutzutage noch relevant, und in allen Sprachen oder nur in ein paar? Begründen Sie Ihre Antwort.	What is your personal opinion of *Sprachgesellchaften*? Are they still relevant today, and in all languages or just a few? Substantiate your answer.
Topic 13 (01.10.2014)	In Südafrika gibt es 11 Amtssprachen. Gibt es auch hier, wie in Deutschland, Dialekte? Sie können sich auf eine der Sprachen Südafrikas konzentrieren.	There are 11 official languages in South Africa. Are there dialects here, as there are in Germany? You can concentrate on any one of the South African languages.
Topic 14 (13.10.2014)	a) Sie schreiben seit einem Jahr auf Deutsch. Für diese letzte Aufgabe reflektieren Sie über diesen Prozess. Wie unterscheidet sich das Schreiben auf Deutsch vom Schreiben auf Englisch? Wie war diese Erfahrung für Sie persönlich: eine Herausforderung, eine Freude, ein Alptraum? b) Zu guter Letzt überlegen Sie, für welche Zwecke Sie Ihre Deutschkenntnisse später einmal einsetzen können.	a) You have been writing in German for the past year. For this last assignment, reflect on this process. How does writing in German differ to writing in English? How has it felt for you personally: a challenge, a joy, a nightmare? b) Lastly, consider the flowing: For which purposes can you use your German language skills in later life?

During the course of 2014, 13 participants completed the above seven writing tasks (a total of 88 paragraphs, as P10, P14 and P19 each did not hand in one assignment). Two of the participants in 2014 (P2, and P6) were participants in the original pilot study in 2013, having continued from third year to do Honours in German Studies. Five students were in third year-, and six in second year-German Studies in 2014. As stated above, three participants (P8, P9 and P17) were excluded from the study, as they did not meet the learner corpus design criteria for the RUDaF corpus.

Participant submissions may be represented as follows:

Table 5, Participant Submissions 2014[17]

Participant No	Term 2	Term 2	Term 3	Term 3	Term 3	Term 4	Term 4
GS2	Topic 8	Topic 9	Topic 10	Topic 11	Topic 12	Topic 13	Topic 14
P10	1	1	1	0	1	1	1
P11	1	1	1	1	1	1	1
P12	1	1	1	1	1	1	1
P13	1	1	1	1	1	1	1
P14	0	1	1	1	1	1	1
P15	1	1	1	1	1	1	1
GS3							
P16	1	1	1	1	1	1	1
P18	1	1	1	1	1	1	1
P19	1	1	1	1	0	1	1
P20	1	1	1	1	1	1	1
P21	1	1	1	1	1	1	1
GSH							
P2	1	1	1	1	1	1	1
P6	1	1	1	1	1	1	1

2015 Writing Topics

In 2015, all writing topics ran simultaneously to writing course which had begun in 2014. Only four[18] participants were involved with the writing course in 2015, of these three were third-year students (p12, p13, p15), and one an Honours student (p21), all having participated in the 2014 study. The procedure for marking the writing followed on from the 2014 methodology as described above. Topics were again designed to supplement one of the relevant ongoing modules in German Studies at the time (either literature or *Landeskunde*).

[17] Shaded areas are representative of submissions made during the writing course, non-shaded areas are representative of submissions made before the writing course

[18] A fifth student, P9, also completed all writing tasks and questionnaires in 2015, however, as stated previously this data has been excluded as the participant has almost mother-tongue German language competency and thus did not meet the design criteria requirements set out for the RUDaF learner corpus.

2015 writing topics were as follows:

Table 6, 2015 Writing Topics

German Studies (GS3/GHons)		
Topic (date)	Thema (Deutsch)	Topic (English)
Topic 15 (03.03.2015)	Österreichische Kultur: was ist gleich oder anders als in Deutschland?	Austrian culture: what is similar or different to that of Germany?
Topic 16 (17.03.2015)	Wie kann es heutzutage noch Analphabetismus in Deutschland geben?	How can it be that there is still illiteracy in Germany in this day and age?
Topic 17 (28.04.2015)	Die Deutschen sind krimibesessen – siehe *Cicero* „Warum der Tatort glücklich macht.“ Warum? Wie ist die Situation in Südafrika?	The Germans are obsessed with crime – see *Cicero* “Why the *Tatort* makes people happy”. Why? What is the situation in South Africa?
Topic 18 (05.05.2015)	Sie bekommen das „offizielle“ letzte Kapitel von „Liebe bis in den Tod“, also den Schluss, so wie er im Buch steht. Lesen Sie dieses Kapitel 13 bitte und schreiben Sie ca.350 Wörter: Warum hat Ihnen der offizielle Schluss gefallen (nicht gefallen)?	You have received the “official” last chapter of “Liebe bis in den Tod”, as well as the ending as it is in the book. Please read this Chapter 13, and write approx. 350 words: Why did you enjoy (not enjoy) the official ending?
Topic 19 (12.05.2015)	a) Beschreiben Sie kurz eine Kriminalgeschichte oder einen Krimi (Film) oder eine Krimi-Serie, die/ der Ihnen besonders gefallen hat/ gefällt. b) Inwiefern passt sie/ er in ein Krimi-Schema, das wir besprochen haben, inwiefern nicht? c) Was genau gefällt Ihnen an Ihrem Krimi?	a) Shortly describe a crime novel or a crime film or a crime series that you particularly enjoyed/ did not enjoy. b) To what extent does this crime novel/film/series fit within the crime fiction schema that we have discussed? c) What exactly do you like about your crime novel/film/series?

The five topics above were completed by the four students. Thus, a total of 18 writing pieces were submitted during the writing course for corpus analysis in 2015[19], as shown in the table below.

Table 7, Participant Submissions 2015

Participant No	Topic 15	Topic 16	Topic 17	Topic 18	Topic 19
P12	1	1	1	1	1
P13	1	1	0	1	0
P15	1	1	1	1	1
P21	1	1	1	1	1

Re-worked writing pieces

In addition to these new writing tasks in 2015, participants were also tasked to rewrite their favourite two writing pieces from the past year. This task encouraged students to engage with previous writing pieces and go back over the feedback given, and institute it. As the primary marker of the writing pieces, Dr Weber had noted that students produced recurrent errors (see Ortner, 2013), and it occurred to her that they may not be reading the feedback given, or, if read, not internalised. By reworking the writing pieces students were able to engage deeply with the feedback, as well as to gain a sense of agency over their own work. The reworked writing pieces were not remarked, they were accepted as controlled by the participants themselves. These were compiled into booklet form, titled "Writing for RUDaF" and copies printed for the students and lecturers to keep as a showcase of their work (see Appendix 1).

4.3.2) Creating the RUDaF Learner Corpus

A learner corpus, titled the Rhodes University *Deutsch als Fremdsprache* (RUDaF) learner corpus, was created using the data elicited for the German paragraphs exercises in order to analyse our students' writing in German. The principled design criteria and extensive metadata (elicited through questionnaires and interviews) which accompany this corpus, ensured valid and reliable data for analysis and interpretation (Flowerdew, 2012; Granger, 2004; Granger, 2003). The following section outlines the procedure for designing and capturing student writing for the RUDaF learner corpus

[19] As can be seen in the table, P13 neglected to hand in two of 5 assignments. This shows that even where marks are assigned, there is no guarantee that students will indeed engage with all the required tasks. This variable causes a slight imbalance within the corpus in terms of participant representation, but on the whole is balanced out over the whole corpus and remains acceptable within the bounds of the learner corpus design criteria.

Considerations

A researcher and corpus linguist must take into account many factors when deciding to build a new corpus, as seen in Chapter 3, with reference to creating or choosing a corpus to use as a language learning aid. The first consideration in building any corpus is what the purpose of the corpus will be, and if there is already one in existence which can fulfil the requirements of the need. With regard to my specific research goal of assessing the language use of the German Studies students at Rhodes University in South Africa before and after a corpus-based writing course, and then comparing them to mother-tongue writers of German at a similar institution in Germany, there existed no such corpus. In fact, as stated in Chapter 3, there are very few learner corpora which have been developed in South Africa as a whole. Goossens & Granger (2013) of the *Universite Catholique de Louvain* (UCL) have been working on creating a comprehensive list of learner corpora around the world, in order to assist researchers interested in this field (Granger, 2003; Goossens & Granger, 2013). To my knowledge, the Tswana Learner English Corpus (Van Rooy & Schäfer, 2002) is the only South African learner corpus about which there have been any publications, and the only South African learner corpus on the UCL list.

Whilst traditionally corpus linguists have strived for the motto "bigger is better" (see Chapter 2), in Foreign Language Teaching (FLT) as pointed out by Ragan (1996, p. 211) "small corpora compiled by teachers of their own pupils' work are of considerable value: 'the size of the sample is less important than the preparation and tailoring of the language product and its subsequent corpus application to draw attention to an individual or group profile of learner language use". As seen in the discussion in Chapter 3, according to Granger, a leading expert in the field of learner corpora, size is only really useful when a corpus is collected according to strict design criteria (Granger, 2004, p. 125). Having a good set of principled design criteria can in part determine what this size should be.

The following criteria (following on from Granger, 1998, p. 7) were controlled for in the RUDaF learner corpus design and recorded:

Factors pertaining to the learner

- Age: early 20s
- Sex: male/female
- Mother tongue: English/Afrikaans/African Languages (mainly Nguni languages)
- Region: Southern Africa
- Other foreign/second languages: French/Afrikaans/African Languages

- Level: Intermediate learners of German from Southern Africa, with a minimum CEFR level A2, and a maximum of B2 (Council of Europe, 2012). Two to four years prior exposure to German[20].
- Learning context: German as a Foreign Language in South Africa

Factors pertaining to the task setting

- Type written texts of approximately 300 words each. No spoken data.
- Argumentative/opinion based essay topics and some creative writing pieces. (See tables in section above)
- Written at home, and the use of dictionaries and web-based sources and resources permissible. Students required to reference direct sources, and submit writing tasks electronically via e mail.
- Files saved as UTF 8 text format.
- No annotation[21]
- Diachronic data collection: text collection covers a period of time with multiple sets of data collected from the same learner over the course of a year, allowing for the comparison of writing before and after DDL exercises.

For both the pilot study, as well as the main study, the raw data, submitted electronically, was cleaned and converted to text format, UTF 8 [22]. Metadata included with each file included the scores assigned for each writing piece, as well as participant questionnaire information (age, years of study, reasons for studying German) and interview information, and the week of submission[23].

[20] The contact hours which the German Studies Section at Rhodes University has with its students are as follows: first year, 130 hours (116 language teaching); second year, 143 hours (91 language teaching; excl. translation); third year, 146 hours (78 language teaching; excl. translation); fourth year, 156 hours (78 language teaching; excl. translation).

[21] Given the high rate of errors, both morphological, lexical and syntactical, particularly in the first set of writing tasks, it was decided not to annotate the RUDaF learner corpus. This decision is based on insights provided by Van Rooy & Schäfer, (2002) who found that the success rate of a POS (Part Of Speech) annotaters decreases as the number of learner errors increases. As the RUDaF learner corpus is relatively small in size, it was possible to manually observe these errors in the analysis which focusses specifically on the use of TAG words introduced in the writing course.

[22] Data was taken in its original form, without the corrections made in track changes.

[23] Metadata is an invaluable accompaniment to the researcher using corpus linguistic methodology, as this information becomes a kind of bibliography, so that the original text can easily be found and referred to within any study of the corpus as a whole (Granger, 1998).

Corpus composition

As the corpus contains a diachronic collection of data, the learner corpus is divided into two parts in order to draw internal comparisons between participants. The first section of the corpus (RUDaF PreWIC[24]) is made up of the 95 pre-writing course text submissions and includes the 45 texts collected in 2013 for the pilot study, as well as 50 texts collected in 2014 before the writing course (Topics 8 to 11). The second part of the corpus (RUDaF DuringWIC) consists of a total of 56 texts, of which 38 texts were collected in 2014 during the pilot writing course (Topics 12 to 14) and 18 texts collected during the 2015 corpus-based everyday academic German writing course. This may be summarized in Table 8 below.

Table 8, RUDaF Corpus Composition

Corpus/Subcorpus	Total no. of files	Total no. of tokens
RUDaF	151	33987
RUDaF PreWIC	95	17998
RUDaF DuringWIC	56	15989

While the learner corpus may appear to be imbalanced with reference to the number of files in each section, the number of tokens (words) in each sub-corpus differs by only 2009 tokens. The total size of the RUDaF corpus overall is 33,987 tokens, of which there are 5,500 word types. This may be divided into 17,998 tokens for RUDaF PreWIC, and 15,989 tokens for RUDaF DuringWIC. This is because the length of the writing pieces which participants submitted in 2013 and 2014 before the writing course were much shorter than those submitted in 2014 and 2015 during the writing course. This may be attributed to the fact that as students had more practise with writing in German, they felt more comfortable and were able to produce longer, more academic texts (See Chapter 5). The length of the writing pieces may also be attributed to the nature of the topics as listed in the section above. As the students showed progression in their writing ability, so the difficulty of the subject matter at hand was enhanced. The more advanced the topics became, the more engagement they required on the part of the participants. As is shown in Chapter 5, the nature of the participants writing changed as they entered the writing course. Aside from their writing pieces becoming longer, participants started to

[24] RUDaF PreWIC is so named as it consists of writing collected before the writing (intensive) course. RUDaF DuringWIC is so named as it consists of writing collected during the writing (intensive) course.

use more sophisticated academic phrases and construct better short essays in which they were able to present an argument rather than simply describe a topic.

Corpus analysis

This is shown in Chapter 5, where data submitted before the writing course beginning in the second semester of 2014 (RUDaF PreWIC) is compared, using corpus tools, to that collected during the writing course (RUDaF DuringWIC). This is achieved by comparing the use of specific lexical items and academic collocations associated with everyday academic German (which were introduced during the writing course, see sections below) in the two data sets. This aims to show whether or not CL methods have helped with the writing process, and offers an objective method of assessing the improvement of students' academic writing skills in German. Whilst our learner corpus may be viewed as relatively small at a total size of 33987 tokens and 5500 types, it is sufficient for the purposes of examining the use of the high frequency TAG words introduced during the writing course, as Granger states "[f]or some linguistic studies, for instance those involving high frequency words or structures, relatively small samples of c. 20 000 words may be sufficient" (Granger, 1998, p. 9).

Lastly a comparison of RUDaF and an existing mother-tongue corpus of German academic writing (FALKO L1) is made, in order to assess the use of specific lexical items and academic collocations associated with everyday academic German in the two data sets (Granger, 1998; Zeldes, *et al*., 2003). This section of analysis draws on elements of a Contrastive Interlanguage Analysis (Granger, 1998, p. 10), and aims to show how our learner writing differs from, or is similar to, mother-tongue writing with regard to specific academic German formulaic sequences. FALKO L1 was chosen as a mother-tongue control corpus as it consists of circa. 70 000 tokens of mother-tongue German, written by school leavers who were tasked to write argumentative essays. While not directly comparable to the RUDaF learner corpus, FALKO L1 is the only available corpus of German argumentative writing and thus has been widely used as a mother-tongue reference corpus for many projects (Krummes & Ensslin, 2012; Jaworska, *et al*., 2015; Lüdeling, *et al*., 2008). Given more time and better resources, a truly comparable corpus of mother-tongue German writing could be formed.

The free corpus linguistic software programme, *AntConc 3.4.4*, was used to perform all operations (Anthony, 2014).

4.3.3 & 4.3.5) Questionnaires and Interviews

In order to engage with our students' experiences of writing and to qualitatively assess their attitudes towards writing a questionnaire and an interview were performed in 2014. A further questionnaire was administered in 2015 at the end of the corpus-based writing course in everyday academic writing in German to explore the usefulness of, and attitudes towards, corpus-based learning materials.

Questionnaires[25]

Pre-writing course questionnaire

The first questionnaire allowed for a more in-depth knowledge of participant background which provided valuable metadata for the creation of the learner corpus, as well as a broad spectrum analysis of attitudes towards writing and German Studies. This was achieved mainly through the use of open-response items in the form of fill-in items, short-answer items, and few longer exploratory open-response items (Brown, 2009). Open-response questions were employed in this case as I wanted to form an idea of what the participants own views were, and did not wish to restrict the responses able to be given by participants, but rather to allow for participants to engage with the question and "express their own ideas more fully" in their own words (Brown, 2009, p. 202).

This type of questionnaire, which is usually time consuming both to complete as well as to analyse (Brown, 2009), was possible in the case because of the small sample size. The first questionnaire conducted in the first half of the writing course was completed by 13 of the 18 participants (the other 5 having graduated in 2013 and as such no longer present at the time at which the questionnaires and interviews were conducted). As the first questionnaire aimed to capture information relating to the individual participants, in particular to their language background and their personal experiences of being a German Studies student, I was able to form a profile of the individual participant, and the group as a whole, which enabled a better qualitative understanding of their writing.

Fill-in items pertaining to the participant's biodata (Brown, 2009, p. 202) included age of the participant, years of exposure to German and other subjects enrolled for at university. Short answer items included what type of individual reading, and other engagement with the German language, is undertaken by the participant[26]. Broad open questions included what processes and techniques are followed by participants when writing in German, and what problems or limitations students experience when working with the dictionary.

Care was taken to keep both the number of questions, and the questions themselves, short and succinct, in order to avoid participant fatigue. Time was given in class to fill out the questionnaire, to ensure that the participants did not feel rushed, and were able to give their full attention to the questions. The interviews (see next section) which were undertaken followed shortly after the first questionnaire, and some of the more interesting topics from the questionnaire were expanded upon.

[25] Please see Appendix 6 for a copy of the questionnaire.
[26] Research has shown that the amount of individual reading undertaken impacts strongly upon vocabulary knowledge in language acquisition (Cunningham & Stanovich, 1998).

Post-writing course questionnaire

The second questionnaire was conducted at the end of the writing course, at the end of the first semester, in 2015. The five participants who took part in the writing course (the third year students of 2015) were asked to complete the second questionnaire via the online Rhodes University student website, RUconnected. This 2015 questionnaire was issued through the online platform in order to ensure anonymity of the participants (should they wish to comment negatively, and perhaps be scared to do so because of the small size of the class and fear of personal retribution), and allow ease of analysis (the paper-based responses from the 2014 questionnaires all had to be individually typed out and were thus time consuming to capture and analyse). The 2015 questionnaire aimed to explore the participants' attitudes towards the course. A copy of questionnaire can be found in Appendix 6. This questionnaire made use of more closed-response items. Closed-response items were used to collect numerical data from questions presented with a range of scaled responses, pertaining to participants' evaluations of certain aspects of the course, and of aspects of their own personal writing development. These closed response items included questions such as whether the participants felt their writing had improved since 2014, and whether they felt the writing course was helpful. The responses choices included the following items: 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree). These closed-response questions allow for an objective statistical analysis (Brown, 2009, p. 202) of the groups subjective feelings, and thus are a useful tool. Each closed-response item was followed by an open-response question which allowed the participants to elaborate on the answer given and provide an explanation thereof. An example of this type of question would be "why do you think this is the case?" (see Appendix 6).

Follow-up Questionnaire

A follow-up questionnaire was delivered in the second semester of 2015. This short questionnaire sought to explore whether participants had perceived an impact on their writing in other subject and other languages as a result of participating in the corpus-based German writing course. Participants were asked to state (a closed-response question item) whether they thought their writing had improved (as a consequence of producing texts in German) with regard to writing in an academic way, as well as structuring assignments. They were then given the space to elaborate as to why they thought this was the case for each of the above questions (an open-response question).

Interviews

The interviews which were undertaken at the end of the first semester of 2014 allowed an in depth exploration of certain themes from the first questionnaire which proved to be interesting and needed to be elaborated on. As with the first questionnaire, the first interview was completed by 13 of the 18

primary participants. Each participant was given a 15 minute slot, which they signed up for at a time convenient for them on a Friday. Students were informed that the interviews would be recorded, but that their identities would remain anonymous. The audio recordings were later transcribed, and analysed.

According to Richards (2009, p. 183), interviews allow one to "probe beneath the surface of things and try to see things from the students' perspective". I wanted to investigate the German Studies Students' perspectives of writing in German, as well as other subjects. Of the three types of interviews (structured, unstructured and semi-structured), I opted for a semi-structured approach. I had a pre-set of questions (an "interview guide" (Richards, 2009, p. 185)) to guide the various aspects which I wanted to discuss, but this was not rigidly structured, as I allowed room for interaction where I felt further probing was necessary, or where I sensed students wanted to share more than what my pre-set questions allowed for (Richards, 2009, p. 185). The interviews thus provided a good platform for an in depth qualitative analysis of the students experience of writing in German.

The following table illustrates the participation in the above questionnaires and interviews in 2014 and 2015. The small pool of participants in 2015 is immediately noticeable. This is as many students who participated in 2014 did not carry on with German studies in 2015, resulting in a limited pool of participants who partook in the full corpus-based writing course.

Table 9, Participation in Questionnaires and Interviews 2014-2015

Participant No.	Consent Forms	Questionnaire 1, 2014	Interview 1, 2014	Questionnaire 2, 2015
P2	1	1	1	
P6	1	1	1	
P8	1	1	1	
P9	1	1	1	1
P10	1	1	1	
P11	1	1	1	
P12	1	1	1	1
P13	1	1	1	1
P14	1	1	1	
P15	1	1	1	1
P16	1	1	1	
P17	1	1	1	
P18	1	1	1	

P19	1	1	1	
P20	1	1	1	
P21	1	1	1	1

4.4.4) Designing and Evaluating a Corpus-Based Writing Course

In designing an everyday academic writing course for German Studies at Rhodes University, I sought to harness the current and existing findings relating to the understanding and teaching of formulaic language in German as a Foreign Language, as significant recent contributions to this field have been made by Krummes & Ensslin (2012), Krummes & Ensslin (2014) and Jaworska *et al.*, (2015) as outlined in Chapter 3. These researchers have made use of corpus-driven data analysis techniques to create corpus-based learning materials (Krummes & Ensslin, 2014) which are on paper, and therefore do not require learners to be involved directly with a corpus. An overview of the research relating to the WHiG project was given in Chapter 3, under the heading *Applications of the theory: examining what's hard in German (WHiG).*

The aim of the writing course

The specific aim of the writing course was to introduce our second-, third-, and Honours- level students to academic German vocabulary and the collocations associated therewith, which they have little opportunity to engage with on an everyday basis, given the relatively small amount of academic reading undertaken in German by the students due to curricular and time constraints. Mother-tongue corpora can offer "condensed exposure to language patterns" (Gabrielatos, 2005, p. 10) and so help language learners to "recognise recurring patterns of structure and meaning" (Gabrielatos, 2005, p. 6) (See Chapter 3). The learning outcomes which will be outlined in later sections reflect this goal, and expand on the specific facets of learning incorporated into the writing course. As stated in Chapter 3, the writing course developed makes use of insights from both facets of the direct approach to implementing corpora in the language learning classroom, namely teaching to exploit and exploiting to teach.

As outlined in Chapter 3, there are three important considerations for creating DDL exercises, and these shaped the decisions taken in creating the German Academic Writing Course. To reiterate, as set out by Oghigian & Chujo (2010) the first consideration lies in choosing a corpus (or corpus resources), the second in controlling the tasks set and the third in bridging the gap between language research and language production.

Considerations of accessibility, relevance and size when choosing the writing course corpus resources

In designing the writing course I chose to make use of already existing corpus materials as the task of creating a new corpus of mother-tongue academic German is one which exceeds the scope of this study, given time constraints, copyright issues and the scarce availability of relevant texts. The writing course made use of the exercises developed by the authors of the WHiG project which are freely available online, and supplemented this with an exploration of real life mother-tongue German corpus data obtained from the FALKO L1 corpus, available under a Creative Commons License, and the freely available online parallel corpus software Linguee.de.

The FALKO L1 corpus (Lüdeling, *et al.*, 2008) was deemed appropriate as it was perceived as neither too large, so as not to drown the learners in data, nor too small that the instances we were searching for would not be present[27]. Both the WHiG handouts and the FALKO L1 corpus were chosen for their high degree of relevance to the learning context, as they provide evidence of everyday academic German written by mother-tongue speakers of German. The web-based parallel corpus Linguee.de, while not only composed of academic texts[28], is an additional useful resource as it provides a general translation of a search item, and highlights the key term in the English text and the corresponding German text which are viewed alongside one another as direct translations (see Appendix 4 for an example included in the second WHiG handout). As a class exercise, I also sought to have the participants create their own small corpus of German writing in order to familiarize them with corpus software, as well as to provide an extra resource for learning (see learning outcomes for 2015).

The corpus-based writing course took place over two semesters, the first section consisting of three lectures in the second semester of 2014, and the second section consisting of 10 lectures in the first semester of 2015. An important aspect of the writing course was that writing tasks (see the sections above for comprehensive lists of topics) were issued each week and students were able to put the newly learnt collocations into practise. This allowed the students to make the move from language research to language production, a vitally important aspect of language learning which is often underscored in corpus-based learning, as stated in Chapter 3 (see Oghigian & Chujo, 2010, p. 202).

[27] The corpus was perceived to be big enough at the beginning of the course, however it was in fact limited in some instances, and in future I would recommend making use of a larger corpus in order to see sufficient examples of specific collocations.

[28] This did not prove to be too great an obstacle as the nature of the search term often determines the nature of the texts which appear. For example, a search for an academic collocation such as “according to” will often bring up texts which of a more academic nature. The source of the texts is also visible, and so participants were taught to be able to discern what sources were of a more academic nature, and more likely to be reliable.

2014

In 2014, a series of three lectures was given on corpus linguistics and academic literacy in German Studies (22.09.2014; 29.09.2014; 06.10.14). This was presented to both second-year and third-year students. The two Honours students joined the third year classes for this lecture series[29].

The learning outcomes for the participants of this short course were as follows: participants should

- understand what a corpus is, and how this can be used to study language;
- understand why formulaic language is important;
- recognise that grammatical sentences may not be idiomatic;
- be able to use one new German 'academic' word/ collocation in context per lecture;
- gain a better understanding of everyday academic German.

Having clear learning goals or outcomes is important for course-based assessment which "links student performance to specific learning outcomes in order to provide useful feedback to the instructor and students about how successfully students are meeting specific outcomes" (Doyle, *et al.*, 2000, p. 19). In answering the question of whether the corpus-based writing course was effective, the course goals, objectives, and content can be returned to in order to "gauge the extent of the learning that is taking place". This is undertaken in Chapter 5.

The short lecture series consisted of an introductory lecture, and a further two lectures which were spent exploring the WHiG handouts. An introductory lecture sought to introduce the class to some of the basic concepts: what is a corpus, what is academic literacy, how does corpus linguistics aid the learning of German studies, and what does learner writing in German Studies at Rhodes look like, in comparison to writing of mother-tongue German writers (based on the Honours research paper, Ortner, 2013).

In the second and third lectures the class began to explore the WHiG handouts, and we focussed on three specific lexical items in context (*Beispiel*, *Meinung*, *Laut*)[30]. Due to the time constraints (two lecture periods) I opted for this type of paper-based approach as an introduction to corpus-based learning materials rather than an intensive computer-based course where time is needed to acquaint

[29] Whilst it is not usually the case to have such a diverse group of students attending the same class, due to the small size of the year groups, and the introductory nature of the lecture series, this was deemed appropriate by myself, as well as the head of Section, Dr Weber. The learning outcomes, assignments and assessments for each group were stratified, and the nature of their writing pieces reflects this. As none of the students had previously engaged in any courses of this nature, it was agreed that they would all benefit from this type of learning.

[30] For a review of the WHiG handouts please refer to the discussion in Chapter 3, and see Appendix 4.

participants with software, as well as a venue in which to perform these operations. Certain sections of the WHiG handout were not at an appropriate level for our second-year learners. For example, the vocabulary in the sections which task students to "improve these German sentences" was often overwhelming for them, with the result that certain sections of the handout had to be re-worked. Nevertheless, they provided a good base from which to introduce the participants to formulaic German, and enhance their understanding of academic colocations. In addition to the paper-based WHiG exercises, participants were introduced to the online parallel corpus Linguee.de as a tool for checking collocations in writing assignments. While the participants had been familiar with online dictionary websites such as Leo.org, they were unaware of Linguee.de and the potential for discovering mother-tongue German collocations which they would be able to use in their own writing, in addition to those collocations introduced in class.

2014

GS2/GS3/GSHons students

22.09.14 Lecture 1: Introduction

29.09.14 Lecture 2: WHiG Handout, *Beispiel*/*Meinung*

06.10.14 Lecture 3: WHiG Handout *Laut,* introduction to Linguee.de

2015 Lecture series

In 2015, a series of ten lectures was delivered in the first semester. This was undertaken with the four third-year students and the one Honours student enrolled for German Studies. All participants taking the course in 2015 had also been involved in the 2014 writing assignments and course. As the participants were therefore acquainted with the project and had linguistically and methodologically advanced, the learning outcomes for 2015 were adapted to reflect this. The 2015 lecture series began with a short overview of academic writing and corpus linguistics, and how corpora can provide useful learning tools, especially for the specific context of academic writing in German. The aims of the course were outlined and the first writing topic was issued. Participants were encouraged to discuss academic literacy and how this may differ or be similar across languages, and to reflect on their experiences from 2014. All students for this lecture series possessed personal laptops, which enabled a far more in-depth approach than the 2014 lecture series, as the class was able to perform computer-based corpus exercises using the FALKO L1 and Linguee.de corpus resources in the lecture venue individually. The learning outcomes for the 2015 lecture series were as follows: participants who complete the course should

- gain a better understanding of academic literacy;

- understand what a corpus is, and how this can be used to study language;
- understand why formulaic language is important;
- recognise that grammatical sentences may not be idiomatic;
- form an understanding of academic German collocations, and how to find them using online corpus resources;
- be able to use academic German collocations within their own writing pieces;
- be able to use AntConc concordancing software;
    - generate a word-list;
    - examine keywords in context;
    - be able to normalise the frequency of a keyword when comparing two corpora
    - understand the notion of 'keyness';
    - be able to sort a concordance list in meaningful ways;
- formulate grammatical rules based on corpus evidence;
- compare a corpus of learner writing to a corpus of mother-tongue writing.

After the introductory lecture, in which handouts were distributed, aims of the course were clarified, a recap of terms and concepts introduced in the 2014 course was performed, and CEFRL guidelines for the marking of texts were clarified, I continued to explore the WHiG handouts in class, focussing on the introductory item *Zweck* and the conclusionary item *Fazit*. The Bangor Workshop handout, also created by the WHiG project authors, was explored and the structure and functions of the three main parts of an essay discussed. The academic vocabulary items *Thema*, *Frage* and *Erachten* and their collocations were explored. As participants had thus far only indirectly interacted with corpora via paper-based exercises, in the third and fourth lectures participants were introduced to direct corpus exploitation. Participants were explicitly taught about corpora and their uses, and shown in class how to make and manipulate their own corpus. As all participants of the course possessed their own personal laptops, each student was able to download free online concordancer, AntConc, for their own personal use. As a class we explored "Projekt Gutenberg" (gutenberg.spiegel.de) as a resource for German corpus creation, and via class consensus chose to download the text "*Der kleine Herr Friedemann*" as this had been a previous set work. The participants cleaned the text and uploaded it to AntConc. Participants were taught simple operations such as how to create a word list, how to examine keywords in context, how to do a simple sort of the keywords in context. In the following lecture, participants were tasked to identify the top ten nominal items in the frequency list, and

discuss what this may tell one about the content of the story. This follows the basic procedure for corpus linguistic text analysis, as stated in Chapter 2, a frequency list is usually the first step of analysis for any corpus linguist, as it provides one with an idea of what to focus on, particularly when focussing on lexical items (Baker, 2006, p. 121). This may then be followed up by looking at the collocations and keywords in context (KWIC) in the concordance lines, as well as by comparing the term to other similar lexical items to provide a comparison (Baker, 2006, p. 121).

Lecture 5 involved finding academic German sources online for a specific assignment, as students had expressed that they found it difficult to find appropriate materials to reference in their weekly paragraphs. As a class we researched the topic of illiteracy in Germany (in relation to the literature module at the time *Der Vorleser*/*The Reader*) in order to discuss how such a phenomenon could still exist in a "developed" country such as Germany, who have an excellent public education system. Participants were directed to two online German sources, in this case articles published online, relevant to the topic. These were found by typing in the topic sentence, in German, into the google search engine. This was a novel experience to some of the participants who had never thought to search online in German before. Planetwissen.de and Welt.de both had recently posted (2014/2015) articles on the topic of *Analphabeten* (illiterates) in Germany[31]. These articles were chosen as they are recent publications on the topic from reputable sources (though online). As they are short web-based articles, they are not strictly speaking academic, but not as overwhelming for the students as a German academic journal article would be in terms of vocabulary, structure, and length.

The participants were asked to discuss ways in which to present the information gleaned from the websites, using the academic collocations we had learnt in class in previous lectures (e.g. *Thema*, *Frage*, *Erachten*, *Fazit* etc.). In lecture 6, time was provided in class to start revising a chosen paragraph the students themselves had submitted during the module and making corrections based on the feedback they had received by the course convenors. This was as the lecturers had noticed a recurring pattern of errors in certain students' work, and there was the feeling that students were not internalising the feedback given on assignments. It was thought that by having students rewrite some of their paragraphs, they would be able to internalise better the formative feedback given. The course convenors decided that a booklet of the class' re-written paragraphs would be created and printed, in order to heighten motivation for the task, as well as to create a feeling of achievement and pride among the students for their own written work. The booklet also includes students' reflections on the

[31] See the following links: http://www.planet-wissen.de/alltag_gesundheit/lernen/analphabeten/ http://www.welt.de/wissenschaft/article133584228/1-5-Millionen-junge-Analphabeten-in-Deutschland.html

process of writing more, and learning to write better, in German. This booklet may be found in Appendix 1 and is titled “Writing for RUDaF”.

Participants were tasked to look up incorrect collocations and terms on Linguee.de in class. We returned to the WHiG handouts for guidance during this process, and re-capped what was learnt and how the collocations are best used for each section of an essay. Following this, a ‘pop-quiz’ was issued in lecture 7, tasking the participants to write down one of the collocations learnt in class for each 5 sections of an essay, namely: introduction, providing an example, putting forward your opinion, presenting somebody else’s opinion, and concluding a paragraph or an essay. At the end of lecture 7, and for the duration of lecture 8, the participants examined and compared the RUDaF and FALKO L1 corpora using the concordancing software, AntConc, introduced in lecture 3. Participants were taught how to normalize the frequency of two lexical items as they appear in a frequency list, in order to make an accurate claim about the ‘keyness’ of a term in a corpus. Participants were tasked to specifically focus on the items from the pop-quiz, and to find a good example collocation from each corpus and to write it down underneath their own made-up example. In keeping with current research, the corpus exercises were structured to create a “guided dynamic partnership” (Oghigian & Chujo, 2010, p. 202) between myself as the language teacher, and the students as language learners. I sought to provide guided searches which would be fruitful and informative revealing patterns which could be explained in class and fit into the context of the learning outcomes. Participants noted the interesting similarities and differences in the two corpora with regard to usage. These similarities and differences are expanded on in Chapter 5, where I discuss how South African learner German writing at Rhodes University compares to mother-tongue German writing.

In lecture 9 participants were tasked to evaluate and improve upon the handout “*Redemittel für wissenschaftliche Texte”.* This handout which can be found online as a resource for academic writing at a German University (Schultis, 2011) includes only vocabulary items in isolation with no reference to meaning or usage, no explanations, or examples of collocations. The participants were given the task of improving this worksheet using their knowledge of corpus tools and resources to better the worksheet. They completed this exercise in class, working in pairs, and using their personal laptops to access Linguee.de and FALKO L1 as references. The original worksheet as well as the improved worksheet created by the participants may be found in Appendix 4. The participants completed this task in the final lecture of the series, lecture 10. Also addressed during this lecture was nature and purpose of the online course evaluation, which participants were asked to complete at home. For an overview of the course evaluation, please see the sections above on questionnaires. Chapter 5 expands on the responses provided by the participants in this questionnaire. A brief summary outline of the course is included below.

2015

GS3/GSHons students

24.02.15 Lecture 1: Introduction, review of aims and previous writing, hand out notes, recap of terms. Recap of CEFRL guidelines for marking texts.

03.03.15 Lecture 2: WHiG handouts were further explored and sections on *Zweck* and *Fazit* completed. German academic article read in class.

10.03.2015 Lecture 3: Direct corpus exploitation. Downloading AntConc, exploring project Gutenberg, downloading and creating a classroom corpus.

17.03.15 Lecture 4: Continuing to explore our *Friedemann* corpus, using corpus tools. A discussion of the result for the *Friedemann* corpus ensued.

24.03.2015 Lecture 5: Peer evaluation of texts, finding good academic German sources online relating to the topic of *Analphabetismus*. Following this a specific reading of academic texts with an eye for collocations. Texts read in class include the following:

http://www.planet-wissen.de/alltag_gesundheit/lernen/analphabeten/

http://www.welt.de/wissenschaft/article133584228/1-5-Millionen-junge-Analphabeten-in-Deutschland.html

21 04.2015 Lecture 6: Time in class to start revising a chosen paragraph and making corrections based on feedback. Look up incorrect collocations and terms on Linguee.de in class. Return to WHiG handouts and re-cap what was learnt and how the collocations are best used.

28.04.2015 Lecture 7: Pop Quiz, followed by practical exercises in AntConc on FALKO L1.

05.05.2015 Lecture 8: FALKO L1/RUDaF cross comparison. Participants asked to examine interesting similarities/differences which are present in the two corpora with regard to usage. Exploring notions of "keyness".

19.05.2015 Lecture 9: Participants tasked to evaluate and improve upon the handout "*Redemittel für wissenschaftliche Texte*".

26.05.2015 Lecture 10: Finishing worksheet in pairs. Online course evaluation set up explained. Participants thanked for their effort and participation.

In summary, the importance of a multi-faceted approach to exploring the effects of a corpus-based writing course

One of the main objectives of this study is to explore the usefulness of the corpus-based approach to teaching everyday academic German vocabulary and writing skills, as outlined above. Traditionally, one would look at scores in order to examine improvement. However, if one examines the assigned scores for the assignments, one cannot note an obvious improvement, and this may be because the marks allocated were assigned for both grammar and content, not only for good use of specific academic collocations (see Appendix 1). Furthermore, an assigned score over a longitudinal study cannot be said to correlate specifically and exclusively with the writing course for the same reason that students are exposed to more German in the form of other modules over time, and the aim of their German course on the whole is of course improvement in German language competency[32]. There is also the additional concern to be raised here that an assigned score is partly subjective on the part of the marker (and could potentially be adapted to show favourable progress).

A second approach is then to ask the participants to reflect on whether they have perceived an improvement in their writing, an approach adopted by this study as outlined above and further discussed in Chapter 6. While participants do voice that they feel they have improved specifically because of the writing course, this is of course a subjective viewpoint on their part, and whilst a valuable finding for the aims of this study, one is not able to assess through questionnaires the extent to which this improvement has taken place. Methodological concerns such as this have plagued other researchers undertaking similar projects which seek to introduce writing intensive course (Homstad & Thorson, 1996). These concerns also trouble those who use corpus-based approaches to improving writing. According to Varley (2009, p. 134), "Many studies point to the advantages of this [corpus-based] approach but few have researched the extent to which learners actually benefit, particularly from exposure to extended and large-scale incorporation of corpus consultation into the language syllabus".

Overcoming methodological obstacles using corpus methods

In this study, the methodological obstacle described by other researchers is overcome by utilizing a corpus-based approach to language assessment, which is linked to the learning outcomes of the course (a course-based assessment). A course-based assessment assesses student learning within the classroom environment by using course goals, objectives and content to "gauge the extent of the

[32] Researchers such as Homstad & Thorson (1996) "found it impossible to do a systematic analysis of the graduation proficiency exams, as there were too many variables to isolate the effects of intensive writing". Although Homstad & Thorson were unable to quantify their results, their report states that the general impression formed by the researchers was that there was in increase in fluency and confidence from previous years (Homstad & Thorson, 1996, p. 6).

learning that is taking place" (Doyle, *et al*., 2000, p. 19). Course-based assessment links student performance to specific learning outcomes, in order to provide information to teachers and students about how successfully students are meeting specific outcomes (Doyle, *et al*., 2000, p. 19). The corpus-based writing course had specific objectives which participants were expected to fulfil such as: "to be able to use one new German 'academic' word/ collocation in context per lecture" in 2014, and in 2015 to "form an understanding of academic German collocations, and how to find them using online resources"; and to "be able to use academic German collocations within their own writing pieces".

Accordingly, one of the main aims of the writing course (see sections above) was to introduce typical everyday academic German words, collocations and phrases identified by the WHiG project as the basic building blocks for everyday academic German writing, and for students to implement these successfully in their writing. These basic building blocks (TAG words) were identified by WHiG project authors (Krummes & Ensslin, 2012) by examining German mother-tongue writing, and sought to provide learners of German with idiomatic alternatives to raw and ineffective (if grammatical) translations when writing an academic assignment. The TAG words were introduced to participants during the course through paper-based corpus exercises based on the WHiG findings, as well as through computer-based corpus consultation. Participants were encouraged to use these newly learnt collocations in the weekly/fortnightly paragraphs. The typical academic words and phrases were not introduced in other lessons, and so their use in student writing may be seen as either a product of the writing course, or a result of self-study. By examining the participants' use of these TAG words before and during the writing course, one may gain an idea of how the writing course impacted upon their knowledge and ability to use these words. This analysis is undertaken in Chapter 5.

Thus, this methodology seeks to take a multifaceted approach to exploring the effects of a corpus-based everyday academic writing course in German studies. By approaching the analysis of the learners' writing from many angles, by comparing the use of typical academic words taught in the course before and after the course, and then by seeking out participant opinions, and examining scores, one is able to form a holistic perspective of learner German writing at Rhodes University and how this has been affected by the corpus-based writing course.

CHAPTER 5. QUANTITATIVE ANALYSIS: EXPLORING THE RUDAF LEARNER CORPUS

This chapter explores how the writing course impacted on the participants' use of the everyday academic German collocations, with reference to five types of typical (everyday) academic German (TAG) collocations taught during the course. The analysis is furthered by including a comparison of the use these same TAG words in a corpus of mother-tongue German writing, FALKO L1. An exploration of the implementation of taught TAG words is undertaken in order to provide quantitative evidence to support qualitative perceptions of participant progression, which will be expanded upon in Chapter 6. By approaching the analysis of the learners' writing from this angle, as well as by seeking out participant opinions, and examining scores, one is able to form a rounded perspective of learner German writing at Rhodes University and how this has been affected by the corpus-based writing course. This multi-faceted analysis is learner centred and while focussed on the texts produced by learners, does not lose sight of the people behind the words; as McCarthy (1998, p.23) states, "As long as we keep a cool head in the face of the exhilaration of computer power and vast arrays of text, we will not fall into the temptation of substituting cold numbers for the real people who actually produced the words" (McCarthy, 1998, p. 23).

Comparing the use of TAG words prior to and during writing course

As stated in the methodology (Chapter 4) the first comparison seeks to compare the use of TAG words in the writing of our participants before the writing course (RUDaF PreWIC), to the use of TAG words in the writing which took place during the writing course (RUDaF DuringWIC). The analysis is focussed on the following eight basic TAG words which were introduced in the first handout in 2014 and 2015. The TAG words were namely *Zweck* "aim", *Beispiel* "example", *Erachten* "opinion", *Ansicht* "opinion", *Meinung* "opinion", *laut* "according to", *Zufolge* "according to", and *Fazit* "conclusion" (Krummes & Ensslin, 2012; Krummes & Ensslin, 2014). The use of the TAG words *Meinung* and *Beispiel* are included in this analysis, as although they were known and used by participants before the course, their usage changed as a result of the writing course which emphasised using variations of these TAG words. The eight typical everyday academic German words were introduced to the participants in corpus-based exercises (on paper) which were performed in class (see Appendix 4 for a copy of the handouts). The purpose of these handouts was not to restrict learners in their writing, but rather to give them authentic collocational choices for the various essential sections of an essay, and help participants to write more fluently and sound more like mother-tongue writers. Many variant collocations of each of these TAG words were presented to the participants of our writing course, who were also shown what forms British learners tended to over-use or under-use in their essays based on research by Jaworska, Krummes & Ensslin (Jaworska, *et al*., 2015).

An explanation of simple corpus-based statistics

The comparison below seeks to use basic corpus linguistic statistical methods in order to systematically analyse the use of these TAG words in the two RUDaF sub-corpora and in the FALKO L1 control corpus. As stated in Chapter 4, the RUDaF learner corpus is comprised in total of 151 files which result in 33,987 tokens. This can be further divided into 95 files submitted before the writing course (RUDaF PreWIC, 17,998 tokens) and 56 files submitted during the writing course (RUDaF DuringWIC 15,989 tokens). The FALKO L1[33] corpus (to which access was granted for this research) consists of 116 files, and a total of 76,957 tokens[34].

The most basic statistical measure in corpus studies is a measure of frequency, or a frequency count. Frequency in corpus linguistics refers to the total amount of times a word type occurs in a corpus of a total amount of tokens (normally the total word count). For example, the most frequent 'type' that occurs in most corpora is the definite article *the* (Baker, 2006, p. 47). A frequency count involves counting how many times a word type occurs in a corpus, and all corpus related software will perform this task with ease. As there is a considerable size difference between RUDaF and FALKO L1, frequency of TAG words need to be normalised, for example to a base of 1000, to indicate the saliency of a term within a corpus. The normalised frequency (NF) will show how many times one can expect to see a given word in a thousand words of running text, and so allows one to provide a good comparison of word usage between corpora. A further test of log-likelihood can be performed to show the statistical significance of the difference in frequency of a word used in two corpora (such as RUDaF and FALKO L1, or RUDaF PreWIC and RUDaF During WIC) (McEnery & Hardie, 2012(a)). Measuring statistical significance is a way of saying that a word choice has, typically, a less than 5% chance of appearing as a result of chance or coincidence (i.e. 95% certainty it is not a coincidence). The UCREL log-likelihood wizard (created by Paul Rayson) is used in this analysis as it allows one to perform tests for a significant difference in frequency (of a specific word or sentence) between two corpora, and is easily accessible online[35]. It is based on four simple figures: the frequency of a word in each of the two corpora, and the total frequency of each corpus (number of opportunities that the term could appear) (McEnery &

[33] See Chapter 4 for a rationale of why this corpus was chosen.

[34] The type/token ratio in a corpus is a rather basic statistic which tells one about the range of vocabulary used in that corpus. This is calculated by dividing the number of types (unique word-forms) in a corpus by the number of tokens (individual words) and multiplying this figure by 100 (the closer the figure is to 100 the more dense the text). For RUDaF this is 5500/33987*100= 16.18%. For FALKO L1 this is 10009/76957*100= 13.005%. The lexical density of these two corpora is thus very similar, and the slight advantage shown for the RUDaF learner corpus may be attributed to the large number of topics which students were required to write about within the corpus.

[35] See the Lancaster University website: http://corpora.lancs.ac.uk/clmtp/2-stat.php

Hardie, 2012 (b)). The log-likelihood (LL) score must be "above 3.84 for the difference to be significant at the $p < 0.05$ level" (McEnery & Hardie, 2012 (b)). These simple corpus statistics are combined with an in depth analysis of the keywords in context (KWIC) in the analysis below to provide insights into learner German writing at Rhodes University.

The raw frequencies and normalised frequencies (rounded off to two decimal places) of the items which will be discussed below may be represented in the following table.

Table 10, Raw and Normalised Frequencies of TAG Words

TAG word	RF PreWIC	NF PreWic	RF DurWIC	NF DurWIC	RF RUDAF	NF RUDaF	RF FALKO L1	NF FALKO L1
Zweck	1	0,06	16	1	17	0,5	4	0,05
*Beispiel**	21	1,17	38	2,38	59	1,74	166	2,16
*Meinung**	29	1,61	13	0,81	42	1,24	106	1,38
*Erachten**	0	0	11	0,69	11	0,32	6	0,08
*Ansicht**	1	0,06	17	1,06	18	0,53	16	0,21
Laut	2	0,11	12	0,75	14	0,41	7	0,09
Zufolge	0	0	4	0,25	4	0,12	2	0,03
Fazit	0	0	8	0,5	8	0,24	5	0,06

The raw frequencies listed above and the total raw frequency of each corpus were used to calculate the log-likelihood scores of each TAG word in order to test the significance of each words usage in 3 separate comparisons: 1) in leaner writing before and during the writing course (see LL Pre/During) in order to see if participants used TAG words significantly more during the writing course; 2) before the writing course as compared to mother-tongue writing (See LL Pre/Falko L1) in order to see if there was a significant difference in learner writing before the writing course as compared to mother-tongue writing; and lastly 3) writing during the writing course as compared to mother-tongue writing in order to see if the writing course caused learner writing to be significantly more or less like mother-tongue writing.

Table 11, Log Likelihood Scores[36]

TAG word	LL Pre/ During	LL Pre/ Falko L1	LL Dur/ Falko L1
Zweck	17.8	0.00	37.82
*Beispiel**	7.19	8.24	0.28
*Meinung**	4.51	0.54	3.69
*Erachten**	16.59	2.52	18.91
*Ansicht**	19.19	2.44	20.17
Laut	9.16	0.06	19.88
Zufolge	6.03	0.84	7.20
Fazit	12.07	2.10	12.73

There are some immediate trends apparent in the two tables above. Almost all (87.5%) of the TAG words introduced during the writing course were employed by the learners significantly more often during the writing course as opposed to before the writing course, with the exception of *Meinung*, which was used significantly less often during the writing course as compared to before the writing course (most probably because of an increased use of the synonyms *Erachten* and *Ansicht* during the writing course). As seen in Table 10, many of the TAG words were used very infrequently, or not at all, in the learner writing before the writing course, and in most cases less than in the mother-tongue reference corpus. However, the differences in the use of 87.5% of the typical everyday academic German words in RUDaF PreWIC as compared to FALKO L1 was not statistically significant, other than in the case of the keyword *Beispiel** which was used significantly less in the learner writing before the writing course, as compared to mother-tongue writing. Lastly, there is a significant increase of the use of 75% of the TAG words in RUDaF During WIC as compared to in FALKO L1, with the exception of *Beispiel** and *Meinung* whose difference in usage (see NF scores) across the two corpora are not significant in this case ($p > 0.05$). These results are elaborated upon in the discussion below, where each TAG words usage is examined in the context of each corpus (RUDaF PreWIC, RUDaF DuringWIC, FALKO L1).

[36] Dark shaded squares in this table indicate LL scores which are greater than 3.84 and thus indicate statistical significance at the $p < 0.05$ level (McEnery & Hardie, 2012).

Introducing a topic: Zweck

Zweck and its accompanying phrases was one of the first TAG words to be introduced to participants during the writing course. Phrases with *Zweck* are used in everyday academic writing to discuss the aim of an essay or introduce the topic of an essay. Phrases with *Zweck* are used in introductions (Krummes & Ensslin, 2012). As seen in the tables above *Zweck* was used significantly more during the writing course as compared to before the writing course, and significantly more during the writing course as compared to mother-tongue writing, which reveals that participants significantly incorporated this term in their writing as a result of the writing course which highlighted this word as an introductory word.

If one examines the keyword in context of RUDaF PreWIC, only one participant (P6) made use of the word *Zweck*. This was in the pilot study in 2013, where P6 writes "*Jede Organisation hat eine primaere Zweck und verschiedene andere Ziele*". While this use of *Zweck* is correct and apt here, it does not reflect the collocation of the "purpose of the essay" (which is what we aimed at in RUDaF DuringWIC), and it may be assumed that the participant has looked up *Zweck* as a translation for "aim". In RUDaF DuringWIC however, *Zweck* is used 16 times, all but once in the context of introducing the aim or purpose of the essay, showing an internalization of the use of *Zweck* as an introductory word. Furthermore, participants do not only use one form of the collocation, but employ variation in their use of *Zweck*, as highlighted in the handout (Krummes & Ensslin, 2012). Participants used variant forms such as: "*Der Zweck dieses Aufsatzes besteht darin*", "*Der Zweck dieses Aufsatzes ist* (+Noun Phrase)" and "*Es ist der Zweck dieses Aufsatzes* (+Noun Phrase)". Six out of 13 participants (P13, P15, P16, P18, P19, P20) incorporated *Zweck* into their assignments in this way, with P15 and P14 showing the highest usage of the term. This can be seen in the concordance lines in Figure 3 below.

Figure 3, Concordance of Zweck *in RUDaF During WIC*

Concordance Hits 16

Hit	KWIC	File
1	ıd insbesondere krimibesessen. Der Zweck dieses Aufsatzes besteht darin, die Art	P15_T17_2015.txt
2	eine nette, kleine Krimi und der Zweck dieses Aufsatzes besteht darin zu bespreche	P15_T18_2015.txt
3	Deutschen es nennen, Navy CIS. Der Zweck dieses Aufsatzes besteht darin, meine liebe	P15_T19_2015.txt
4	Der Zweck dieses Aufsatzes besteht darin, ob suedafrik	P20_T13_2014.txt
5	es diese Sprachen ist Afrikaans. Der Zweck dieses Aufsatz ist, Afrikaans und die	P15_T13_2014.txt
6	Der Zweck dieses Aufsatzes ist, die Unterschiede zwisc	P13_T15_2015.txt
7	Der Zweck dieses Aufsatzes ist, einen Ueberblick ueber	P13_T13_2014.txt
8	Der Zweck dieses Aufsatzes ist es, mein deutsches	P13_T14_2014.txt
9	Der Zweck dieses Aufsatzes ist es, die aktuelle	P13_T16_2015.txt
10	Der Zweck dieses Aufsatzes ist es, das Ende	P13_T18_2015.txt
11	r Amtssprache des Suedafrikas. Der Zweck dieses Aufsatzes ist es, zu erklaehren,	P18_T13_2014.txt
12	nalphabeten sind (Witte, 2014). Der Zweck dieses Aufsatz ist, in Frage zu	P15_T16_2015.txt
13	Der Zweck meines Aussatzes ist ueber die Dialekten	P16_T13_2014.txt
14	Es ist der Zweck dieses Aufsatzes meine Erfahrung mit der	P19_T14_2014.txt
15	ber verschiedene Themen. Es ist der Zweck dieses Aufsatzes zu reflektieren ueber diese	P15_T14_2014.txt
16	ng der Standardsprache fuer dieser Zweck ist nicht der Schriftsteller, um Treue	P14_T13_2014.txt

Zweck *in FALKO L1*

In FALKO L1, *Zweck* appears only 4 times (NF 0.05, as compared to 0.5 in RUDaF), none of which are in the context of introducing the topic, aim or purpose of an essay. This would suggest that the mother-tongue writers who contributed towards FALKO L1 did not make use of the collocation *der Zweck dieses Aufsatzes* in their introductions, but made use of alternative collocations (such as those using the TAG words *Thema* and *Frage*) when introducing their topics. This may also suggest that as a result of the writing course, based on the WHiG worksheets, our participants significantly ($p < 0.05$) overuse *Zweck* as compared to mother-tongue German speakers, possibly as they purposefully attempted to incorporate the newly learnt TAG word in their writing.

Giving an example: Beispiel

Phrases with *Beispiel** are used to give examples in an essay; this in normally done in the body of an essay. According to findings from the WHiG project, the phrase *zum Beispiel* is overused by learners of German in Britain (Krummes & Ensslin, 2014). Mother-tongue speakers show more variation in their use of phrases with *Beispiel**, for example they may use any of the following: z.B., *beispielsweise, nehmen wir beispielsweise*, *nehmen wir als Beispiel*, *nehmen wir das Beispiel*, *nehmen wir ein Beispiel*, or *nehmen wir zum Beispiel* (Krummes & Ensslin, 2014).

In RUDaF PreWIC *Beispiel** is significantly underused as compared to in FALKO L1. *Beispiel** is used in 21 instances, 13 of which are in the context of the phrase *zum Beispiel*, 4 in the context of *ein* Beispiel* von*, 2 instances of *Beispiele*, one instance (spelt incorrectly as *"beispeileweise"*

P20_T10_2014) of *beispielsweise*, and lastly one incorrect use of "V*on beispiel*" (P2_T7_2013). There are no instances of the abbreviation *z.B.* in the RUDaF PreWIC learner corpus.

As seen in the use of *Zweck*, there was a significant increase in the use of *Beispiel* in all its forms in the RUDaF DuringWIC learner corpus ($p < 0.05$). In RUDaF DuringWIC there are 38 instances of *Beispiel** and seven instances of the abbreviation *z.B.* (used by P11 on one occasion, and P21 the latter 6 occurances). The most frequent collocation in RUDaF DuringWIC was (as in RUDaF PreWIC above) *zum Beispiel*, with 17 instances of the word found in this context (one *nehmen wir zum Beispiel* P13_T16_2015). However, there was also more variation to be found with 13 instances of *beispielsweise* (of which eight spelt incorrectly: five *beispielweise* [37], three *beispieleweise* P20_T13_2014), three instances of *als Beispiel* (including one instance of *nehmen wir als Beispiel* and one instance of *als Beispiel nehmen*), and five instances of *Beispiel** preceded by an adjective, specifier or preposition.

In relation to the findings of the WHiG project (Krummes & Ensslin, 2014) as discussed above, our learners of German display similar tendencies to other learners of German in their preference for collocation *zum Beispiel* rather than any of the other variant forms. The participants of the writing

[37] This instance of the incorrect spelling of *beispielsweise* as *beispielweise* is in fact noteworthy as on four occasions this was used by P13 who consistently miss-spelt the word in T14 in 2014, and continued to repeat this mistake in T15 and T16 in 2015. This serves to illustrate a point made in Chapter 6: the lecturers noted that students did not seem to be reading and internalizing feedback, and thus were producing recurring errors. This misspelling was corrected in each instances via track changes by Dr Weber, however P13 continued to use it incorrectly in each assignment.

course showed a marked increase in their use of variant forms during the writing course, particularly the form *beispielsweise* and the abbreviation *z.B.*.

Figure 4, Concordance of Zum Beispiel *in RUDAF DuringWIC*

Concordance Hits 38

Hit	KWIC	File
22	kzente der anderen Sprache (zum Beispiel Afrikaans) gesprochen werder	P6_T13_2014.txt
23	nglisch zu verbieten, Canada, zum Beispiel (Anon.). Ich glaube, dass Gesel	P20_T12_2014.txt
24	oerterbuecher kontrollieren. Zum beispiel, Der Rat fuer deutsche Rechtsc	P12_T12_2014.txt
25	erlaendisch hat als Afrikaans (zum Beispiel die Aussprache von 'ek' als 'ik'	P6_T13_2014.txt
26	icht relevant in allen Sprache, zum Beispiel die tansanische Menschen, die	P13_T12_2014.txt
27	. (Wolters:2005) Nehmen wir zum Beispiel eine Frau, die in einer Kueche	P13_T16_2015.txt
28	n. Dies koennte eine Sprache (zum Beispiel Englisch) mit einer der Akzent	P6_T13_2014.txt
29	chaften fuer andere Sprache. Zum beispiel, Franzoesisch hat zwei Sprach	P12_T12_2014.txt
30	ich als ein Dialekt qualifiziert. Zum Beispiel gibt es Unterschiede zwischen	P6_T13_2014.txt
31	esotho Dialekt aus Hoopstadt zum Beispiel hat die Elemente des Afrikaans	P2_T13_2014.txt
32	en beeinflusst zu integrieren. Zum Beispiel, hat das Wort 'Zeitgeist' keine	P6_T12_2014.txt
33	schiedene regionale Dialekte. Zum Beispiel in Bezug auf die Sprache Afrik	P6_T13_2014.txt
34	oreinen gewonnen koennen. Zum Beispiel kann man eine allgemeine Voı	P20_T12_2014.txt
35	aechtnis verlassen muessen – zum Beispiel, koennen wir einen Einkaufsze	P21_T16_2015.txt
36	ecke sondern auch im Alltag. Zum Beispiel koennte man seine Deutchken	P2_T14_2014.txt
37	nutzt „vandaga" in Hoopstadt zum Beispiel. Man koennte nur die Auswah	P2_T13_2014.txt
38	gemeinsame Sprache haben. Zum Beispiel, wann man von Sueddeutschla	P10_T12_2014.txt

Unlike *Zweck*, and in fact all other TAG words introduced during the writing course, the increase of *Beispiel** during the writing course served to make the learner writing more similar to (not significantly different from) the mother tongue writing. In the FALKO L1 corpus, mother tongue writers make use of the TAG word *Beispiel** 2.16 times per 1000, and during the writing course our learners made use of *Beispiel** 2.38 times per thousand. Providing examples thus appears to be an important facet of everyday academic writing in German.

Giving an opinion, exploring the use of the TAG words Meinung, Erachten and Ansicht

As with the phrase *zum Beispiel*, Krummes & Ensslin (2014) found that learners of German overuse the word *Meinung* (as used in the dative phrases *meiner Meinung nach/ der Meinung sein, dass*) in their essays when trying to express an opinion. In the handout, the TAG words *Erachten** and *Ansicht* are presented as alternative forms to use when presenting one's own opinion. *Erachten** is used in the genitive form *meines Erachtens* and the abbreviated form *m.E.* when one wants to express ones opinion in the main body of an essay (Krummes & Ensslin, 2014). *Ansicht* is also used to express one's

own opinion in the body of an essay, and may be used in a number of ways: e.g. *meiner Ansicht nach*/ *der Ansicht sein, dass*/ *die Ansicht vertreten, dass* or *die Ansicht teilen, dass*.

Participants express their opinions in a limited way before the writing course

In RUDaF PreWIC participants made use of the word *Meinung** on 29 occasions (1.61 times per thousand) when trying to express their own opinion. The word *Meinung** was used in various phrases, with the most prevalent collocation being those identified by Krummes & Ensslin (2014) as overused by learners namely: *nach meiner Meinung*/ *meiner Meinung nach* (four instances) followed by *ich bin der Meinung, dass* (three instances). However, aside from these, participants used many other phrases with *Meinung* before the writing course, most of which were incorrect, illustrating a need for direct clarification of the correct collocations of *Meinung*, as was performed during the writing course.

For example, P2 made use of the incorrect collocation *ich bin an der Meinung, dass* in three different writing pieces, inserting the preposition *an* unneccesarily. *Ich bin/er war von der Meinung, dass* also makes an appearance in the RUDaF PreWIC (P3 and P7, T6_2013), which can be seen as a direct translation from the English phase "I am/he is of the opinion that". Variations of the incorrect phrase *in meine/r Meinung* are the most prevalent learner errors involving *Meinung* in RUDaFPreWC (four instances). The incorrect phrases *Meine [] Meinung über* (P14_T11_2014) and *in meine Meinung auf* (T6_2013) are also used by participants as alternatives to *meiner Meinung nach*. P20 (T9_2014) makes use of the phrase *streit für die Meinung, dass* (a direct translation of the English: to argue for the opinion that).

P15 (T10_2014) used *Meinung* in the following way to introduce their essay: „*Ich versuche mit diesem Aufsatz meine eigene Meinung vorzubringen, wie es zu grammatischen Geschlechtern im Deutschen gekommen ist*".The correct use of prepositions thus appeared to pose a particular difficulty for our learners (as seen above), and this was addressed in the writing course.

The TAG word *Ansicht* is used only once in RUDaF PreWIC (P1_T7_2013): „*Das ist meine allgemeine Ansicht und nicht unbedingt die Norm heute"*. The TAG word *Erachten*, a good alternative to *Meinung*, was not used at all before the writing course.

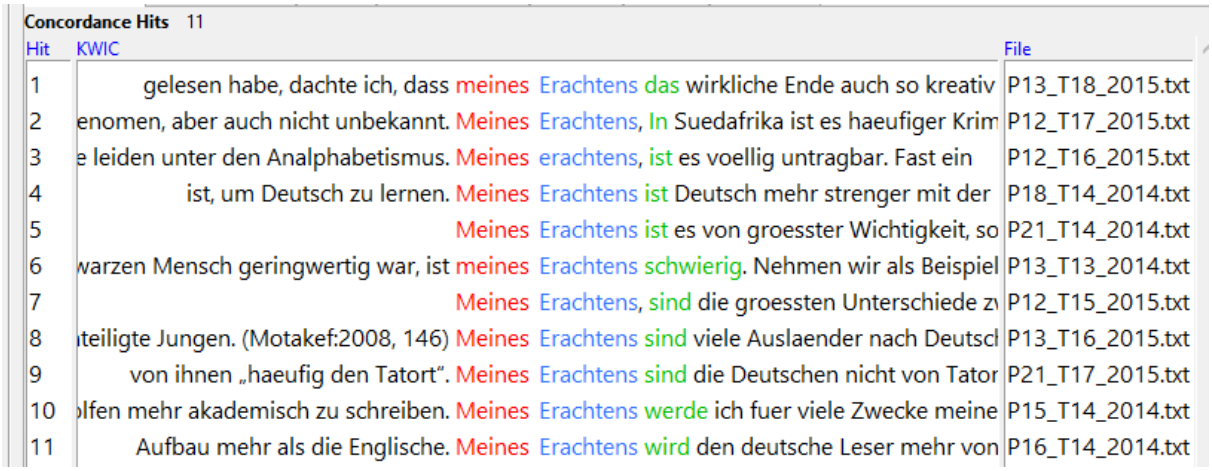
Participants express their opinions more correctly, and more diversely, during the writing course

During the writing course, participants were able to express their opinions more clearly, by making use of more diverse TAG words, and making use of more mother-tongue like, idiomatic collocations significantly more frequently. The first instance in which this can be seen is in the use of the word *Meinung** which was used significantly less during the writing course (13 as opposed to 29 times

PreWIC), and with a far smaller margin of error. The most frequent (and correct) collocational pattern of *Meinung** was *meiner Meinung nach/ nach meiner Meinung*.

While *Meinung** was significantly overused in RUDaF PreWIC, *Erachten** was not used at all and was presented as an alternative form during the writing course. The phrase *meines Erachtens*, which was absent in RUDaF PreWIC, was used during the writing course a total of 11 times by six of the participants (P12, P13, P15, P16, P18, P21) who used this phrase to express their opinion in their writing. The abbreviation *m.E.* was used twice by P13 in 2015 (T16, T18). What is more, this phrase was only ever used in its correct genitive form as seen below.

Figure 5, Concordance of Meinung *in RUDaF DuringWIC*

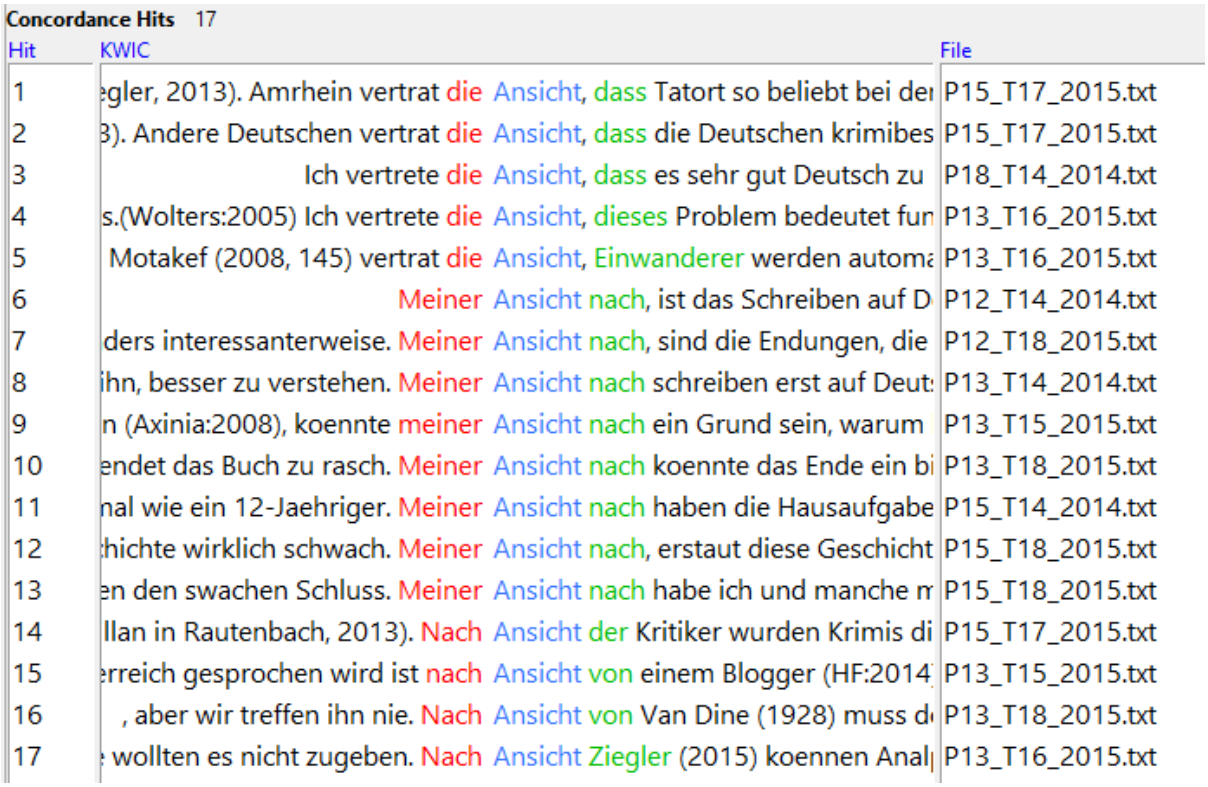


Concordance Hits 11

Hit	KWIC			File
1	gelesen habe, dachte ich, dass	meines	Erachtens das wirkliche Ende auch so kreativ	P13_T18_2015.txt
2	enomen, aber auch nicht unbekannt.	Meines	Erachtens, In Suedafrika ist es haeufiger Krim	P12_T17_2015.txt
3	e leiden unter den Analphabetismus.	Meines	erachtens, ist es voellig untragbar. Fast ein	P12_T16_2015.txt
4	ist, um Deutsch zu lernen.	Meines	Erachtens ist Deutsch mehr strenger mit der	P18_T14_2014.txt
5		Meines	Erachtens ist es von groesster Wichtigkeit, so	P21_T14_2014.txt
6	warzen Mensch geringwertig war, ist	meines	Erachtens schwierig. Nehmen wir als Beispiel	P13_T13_2014.txt
7		Meines	Erachtens, sind die groessten Unterschiede z	P12_T15_2015.txt
8	teiligte Jungen. (Motakef:2008, 146)	Meines	Erachtens sind viele Auslaender nach Deutscl	P13_T16_2015.txt
9	von ihnen „haeufig den Tatort".	Meines	Erachtens sind die Deutschen nicht von Tator	P21_T17_2015.txt
10	lfen mehr akademisch zu schreiben.	Meines	Erachtens werde ich fuer viele Zwecke meine	P15_T14_2014.txt
11	Aufbau mehr als die Englische.	Meines	Erachtens wird den deutsche Leser mehr von	P16_T14_2014.txt

Similarly, the word *Ansicht* is used significantly more often (on 17 occasions as opposed to once) in RUDaF DuringWIC than in RUDaF PreWIC. The collocation *meiner Ansicht nach* was used in 8 instances of 17 (47.05%), *Ich vertrete/ [subject] vertrat die Ansicht* (5 of 17 instances or 29.41%); and *nach Ansicht [subject]* (4 instances or 23.52%). This shows a greater diversity in the students' ability to express their opinions, and a good implementation of the opinion-related TAG words taught using corpus-based methods in the writing course.

Figure 6, Concordances of Ansicht *in RUDaF DuringWIC*



Concordance Hits 17

Hit	KWIC	File
1	egler, 2013). Amrhein vertrat die Ansicht, dass Tatort so beliebt bei der	P15_T17_2015.txt
2	3). Andere Deutschen vertrat die Ansicht, dass die Deutschen krimibes	P15_T17_2015.txt
3	Ich vertrete die Ansicht, dass es sehr gut Deutsch zu	P18_T14_2014.txt
4	s.(Wolters:2005) Ich vertrete die Ansicht, dieses Problem bedeutet fun	P13_T16_2015.txt
5	Motakef (2008, 145) vertrat die Ansicht, Einwanderer werden automa	P13_T16_2015.txt
6	Meiner Ansicht nach, ist das Schreiben auf D	P12_T14_2014.txt
7	ders interessanterweise. Meiner Ansicht nach, sind die Endungen, die	P12_T18_2015.txt
8	ihn, besser zu verstehen. Meiner Ansicht nach schreiben erst auf Deut	P13_T14_2014.txt
9	n (Axinia:2008), koennte meiner Ansicht nach ein Grund sein, warum	P13_T15_2015.txt
10	endet das Buch zu rasch. Meiner Ansicht nach koennte das Ende ein bi	P13_T18_2015.txt
11	nal wie ein 12-Jaehriger. Meiner Ansicht nach haben die Hausaufgabe	P15_T14_2014.txt
12	chichte wirklich schwach. Meiner Ansicht nach, erstaut diese Geschicht	P15_T18_2015.txt
13	en den swachen Schluss. Meiner Ansicht nach habe ich und manche m	P15_T18_2015.txt
14	llan in Rautenbach, 2013). Nach Ansicht der Kritiker wurden Krimis di	P15_T17_2015.txt
15	rreich gesprochen wird ist nach Ansicht von einem Blogger (HF:2014	P13_T15_2015.txt
16	, aber wir treffen ihn nie. Nach Ansicht von Van Dine (1928) muss d	P13_T18_2015.txt
17	wollten es nicht zugeben. Nach Ansicht Ziegler (2015) koennen Anal	P13_T16_2015.txt

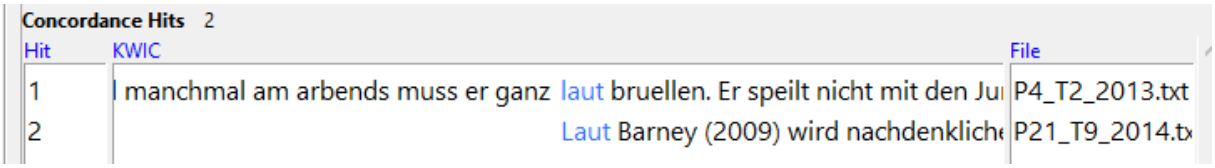
Expressing one's own opinion in FALKO L1

Both *Ansicht* and *Erachten** were used significantly more often during the writing course as compared to in FALKO L1. Mother-tongue writers in FALKO L1 appear to rely on the word *Meinung** in FALKO L1 (106 instances) and using the alternatives *Erachten** and *Ansicht* significantly infrequently (only 6 and 16 times respectively) as compared to our learner writers during the writing course who were encouraged to use these variations.

Presenting the opinions of others: exploring the use of laut *and* zufolge

The TAG words *laut* and *zufolge* were presented to the participants as a means with which to express the opinions of others, or to quote from other people and sources. This is a useful tool in the main body of an essay, where one may want to provide supporting evidence for an argument. In RUDaF PreWIC, participants included very little in the way of supporting quotes and opinions of others, and this may be because they lacked the vocabulary to do so or because the topics did not require them to provide much supporting evidence. *Laut* was used only twice in RUDaF PreWIC, with the first occurrence used in the sense of 'loud' (sometimes the lion must roar loudly in the evening), and in the second instance it was used correctly by P21. *Zufolge* was not used at all.

Figure 7, Concordance of laut *in RUDaF PreWIC*



Concordance Hits 2

Hit	KWIC	File
1	manchmal am arbends muss er ganz laut bruellen. Er speilt nicht mit den Ju	P4_T2_2013.txt
2	Laut Barney (2009) wird nachdenkliche	P21_T9_2014.tx

RUDaF DuringWIC

Laut was used significantly more often during the writing course (successfully on most occasions, 12 in total) by six participants (P12, P13, P15, P18, P19, P21). *Laut* can be used without inflecting the subject (in the case of names), or alternatively as a dative or genitive construction. As can be seen below, this was not always done correctly by participants (See for example P19 and P18 who both make mistakes with the dative *Laut dem Goethe Institute*. P12 uses the genitive incorrectly where the phrase should be *Laut Artikel*, or *laut dem Artikel*) and it may be the case that further instruction into the use of this term is needed.

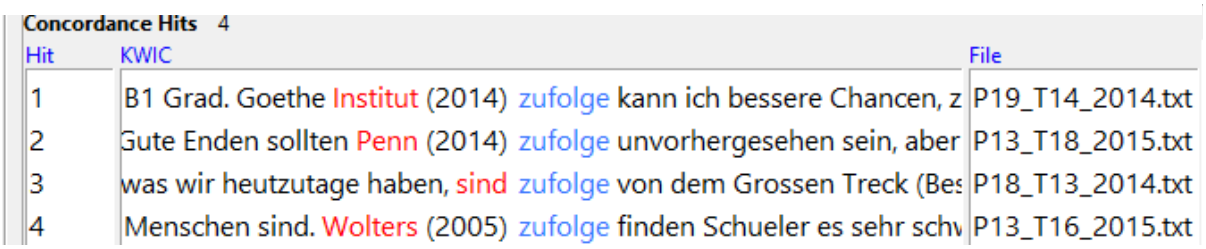
Figure 8, Concordance of laut *in RUDaF DuringWIC*

Concordance Hits 12

Hit	KWIC	File
1	(Goethe Institut, 2014). Laut das Goethe Institut, wenn mans Deu	P19_T14_2014.txt
2	er mich ist das wichtig. Laut dem Schema von Cawelti (1976) pas	P21_T19_2015.txt
3	erper entdeckt hat; aber laut den Autoren war es wichtiger, ueber	P21_T18_2015.txt
4	eutschland reisen, weil, laut der Goethe-Institut, ich ein gutes	P18_T14_2014.txt
5	ch, ob es denn wahr ist? Laut des Artikels, kann ich nur "ja"	P12_T17_2015.txt
6	ches Deutsch. (HF:2014) Laut diesem Blogger (HF:2014) sind ihre	P13_T15_2015.txt
7	iktionale Analphabeten. Laut Jimenez (2012) gibt es 7.5 Millionen	P13_T16_2015.txt
8	ein B1 Zertifikat haben. Laut Johnson Sprachekenntnisse kann ma	P19_T14_2014.txt
9	weil es zu aufrichtig ist. Laut Patterson (2013) muessen wir am Er	P13_T18_2015.txt
10	ist nicht immer der Fall. Laut Studien, haben Analphabeten besser	P21_T16_2015.txt
11	ersitaet in Deutschland. Laut Vorschrift muss mann erst das Goetl	P15_T14_2014.txt
12	Typische Krimi-Schema, laut Wigbers (2006: 207) besteht "den idy	P12_T19_2015.txt

Zufolge, like *Erachten** which was not seen at all in RUDaF PreWIC, was used significantly more often (in four instances) during the writing course in the correct form, as can be seen in the collocations below.

Figure 9, Concordance of zufolge *in RUDaF DuringWIC*



Concordance Hits 4

Hit	KWIC	File
1	B1 Grad. Goethe Institut (2014) zufolge kann ich bessere Chancen, z	P19_T14_2014.txt
2	Gute Enden sollten Penn (2014) zufolge unvorhergesehen sein, aber	P13_T18_2015.txt
3	was wir heutzutage haben, sind zufolge von dem Grossen Treck (Bes	P18_T13_2014.txt
4	Menschen sind. Wolters (2005) zufolge finden Schueler es sehr schv	P13_T16_2015.txt

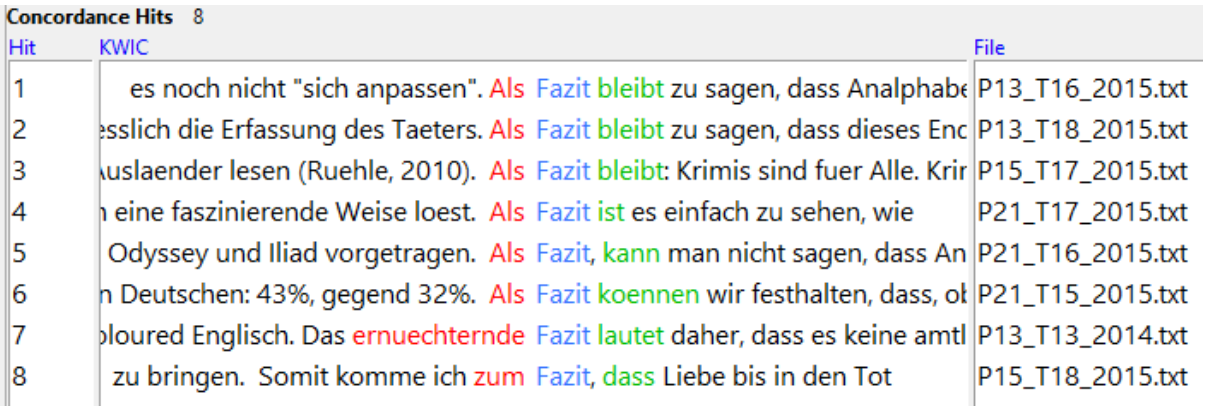
As with the alternate opinion TAG words *Ansicht* and *Erachten*, Laut* and *Zufolge* were both used significantly more often ($p < 0.05$) by participants of the writing course than by mother-tongue writers in FALKO L1 who only used *Laut* seven times (NF 0.09) and *Zufolge* twice (NF 0.03). This may be as a result of the writing topics which the mother-tongue participants had to engage with for FALKO L1, which did not explicitly call for participants to reference others (a set argumentative topic for which participants had one hour to complete). Further research could be done to assess the use of *Laut* and *Zufolge* in other German mother-tongue reference corpora of academic writing.

Concluding an essay, TAG word Fazit

Fazit was introduced during the writing course as a TAG word to use in conclusions to essays or in concluding a paragraph, as we would say in English "in conclusion" so *Fazit* is used in German in the following collocations: *Als Fazit bleibt*, *Fazit ziehen*, *das Fazit lautet* and finally *zum Fazit kommen.*

As with *Erachten** and *zufolge*, *Fazit* did not occur in RUDaF PreWIC, and occurred significantly more often during the writing course ($p < 0.05$). During the writing course three participants (P13, P15, P21) made use of the word *Fazit*, eight times as can be seen in the table below. These participants experimented with some of the above collocations, with the most common usage being *Als Fazit bleibt*.

Figure 10, Concordance of Fazit *in RUDaF DuringWIC*



Concordance Hits 8

Hit	KWIC	File
1	es noch nicht "sich anpassen". Als Fazit bleibt zu sagen, dass Analphabe	P13_T16_2015.txt
2	esslich die Erfassung des Taeters. Als Fazit bleibt zu sagen, dass dieses Enc	P13_T18_2015.txt
3	uslaender lesen (Ruehle, 2010). Als Fazit bleibt: Krimis sind fuer Alle. Krir	P15_T17_2015.txt
4	n eine faszinierende Weise loest. Als Fazit ist es einfach zu sehen, wie	P21_T17_2015.txt
5	Odyssey und Iliad vorgetragen. Als Fazit, kann man nicht sagen, dass An	P21_T16_2015.txt
6	n Deutschen: 43%, gegend 32%. Als Fazit koennen wir festhalten, dass, ob	P21_T15_2015.txt
7	bloured Englisch. Das ernuechternde Fazit lautet daher, dass es keine amtl	P13_T13_2014.txt
8	zu bringen. Somit komme ich zum Fazit, dass Liebe bis in den Tot	P15_T18_2015.txt

Drawing conclusions in FALKO L1

Fazit is also used significantly more often during the writing course as compared to in FALKO L1 (NF 0.06 as opposed to 0.5 in RUDaF During WIC). Similarly, the phrase *Lässt sich sagen, dass* appears only 4 times. This shows a similar trend to the use of *Zweck*, in that our participants have implemented the TAG word taught to an extent that exceeds the "natural" use of the same word by mother-tongue speakers, an artefact of the writing course.

Summary of the results

Given the above evidence it can be seen that the qualitative perceptions of improvement revealed through questionnaires and self-reflection of participants, do indeed match up with quantitative evidence provided through corpus-based (and course-based) analysis of the learner corpus, as 87.5% of some of the typical everyday academic German words introduced during the writing course were used significantly more often during the writing course as compared to before the writing course, with the exception of *Meinung** whose use was significantly less as a result of the writing course which encouraged variation of opinion giving collocations.

In the case of 75% of TAG words (as with the introductory word *Zweck*, the opinion words *Erachten**, *Ansicht**, *laut* and *zufolge*, and the conclusionary word *Fazit*) the participants have instituted the taught TAG words and collocations thereof with significantly far more regularity than is seen in the mother-tongue German corpus, FALKO L1. This reveals that learners may overuse or rely on set phrases taught especially where they had no alternatives beforehand. However, this is not seen as a problem in our context as the participants did not use these words at all before the writing course, and experienced anxiety at the thought of writing in German, relying on translations from English (see Chapter 5). If using the TAG words and their collocations has helped our participants to become more confident in their writing, and given them the tools to express their thoughts in German, then this slight overuse is not problematic and can be seen as a building block to further competency. With further instruction, further alternatives can be taught and participants will be able to employ greater variation within their writing and be able to formulate even more mother-tongue like short academic essays.

This process of learning new variants was seen in the overuse (and incorrect use) of *Meinung* and its collocations before the writing course, and the more correct (and dimished to the point of becoming insignificantly different from mother-tongue usage) use of *Meinung* during the writing course, with alternatives *Erachten* and *Ansicht* used equally as often where they were not used at all before the writing course. It is a great achievement for our learners to be writing longer academic texts when they have previously not written any longer texts in German of a more academic nature before 2013.

Where before the writing course participants were not able to introduce a topic, eloquently put forward their own opinion or the opinions of others, and conclude an essay, during the writing course they were able to make use of the TAG words and collocations taught in order to do all three.

Even without further explicit instruction, participants themselves can investigate other alternative words and collocations to use in their writing with the help of corpus tools and techniques imparted during the writing course (for example by using linguee.de) thus creating long-term benefits for participants through understanding corpus concepts. Thus, as Yoon (2008) highlights, through a corpus-based approach to writing students are able to take more responsibility for their writing.

CHAPTER 6. ANALYSIS OF QUESTIONNAIRES AND INTERVIEWS: THE PARTICIPANTS' PERSPECTIVES

Introduction

Questionnaires and interviews played an integral role in understanding our participants' attitudes towards German Studies and approaches towards writing in 2014. This informed the way in which the writing course was structured, in order to tailor the content specifically to the participants' needs. The second questionnaire delivered at the end of the writing course in 2015 assessed participants' attitudes towards the writing course and the use of corpus tools in 2015. The rationale, design and content of the questionnaires and interviews are outlined in Chapter 4 above. This chapter seeks to explore the topics upon which the questionnaires and interviews were founded, and to engage with the responses which the participants provided. The participant responses inform the qualitative evaluation of whether or not the writing course was successful, and add to the quantitative analysis of TAG words discussed in Chapter 5.

Insights from the first questionnaire 2014[38]

The first questionnaire (2014) was based on the questionnaire created and developed in 2013 for the pilot study. Insights from the 2013 questionnaire shaped the way in which the 2014 questionnaire was formed. The 2014 questionnaire was divided into 3 sections, and these will be discussed below. Section 1 was titled "language background" and aimed to identify participants' motivations towards studying German, as well as towards the specific modules of German Studies, namely grammar, literature, translation and cultural studies. This section also aimed to identify which modules participants found particularly challenging. Section 2 of this questionnaire investigated how students approach writing in German, and aimed to form an understanding of how students use the dictionary, whether they find the dictionary to be helpful, and with what kind of challenges they are faced when using the dictionary. Section 3 investigated participants' engagement with texts and other German sources outside of the language learning classroom, in order to form a picture of the type of textual input which participants have in their journey to German literacy.

Why German?

The first indicator of motivation may be seen in why students chose to begin learning German. In response to questionnaire item 1.5 "Why did you decide to begin learning German?" the dominant

[38] The dominant trends from the questionnaire are represented here. For the actual answers and a summary of the trends with participant information attached, please refer to Appendix 6.

trend (displayed by 50% of participants) was simply a particular interest in the German language (p06, p10, p11, p12, p13, p16, p18, p20), with further trends including a prospective economic/international/marketability benefit (31%), a love of languages in the general sense (25%), a particular interest in the German culture (25%), and lastly a feeling that German was a convenient degree-supplementing subject (19%)[39]. These trends are not unique to our learning context, and indeed mirror the categories of motivation described by other researchers in other parts of the world (see Masgoret & Gardner, 2003).

All sections of German are equally challenging

Question 1.6 asked participants to identify which section of German Studies they found most difficult. There was an almost exactly equal distribution between the four sections of grammar, literature, translation and cultural studies, with p16 identifying more than one section as 'difficult'. Question 1.7 required students to elaborate as to why this was the case. The main concerns presented by participants concerning grammar, translation and literature, had to do with feelings of insecurity regarding German language 'rules' (such as word order) and an inadequate grasp of the vocabulary which was regarded as a barrier to these three areas of learning. Cultural studies on the other hand presented the participants with a set of challenges regarding the application of theories and conceptual frameworks which some participants felt unfamiliar with.

This particular set of responses alerts one to the fact that our participants really are still learners of German, and they face constant challenges regarding their language competency. The writing course sought to address these issues of rules and vocabulary by explicitly teaching collocations (see Chapter 4, "Creating a corpus-based writing course"), and providing a clear set of examples of the patterns which German follows in everyday academic language.

Grammatical and literary favouritism

Question 1.8 asked participants to identify what their favourite section of German Studies is. Grammar (39%), followed by literature (31%), proved to be favourites among the class, whereas translation (15%) and cultural studies (15%) were found to receive far less popularity. Question 1.9 again addressed the question of why this was the case, and an interesting connection to the responses to question 1.7 emerged. Participants who enjoyed grammar answered that they enjoyed the rules and formulas of grammar, and the opportunity to enhance their language competency. Participants who

[39] As is evident, participants did not each identify only one reason, but often supplied a combination of reasons for learning German

showed a preference for literature stated that they enjoyed reading, and enjoyed German literature as it afforded them the opportunity to engage with German texts written by German authors, an opportunity which they found to be "intellectually engaging"(P10), "interesting and challenging" (P12) and "intriguing" (P14).

The participants who preferred translation enjoyed that translation requires contextual language awareness, and an application of knowledge. As P16 stated: "One has to engadge [*sic*] with one's knowledge of the language and find new ways of expressing an English sentence. It also helps one learn idioms and phrases. It is nice to use a language in context as in a conversation where one can practise grammatical structures rather than isolated and unlikely grammar exercises". This response highlights the type of language competency which the writing course would later address specifically (see Chapter 4. "Creating a corpus-based writing course"). Participants who enjoyed cultural studies enjoyed the opportunity afforded to engage with historical and cultural topics that they were previously unfamiliar with[40].

The writing assignments sought to provide an integrated learning opportunity. In instituting the writing exercises as a part of the writing course, I sought to provide a platform for those who enjoy the 'rules' of German to put their learning into practise, with an opportunity to fully engage with language structure and vocabulary, while also requiring students to engage with concepts and ideas which they may find challenging.

Writing in German, taking the "plunge"

Question 2.1 addressed the important question of how participants approach the writing tasks, and whether they write directly in German, or whether they write their work in English first and then proceed to translate it into German. This proved a be a pivotal question, as responses revealed almost a third of the class relied on a translation from English to write their set paragraphs. 23% stated that they first begin their assignments by writing out their thoughts or ideas in English, and then translating this into German, showing no awareness of the effect which this may have upon their writing. 15% of participants stated that they did both, depending on their level of comfort with the vocabulary. For example, P18 states: "I tend to do both. If I know the words, then I plunge straight into German, but if I am unsure, then I do it in English first". P18's use of the collocation to "plunge into German" displays a certain fear or apprehension at the thought of writing directly in German. Given that we know that

[40] Evidence provided by a course convenor in the form of e-mails from previous students, also suggests that students who have completed German Studies and are now working can attest to how useful the cultural studies modules are proving to be in their current professional and personal lives. This suggests that while cultural studies may not prove to be the most popular module in this questionnaire, its effects are far-reaching.

the respondent is a competent English mother-tongue speaker, we can assume that they are employing the word “plunge” correctly.

Reassuringly, the majority of participants (62%), stated that they write their work directly in German. Participants elaborated as to why this was the case, stating that writing in German first is essential as otherwise one runs the risk of ‘becoming confused’ (P12, P20), losing the ‘tone/feeling of the original idea’ (P15), or ‘falling into the trap of direct translations’ (P19). P13’s response encapsulates these feelings as she writes “I proceed to write straight into German because I know things won’t translate perfectly and it is easier thinking about the things you can say in German (i.e. vocab you already know (also Grammar))”. A quarter (25%) of this group of participants admitted that they do think things through in English first, whereas others prefer to think it through in German too, to ‘get into the habit of thinking in German’ (P19).

The answers to this question reveal an interesting pattern. If participants feel comfortable with vocabulary, then they feel confident to write directly in German, however a large number of students (38%) did not have the confidence to write directly in German, a possible reason for inauthentic German phrases, and nonsensical convoluted sentences which were a common feature of the first set of writing assignments. Using these insights, the writing course served to emphasise the importance of training oneself to think in the target language, and to use resources to find and check vocabulary items and collocations while writing, or before writing (by reading relevant German sources). Writing out an assignment in English first and then translating it, as done by 23% of participants, was highly discouraged, as this leads to inauthentic writing as a result of direct translation. Participant answers in the second questionnaire serve to illustrate that this was internalised and that participants did indeed approach writing tasks differently after the writing course. See the sections below for an elaboration on this point.

Dictionary use and preferences

All participants answered “yes” to the question as to whether they use a dictionary (Q 2.2). When asked to elaborate upon whether they used an online or print source dictionary (Q 2.3), the majority of participants, 85%, answered “both” and 15% answered “online”. Participants were then asked to name the dictionary/s which they use to perform these operations. Here, the vast majority (85%) listed Leo.org as their online dictionary resource, with Dict.cc/Dict.co (23%) and Reuerso.com (8%) only receiving small mention. The popularity of Leo.org is unsurprising as this resource had been introduced in class. However, students are also introduced to dictionary work in translation classes and taught dictionary skills using the *Langenscheidt* dictionary, and one would surmise that this would be a popular choice. Students had varying responses when listing the print dictionary/s that they used.

3 participants (23%) listed a form of the Oxford dictionary: P14 “Oxford English – German dictionary”, P21 “Concise Oxford German Dictionary” and P6 “Oxford”. Other participants made mention of the following: P10 “Collin’s Easy Learning German Dictionary”, P11 “Berlitz German-English Dictionary”, P12 “bilingual dictionary”, P15 “Langenscheidt”, P16 “Duden”.

The preference for online sources may be attributed to a shift in popularity from print sources to online sources throughout tertiary settings (Chauhan, *et al*., 2012). Online sources tend to be viewed as more convenient and more accessible than print sources. This is shown in P13’s response: “I use Leo.org and I also have the app for my I-pad”. The writing course sought to harness this trend, and provide students with further online resources to aid with writing.

Participants were then asked (Q 2.5) whether or not they perform a cross-check of the words they look up in the dictionary. This was clarified as follows “I.e. when unsure of a selected word, do you look up that word in a monolingual dictionary, or look up the word found from the other side of the dictionary in order to check that the selected word matches the meaning?”, to which 77% answered “yes” and 23% answered “no”. Cross-checking words is highly important, as often student will settle on the first translation given in the dictionary, and this is often used in the incorrect context, as shown in the writing submitted for the pilot study (Ortner, 2013). This was emphasised in the writing course, and further online tools for checking words in context (within a corpus of mother-tongue writing) were provided.

Question 2.6 asked participants to identify whether they searched for single words or whole phrases. Participants tended to express the view that their searches were dependent on the context (31%). 23% stated that they searched only for single words, 23% for whole phrases, and 38% stated that they searched for both. Responses indicate that participants tend to look for whole phrases if they are uncertain of how to use a word in context: for example P14 states: “I look for whole phrases when I am not sure of the word order in a sentence”. Similarly, P16 states: “Both starting with words and then a phrase ideally as words (alone) can be out of context”. Whole phrases are however not easy to search for in dictionaries, a point noted by P13 who stated: “I like to find example phrases but these are not always readily available”.

These responses displayed that some participants already had an idea of the importance of collocations, however, they experienced difficulty locating phrases in the dictionary. The writing course emphasised the importance of whole phrase searches, and of checking a single word translation in context, as stated above. The writing course also introduced participants to tools such as Linguee.de which allow one to perform a search of a whole phrase which will draw a corpus

comparison even when there are conjugated verbs within the phrase, something which dictionary resources are often unable to process.

When asked whether they found the dictionary to be helpful (Q2.7), 98% of participants answered "yes", with the exception of P02, who showed uncertainty and answered "Yes/No" – which the participant clarified by stating that the dictionary is helpful but "online is far quicker", providing further evidence for a preference of online resources as discussed above (Chauhan, *et al*., 2012). Question 2.8 addressed why participants thought the dictionary was helpful. Participants responded that they found the dictionary useful as a vocabulary expansion tool, as well as a grammatical helper, allowing students to see, as P13 states, "[..] all the information about the word (i.e. conjugations in all tenses, plural form, der, die, das, etc.)". Two participants (15%) referred to the dictionary as particularly helpful because they were able to look up phrases. As P19 states, "I find it helpful because it also has forums where certain words or phrases are discussed and it also has a pronunciation option".

Lamentations on lexicographical limitations

While all participants found the dictionary to be helpful (see above), many found their use of the dictionary to be a challenging experience. This was revealed in the answers to Question 2.9 which asked participants to address the problems or limitations they may have experienced when working with the dictionary. The first issue identified by participants is that of being unable to search for words in context within a traditional dictionary. For example, P10 states "Sometimes dictionaries do not have the words in the particular context I need to use them in". Similarly, P13 states "The dictionary does not always have example sentences which can give the context the word is used in, because it may not be an exact translation from English as we expect". P20 further notes "Often dictionaries do not provide the context that a word is found in or do not demonstrate the different meanings of the word".

Along with context, participants note that 'natural' phrasing is important, and that the dictionary does not always assist with this. As P12 states "Using a dictionary does not assist with phrasing a sentence naturally, and therefore I may still write in a stilted or difficult to comprehend manner. Also it is sometimes difficult to establish exactly which word to use". P19 on the other hand holds a different view: "Sometimes, I prefer to use the printed dictionary because it tells you exactly how certain words are used in that it gives you an idea of the connotations certain words carry". P11 makes note of the fact that "Figurative/Colloquial language is very difficult to translate". P16, who had some experience of working with a corpus, specifically mentions that "It [the dictionary] doesn't show collocations of words often or contexts where that word is most often used, as in corpus". Searching for single words

is not always possible, P19 notes that "Certain phrases do not exist in German and having to break the phrase down can be a bit challenging".

Which word to choose from a list of possible translations also presents significant challenges for the participants, and is viewed as a time consuming process. For example, P02 states "[It] takes a long time to find words and [I] often need to find the specific form of the word which the dictionary uses". P06 "Word choice/ variation. It is also time consuming and sometimes lacks specific contextual phrases". P21 similarly lists "Sifting through numerous meanings and being unable to find the one I need" as a particular issue. P16 is a dyslexic student, and further notes that "Print dictionaries are slow especially if ones battles with alphabet memorisation". This may also contribute towards the popularity for online resources, as one does not have to manually search through pages of text, rather the word for which one is searching appears automatically on screen – provided it is spelt correctly, and in the case of print and online dictionaries, that one has the correct form of the word, particularly in the case of conjugated verbs. This particular challenge can be seen in P14's response: "Sometimes the word I am looking for is not there for some particular reason and I have to use online sources". P18 lists some an additional concern with print dictionaries: "The terms/symbols are always confusing. They are not always the same, plus often the printing is so small that reading it becomes a problem and you muddle up explanations".

The responses, listed above, are poignant and reveal a serious need for better, context-based resources, which are easy to use and navigate, an area which was focussed on specifically in implementing corpus-based tools in the writing course. In order to combat this challenge of decontextualized dictionary searches, participants were introduced to the online corpus resource, Linguee.de which shows a search term in context in two parallel corpora. Participants found this particularly helpful, as will be shown in the sections below where the post writing course questionnaire responses are discussed.

Evidence of external reading

The third and last section of the pre-writing course questionnaire (2014) dealt with the participants' engagement with German texts outside of the classroom. As stated in Chapter 3, the aim of the language learner is to develop "a feel for the language" or language "intuition" (Gabrielatos, 2005, p. 6), something which mother-tongue speakers develop through exposure to language (Gabrielatos, 2005, p. 6). Through exposure to language use, mother-tongue speakers develop a mental corpus which is informally examined in the mind to produce these intuitions (Gabrielatos, 2005, p. 6). In countries where the target language is not widely spoken (as is the case of German in South Africa), learners may lack sufficient exposure to be able to form and recognise patterns. Reading facilitates

language learning, however, this is often expected to be undertaken outside of the classroom, given the limited amount of contact hours in a language classroom (see Chapter 4 for a summary of the German Studies contact hours at Rhodes University). The pilot study revealed that students had very little engagement with German texts outside of the classroom, and that when they did the nature of these texts was not academic (Ortner, 2013). The results from this study show similar findings.

While the majority of participants (77%) answered that they did engage with German texts outside of the classroom, 33% of participants admitted that they did not engage with German reading privately (Q3.1). Those who answered that they did engage with texts outside of the classroom were asked to identify the type of reading matter with which they engaged (for example: novels, recipes, academic texts, news articles, etc). News articles (70%), followed by novels and short stories (60%), were the most popular reading matter choices. In addition to this, participants listed less formal social media sites such as Facebook (20%), and social activities such as e-mails (P16), along with engagement with German songs (P06), and recipes (P02). P13 expressed "I have no time but I will start soon". The majority of the participants (80%) engaged with these resources on a weekly basis (Q 3.3), with P15 listing "daily" engagement, and P16 "weekly-monthly" engagement. All participants stated (Q 3.4) that they did engage with German via other sources such as movies, radio-podcasts, audio books, music, TED talks, YouTube clips and Facebook. Again, there is a noticeable trend towards the use of online resources. For example, P12 states "I watch German movies and listen to German radio online as well as the German news on Tagesschau.de" (Q 3.5). 23% of participants engaged with these types of sources "daily", 54% "weekly", and 23% "monthly" (Q 3.6).

While it is heartening to see that participants do engage with German outside of the classroom, the type of engagement listed above (with the exception of news articles) is not necessarily well suited to becoming familiar with everyday academic German writing, which revealed a need for condensed engagement with this sort of writing genre in class. The corpus-based resources introduced in the writing course sought to do exactly this, and to provide participants with condensed language exposure to help them become familiar with the mother-tongue collocational patterns used in everyday German academic writing.

The last section of the 2014 questionnaire provided participants with space to give any additional comments. P16 and P21 both make reference to spoken language engagement, an area which I had not focussed on in the questionnaire, as I had specifically wanted to gain insights into written language engagement. However, this was an oversight on my part as spoken language competency is of course an essential facet of learning to speak and write in a new language. P16 made note of this oversight directly: "Possibly it would be useful to asses [*sic*] German speaking frequency eg: skyping, telephoning or practicing with friends, as well as tuts in cementing skills learnt". P21 simply added/reiterated that

their engagement with the language extends beyond the categories listed in the sections above: “My best friend is German and we often speak in German together. I made friends in Germany, learned a bit of colloquial German and correspond with them regularly on Facebook (in German)”.

Participants further noted their appreciation for a holistic approach to learning German in the following comments: P13 “I thoroughly enjoy learning German and about Germany and its development. I do feel that all the sub-topics we cover help strengthen what we know about the language”. Similarly, P19 states: “Languages are easier to learn when you are really curious and passionate about them. I guess it was a different experience to have to learn about the culture and the history of a language. It would not be complete if we did not learn about the language holistically, it adds a new dimension to the learning process which can be challenging”. Lastly, P14 expressed their enthusiasm for German, despite individual challenges: “I have found German particularly difficult this year mainly because of all my constraints this year. It is still a lovely language that I am still keen on learning. Doing linguistics has benefited my understanding of language, helps me with grammar and makes me more aware of my learning process”. This comment also highlights the importance of interdisciplinary knowledge. Certain skills are transferrable across subjects. Just as the metalanguage learnt in linguistics can help the language learner to understand grammar, so the writing skills learnt in a foreign language can aid writing in all subjects (Canagarajah, 2011, p. 113). The questionnaire in 2015 addressed this aspect and will be discussed at the end of this chapter.

Insights from interviews, 2014, uncertainties about academic literacy

Audio interviews followed on form the first written questionnaire in order to gain a deeper understanding of issues surrounding participant attitudes towards reading and writing at university, reading and writing in German and the nature of academic literacy. I will only address certain themes discussed in the audio interviews as not all of the content is pertinent to this discussion.

A specific aim of the writing course was to help participants become more academically literate in German. During the interviews, I engaged with participants on the topic of academic literacy. One of the first questions participants were asked was to explain what they understood by the term “academic literacy” in order to gauge the level of understanding that participants had of this term. This question was introduced as a means of gauging whether students had an awareness of the goal towards which they should be striving, and an understanding of the metalanguage used in the writing course. Participants had varying responses which ranged from considerable uncertainty to a deep understanding of academic literacy.

The level of understanding appeared to correlate with the level of study of the participant. The second-year participants expressed the most uncertainty, see for example P10’s very flustered response: “Um,

oh wow! Um, (giggles) Um, like wow it's hard to explain, like knowing the, oh god, um, um…". Whilst not quite so uncertain as P10, P12 said "I'm not sure I have actually come across that term before. I suppose that if one had to take it literally it would be literacy in academic texts?". P13 asked whether academic literacy referred to both reading and writing, (this was confirmed to apply to both) to which P13 replied "so probably just in a more formal formal [*sic*] way". P14 and 15 also expressed uncertainty in their responses as follows: P14 "Academic literacy... um, like wouldn't it be like writing and reading correctly or something?". P15, "Well, I'm thinking that would mean like academic articles...". Whilst all of these participants tend towards an understanding of academic literacy, only P11's response gets closer to a deeper understanding of academic literacy: "I guess it's being able to read and write […] academically without too much difficulty. So beginning to make sense of say, academic textbooks, being able to write academically and being able to convey your thoughts… ah… effectively".

The third-year participants on the whole showed a simple understanding of academic literacy as being able to read and write in an academic manner, but did not expand on what this may mean and some participants showed uncertainty relating to their own understanding of the term. For example: P20 "I think I understand it, I'm not overly familiar with it. I would think it was like how you understand academic literature and interpreting that and things". P16, P17, P18 and P21 gave a simple surface understanding of academic literacy as a term: P16 "Like writing, knowing how to write in an academic context"; P17 "Well, being able to write academically?"; P18 "I assume it means when someone can understand academic writing" and P21: "It's being able to read and write in an academic manner, like research reports and reading academic articles and stuff". P19 provided some examples and made inferences towards discipline specific knowledge, but did not expand on what knowledge would be needed in order to write on a particular topic, P19: "I think it means literacy on a particular subject, like writing on a particular subject. For instance an article, maybe a journal article about interest rates, or a journal article about how to, let me think, maybe about a particular subject like um accounting."

The Honours-level students had less difficulty and displayed far more certainty when providing an explanation for the term. P06's understanding of the term academic literacy was a "[…] basic understanding of how to read and write texts in an academic environment". Similarly, P02 stated "I think it's just knowing how to write in the correct register, using the right tone [and] structure to form an opinion, something you do almost unconsciously by the time you get to […] postgrad". These responses reflect the definition of academic literacy presented in Chapter 1, "The concept called literacy involves knowing how to speak and act in a discourse. Academic literacy involves knowing how to speak and act in academic discourses" (Boughey, 2000, p. 283). After talking about the writing course and the aims of the course, and about the expectations of students with regard to learning to write in an academic way, P02 stated that from her experience of the English department and from

tutoring English she had gained the following insight: P02 “Though I see in English, the way we guide our first years, it’s very subtle you know teaching them how to do academic work. So we don’t sit them down and go this is academic literacy, this is what you need to know. We make them write and then we say this is where you’re going wrong, this is where you need to fix up”.

A brief summary of the insights from the 2014 questionnaire and interviews

The answers above, delivered in a questionnaire and interview issued in 2014 before the writing course, thus offered many insights into the participant attitudes and thought processes. In summary, the overarching understandings which emerged from the questionnaires and interviews were as follows: participants felt motivated to learn German, and found all sections of German equally challenging. Participants expressed a certain amount of apprehension at the thought of writing in German, and 23% of participants stated that they write their assignments in English and translate them thereafter into German. By analysing the participant answers, one may note that writing in German appears to be related to how comfortable participants felt with the vocabulary. Participants stated that they did make use of the dictionary (particularly online dictionary resources) but noted many problems or limitations experienced in their dictionary use. While participants did engage with German outside of the classroom, the engagement was not with academic German texts. Lastly, undergraduate participants expressed much uncertainty relating to academic literacy. The insights gleaned from this questionnaire informed the structure and content of the writing course, in order to tailor the exercises specifically to the participants needs[41].

Insights from the post-writing course questionnaire

Four third-year students and one Honours-level student took part in the corpus-based writing course and answered the second questionnaire in 2015[42]. All participants had been involved with the writing tasks since 2014, and had undertaken both sections of the writing course (2014 and 2015). This questionnaire was set up via the Rhodes University online student services platform, RUconnected, and ensured anonymity of the participants in order to encourage them to answer as honestly as possible. Unfortunately this also limits analysis as cross-comparisons to individual answers from the

[41] An overview of the rationale and structure of this questionnaire may be found in the Chapter 4, “Methodology”.

[42] A note on our Honours level participant: This participant felt that the writing course was beneath them as they had recently completed a C1 level German exam at the end of 2014. This participant had also been writing paragraphs since the original pilot study in 2013. One may notice a that there is consistently a negative response set in the answers below, and while the questionnaire is anonymous, it is useful to note that this particular respondent did not feel that the course helped to improve their writing as they felt their writing was already good, a sentiment expressed by our honours level participant.

first questionnaire cannot be drawn. For further discussion on this particular challenge, see Chapter 7, “Challenges to the project”. The insights are nonetheless valuable, as they serve to provide further evidence of the effectiveness and usefulness of the corpus-based everyday academic writing course in German.

Participants perceive an improvement in their writing

80% of participants felt that their writing had improved since 2014 (75% of this number “strongly agreed” that their writing had improved), while 20% of participants were “neutral”. Question 2 addressed the question of why participants thought that this was the case. The participant who showed neutrality expanded on this view: “I don't feel there has been a significant improvement in my writing, nor a significant decrease. I have not had as much time to dedicate to German in my free time this year as I would have liked, and I think my writing reflects this”. Thus we can see that for this student, perhaps even further German writing tasks were necessary. On the other hand, the participants who expressed that they felt their writing had improved specifically attributed this to having to write more, more often, in German. For example, a participant states: “I think this is really because of the paragraphs we had to write in German. At first the writing pieces I did were just paragraphs, but they soon became so much more and started to represent real (short) essays”. Similarly, a second participant makes the claim that: “Due to the constant practice (paragraphs) and feedback [my writing has improved]”. The mention of feedback here is important, as it is an important part of the writing process to review ones mistakes. Another participant notes that while they perceive that their writing in German has always been good “[…] [T]he paragraphs have helped me to write about topics with which I am not necessarily comfortable and in which I am not necessarily interested – this is useful because, if I have a German related career one day, I won't be able to stay in my comfort zone”. Finally, a participant makes specific mention of the writing course content as helping to improve writing: “With time, my vocabulary has grown, furthermore, my sentence structure is more stable, i.e. using the correct articles for the specific case (nominative etc.). Coming into contact with linguee and getting in more depth with the idiomatic type phrases German makes use of in academic writing has really helped to make writing easier and comprehensible”. This serves to illustrate that the knowledge of specific everyday academic German collocations was particularly useful for this participant. As seen in all the responses above, the majority of the participants express the view that their writing has improved as a direct result of the writing course.

Writers feel more confident, and have a changed approach to writing

One issue raised in questionnaire 1 (2014) was a lack of confidence on the part of students to “take the plunge” (P18) and write directly in German. 80% of participants felt more confident in their writing

after the writing course (Question 3), a heartening result. Question 4 of the post-writing course questionnaire asked the important question of how participants approach writing now (after the writing course) and whether this is this different to how participants used to approach writing tasks before they started writing regularly in German. The responses illustrate that the learning undertaken in the course was internalised, and display a great degree of self-awareness on the part of the participants. As seen in the first questionnaire (2014), there were still a number of students who proceeded to write assignments in English first, and then in German. These responses show an awareness for the importance of trying to "think" in German, as well as an increased understanding and ability to structure paragraphs and mini essays. The responses are listed below as they speak for themselves, and provide evidence for the successfulness of the writing course in facilitating a more mother-tongue-like approach to writing. Once again it may be noted that the analysis of these responses is limited by the anonymity of the questionnaire, as a cross comparison of participant answers is not possible.

"I do approach it differently, as I try not to structure my writing in the same manner as I would structure English. I also begin in German now, as opposed to starting with an English base and trying to shoehorn it into German."

"Yes with regards to how I used to approach writing, I feel that before I had no structure and struggled to start, however, now I have more knowledge on how to structure my writing pieces and it is indeed much easier to start and structure one's essay."

"I find that I approach writing with more skilfully [*sic*], Instead [*sic*] of just writing whatever comes to mind, I try and structure my senetences [*sic*] more carefully with the learnt academic phrases in mind."

"Yes, it is much different. I find that now I tend to focus more directly on the German language. By this I mean that I try to not first think in English and then look up the German equivalents anymore, but rather try to switch on the tiny German part of my brain when I start writing in German and only look up things that I am really not sure about. This is still not easy, but I am so much better at it than I was before."

"Actually, my writing approach is no different. Obviously I formulate a vague idea in my head in English about what I'm going to write, but I NEVER write a paragraph in English and translate it into German. More and more, however, I am able to think in German and write much faster."

While 62% of participants prior to the writing course wrote their assignments straight away in German, by the end the writing course 100% of participants approached their writing tasks by writing directly in German. As seen in the responses above, participants felt that they had a better grasp of the German language, and a better ability to structure writing assignments in German. Participants of the

writing course no longer attempted to “shoehorn” their English-based writing into German, but rather “switched on” the “tiny German parts” of their brains, and were able to “think in German and write much faster”. While they expressed that this was “still not easy” participants felt a gained sense of confidence, expressing “I am so much better at it than I was before”

Writing more, more often, is most helpful

In response to whether the participants found the writing course helpful (Question 5), 60% strongly agreed, 20% agreed, and 20% were neutral. The answers to question 6, which asked the participants to identify what aspects of the course had been most helpful, offer some enlightening insights into the usefulness and efficacy of a corpus-based writing course. Responses given by our students reflect the findings of those discussed in Chapter 3 (Yoon, 2008; Oghigian & Chujo, 2010; Varley, 2009), as 60% of the participants voice that they found the new knowledge of academic collocations to be very helpful, and that seeing the vocabulary in the context of a corpus “really cemented the knowledge”. Response 1: “I definitely found the discovery of linguee to probably be the best part of this course. I also found the classes in which we discussed how one would use a certain phrase ("Der Zweck diesels Aufsatzes ist", "laut", "beispielsweise" etc.) very helpful, as well as seeing it in context by viewing the corpora. This really cemented the knowledge”.

In the second response below, a participant notes that the writing exercises gave them the opportunity to practise the vocabulary learnt in class: “It was very useful to be able to learn more sophisticated vocabulary to enhance my writing, and to do in [*sic*] exercises in which I could practise this vocabulary”. This response provides evidence for the validation of the second aim in instituting the writing course as stated in Chapter 4, which was to provide a platform for participants to engage with the vocabulary and phrases learnt in the writing course, to encourage the move from language research, to language production (Oghigian & Chujo, 2010). Moreover, a third participant notes “looking at academic phrases and especially learning how to use antconc [*sic*]” as particularly helpful. This participant further states that AntConc “is a great program!” which shows that the concordancer itself was seen as particularly helpful by some participants.

The next two responses do not mention collocational knowledge, but highlight the importance of the writing tasks in improving German language competence and confidence. These responses support my first aim in instituting the writing exercises, which was to improve participants writing skills in German through the process of writing more, more often (Estes, *et al*., 1998). A participant states: “Being forced to write regularly has been helpful and I have had to tackle different ways of thinking in order to express my thoughts in German”. In addition to this, another participant states: “I found the actual writing pieces to be the most helpful, although the corpus exercises were very interesting. This

is partly because when I started writing in German I started looking up more German sources. Which meant that I not only learned tow [*sic*] write better in German, but also began reading more in German". As revealed in the first questionnaire (2014), participants did not engage with German articles of a more academic nature very regularly at all before the start of the writing course. As this participant notes, in order to write academically, one has to read, and by reading, one becomes a better writer (Cunningham & Stanovich, 1998). This is a cyclical process and it is a very positive outcome of the course that students were able to engage more with more academic German sources.

The findings above from the post-writing course questionnaire directly mirror Yoon's (2008) findings discussed briefly in Chapter 3. Yoon instituted a writing course with six L2 writers in an English for Academic Purposes writing course. This was not discipline specific. As with the participants in this study, Yoon (2008) found that the six participants in the study noted an improvement in their writing and attributed this to both writing instruction (corpus based methods were included) and writing practise in the first half of the year. In the second half of the year all but one student felt more confident in their writing and again attributed this to: "1) the writing course (partly corpus use), 2) more writing experience, and 3) their enhanced disciplinary/content knowledge" (Yoon, 2008, p. 42). The student who did not feel more confident stated that the main reason was because her course offered few opportunities to write in English, thus providing few opportunities for writing practise. The participant noted that her "writing skills may even have declined because she did not have a lot of writing practise" (Yoon, 2008, p. 42). While Yoon seemed to accept this as a matter of course, this highlights the extreme importance of having writing practise in order to improve. Instruction alone, (even real-life "revolutionary" corpus-based instruction) is not enough, as students need to practise writing in order to improve writing.

Early mornings create sluggish responses: Aspects of the course which needed improvement

Following the above, participants were asked to identify aspects of the course which they felt needed to be improved. The participants on the whole seemed happy with the course. Respondent 1 stated "Nothing I can think of" while respondent 2 stated that: "There was nothing "wrong" with the course – I just felt I would have responded better to it if the lectures were slightly later in the day". While we cannot change the university lecture timetable, it is necessary to note that students tend to be less responsive in early morning lectures. Respondent 5 states: "Overall it was really great. As far as the corpus lectures go, it would have been nice to have had a handout to know what exactly we would be looking at every week and to also keep track of the course/homework that needed to be done. Other than that it was really very interesting and insightful". This is noted, and a handout which outlines the homework tasks in advance can be easily implemented in future courses. However, this comment also

raises the question (which cannot form part of this study) as to how independent our students are regarding note-taking and list-making.

Respondent 3 noted that it was difficult to apply the knowledge learnt in class to their own writing, highlighting the difficulty which students face in making the move from being a language researcher to a language producer (Oghigian & Chujo, 2010, p. 202), as highlighted in Chapter 3. The respondent states: “Perhaps a different approach to structuring German texts, as sometimes I found it tricky to apply the essay writing skills from the WHIG worksheet in my own writing”. Perhaps an even greater emphasis on practising writing in class using the collocations learnt could be emphasised in future courses in order to facilitate the application of knowledge.

Lastly, respondent 4 states: “If anything, perhaps we could also go through how one would go about referencing in German, since we are writing academically and need to include references. I know that in-text referencing would be the same, but the terms one would use for “Accessed” or “Online”. It really doesn’t need any other kind of improving”. Referencing German academic texts was not addressed in the 2015 writing course which focussed on learning new academic collocations and structuring of assignments; however, this is something which could certainly be added to further German Studies writing courses.

Corpus tools are seen as helpful, increasing “diversity” in writing

As seen in the responses above, there was a generally positive attitude to the use of corpus tools, and participants found the knowledge of collocations and the application of knowledge in regular exercises to be particularly helpful. This was confirmed in the responses to Question 8 which asked participants directly whether they found the corpus tools to be helpful. 80% of participants (i.e. all but one) found the corpus exercises to be helpful tool. The general trend shown in participant responses was that participants felt they were able to revise the academic phrase learnt within a corpus, as well as learn new phrases in context. A participant responded: “I found them useful as I was able to revise many academic phrases which have helped me improve my writing in that it has become more diverse (if you can put it that way). When writing I now have more options and more ways of structuring my argument”. This notion of “diversity” in writing is essentially what the WHiG authors (Krummes & Ensslin, 2012) aimed to institute in their creation of the WHiG worksheets, by giving students valid and diverse options to use in academic writing. Similarly, a second participant notes specifically that: “Yes, definitely very useful, as I've mentioned before, it was one of the very useful things about the course because you get to see the general context in which the phrase/term is used. This is helpful, because it means we can show more variation in our essays by using those terms or phrase with, for example, adjectives”.

Further responses revealed that participants found it particularly useful and interesting to compare their own writing in the RUDaF corpus to the writing of German mother-tongue speakers in the FALKO L1 corpus. A participant responded: “I found it interesting to see the word choices made by German speakers as opposed to non-German speakers”. Moreover another participant notes: “The corpus exercises gave us a chance to look at our writing pieces as a whole and also gave us an idea of how German students write”. The participant who stated that they did not find the corpus exercises useful stated “I did not find the corpus exercises useful in terms of improving my writing skills - however, I did find it rather fun and interesting to examine the various corpora and learn how to use a new computer programme!”

Rating methods for learning new academic words/phrases: Looking at a German corpus is most helpful

Question 10 asked participants to identify how helpful various methods of learning new academic words or vocabulary items were, according to a ranked scale (1= not at all helpful, 5= extremely helpful). Seeing a word/phrase in the context of a German corpus was rated with the highest average score of 4.0/5, (helpful) with 60% of participants rating this as “extremely helpful”, 20% “neutral” and 20% “somewhat helpful”. Following this, participants perceived looking up the word/phrase in a parallel corpus (German with English translations such as Linguee.de) to be helpful with an average score of 3.8/5. These two high scores show that participants valued a context-based approach to learning new vocabulary. Also viewed by participants as “helpful” with lower raking scores were seeing the word/phrase as it appears in a list of possible words to use (3.4/5), and completing corpus-based exercises such as the TAG words worksheet (3.2/5). Participants rated using “a dictionary definition” as a tool to learn a new academic word /phrase with the lowest average of 3.0/5.

Final comments on this questionnaire included the following positive response: “Thanks so much for the course. It is exciting to know that I can now write and understand German better and also write German in an academic manner (or at least in a more academic manner)”, as well as a participant suggestion for further course content: “It might be a good idea to discuss hedging, not everyone knows what it is and revision would be good. (See what I just did there :)) This means going over konjunktiv 2 and other related words like perhaps, possibly, it can be said etc. Just an idea, I have no idea if hedging is that important in German, I just assume it applies. Bitte schön!”.

Learning to write better German helps writing in other languages

While the above questionnaire addressed participants’ views on the development of their writing in German, a follow up questionnaire sought to investigate whether participants of the writing course had consciously perceived any impact on their academic writing in other subjects as a consequence of the knowledge acquired during the writing course in German Studies. As mentioned before, certain

skills which are developed while learning a new language can be transferred to other realms of learning, and this applies specifically to developing ones academic literacy. This follows on from the work of Canagarajah (2011) who writes on the topic of writing to learn and learning to write by shuttling between languages, as well as the work of Yoon (2008) whose findings revealed that a corpus approach to learning to write academically in English as a second language "heightened students' language awareness, which, in turn, affected their approaches to writing and the writing process" (Yoon, 2008, p. 43). Yoon further asserts that "Regardless of their frequency of corpus use, their exposure to it made them aware of the importance of common usage and collocation in writing" (Yoon, 2008, p. 44). Becoming aware of the importance of collocational patterns may help students to reflect on their language use in their mother-tongue or other additional languages. As Yoon states, a corpus can serve both as a "meaningful reference for language input" and also as a catalyst in helping students to "become more attentive to their writing" (Yoon, 2008, p. 44), and this can extend across language barriers.

This follow-up questionnaire was divided simply into two questions asking participants to identify whether they thought their writing (of assignments) for their other academic courses/subjects has improved as a consequence of producing texts in German 1) with regard to writing in an academic way in general, and 2) with regard to structuring an assignment. Participants were then asked to elaborate on their answers for each point. The responses show an overall trend that students perceived that the writing course in German helped them to better grasp academic writing across all subjects, but did not in all cases necessarily help them to better structure assignments.

P21 who had been involved with the writing tasks since 2013 stated that the writing assignments had helped her develop her academic writing skills in English, but not her essay structuring skills. P21 elaborated by saying that "German academic writing has still made me think about my English academic writing, because the editing process that goes into academic writing is fundamentally the same in any language. One has to think: Is this in proper academic register? If not, how can I resolve this? Are my sentences structured in a way that is readable, but not too simplistic. This applies to any language, so the more one thinks about academic writing the better. Punkt".

P13, who was a weaker student (according to their marks), states that they felt that both their academic writing skills and essay structuring skills improved as a direct consequence of the every academic German writing course. P13 writes: "I believe that, because of writing academic essays in German, I have become more aware of how to write academically in English. I am more aware of referencing and writing more formally. I also feel that my vocabulary improved because of the academic word use I experienced in German, and I then implemented [E]nglish equivalents (of certain academic jargon) into my English essays". P13 continues by elaborating of the effect of the writing

course on the structuring of their assignments: "I think that because we were given tools (i.e.) "*der Zweck*" or "*zufolge*"), it confirmed my knowledge of how one should structure an essay. I always struggled to start an essay but it has become easier by considering how I would do it in German and also what should come next and providing examples has a more [*sic*] clearer place in my essays. Before, I used to have to write parts and make the essay flow afterwards, but now I can write an essay that flows the first time".

P09, an academically strong student, whose writing was excluded from the learner corpus as their German language level was above B2, but whose response is still worth recording, makes note of the fact that again writing more, more often in all subjects (including German) has been significantly helpful in developing her academic writing skills. P09: "Firstly, my academic writing has improved due to the amount of academic texts written in German as well as in my other subjects. The continued and consistent assignments and essays have thus all helped. Specifically German has helped me to pay more attention to the meaning of meanings of words. I have found the translation module particularly helpful in this regard. Instead of just looking at my thesaurus and using a word because it sounds good, I am more aware of the consequences of using a particular word in the given context. This has helped me to make my writing more concise and relevant". This reflects what Canagarajah's viewpoint that "multilingual students can become sensitive to the ways in which writing shapes their thinking and knowledge representation as they learn to write in different languages" (Canagarajah, 2011, p. 113). P09 displays this heightened sensitivity towards the effect of word choices in her writing.

Similarly, Canagarajah (2011) states that "the possibility of looking at the conventions of writing in one language through the spectacles of another provides an effective form of defamiliarization of writing conventions, enabling multilingual writers to explore alternate textual forms and alternate ways of representing knowledge" (Canagarajah, 2011, p. 113). This deeper awareness of writing conventions is shown by P15:

> Writing the paragraphs in German while learning to do so in a more academic manner made me analyse the way I write academically in English and Afrikaans. While we were doing the course on German academic writing we learned many ways of forming an introduction, different means of providing examples and a variety of ways to conclude our short essays. This made me think of how I form introductions and conclusions in my other academic writing. I am now trying more variations in my English and Afrikaans academic writing when it comes to conclusions and introductions.

Participants appear to perceive learning how to structure an assignment as something which one acquires over time though being at university, and through practising writing. This view was put forward by four of five participants (all but P13, see above) as follows: P09 "I don't think German is the only reason that I have improved with regard to structuring my assignments. I think it is rather a

combination of all the subjects and the fact that I have now written many essays since first year". Similarly, P15 writes "I have been at University for quite a long time which means I have written many academic essays. When we started writing the paragraphs I immediately started writing them as essays with an introduction, body and conclusion. So I used my previous knowledge to write/structure the little German paragraphs". P12 does not state where or when they learnt to structure an essay, but writes "I don't believe the way that I structure my assignments has altered significantly as I have consistently worked in a similar way for many years". Lastly, a rather peculiar answer was delivered by P21 who felt that their assignment structure had not improved, nor had it needed to: "In terms of structuring an assignment I already knew what the course taught. My academic register is a bit off sometimes, but I don't think I've ever had an issue with structuring an essay. My point is not to be arrogant or complacent but merely to illustrate that what I've gained from the course had more to do with academic discourse. Structuring one's thoughts requires logic, which one either has or doesn't"[43].

P12 was the only student who did not perceive an effect on their writing in English as a result of the writing course: "The style that I employ when writing in German differs from the style employed in my other academic writing. As I am currently writing a [long research essay], it is difficult to track the influences on my writing as I have been reading very broadly across many fields of literature". This response shows that P12 perceives German academic writing as completely separate and different from other (in this case, English) academic writing.

Conclusions

The above responses illustrate how our students approach writing in German, and engage with the German language both within the language classroom as well as outside of it. By analysing the participant responses before and after the writing course, one may have a clear view of how and why the writing course was structured, and what effect the writing course had on the students' perceptions of their own writing skills. The responses to the second questionnaire in particular reveal on the whole a very positive qualitative evaluation of the everyday academic German writing course on the part of the students. Most participants perceived an increase in both confidence and writing ability and attributed this specifically to the writing course; findings which mirror the results of other, similar research undertakings in foreign contexts (Yoon, 2008). While the writing course was perceived to be very helpful, this was perhaps equally due to the corpus-based instruction, as to the intensive writing

[43] I do not agree with the latter part of P 21's statement, as structuring an assignment does not come naturally to many students as a matter of logic. Logical thought can, and is, often taught, and the same is true of essay structure.

exercises instituted. By learning more about academic writing in German, our participants on the whole also became more aware of their own academic writing in English. This points to the value of a multilingual education which helps one to become more aware of one's own language use in the mother-tongue.

CHAPTER 7. CONCLUSIONS AND RECCOMENDATIONS

Final Summary

The principal aims of this study were to better understand how learners of German at Rhodes University write in German and approach writing in German, in order to implement a course that uses and adapts the tools and methods of corpus linguistics in order to improve the writing of students in German at Rhodes University, as well as to explore the efficacy of instituting a corpus-based approach to teaching and learning in the South African tertiary context.

In Chapter 1 the challenge of developing academic literacy in in the South African tertiary context (in first, second and foreign languages alike) was outlined. The role of German Studies in South Africa is an important one, with benefits spanning the realms of economic development, to inter-cultural competence. While there are shared values norms and etiquettes across the realms of both English and German everyday academic writing, the formal language employed in academic discourse is often used metaphorically, and is not directly translatable, and therefore needs to be explicitly taught. As then shown in Chapter 2 and 3, teaching academic writing can be achieved through a combination of writing more, more often, coupled with language instruction based on insights gained from mother-tongue and learner corpora.

The theoretical underpinning of this thesis has rested upon recent corpus linguistic theory developed in the past 20 years as technological advances have made the systematic analysis of large bodies of text possible through computer software (Bonelli, 2010; McCarthy & O'Keeffe, 2010; McEnery & Hardie, 2012(a)), as described in Chapter 2. The pedagogical concerns which have arisen from this development underline the importance of teaching language based on real-life data or actual language use (Sinclair, 1997, p. 32-34) rather than on a teacher's linguistic intuitions which are seen to be helpful, but not always accurate or reliable (Gabrielatos, 2005).

As these concerns have been realised, many researchers have championed corpus-based approaches (Sinclair, 1997; Leech, 1997), and research into the many facets of teaching and learning which corpus linguistics can impact (such as materials design and assessment, see Chapter 3) have been researched in English (Granger, 2004). Collocational patterns in corpora have been found to be highly important as researchers such have Sinclair (1987) have demonstrated that collocations are indicative of psychologically related, semi-preconstructed phrases (or 'chunks' of language (Ellis, 2003)) which are now considered to constitute single choices for the mother-tongue speaker (Sinclair, 1987, p. 320). This has implications for teaching foreign languages, and everyday academic language in a foreign language in particular as many of the phrases that we use in academic language are not directly translatable, with the result that language learners may make use of grammatical, but non-idiomatic

phrases (Krummes & Ensslin, 2014). Recent research undertaken for teaching German as a Foreign Language has drawn on the creation of learner corpora to enhance understandings of how learners of German write in the target language, and corpus-based materials created for teaching everyday academic German.

These understandings informed the methodology for the research, and the formation of the research goals, outlined in Chapter 4. These were to 1) identify the writing attitudes, methods and strategies employed by our students when writing in German, 2) to – through the implementation of a course that uses and adapts the tools and methods of corpus linguistics – improve the writing of students in German at Rhodes University, 3) to assess the attitudes of German Studies learners towards the use of corpus linguistic tools in developing academic literacy, and 4) to evaluate the corpus-based course in terms of its usefulness and effectiveness by comparing the writing of our students both before and during the writing course with reference to five key everyday academic collocations, as well as to mother-tongue writers of German.

Addressing the research goals

1. *Identify the writing attitudes, methods and strategies employed by our students when writing in German*

In Chapter 6 it was shown that the writing course had a great effect on participants' confidence and their approach to writing in German, and even English. Insights from the 2014 questionnaire revealed that before the writing course, almost a third of participants were writing their assignments in English first and then translating this into German. Some participants displayed feelings of fear and anxiety at the thought of writing in German. This fear and lack of confidence was found to be correlated to an inadequate grasp on the necessary vocabulary needed to write an assignment in German. Where one lacks necessary vocabulary, the place one normally turns to for help is the dictionary. Responses to the 2014 Questionnaire showed that participants did make use of the dictionary, but displayed a marked preference for online sources, a trend which mirrors recent findings relating to tertiary education worldwide (Chauhan, *et al*., 2012). While participants did find the dictionary helpful, they noted many limitations from their experience of using the dictionary, and these largely centred on a lack of contextual information, leading to infelicitous word choices when writing assignments. This was addressed by introducing participants to online corpus-based resources such as linguee.de which were evaluated by the participants during the course to be very helpful (see below).

2. *Through the implementation of a course that uses and adapts the tools and methods of corpus linguistics improve the writing of students in German at Rhodes University*

After the writing course, 80% of the participants perceived that their writing had improved and specifically attributed this to the writing course, and regular writing in German in the form of the writing paragraphs, in particular. Where a lack of confidence to write in German was raised as an issue in the 2014 questionnaire, 80% of participants felt more confident in their writing after the course, and displayed a change in approach to their writing tasks with a far lesser reliance on English than in 2014. Participants report a self-awareness of trying to think in German, and focus more directly on the German language.

3. *Assess the attitudes of German Studies learners' towards the use of corpus linguistic tools in developing academic literacy*

Participants reported that they found the knowledge of academic collocations to be helpful and the use of corpus tools such as linguee.de to be enlightening. In the post writing course questionnaire, participants evaluated looking at a German corpus and looking at a word/phrase in the context of a German/ English parallel corpus as the most useful tools for language learning (even more useful than the dictionary). Participants also particularly appreciated that corpus tools allowed them to create more diversity/variation in their writing. Given the expected CEFRL levels for our participants' level of study, these writing tasks pushed them to achieve above and beyond the normal expectations. As the students themselves explained in their interviews, while the writing tasks were challenging, they were also extremely helpful, as they forced the participants to think and write in the target language. Moreover, participants noted that being able to put this knowledge into practise in the writing exercises and being forced to write more, more often enabled them to change set patterns of thinking and apply knowledge in context.

This is perhaps the most important finding pertaining to this research thesis, as it illustrates that corpus tools, while a useful aid, are not by themselves what the students found to be most helpful. Corpus tools were an effective aid to teach new vocabulary in context, and combined with a course which enforced regular writing in which this new learning could be implemented, proved to be very effective. Perhaps either method in isolation would not have produced as good results as were found in this case study, and comparisons may be drawn to research undertaken by Yoon (2008) who found that a participant in their writing course whose coursework did not provide many opportunities for practising writing (and applying corpus-based knowledge learnt in the writing course) in fact lost confidence in her writing ability, and perceived a decline in writing skills. Writing more, more often, and receiving specific corpus-based instruction and tools to facilitate this process proved to be a successful combination for teaching everyday academic writing in German Studies at Rhodes University.

Furthermore, it is important to note that the more integrated into the course work the assignments were, the more participants engaged with the topics. During the pilot study, participation in the writing tasks was voluntary, and assignments were not included within the normal German Studies curriculum, nor did they tie in with any topics being currently taught. All third year students signed up to take part, as they perceived that writing more regularly in German would probably aid their language skills as they would receive extensive feedback on their writing pieces (i.e. they perceived a benefit). It was made clear that if participants signed up, they would have to complete all assignments. Participants agreed. However, when it came to handing in assignments, students often handed them in late, and on occasion not at all. Even though there was a perceived benefit, participants did not fully engage with the writing tasks. This could be attributed to the assignments not being awarded marks which could count towards course credits. It could also possibly be a product of the first topics having little to do with the course topics being currently taught at that time, and thus did not require deep engagement. In 2014, after the pilot study but before the writing course was instituted, the paragraph submissions were integrated into the modules being taught at the time. This immediately facilitated engagement with the topics at hand, and the writing of longer paragraphs. While this is no revolutionary finding, it is important in the context of designing further writing intensive courses. Participants appear to be far more engaged and willing to learn when writing within the subject for topics which are currently being taught, rather than writing outside of the curriculum.

4. *Evaluate the corpus-based course in terms of its usefulness and effectiveness by comparing the writing of participating students both before and during the writing course with reference to five key everyday academic collocations, as well as to mother-tongue writers of German*

Where questionnaires and interviews allowed a deep qualitative engagement with the participants experiences of writing in German, and learning to write using corpus tools, investigation into the use of specific academic words used in the RUDaF learner corpus enabled a quantitative evaluation of the impact of the corpus instruction on participant writing.

The corpus-based writing course can be said to be both very useful and effective, as in Chapter 5 it was shown that participants showed a significant increase in the use of 87.5% of the key academic words and collocations taught during the writing course, with the exception of the word *Meinung** whose usage was significantly overused before the writing course, and significantly decreased as a result of the writing course as participants were taught to use alternative opinion related TAG words, such as *Ansicht* and *Erachten**. However, where participants did not use a TAG word at all or very often before the writing course, they tended to overuse the TAG word taught during the writing course as compared to mother-tongue speakers. This reveals that learners in this context may rely on collocational patterns taught. This is not problematic in the South African German as a Foreign

language context, as it was shown that the taught collocational patterns may serve as stepping stones which learners must overuse in order to advance competence.

Challenges to the project

As discussed in Chapter 3, the experience of the teacher is important when instituting corpus-based teaching and learning methods and there are various challenges which need to be overcome in order to successfully institute corpus methods in the classroom. As this thesis seeks to provide a case study of the successful implementation of corpus methods in a foreign language classroom in the tertiary setting in South Africa, I will outline the main challenges to the project, and how these were overcome, below as readers who wish to institute similar methods within their own classrooms may find themselves faced with some of the same obstacles.

Sourcing relevant corpus materials

It was a particular challenge of this project to source pre-ordained corpus-based exercises for teaching German, a challenge faced by many researchers as outlined in Chapter 3 (Al Saeed & Waly, 2010; Oghigian & Chujo, 2010; Breyer, 2009). This is as "currently, direct corpus applications are neither a standard component of curricula nor textbooks" (Breyer, 2009, p 155). In South Africa, corpus-based exercises are certainly not the norm for teaching language. The discovery of the materials developed by the WHiG project was instrumental in designing and implementing the writing course in German Studies at Rhodes University. While not completely suited to the level of our learners, (as they were designed for advanced learners of German) they proved most helpful in teaching the five key everyday academic German words and collocations outlined by the WHiG researchers, as was shown in Chapter 5 and 6. Access to FALKO L1 was also very helpful in this regard, as participants were able to perform searches through this mother-tongue corpus and developed linguistic insights through this process. While there are many corpora which are advertised as being freely accessible online, there are few for German and this calls for further research to be done in this area, as well as for other languages such as the indigenous languages of South Africa. The introduction of the online corpus-based tool Linguee.de was most helpful to learners. Subsequently, we have also discovered a new online corpus-based tool dwds.de which has also been introduced to the learners, and we are hopeful that this will provide them with even more tools for independent language research.

Small number of participants

The German Studies section of the School of Languages at Rhodes University is not large. The findings of this research are based on a very small group of participants (21 in total). This is not by design, but by circumstance, as all students enrolled for German Studies from 2013 to 2015 who fit the design criteria were included in this research. Were more participants available, this type of study (which

aims to explore the usefulness of a corpus-based writing course), would benefit from an experimental design, where one group of participants is taught academic vocabulary using corpus-based methods, and another is taught the same vocabulary using, for example, CALL methods, as was done by Oghigian & Chujo (2010).

Assembling learner data for corpus analysis

"Is learner data easy to acquire? The easy answer is definitely 'no'". So states Granger (1998), and while there have been many advances in the past decade and a half which have made data collection a much easier task than it was then, it is still not an effortless process. In the past, researchers who wanted to create learner corpora were constrained by the fact that all assignments were handwritten, and so had to be typed out by the researcher (Granger, 1998). This was an extremely time-consuming process and left much room for error of representation as the researcher would have to correctly reproduce learner errors (Granger, 1998). This is no longer the case, as technology is advancing it is becoming rare to receive hand-written assignments, and typed texts are seen as the norm in many tertiary institutions, with Rhodes University providing no exception. However, cleaning and converting texts received as word documents from students is still a time consuming process. In the case of German, one has to search for and remove/replace all umlauts (ä, ü, ö, ß,) from each document, and then convert the word document to a text document and save this in UTF-8 encoding, as corpus concordancing software is idiosyncratic and often will not accept unfamiliar characters, converting them into strings of symbols which create chaos in data analysis. Our relatively small learner corpus is comprised of 151 text submissions. Performing these operations entails 9 to 10 hours of work, time which a teacher would normally not have. This calls for better reformed methods of data collection, such as perhaps tasking students to type their assignments in notepad, using no umlauts.

Anonymity

A further challenge experienced during data collection for this research was anonymity. The first questionnaires and interviews were not undertaken anonymously. Whilst participant themselves remain anonymous through being assigned numbers and not being mentioned by name, their identity at the time of the interviews and first questionnaires was known to the researcher. As I was not lecturing the students at the time of these first interviews and questionnaires in 2014, I felt that this would not cause an issue, as students would be able to respond honestly. However, the second questionnaire at the end of 2015 sought to evaluate the writing course which I had taught. I did not want the participants to feel unable to answer honestly for fear of retribution in any way, or for fear of hurting my feelings. I thus decided to make the online questionnaire anonymous. Creating an environment where participants felt able to give honest feedback on the course proved to provide

challenges for analysis, as comparisons could not be drawn to earlier data collected in 2014. The answers given in this questionnaire determined in part how effective and useful the course was evaluated to be (Chapter 6), and a change in personal opinion would have been an interesting facet of analysis, however I was unable to draw such comparisons.

Final conclusion

This thesis has provided insights into instituting a corpus-based writing course in German as a Foreign Language in German Studies at Rhodes University. It has been shown that corpus-based methods of instruction coupled with regular writing exercises, are highly effective in this setting, causing both an increase in learner confidence in writing and their perception of their writing skills, and well as providing good stepping stones for learners to become aware of the importance of common usage and collocation patterns, providing them with tools for further language learning. As a cause of the writing course, the participants became more aware of the conventions of academic writing and this awareness was transferred to their writing in English. This highlights the significance and benefits of becoming a multilingual writer, and of fostering multilingualism within the tertiary context.

I foresee that the methodology outlined for creating this writing course would be a particularly valuable tool for other language departments and sections in the South African Tertiary Setting. Whilst there are challenges to overcome as pre-ordained corpus-based language learning materials are not available for every language, these challenges may be overcome by a motivated teacher with the help of online tools.

REFERENCES

African Languages, 2015. *Rhodes University, School of Languages.* [Online]
Available at: http://www.ru.ac.za/africanlanguages/research/
[Accessed 26 July 2015].

Aijmer, K., 2002. Modality in advanced Swedish learners' written interlanguage. In: S. Granger, J. Hung & S. Petch-Tyson, eds. *Computer Learner Corpora, Second Langauage Acquisition and Foreign Language Teaching.* Amsterdam: John Benjamins, pp. 55-76.

Aijmer, K., 2010. *Corpora and Language Teaching.* Amsterdam: John Benjamins.

Al Saeed, N. & Waly, S., 2010. *Proceedings of the 10th METU ELT Convention.* Ankara, Turkey, http://dbe.metu.edu.tr/convention/proceedingsweb/Corpus.pdf.

Annas, R., 2002/2003. Zur Situation des Faches Deutsch an südafrikanischen Universitäten. *Acta Germanica,* Volume 30/31, pp. 181-192.

Anthony, L., 2014. *AntConc (Version 3.4.3) [Computer Software].* Tokyo, Japan: Waseda University. Available from http://www.laurenceanthony.net/.

Baker, A., 2007. Teaching German as a foreign language in a multilingual South African context. *Per Linguam,* 23(1), pp. 30-44.

Baker, P., 2006. *Using Corpora in Discourse Analysis.* London/New York: Continuum.

Baker, P., 2010. Corpus methods in linguistics. In: Litosseliti, ed. *Research Methods in Linguitics.* Oxford: Oxford University Press, pp. 94-111.

Bonelli, E. T., 2010. Theoretical overview of the evolution of corpus linguistics. In: A. O'Keefe & M. McCarthy, eds. *The Routledge Handbook of Corpus Linguistics.* London: Routledge, pp. 14-27.

Boughey, C., 2000. Multiple metaphors in an understanding of academic literacy. *Teachers and Teaching: Theory and Practice,* 6(3), pp. 279-290.

Boughey, C., 2005. Epistemological access to the university: an alternative perspective. *South African Journal of Higher Education,* 19(3), pp. 638-650.

Boughey, C., 2012. *Undergraduate Support and Success,* Grahamstown: Rhodes University.

Boulton, A., 2008. DDL: Reaching the parts that other teaching can't reach?. *Teaching and language corpora,* Volume 8, pp. 38-44.

Braun, S., 2008. From pedagogically relevant corpora to authentic language learning contents. *ReCALL,* 17(1), pp. 47-64.

Breyer, Y., 2009. Learning and teaching with corpora: reflections by student teachers. *Computer Assisted Language Learning,* 22(2), pp. 153-172.

Brown, J. D., 2009. Open-response items in questionnaires. In: J. Heigham & R. A. Croker, eds. *Qualitative Research in Applied Linguistics: A Practical Introduction.* New York: Palgrave Macmillan, pp. 200-219.

Bryman, A., 2012. *Social Research Methods.* 4th ed. Cape Town: Oxford University Press.

Campbell, C., 1998. *Teaching Second-Language Writing: Interacting with Text.* Canada: Heinle & Heinle.

Canagarajah, S., 2011. Writing to learn and learning to write by shuttling between languages. In: R. M. Manchón, ed. *Learning-to-Write and Writing to Learn in an Additional Language.* Amsterdam: John Benjamins, pp. 111-132.

Chambers, A., 2005. Integrating corpus consultation in language studies. *Language Learning & Technology,* 9(2), pp. 111-125.

Chambers, A., 2007. Popularising corpus consultation by language learners and teachers. In: E. Hidalgo, L. Quer & J. Santana, eds. *Corpora in the Foreign Language Classroom: Selected Papers from the Sixth International Conference on Teching and Language Corpora (TaLC 6).* Amsterdam: Rodopi, pp. 3-16.

Chauhan, J. K. S., Sharma, L. & Tomar, M., 2012. Dynamic shift from print to electronic resources. *International Journal of Information Technology and Business Management,* 1(1), pp. 9-15.

Corder, S. P., 1981. *Error Analysis and Interlanguage.* Oxford: Oxford University Press.

Council of Europe, 2012. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR).* [Online]
Available at: http://www.coe.int/t/dg4/linguistic/cadre1_en.asp
[Accessed 14 June 2013].

Coxhead, A., 2000. A new academic word list. *TESOL Quarterly,* 34(2), pp. 213-238.

Cunningham, A. E. & Stanovich, K. E., 1998. What reading does for the mind. *American Educator,* 1(1 & 2), pp. 8-12.

DAAD, 2014. *DAAD-Länderinformationen Südafrika.* [Online]
Available at: https://www.daad.de/laenderinformationen/suedafrika/de/
[Accessed 08 October 2014].

Dagneaux, E., Denness, S. & Granger, S., 1998. Computer-aided error analysis. *System,* Volume 26, pp. 163-174.

Department of Basic Education, 1997. *Language in Education Policy.* [Online]
Available at: http://www.gov.za/sites/www.gov.za/files/38679_gen325.pdf
[Accessed 14 April 2015].

Department of Education, 2015. *Staatskoerant.* [Online]
Available at: http://www.gov.za/sites/www.gov.za/files/38679_gen325.pdf
[Accessed 14 April 2015].

Dodd, B., 1997. Exploiting a Corpus of Written German for Advanced Language Learning. In: W. e. al., ed. *Teaching and Language Corpora.* London: Longman, pp. 131-145.

Doyle, L. et al., 2000. *Establishing Learning Objectives: Applications for Course Planning and Assessment,* Massachusetts: University of Massachusetts.

Durrant, P., 2009. Investigating the viability of a collocation list for students of English for academic purposes. *English for Specific Purposes,* 28(3), p. 157–169.

Ehlich, K., 1995. Die Lehre der deutschen Wissenschaftssprache: Sprachliche Strukturen, didaktische Desiderate. In: H. L. Kretzenbacher & H. Weinrich, eds. *Linguistik der Wissenschaftssprachen.* Berlin: de Gruyter, pp. 325-351.

Ellis, R., 1994. *The Study of Second Language Acquisition.* Oxford: Oxford University Press.

Ellis, R., 2003. *Task-based Language Learning and Teaching.* Oxford: Oxford University Press.

Estes, G. B., Lopez-Mayhew, B. & Gardner, M. T., 1998. Writing In the foreign langauges department. *Writing Across the Curriculum,* Volume 9, pp. 68-81.

Fairclough, N., 2001. *Language and Power.* 2nd Edition ed. Harlow: Pearson Education Limited.

Fandrych, C., Meißner, C. & Slavcheva, A., 2012. The GeWiss corpus: comparing spoken academic German, English and Polish. In: T. Schmidt & K. Wörner, eds. *Multilingual Corpora and Multilingual Corpus Analysis.* Amsterdam: John Benjamins, pp. 319-337.

Firth, J. R., 1968. A synopsis of linguistic theory, 1930-55. In: F. R. Palmer, ed. *Selected papers of J.R. Firth 1952–1959.* Harlow: Longman, pp. 168-205.

Fleisch, B., 2008. *Primary Education in Crisis: Why South African School Children Underachieve in Reading and Mathematics.* Cape Town: Juta & Co.

Flowerdew, L., 2008. *Corpus-Based Analyses of the Problem-Solution Pattern: A Phraseological Approach.* 2nd ed. Philadelphia: John Benjamins Publishing.

Flowerdew, L., 2009. Applying corpus linguistics to pedagogy. *International Journal of Corpus Linguistics,* 14(3), pp. 393-417.

Flowerdew, L., 2012. *Corpora and Language Education.* 1st ed. London: Palgrave Macmillan.

Gabrielatos, C., 2005. Corpora and language teaching: just a fling, or wedding bells?. *TESL-EJ,* 8(4), pp. 1-35.

Gardner, D., 2007. Validating the construct of 'word' in applied corpus-based vocabulary research: a critical study. *Applied Linguistics,* 28(2), pp. 241-265.

Gardner, D. & Davies, M., 2014. A new academic vocabulary list. *Applied Linguistics,* 35(3), pp. 305-327.

German Missions, 2015. *German South African Business Relations.* [Online]
Available at:

http://www.southafrica.diplo.de/Vertretung/suedafrika/en/06__Economy/1__Ger__SA/__Ger__SA.html
[Accessed 1 October 2015].

German Studies, 2015. *German Studies Handbook,* Grahamstown: Rhodes University.

Gethin, A. & Gunnermark, E. V., 1997. *The Art and Science of Learning Langauges.* Google eBook: Intellect Books.

Goethe Institut, 2015. *The Goethe-Zertifikat B2.* [Online]
Available at: http://www.goethe.de/lrn/prj/pba/bes/gb2/enindex.htm
[Accessed 1 October 2015].

Goossens, D. & Granger, S., 2013. *Learner Corpora Around the World.* [Online]
Available at: http://www.uclouvain.be/en-cecl-lcworld.html
[Accessed 01 11 2014].

Gordhan, P., 2013. *National Treasury, Budget Speech 2013.* [Online]
Available at: http://www.treasury.gov.za/documents/national%20budget/2013/speech/speech.pdf
[Accessed 13 May 2013].

Granger, S., 1998. The computer learner corpus: a versatile new source of data for SLA research. In: S. Granger, ed. *Learner English on Computer.* London: Addison Wesley, pp. 3-18.

Granger, S., 2002. A birds-eye view of learner corpus research. In: S. Granger, J. Hung & S. Petch-Tyson, eds. *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching.* Amsterdam: John Benjamins, pp. 3-33.

Granger, S., 2003. The international corpus of learner English: a new resource for foreign language learning and teaching and second language aquisition research. *TESOL Quarterly,* 37(3), pp. 538-546.

Granger, S., 2004. Computer learner corpus research: current status and future prospects. *Language and Computers,* 52(1), pp. 124-145.

Granger, S., Hung, J. & Petch-Tyson, S., 2002. *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching.* Amsterdam: John Benjamins.

Harden, T., 2000. The limits of understanding. In: T. Harden & A. Witte, eds. *The Notion of Intercultural Understanding in the Context of German as a Foreign Language.* Bern: Peter Lang, pp. 103-123.

Heigman, J. & Croker, R. A., 2009. *Qualitative Research in Applied Linguistics: A Practical Introduction.* New York: Palgrave Macmillan.

Hoey, M., 2005. *Lexical Priming: A New Theory of Words and Language.* London: Routledge.

Homstad, T. & Thorson, H., 1996. *Using Writing to-Learn Activities in the Foreign Language Classroom,* Minneapolis: University of Minnesota.

Howie, S. et al., 2011. *Progress in International Reading Literacy Study: South African Children's Reading Literacy Achievement,* Pretoria: University of Pretoria.

Howie, S. et al., 2006. *Progress in International Reading Literacy Study: South African Children's Reading Literacy Achievement,* Pretoria: University of Pretoria.

Ivankova, N. V. & Cresswell, J. W., 2009. Mixed methods. In: J. Heigman & R. A. Croker, eds. *Qualitative Research in Applied Linguistics: A Practical Introduction.* New York: Palgrave-MacMillan, pp. 135-161.

Jaworska, S., 2009. *The German Language in British Higher Education: Problems, challenges,teaching and learning perspectives.* Wiesbaden: Harrassowitz Verlag.

Jaworska, S., 2011. Der Wissenschaftlichkeit auf der Spur: zum Einsatz von Korpora in der Vermittlung des Deutschen als (fremder) Wissenschaftssprache. *Deutsch als Fremdsprache,* Volume 4, pp. 235-244.

Jaworska, S., Krummes, C. & Ensslin, A., 2015. Formulaic sequences in native and non-native argumentative writing in German: a corpus-driven comparison. *International Journal of Corpus Linguistics,* 20(4).

Johns, T., 1986. Micro-concord: a language learner's research tool. *System,* 14(2), pp. 151-162.

Johns, T., 1988. Whence and whither classroom concordancing?. In: S. Lobbe & H. Wekker, eds. *Computer Applications in Language Learning.* Dordrecht: Foris, pp. 9-27.

Johns, T., 1991. Should you be persuaded: two examples of data-driven learning. In: T. Johns & P. King, eds. *Classroom Concordancing.* Birmingham: University of Birmingham, pp. 1-16.

Johns, T. & King, P., 1991. *Classroom Concordancing.* Birmingham: University of Birmingham.

Jones, C. & Fortescue, S., 1987. *Using Computers in the Language Classroom.* 1st ed. London: Longman.

Krishnamurthy, R. & Kosem, I., 2007. Issues in creating a corpus for EAP pedagogy and research. *Journal of English for Academic Purposes,* Volume 6, pp. 356-373.

Krummes, C. & Ensslin, A., 2012. Formulaic language and collocations in German essays: from corpus-driven data to corpus-based learning materials. *Language Learning.*

Krummes, C. & Ensslin, A., 2014. What's Hard in German? (WHiG): a British learner corpus of German. *Corpora,* 9(2), pp. 191-205.

Laurillard, D., 1993. *Rethinking University Teaching: a framework for the effective use of educational technology.* London: Routledge.

Laurillard, D., 1997. Learning formal representations through multimedia. In: F. Marton, D. Hounsell & N. J. Entwhistle, eds. *The Experience of Learning: Implications for Teaching and Studying in Higher Education.* Scotland: Scottish Academic Press, pp. 172-183.

Leech, G., 1997. Teaching and language corpora: a convergence. In: A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles, eds. *Teaching and Language Corpora.* London: Addison Wesley Longman, pp. 1-23.

Lillis, T. M., 2001. *Student Writing: Access, Regulation, Desire.* 2nd ed. New York: Routledge.

López, W. C., 2009. Error analysis in a learner corpus. What are the learners strategies?. *Center for English Corpus Linguistics, Catholic University of Louvain,* pp. 675-690.

Lotz-Sisitka, H., 2009. Epistemological access as an open question in education. *Journal of Education,* Volume 46, pp. 57-80.

Lüdeling, A. et al., 2008. Das Lernerkorpus Falko. *Deutsch als Fremdsprache,* Volume 2, pp. 67-73.

Masgoret, A. M. & Gardner, R. C., 2003. Attitudes, motivation and second language learning: a meta-analysis of studies conducted by Gardner and associates. *Language Learning,* 53(1), pp. 123-163.

McCarthy, 1998. *Spoken Language and Applied Linguistics.* Cambridge: Cambridge University Press.

McCarthy, M. & O'Keeffe, A., 2010. *The Routledge Handbook of Corpus Linguistics.* 1st ed. London: Routledge.

McEnery, T. & Hardie, A., 2012 (b). *Statistics in Corpus Linguistics.* [Online]
Available at: http://corpora.lancs.ac.uk/clmtp/2-stat.php
[Accessed 15 03 2015].

McEnery, T. & Hardie, A., 2012(a). *Corpus Linguistics: Method, Theory and Practice.* Cambridge: Cambridge University Press.

McEnery, T. & Xiao, R., 2010. What corpora can offer in language teaching and learning. In: E. Hinkel, ed. *Handbook of Research in Second Language Teaching and Learning (Vol. 2).* London & New York: Routledge, pp. 364-380.

Meunier, F., 2002. The pedegogical value of native and learner corpora. In: S. Granger, J. Hung & S. Petch-Tyson, eds. *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching.* Amsterdam: John Benjamins, pp. 119-142.

Mindt, D., 1997. Corpora and the teaching of English in Germany. In: A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles, eds. *Teaching and Language Corpora.* London: Addison Wesley Longman, pp. 40-50.

Morrow, W., 1993. Epistemological access in the university. *AD Issues,* 1(1), pp. 3-5.

Mukherjee, J. & Rohrbach, J. M., 2006. Rethinking applied corpus linguistics from a language-pedagogical perspective: New departures in learner corpus research. In: B. Kettemann & G. Marko, eds. *Planning, Painting and Gluing Corpora. Inside the Applied Corpus Linguist's Workshop.* Frankfurt am Main: Peter Lang, pp. 205-232.

Münkler, H. & Winkler, H. A., 2014. *Facts about Germany.* [Online]
Available at: http://www.tatsachen-ueber-deutschland.de/en/content-home/facts-and-figures/economy.html
[Accessed 08 October 2014].

National Treasury, 2014. *Budget Review,* South Africa: National Treasury.

O'Donnell, M. B., Römer, U. & Ellis, N. C., 2013. The development of formulaic sequences in first and second langauge writing: Investigating effects of frequency, association and native norm. *International Journal of Corpus Linguistics,* 18(1), pp. 83-108.

Oghigian, K. & Chujo, K., 2010. An effective way to use corpus exercises to learn grammar basics in English. *Language Education in Asia,* 1(1), pp. 200-214.

O'Keeffe, A. & McCarthy, M., 2010. *The Routledge Handbook of Corpus Linguistics.* 1st ed. London: Routledge.

Oliver, J., 2009. *SAlanguages.com.* [Online]
Available at: http://salanguages.com/education.htm
[Accessed 08 October 2014].

Ortner, G. J., 2012. *A report on acquisition and development of academic literacy: a case study of first year linguistics students at Rhodes University,* English Language and Applied Linguistics: Rhodes University.

Ortner, G. J., 2013. *A pilot study for the proposed Masters Dissertation, 2014. Examining the formal writing skills of German Studies 3 students, using corpus linguistics, to uncover common lexical and morphosyntactic errors in student writing,* German Studies: Rhodes University.

Ragan, P. H., 1996. Classroom use of a systemic functional small learner corpus. In: M. Ghadessy, A. Henry & R. L. Roseberry, eds. *Small corpus studies and ELT.* Amsterdam: John Benjamins, pp. 207-236.

Reinbothe, R., 2006. *Deutsch als internationale Wissenschaftssprache und der Boykott nach dem Ersten Weltkrieg.* Frankfurt am Main: Peter Lang.

Rhodes University, 2013. *Undergraduate Courses, German.* [Online]
Available at: http://www.ru.ac.za/german/undergraduatecourses/
[Accessed 1 November 2015].

Richards, K., 2009. Interviews. In: J. Heigman & R. A. Croker, eds. *Qualitative Research in Applied Linguistics.* New York: Palgrave Macmillan, pp. 182-199.

S., D., 2014. *Avaaz.org.* [Online]
Available at:
https://secure.avaaz.org/en/petition/University_of_KwaZuluNatals_Executive_management_structures_Allow_the_disciplines_Afrikaans_French_German_and_Italian_to/?pv=10
[Accessed 30 October 2014].

Schultis, N., 2011. *Redemittel für wissenschaftliche Texte,* Freiburg: PH Freiburg.

Sinclair, J., 1991. *Corpus, Concordance, Collocation.* Oxford: Oxford University Press.

Sinclair, J., 1997. Corpus evidence in language description. In: A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles, eds. *Teaching and Language Corpora.* London: Addison Wesley Longman, pp. 27-39.

Sinclair, J. M., 1987. Collocation: A progress report. In: R. Steele & T. Threadgold, eds. *Language topics: Essays in honour of Michael Halliday.* Amsterdam: John Benjamins, pp. 319-331.

Skrandies, P., 2011. Everyday academic language in German historiography. *German as a Foreign Language,* Volume 1, pp. 99-123.

Steinhof, T., 2007. *Wissenschaftliche Textkompetenz. Sprachgebrauch und Schreibentwicklung in wissenschaftlichen Texten von Studenten und Experten.* Tübingen: Niemeyer.

Stevens, V., 1995. Concordancing with language learners: Why? When? What?. *CAELL Journal,* 6(2), pp. 2-10.

Swales, J. M. & Feak, C. B., 1994. *Academic Writing for Graduate Students: A Course for Non-native Speakers of English.* 4 ed. United States of America: University of Michigan Press.

Terre Blanche, M., Durrheim, K. & Painter, D., 2009. *Research in Practice: Applied methods for the Social Science.* 2nd ed. Cape Town: University of Cape Town Press.

Thomson, C., 2013. *Writing Intensive Project: End of Year Report for the Rhodes University Teaching and Learning Committee,* Grahamstown: Rhodes University.

Torgesen, J., Houston, D., Rissman, L. & Decker, S., 2007. *Academic literacy instruction for adolescents,* Florida State University: Centre on Instruction: Florida Centre for Reading Research.

Tribble, C., 1997. Improvising corpora for ELT: Quick and dirty ways of developing corpora for language teaching. In: B. Lewandowska-Tomasczczyk & J. Melia, eds. *Proceeding of the First International Conference on Practical Applications in Language Corpora.* Lotz, Poland: Lotz University Press, pp. 106-117.

Tribble, C., 2002. Corpora and corpus analysis: New windows on academic writing. In: J. Flowerdew, ed. *Academic Discourse.* Harlow, England: Longman, pp. 131-149.

Trupe, A. L., 2011. *Academic Literacy in a Wired World: Redefining Genres for College Writing Courses.* [Online]
Available at: http://people.bridgewater.edu/~atrupe/AcadLit/WiredWorld.htm
[Accessed 15 05 2013].

Van der Walt, C. & Fourie, H., 2005. Developing a learner's corpus: The case of a first-year module in mathematics. *Lexicos,* Volume 15, pp. 242-252.

Van Rooy, B., 2005. The localisation of English resources: spelling checking for South African students. In: W. Daelemans, T. Du Plessis, C. Snyman & L. Teck, eds. *Multilingualism and Electronic Language Management.* Pretoria: Van Schaik.

Van Rooy, B., 2008. A multidimensional analysis of student writing in Black South African English. *English World-Wide,* 29(3), pp. 268-305.

Van Rooy, B. & Schäfer, L., 2002. The effect of learner errors on POS tag errors during Automatic POS tagging. *South African Linguistics and Applied Language Studies,* Volume 20, pp. 325-335.

Van Schalkwyk, S. C., Bitzer, E. & Van der Walt, C., 2009. Acquiring academic literacy: a case of first-year extended degree programme students at Stellenbosch University. *South African Linguistics and Applied Language Studies,* 27(2), pp. 189-201.

Vann, R. J., Meyer, D. E. & Lorenz, F., 1984. Error gravity: a study of faculty opinion of ESL errors. *Teachers of English to Speakers of Other Languages (TESOL),* 18(3), pp. 427-440.

Varley, S., 2009. I'll just look that up in the concordancer: integrating corpus consultation into the language learning environment. *Computer Assisted Language Learning,* 22(2), pp. 133-152.

Vasagar, J., 2013. *The Telegraph: German 'should be a working language of EU', says Merkel's party.* [Online]
Available at: http://www.telegraph.co.uk/news/worldnews/europe/germany/10128380/German-should-be-a-working-language-of-EU-says-Merkels-party.html
[Accessed 08 October 2014].

Weber, U. S., 2015. Can studying a foreign language build or improve (inter-) cultural competence? A preliminary case study of students' subjective impressions. In: A. Witte & T. Harden, eds. *Foreign Language Learning as Intercultural Experience.* Oxford: Peter Lang, pp. 231-242.

Weber, U. S. & Domingo, R. S. C., 2011. Adding another colour to the rainbow: an attempt at imparting German cultural competence in a South African University context. In: A. Witte & T. Harden, eds. *Intercultural Competence : Concepts, Challenges, Evaluations.* Bern: Peter Lang, pp. 177-191.

Weber, U. S., Domingo, R. S. C. & Fourie, R. B., 2015. Beyond language: German Studies in a South African University context. In: R. H. Kashula, P. Maseko & H. E. Wolff, eds. *Multilingualism in the Professions and in Intercultutral Communication: A South African Perspective.* Pretoria: Wits University Press, p. In review.

Weideman, A., 2003. Assesing and developing academic literacy. *Per Linguam,* 19(1&2), pp. 55-65.

Wichmann, A., 1997. General Introduction. In: A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles, eds. *Teaching and Language Corpora.* 1st ed. London: Addison Wesley Longman, pp. xvi-xvii.

Wichmann, A., Fligelstone, S., McEnery, T. & Knowles, G., 1997. *Teaching and Language Corpora.* 1st ed. London: Addison Wesley Longman.

Witte, A., 2002/2003. Germanistik und DaF in Afrika (Subsahara) - Geschichte, Bestandsaufnahme, Aussichten. *Acta Germanica,* Volume 30/31, pp. 169-180.

Wolf, M., 2008. *Proust and the Squid: The Story and Science of the Reading Brain.* UK: Icon Books.

Yoon, H., 2008. More than a linguistic reference: the influence of corpus technology on L2 academic writing. *Language Learning & Technology,* 12(2), pp. 31-48.

Zeldes, A., Lüdeling, A. & Hirschmann, H., 2003. *What's hard? Quantitative evidence for difficult constructions in German learner data.* Helsinki, Humbolt-Universität zu Berlin.

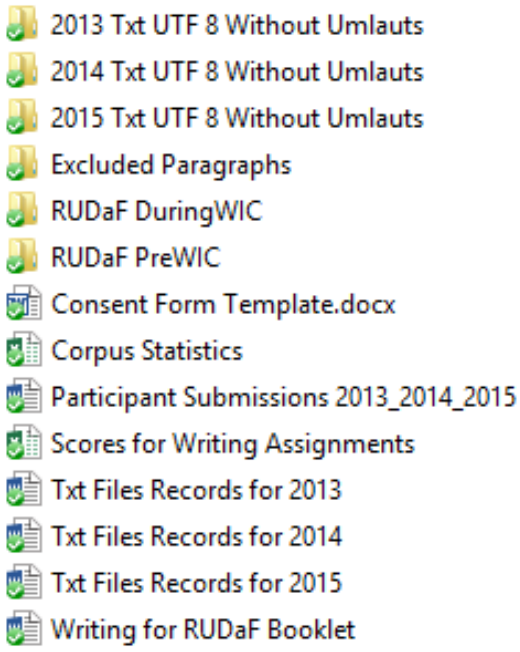
Zipp, G. L., 2013. *A history of the German settlers in the Eastern Cape, 1857-1919, Master's research thesis,* German Studies: Rhodes University.

APPENDICES

Please see the disk in the folder at the back of this thesis for a copy of all appendices.

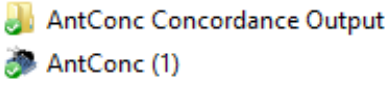
The appendices are organised in the following way:

Appendix 1_RUDaF Learner Corpus and Metadata

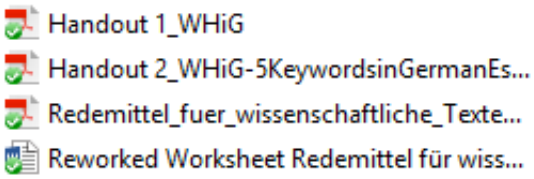


Appendix 2_FALKO L1 Mother Tongue Control Corpus

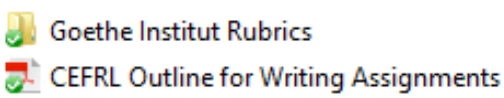
Appendix 3_AntConc Programme and Concordance Output



Appendix 4_Course Handouts



Appendix 5_Rubrics for Assessment



Appendix 6_Questionnaires and Interviews