Tampere University

PAUL RICKMAN

Aspects of Non-finite Verb Complementation in New Zealand English

Tampere University Dissertations 1005

Tampere University Dissertations 1005

PAUL RICKMAN

Aspects of Non-finite Verb Complementation in New Zealand English

ACADEMIC DISSERTATION
To be presented, with the permission of
the Faculty of Information Technology and Communication Sciences
of Tampere University,
for public discussion in Lecture room D11
of the Main Building, Kalevantie 4, Tampere,
on 17 May 2024, at 12 o'clock.

ACADEMIC DISSERTATION
Tampere University, Faculty of Information Technology and Communication Sciences
Finland

Responsible supervisor and Custos	Professor emeritus Juhani Klemola Tampere University Finland	
Supervisor	Professor emeritus Juhani Rudanko Tampere University Finland	
Pre-examiners	Professor Peter Collins University of New South Wales Australia	Docent Turo Vartiainen University of Helsinki Finland
Opponent	Professor Erik Smitterberg Uppsala University Sweden	

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

Copyright ©2024 author

Cover design: Roihu Inc.

ISBN 978-952-03-3402-4 (print)
ISBN 978-952-03-3403-1 (pdf)
ISSN 2489-9860 (print)
ISSN 2490-0028 (pdf)
http://urn.fi/URN:ISBN:978-952-03-3403-1

Carbon dioxide emissions from printing Tampere University dissertations have been compensated.

PunaMusta Oy – Yliopistopaino
Joensuu 2024

ACKNOWLEDGEMENTS

This dissertation has been a long time in the making, and due to the prolonged nature of the project I think I owe a special debt of gratitude to my primary supervisor, Professor Juhani Klemola, for his pragmatic and down-to-earth approach to getting things done, for his patience, and for his willingness to see things through to the end with me, even into retirement. When Juhani and I began discussing my doctoral project in 2012 he may not have suspected that, 11 years later, I would still be at it.

My secondary supervisor, Professor Juhani Rudanko, also retired while I was in the midst of this project, and to him too I extend my warmest thanks for his continued support and supervision. Juhani is also the second author of two of the publications that make up this dissertation, and I thank him for his role in those projects too. It was under Juhani's guidance as an MA student that I discovered my interest in the study of complementation, and since then, his seemingly never-ending infusion of energy into both my own work and our shared projects has been greatly appreciated.

I am also indebted to the second author of another of the publications included here, Dr. Mark Kaunisto. I am fortunate to have collaborated with two excellent scholars on three of the publications for this project, and while all of the experiences have been pleasant and valuable, rarely are such projects brought to completion as smoothly and as amicably as this one was with Mark. Many thanks too to Mark for helping out with the translation of the abstract.

I also thank Dr. Juhani Norri, of the Tampere University English department, for taking on some of my teaching duties in the spring of 2023 and giving me the time to get the final sections of the Introduction written.

I had the benefit of good advice, encouragement and financial assistance in the early stages of this project as a member of the GlobE and ChangE research groups. These projects were led by Professors Juhani Klemola, Markku Filppula and Anna Mauranen, and additionally comprised of top scholars from the universities of Helsinki, Eastern Finland and Tampere. Ray will especially be remembered as a good man with some good ideas.

Professor Johannes Riquet and the participants of the Tampere University faculty of languages doctoral seminar that I attended also provided me with plenty of

information and general good cheer. I am sure that helpful and supportive folk can usually be found in these seminars, but I am happy that I attended alongside this particular group.

The Corpus of New Zealand Newspaper English, which provided data for two of the publications in this dissertation as well as several side projects, and will provide data for future projects, would never have got off the ground had it not been for the help I received from Professor Sebastian Hoffmann during his tenure as visiting professor in 2012–2014. The time and effort he put into patiently devising ways to download and process the Fairfax data is greatly appreciated.

I also thank the two pre-examiners of this work, Professor Peter Collins, of the University of New South Wales, and Docent Turo Vartiainen, of the University of Helsinki, for their careful work and constructive comments. All comments and suggested revisions were accepted in the spirit in which they were given, and carefully considered, and while most did make their way into the manuscript in one form or another, a few have been left as food for thought, possibly to be dealt with in later work.

Finally, I reserve the last paragraph for my family: my wife Annina (who also lent her expertise to the abstract translation – thanks!) and our three children, for making home such a great place to be. They may not care much for complementation, but they were all as relieved as I was to see this project reach its end. Mum and Dad too, obviously, although they didn't really hear about it until it was all over – that's what can happen when you live on opposite sides of the world.

Tampere, March 2024

Paul Rickman

ABSTRACT

The present dissertation is an investigation into one of the lesser-studied areas of the English language, that of non-finite verb complementation, as used in one of the lesser-studied native varieties of English, New Zealand English. New Zealand English is the youngest of the world's main native English dialects, having developed slightly later than its southern hemisphere neighbours, Australian and South African Englishes. Like its neighbours, it is the output of the mixture of British Isles dialects, which, in the case of New Zealand, accompanied the first European settlers to the country during the period of heavy European settlement in the latter half of the nineteenth century. This first main settlement period followed the 1840 signing of the Treaty of Waitangi, an agreement which established British sovereignty over New Zealand.

British influence in New Zealand – and other British colonial territories – remained strong throughout the pre-World War II period, whereafter the United States began to exert a cultural influence, both in New Zealand and in other parts of the world. In addition to the external influence dimension, it has been suggested that, as the natural developmental cycle of a postcolonial variety progresses, the people of the young nations begin to reconstruct their collective identity. Products of this identity reconstruction often take the form of innovations within the linguistic system, making analyses of New Zealand English and other such varieties particularly relevant at the present time.

The focus of the present dissertation is on the non-finite complementation patterns of verbs in a selection of areas. The specific areas covered are verbs of prevention and the shifting preferences of two main forms of gerundial complement; matrix predicates selecting the transitive *into -ing* complement pattern and the nature of innovation in local usage; the predicate pair *promise* and *threaten*, their dual roles as subject control and subject-to-subject raising predicates and comparison of the division of labour between the two roles; and the two main non-finite complements of the matrix predicate *fear*, analysed from the perspective of the recently-postulated Choice Principle. These four areas of study are dealt with in four peer-reviewed articles. Three of the articles incorporate a cross-varietal aspect, with comparison to

British and American Englishes, while the focus of the fourth is on New Zealand English only.

The empirical data for all four articles are taken from the genre of newspaper language. The greater part of the data is provided by a purpose-designed diachronic corpus of New Zealand English, which covers the 15-year period between the mid-1990s and the early 2010s, and the News on the Web corpus, which covers the period from 2010 to the present day. Supplementary data for some of the studies are taken from the British National Corpus and the Corpus of Contemporary American English.

The motivation for the study is the fact that, despite a rise in interest in the grammar of New Zealand English following the post-1980s growth in the field of World Englishes, this particular aspect of the grammar is still an under-researched area. With the distinctive phonological and lexical characteristics of New Zealand English having traditionally been the foci of much of the available research attention, the grammar was not considered a viable research aim until relatively recently. The work that has been done in this area since the 1980s, however, has provided important results that highlight the need for further research.

The study sets out with the hypotheses that there are several significant differences to be found between New Zealand English, British English, and American English, that New Zealand English is beginning to show increasingly more independence in the area of non-finite complementation, and that the chosen areas of study can provide results that contribute to the ongoing discussion of 'colonial lag'. The combined results support these hypotheses, and allow new insight into current effects of the historically strong bilateral influence of the world's two older varieties alongside the possibility of innovation in the younger variety. The results also provide new theoretical and practical data on a semantic aspect of complement selection in New Zealand English, which can be expanded upon in future research.

TIIVISTELMÄ

Tässä väitöskirjassa tutkitaan verbien ei-finiittisiin komplementaatiorakennelmiin uudenseelanninenglannissa. Uudenseelanninenglanti on yksi nuorimmista englannin kielen murteista, ja se on kehittynyt hieman myöhemmin kuin sen läntinen naapuri, australianenglanti. Naapurinsa tavoin se on sekoitus Iso-Britannian murteita. Murteet tulivat Uuteen-Seelantiin uudisasukkaiden mukana 1800-luvun jälkipuoliskolla tapahtuneen voimakkaan muuttoaallon seurauksena, Waitangin sopimuksen allekirjoittamisen jälkeen. Iso-Britannian kulttuurinen vaikutus Uudessa-Seelannissa ja muissa siirtomaissa säilyi vahvana toiseen maailmansotaan asti, minkä jälkeen Yhdysvaltojen vahvistuminen alkoi vaikuttaa kulttuurisesti sekä Uudessa-Seelannissa että muissa osissa maailmaa. Ulkoisen vaikutuksen lisäksi on esitetty, että kun jälkikolonialistisen murteen luonnollinen kehityssykli etenee, nuorten kansakuntien asukkaat alkavat luoda omaa identiteettiään. Merkkejä tästä voidaan usein havaita kielellisesti, minkä vuoksi uudenseelanninenglannin kaltaisten murteiden analyysit ovat nykyään erityisen tärkeitä.

Tutkimuksen erityiskohteina ovat merkitykseltään estämiseen liittyvät verbit ja niiden gerundiaalisen komplementin kaksi päämuotoa; predikaatit, jotka valitsevat transitiivisen *into -ing* -komplementin; predikaattiverbit *promise* ja *threaten* ja niiden kaksoisroolit *subjektikontrolli-* ja *subjekti-subjekti raising*-predikaatteina; ja predikaattiverbin *fear* kaksi tärkeintä ei-finiittistä komplementtia, joita analysoidaan Choice-periaatteen näkökulmasta. Näitä neljää tutkimusaluetta käsitellään neljässä vertaisarvioidussa artikkelissa.

Kaikkien neljän artikkelin empiirinen aineisto on peräisin sanomalehtikielestä. Suurin osa aineistosta on saatu tätä työtä varten suunnitellusta uudenseelanninenglannin diakronisesta korpuksesta, joka kattaa 15 vuoden ajanjakson 1990-luvun puolivälistä 2010-luvun alkuun, sekä News on the Web -korpuksesta, joka kattaa ajanjakson vuodesta 2010 tähän päivään. Joidenkin tutkimusten kohdalla on lisäksi tarkasteltu aineistoja British National Corpus -korpuksesta ja Corpus of Contemporary American English -korpuksesta.

Tutkimuksen taustalla on se, että vaikka kiinnostus uudenseelanninenglannin kielioppia kohtaan on lisääntynyt 1980-luvun jälkeen, tämä kieliopin osa-alue on edelleen vähän tutkittu. Uudenseelanninenglannin fonologiset ja leksikaaliset

erityispiirteet ovat perinteisesti olleet tutkituimpia kohteita, ja kielioppia on pidetty mahdollisena tutkimuskohteena vasta suhteellisen hiljattain. Tällä alalla 1980-luvulta lähtien tehty työ on kuitenkin tuottanut tärkeitä tuloksia, jotka korostavat lisätutkimuksen tarvetta.

Tutkimuksen lähtökohtana olivat hypoteesit, joiden mukaan 1) uudenseelannin-, britti- ja amerikanenglannin välillä on useita merkittäviä eroja, 2) uudenseelanninenglannissa on yhä enemmän omia, itsenäisiä piirteitä ei-finiittisten komplementaatiorakenteiden käytössä ja 3) tämän piirteen tutkimuksen tulokset edistävät meneillään olevaa keskustelua aiheesta ”colonial lag” (ns. ”siirtomaa-ajan viive”). Kokonaisuudessaan tulokset tukevat näitä hypoteeseja ja antavat näin ollen uutta tietoa brittienglannin, amerikanenglannin ja uudenseelanninenglannin suhteista. Lisäksi ne tarjoavat uutta teoreettista ja käytännöllistä tietoa uudenseelanninenglannin komplementtirakenteiden semanttisista piirteistä, joiden tarkastelua voidaan laajentaa tulevissa tutkimuksissa.

CONTENTS

1 Introduction 15
1.1 Aims of the thesis 17
1.2 Structure of the dissertation 19

2 Theory and Background 20
2.1 Complementation 20
2.1.1 Introduction 20
2.1.2 The history of non-finite complements 22
2.1.3 Control and understood subjects 23
2.1.4 Subject control and object control 27
2.1.5 Subject control and Subject-to-Subject Raising 28
2.2 World Englishes 32
2.2.1 Models of World Englishes 32
2.2.2 Dynamics and influence among World Englishes 36
2.2.3 The 'American Invasion' of New Zealand 37
2.2.4 Influence within New Zealand: Māori 39
2.2.5 Colonial lag 41
2.3 A brief historical background sketch 43
2.3.1 History of New Zealand 43
2.3.2 History of New Zealand English 48
2.4 Previous research on NZE 51

3 Data and methodology 56
3.1 Corpus linguistics and newspaper corpora 56
3.2 Two methodological approaches 58
3.3 The Corpus of New Zealand Newspaper English (CNZNE) 58
3.4 The News on the Web Corpus (NOW) 59
3.5 Other corpora 60

4 Results and discussion 61
4.1 Research question 1 61
4.1.1 Verbs of prevention 62
4.1.2 *Promise* and *threaten* 66
4.1.3 Summary of research question 1 67
4.2 Research question 2 68
4.2.1 The transitive *into -ing* pattern 68

4.2.2 *Promise* and *threaten* 71
4.2.3 Summary of research question 2 72
4.3 Research question 3 72
4.4 Research question 4 74

5 Conclusion 77
5.1 Strengths and weaknesses 77
5.2 Future directions 78
5.3 Concluding remarks 79

REFERENCES 81

List of Figures

Figure 1. Kachru's Three Circles Model of Englishes (Kachru 1990: 179) 33
Figure 2. The World System of Standard Englishes (Mair 2013: 261) 35
Figure 3. Map of Polynesia (Walter et al. 2017: 352) 44
Figure 4. Map of New Zealand (https://maps-new-zealand-nz.com) 46

List of Tables

Table 1. NZ population born outside NZ 47
Table 2. Prevention verbs in NZE, BrE and AmE 63
Table 3. *Prevent* and *stop* in NZE (1995–98 and 2010–12) 65

ORIGINAL PUBLICATIONS

Publication I Rickman, Paul. 2017. Aspects of verb complementation in New Zealand newspaper English. In Markku Filppula, Juhani Klemola, Anna Mauranen and Svetlana Vetchinnikova (eds.), *Changing English: Global and Local Perspectives*, 169–190. Berlin, Boston: Mouton de Gruyter.

Publication II Rickman, Paul and Mark Kaunisto. 2018. Aspects of the use of the transitive *into -ing* pattern in New Zealand English. In Mark Kaunisto, Mikko Höglund and Paul Rickman (eds.), *Changing Structures: Studies in Constructions and Complementation*, 71–88. Amsterdam: John Benjamins.

Publication III Rickman, Paul and Juhani Rudanko. 2023. Facing two ways syntactically: on the grammar and use of *promise* and *threaten* in three regional varieties. *English Studies* 104(2), 365–382.

Publication IV Rickman, Paul and Juhani Rudanko. 2024. The Choice Principle and the matrix verb *fear* in recent New Zealand English. *Studia Neophilologica* (Published online: 7 Feb 2024).

AUTHORS' CONTRIBUTIONS

Publication II Both authors worked on the planning of the paper. Rickman wrote sections 1, 2 and 3, and was responsible for the retrieval of the data from the CNZNE. Kaunisto wrote section 4, compiled all results tables, and, with input from Rickman, selected the example tokens to appear in the paper. Both authors wrote section 5, devised the search strings for the corpus data retrieval, and discussed all results. Both authors provided comments on each other's work during the writing process.

Publication III Both authors worked on the planning of the paper. Rudanko wrote section 1 and parts of section 2. Rickman wrote the remainder of section 2, was the primary writer of section 3 and the Conclusion. Rickman was responsible for the statistical analyses and the retrieval of the corpus data. Both authors worked on devising the search strings and analysing the corpus data. Both authors provided comments on each other's work during the writing process.

Publication IV Both authors worked on the planning of the paper. Rudanko wrote Section 1. Both authors wrote the remaining parts of the paper together, with Rickman primarily responsible for the descriptions of the data and methodology (section 2), the statistical analyses, and the retrieval of the corpus data. Both authors worked on devising the search strings and analysing the corpus data. Both authors provided comments on each other's work during the writing process.

1 INTRODUCTION

I am lucky enough to have spent the first 25 years of my life in New Zealand. Some time after that period, I began my studies at Tampere University and took an undergraduate course in introductory linguistics, which appears to have led to this dissertation. In the syntax module of that course I came across a detail that made me realise that my rules of English grammar, which I had never paid much attention to, were somewhat different from others'. The detail that caught my attention, concerning the verb *promise* and the type of complement that a speaker is allowed to use after it, later led me to my MA thesis topic, and from there, on to the present doctoral project.

I was already aware that some other aspects of my variety of English were different from other people's; it soon becomes apparent when people don't understand your pronunciation or your vocabulary choices. Differences in the way sentences are structured, on the other hand, tend to be more subtle, and, unless things are really out of place your conversation partners understand you and any discrepancies between the different sets of grammatical rules often go unmentioned or unnoticed. (This is not just in the case of my own personal experience – as will be discussed below, it was not until the late twentieth century that linguists began to wonder whether there were any differences between the grammars of the native English speakers in New Zealand and Britain.) While all aspects of New Zealand English (NZE) nowadays get some research attention, it seems fair to say that the grammar, in particular the area of predicate complementation, is an area that is still under-researched. It is the aim of this dissertation, therefore, to shed a little more light on this area and hopefully to encourage further research into NZE grammar.

In terms of human habitation, New Zealand is a young country, and large-scale settlement by English speakers took place only relatively recently (Belich 1996; King 2003). It is possible that the first major wave of this anglophone immigration, which began in the 1840s, was still within living memory for the early settlers when visitors to the country first began to notice the different pronunciation features of the colony, and around the same time observers noted the vocabulary, with its borrowings from the indigenous language Māori. Grammar, on the other hand, did

not receive a great deal of attention until much later. Being a nation comprised mainly of immigrants from Britain, it was easy to assume that the grammar was the same as that of the standard model used in Britain, since, as mentioned above, grammatical variation tends to be rather subtle. But with seminal publications from Trudgill and Hannah (1982), and Laurie Bauer (1987)[1], attention began to turn to this area. The age of electronic corpora, which began in the early 1960s with the release of the Brown corpus and the Lancaster-Oslo-Bergen corpus, made it possible to make objective claims about grammatical usage preferences in the different varieties of English, and the study of the grammatical differences between varieties of English shifted gear. NZE, being one of the smaller varieties with fewer than four million native speakers at the time (www.stats.govt.nz), was among the last of the native English varieties to acquire its own corpora – the written and spoken sections of the Wellington corpus, and the NZ section of the International Corpora of English (ICE-NZ). Both of these appeared in the 1990s and are based on data collected in the late 1980s to early 1990s (https://www.wgtn.ac.nz/lals/resources/corpora-default; https://www.ice-corpora.uzh.ch/en.html). A great deal has changed since then, and with the rapid advances in computer technology and data availability, even the modestly tech-savvy researcher can compile massive corpora for specific purposes. The present study primarily makes use of two corpora: one of which was compiled by the (modestly tech-savvy) author (with a little help) for the purposes of this dissertation, and the other is one member of the freely available set of corpora provided by corpora.org. Further discussion of corpora-related matters will be taken up below.

This dissertation reports, analyses, and compares the findings of four peer-reviewed publications, all of which were either published or accepted for publication between 2018 and 2023. Publications I and II appeared in peer-reviewed volumes dealing with the wider themes of variation and change in the English language, and Publications III and IV appeared in, or were accepted by, international peer-reviewed journals with a focus on the English language. The four papers are connected by the common topics of New Zealand English and verbal predicate complementation.

Publication I gives a detailed overview of the compilation of the Corpus of New Zealand Newspaper English (henceforth CNZNE) and contains a pilot study of 14 verbs of prevention in NZE. This publication includes a comparative dimension, with mid-1990s data from British English (BrE) and American English (AmE), as well as a diachronic aspect of the NZE data and coverage of the period 1995 to 2012.

[1] The title of one of Bauer's early (1987) papers, which was one of the earliest to address New Zealand English grammar issues, is 'Approaching the grammar of New Zealand English'.

Publication II adopts a pattern-based approach (see section 3.2 on head-based and pattern-based approaches), taking the productive *into -ing* pattern and mapping out the predicates that are available to select it in NZE. Like Publication I, this publication uses data from the period 1995 to 2012. It is shown that NZE, while for the most part conforming to the combinational patterns of other native varieties, also shows its own innovative tendencies.

Publication III focuses on *promise* and its semantic counterpart *threaten*. It examines the issue of subject control versus subject-to-subject raising uses with these two predicates, each of which can lend itself to use in both types of structure. NZE is contrasted with British and American Englishes, with synchronic data from the News on the Web (NOW) corpus.

Publication IV examines the verbal predicate *fear* and its selectional properties concerning *to* infinitive and *ing* clauses, both of which are available to it. The recently postulated Choice Principle provides a theoretical framework for the discussion of the results, and the synchronic data is taken from the NOW corpus.

1.1 Aims of the thesis

The overall aim of this work is to examine aspects of non-finite verb complementation patterns in NZE, and in doing so, make a contribution to the existing body of work on the grammar of NZE, as well as the existing literature on variation among World Englishes. The reader will note that the four publications comprise work which is either focused on NZE, or which extends to other varieties, but that the common themes connecting them all are verb complementation and NZE. The study therefore brings together two major areas of research: non-finite verb complementation, and World Englishes.

This Introduction brings together the results of the four publications in order to address the following research questions:

1) Does NZE show significantly different complementation patterns of the selected predicates to BrE and AmE?

2) Is there any observable innovation in NZE complementation?

3) How can the research contribute to our current understanding of "colonial lag/innovation"?

4) What insights are gained from the application of the Choice Principle to the verb *fear* in NZE?

The reasons for the focus on these particular areas are as follows: Research questions 1, 2 and 3 are designed to address the complex nature of the influence among the dialects of English, and the debate surrounding the related notion of 'colonial lag'. As is well-known, NZE is a post-colonial language variety, a descendant of BrE and a direct result of British colonisation efforts in the nineteenth century.[2] With Britain's status in the world subsequently overshadowed by the rise of the United States the following century, and spurred on by various events, former British colonies began a process of disconnection from the homeland. A topic of interest among language scholars has since been the interconnected nature of the various different spheres of influence – external and internal, cultural, political, military, etc. – and the resulting effects that can be observed in the language of the different regions. Research questions 1 and 2, therefore, look into the differences that can be observed between NZE and the two older and more influential varieties of BrE and AmE, and consider the reasons for those differences. Research question 3 then reflects on colonial lag, and the ways in which the present research can complement our current understanding of the issue.

Research question 4 addresses an issue that is more central to complement selection than it is to inter-dialectal influence, as it discusses the topic primarily from the NZE perspective. The question is grounded in the choice speakers make between two sentential complements of the verb *fear*, and how the semantic role of the subject affects the selection. Research question 4 discusses the contribution of the study to the existing literature on selectional restrictions between co-exisiting sentential complements. Although the study does not take other varieties into account, the stage is nevertheless set for future comparative studies.

[2] This is perhaps a little misleading. It should be pointed out that in 1839, shortly before the signing of the 1840 Treaty of Waitangi, which saw New Zealand become a British colony, New Zealand was not yet under British rule and at that point "the British government […] had shown little inclination to acquire another colonial possession" (Gordon et al. 2004: 39). The colonisation 'efforts', though, were soon to follow, with active measures taken by both government and a private party to encourage people to move from Britain to New Zealand in large numbers.

1.2 Structure of the dissertation

The dissertation is structured as follows: Chapter 2 discusses aspects of the theoretical frameworks used in the publications that are not discussed at length in any of them. The chapter first covers all topics that fall within the area of complementation, including control theory, the postulation of null subjects, the differences between the various types of control and raising structures covered in the publications, as well as a brief history of the development of the non-finite complements of English. Following this, the chapter introduces the field of World Englishes, the various theoretical models that have been postulated in recent decades, aspects of the power dynamics among dialects of English, and the notion of colonial lag. Following this, the chapter provides the necessary historical background information on New Zealand and New Zealand English, with a review of previous research done on NZE completing Chapter 2.

Chapter 3 describes the materials and methods used in the four publications. This includes discussion of corpus linguistics past and present, justification for the use of newspaper material as a source of data, and a description of all the corpora used in the publications.

Chapter 4 comprises the results and discussion, addressing the research questions that this dissertation aims to answer.

Chapter 5 is the final chapter of the Introduction to the dissertation, and it provides closing remarks and notes future research possibilities that the present work has opened up.

2 THEORY AND BACKGROUND

This section introduces the theoretical frameworks used in the four publications, and expands upon theoretical issues that were not possible to discuss in detail in the publications due to space restrictions. Section 2.1 focuses on the syntactic aspects of the theoretical framework, section 2.2 with the World Englishes aspect, section 2.3 provides historical information on New Zealand, and section 2.4 reviews the existing literature on NZE.

2.1 Complementation

2.1.1 Introduction

A section whose purpose it is to introduce the reader to the venerable field of complementation ought to begin with a general description taken from a reliable source, and one could certainly do worse than the oft-quoted formulation from Noonan (1985: 90), which states that "[c]omplementation is basically a matter of matching a particular complement type to a particular complement-taking predicate." This dissertation is only concerned with verbs as complement-taking predicates, but it is also possible for other parts of speech to license complements. Rudanko (2017: 2) points out that, as well as verbs, "adjectives and nouns may, in their capacity as heads, syntactically subcategorize for a certain type of phrasal constituent or for certain types of phrasal constituents."

Both Noonan's and Rudanko's quotes above highlight a key element of complementation; that the predicate and the complement must be compatible with one another in order to be in line with the grammatical rules of the language or dialect. In standard English, for example, the verbal predicate *promise* can subcategorize for a *to* infinitive clause, allowing a sentence such as *Michael promised to fix the leaky sink*, but it does not subcategorize for an *-ing* clause, disallowing a sentence such as **Michael promised fixing the leaky sink*.

Naturally, the acceptability of different patterns of complementation varies, both diachronically within one variety, and synchronically across varieties. A statement on the degree of variety that is to be found among patterns of complementation notes that "[c]omplement constructions are as intricate and varied as the patterns on a Persian tapestry" (Ransom 1986: 1). The author was referring to standard English, most probably American English, and not to the rich variety of complementation possibilities across the native and non-native dialects of English worldwide – the study of World Englishes was at the time in its infancy – but the statement remains true in either case. Not only is it normal for most predicates to have a wide range of subcategorization possibilities, but speakers of different sociolinguistic groups within the one dialectal region can have varying preferences for different complementation patterns, a situation which reflects the diachronic shifting of grammatical norms.

Speakers of different dialects, too, are known to differ significantly in the way they go about matching complement types to predicates, and a great deal of effort has gone into the analysis of such cross-varietal complementation differences (see, among many others, Algeo 2006; Fanego 2007; Hoffmann 2018; Mukherjee and Hoffmann 2006; Kaunisto and Rudanko 2019; Mair 2002; Olavarría de Ersson and Shaw 2003; Rickman and Rudanko 2018; Rudanko 1996, 1998; 2006a; Rudanko and Luodes 2005; Rohdenburg 2009; Vosberg 2009). When taken together, the results of these studies reveal at least as many differently-patterned Persian tapestries as there are varieties.

On the relationship between the verb complementation system and the overall character of a given variety, it has been suggested that verb complementation is "an all-pervading structural feature of language and thus likely to be more significant in giving a variety its character than, for example, lexis" (Olavarría de Ersson and Shaw 2003: 138). The lexis of NZE is a somewhat different case than that of many other varieties, and as such it tends to upstage most other features, but otherwise I completely agree with the emphasis the aforementioned authors place on the role of verb complementation within the overall character of a dialect.

Turning to the orientation of the present dissertation within the system of verbal complementation, all four publications deal with aspects of the non-finite complementation of verbs, i.e. the complementation of verbs with *to*-infinitive clauses, *-ing* clauses, and *-ing* clauses within prepositional phrases, in both subject and object control contexts. The purpose of the following sections is to discuss issues relevant to the area of non-finite complementation that are not covered in detail within the publications, namely the historical rise of non-finite complements, control

and the issues of understood subjects, subject control and object control, and subject control and subject-to-subject raising.

2.1.2 The history of non-finite complements

The present dissertation deals exclusively with non-finite complements, and it is worth outlining their history and how they rose to become a major part of the system of present-day English verb complementation.

Of the complement types that fall within the scope of the present study, the *to* infinitive has by far the longer history. It is attested as a fully functioning verbal complement in the earliest Old English (OE) texts (Los 2015: 147). It entered into competition with the pre-existing bare infinitive and the finite *that* clause firstly in subjunctive contexts, and went on to gain ground, mainly at the expense of the finite complement.

The rise of the *to* infinitive was subsequently curbed by the development and introduction into the complementation system of the verbal gerund. The progression of the gerund from an OE "abstract noun of action formed by the addition of the suffixes *-ung* or *-ing* to a verb stem" (Fanego 2004: 7) to a form showing syntactic features typical of a verb is a process unparalleled in other Germanic languages.[3] It was propelled along by a complex set of circumstances beginning at the end of the OE period (for a detailed account see Tajima 1983). The decay of the OE case system paved the way for the increased use of prepositions (and word order) as a means of signalling grammatical relations (Mustanoja 1960: 348–349; Mitchell and Robinson 2012: 61–62), and the ability of the gerund to appear as complement to prepositions – something the *to* infinitive could not do – provided a niche for the gerund among predicates of selected verb classes (Fanego 2004). From this starting point, it worked its way into use as complement among other verb classes. Progress was gradual, and even though the first signs of its verbal properties were evident already in early Middle English (Fanego 1996; 2004; De Smet 2013), the ME period 1100–1500 did not see "regular and systematic use" of the fully-developed gerund (Fanego 1996: 33), and it is not until the early eighteenth century that we see a "sharp rise" in the use of the verbal gerund in a number of areas (Fanego 1996: 36).

The continuing encroachment of the gerund into areas previously dominated by the *to* infinitive – as well as the bare infinitive and finite clause complements – has

[3] Tajima (1983: 2) points out that the gerund has "expanded its syntactic role beyond anything characteristic of its own past history or of the other Germanic languages".

been the subject of a great deal of discussion, with the term "Great Complement Shift" (Vosberg 2006; Rohdenburg 2006) coined to describe the nature of these changes within English complementation system. Often discussed within the context of the Shift is a set of constraints on the use of each of the main complement types, which have been instrumental in providing shelter for certain complements in the face of competition from others. These constraints include: the avoidance of subsequent occurrences of the same form – the *horror aequi* restriction (Rohdenburg 2003; Vosberg 2003b); the preference for a more explicit complement in more complex environments – the Complexity Principle (Rohdenburg 1996); and the preference for a *to* infinitive in cases of complement extraction – the Extraction Principle (Rohdenburg 2006; Vosberg 2003b). Recent work has also postulated a semantic constraint, whereby a certain complement type is selected based on the agency of the understood subject – the Choice Principle (Rudanko 2012; 2014; Rickman and Rudanko 2018; Ruohonen and Rudanko 2021). The latter constraint is discussed in more detail in Publication IV.

Thus, the entry of the two main non-finite complements of Present-day English into the complementation system took place in two distinct stages: firstly with the *to* infinitive in OE, and later with the verbal gerund in ME. Leech et al. (2009: 185) summarize the situation as follows.

> If the spread of infinitival complement clauses at the expense of finite ones is a phenomenon which, in one form or another, can be observed throughout the whole recorded history of English, the rise of the gerund is a more recent phenomenon, dating back to the seventeenth century when the deverbal nouns ending in *-ing* began to take on verbal and clausal properties. The gerund thus emerged as an additional competitor in the domain of clausal subordination and started spreading in its turn, both at the expense of finite clauses and infinitival ones.

2.1.3 Control and understood subjects

The theory of control is central to the study of non-finite sentential complementation. It is closely associated with the different branches of generative grammar, and as such, provides a framework for a wide area of research and is a generally accepted concept in modern linguistics. Control can very briefly be described as "the question of what determines the identity of the unexpressed subject of non-finite verbal forms such as the infinitive or the gerund-participle in constructions such as *Joseph tried to find a quiet place* and *Peter enjoyed going fishing in his*

boat" (Duffley 2014: 13).[4] The unexpressed subject in such non-finite clauses is commonly represented by the label PRO in modern work on control, to denote "a phonetically null […] pronoun that occupies the subject position of infinitives in control structures" (Riemsdijk and Williams 1986, 132; see also Chomsky 1981: 20–26; Chomsky 1986: 119–132; Carnie 2013: 436ff; Landau 2013).

Several types of control are recognised in the literature (see Landau 2013 for a comprehensive overview) but the two main types dealt with in the present dissertation are subject control and object control. The former denotes the coreference of PRO with the subject of the matrix clause in a sentence such as *Watson$_i$ tried PRO$_i$ to build a treehouse*, the coreference indicated with matching subscript symbols; while the latter denotes the coreference of PRO with the object of the matrix clause, in a sentence such as *Watson persuaded Susan$_i$ PRO$_i$ to build a treehouse*. Section 2.1.3 below discusses aspects of the two types in more detail.

The existence of an empty subject slot in non-finite clause structure, and of empty categories in general, is something that is widely (but not universally) accepted today.[5] Landau, for example, states that "despite persistent skepticism, PRO exists, and there are compelling arguments … to show it" (Landau 2013: viii–ix). Landau's statement is bold, but there are good reasons for making it, and historically, a great deal of support for the position has been expressed. Observations by some of the early twentieth century grammarians of English indicate that even in the period pre-dating modern grammatical theories, the notion of understood subjects was recognised, as shown in the following extracts from Poutsma and Jespersen.

> The action or state expressed by an infinitive(clause) is necessarily associated with a person or thing (or group of persons or things) by way of subject. (Poutsma 1929 [1905]: 534)

> "[v]ery often a gerund stands alone without any subject, but as in other nexuses (nexus-substantives, infinitives, etc.) the connexion of a subject with the verbal idea is always implied." (Jespersen 1961 [1940]: 140).

The arguments in support of PRO are not discussed at length in any of the publications comprising this dissertation, and therefore the purpose of the remainder

[4] Duffley's "gerund-participle" is what I (and others) term the *-ing* clause or gerundial complement.

[5] Not all theoretical frameworks accept the idea of understood elements. Construction grammar theorists, for example, differ from mainstream generative theorists on this point in taking a "what you see is what you get" approach to syntactic structure (Goldberg 2003: 219; 2006: 10): The general understanding among theorists of this school is that there exist "no underlying levels of syntax or any phonologically empty elements". While my primary goal here is not to argue against this position, the theoretical basis for the present dissertation is nevertheless situated within mainstream generative grammar and the existence of empty categories is therefore assumed.

of this section is to briefly outline the main considerations, firstly theoretical, then empirical.

Firstly, it is assumed here that infinitival and gerundial complements have clausal status within the sentence structure (Koster and May 1982), and secondly, it is assumed that, as clauses, they have subjects. This second assumption is grounded in two principles that have been widely influential within the generative grammar framework for four decades. The Extended Projection Principle (Chomsky 1982: 10) stipulates that all clauses have subjects. Specifically intended to cover the clausal subject requirement, this is a refinement of the Projection Principle (Chomsky 1981: 29), which states that "[r]epresentations at each syntactic level (i.e., LF and D- and S- structure) are projected from the lexicon, in that they observe the subcategorization properties of lexical items".

Furthermore, the recognition of a subject slot in a non-finite clause makes it possible to fully represent the argument structure of the verb, thereby satisfying the theta criterion. The criterion states that "[e]ach argument bears one and only one θ-role, and each θ-role is assigned to one and only one argument" (Chomsky 1981: 36). Thus, the lower verb *build* in the sentence *Watson tried to build a treehouse* subcategorises for two arguments: an external Agent argument and a Theme. The matrix subject NP *Watson* is theta marked as an Agent of the matrix verb *try*, and in order not to violate the theta criterion it cannot be assigned a second Agent theta-role from *build*. Therefore an implicit subject of the embedded clause is needed.

Turning to empirical evidence for PRO, there have been many arguments made over the years. Many of these require discussion of the various different types of control that are recognised in the literature,[6] and since the present dissertation deals only with subject and object control, we only need to consider one of the pieces of evidence here. The evidence in question draws on the restrictions placed on the use of anaphoric expressions within binding theory (Chomsky 1981). Consider the sentences in (1a–d) with a reflexive pronoun in the embedded nominal clauses, in squared brackets.

(1) a. [Pushing yourself to the limit] was normal in those days.
b. [Pushing himself to the limit] was normal for Mike.
c. [To push yourself to the limit] was normal in those days.
d. [To push himself to the limit] was normal for Mike.

[6] See Landau 2013 for a full discussion of the arguments for PRO.

The evidence for PRO is based on Principle A of binding theory: "an anaphor is bound in its governing category" (Chomsky 1981: 188), i.e. the reflexive in the non-finite clause in each example needs a suitable c-commanding binding element.[7] In (1a) and (1c) there is no such element present in the surface structure for the reflexive *yourself*, meaning that there has to be an implicit syntactic subject of the non-finite clauses in order to bind the reflexives. In (1b) and (1d) the NP *Mike* is present in the surface structure, and, although the reflexive *himself* is co-referential with this NP, it does not c-command the reflexive and therefore cannot act as the binder. The postulation of a syntactically present but implicit subject is therefore the only possibility. The versions with the c-commanding binder PRO shown are given below.[8]

(2) a. [PRO$_i$ pushing yourself$_i$ to the limit] was normal in those days.
b. [PRO$_i$ pushing himself$_i$ to the limit] was normal for Mike$_i$.
c. [PRO$_i$ to push yourself$_i$ to the limit] was normal in those days.
d. [PRO$_i$ to push himself$_i$ to the limit] was normal for Mike$_i$.

The argument for PRO based on the binding of anaphors is generally taken to be one of the strongest. An early version of the argument appeared in Postal's influential (1970) discussion of coreferential complement subject deletion – under the theory then known as Equi-NP Deletion – and while a deletion analysis has been replaced by an interpretive approach to control, Postal's early observations are still relevant,[9] which indicates that evidence presented above is generally accepted and has stood the test of time and the modifications to the generative grammar framework that were made in the following decades.

[7] Jespersen had also noted the role of the reflexive in control structures, writing "[t]hat the primary, though not expressed, is present to the mind is shown by the possibility of using a "reflexive" pronoun, i.e., one indicating identity of subject and object, etc., with infinitives and nexus-substantives: to deceive *oneself* |control of *oneself*" (Jespersen 1992 [1924]: 143)

[8] I use the basic bracketing representations here, rather than tree diagrams.

[9] In that paper, which predates the introduction of PRO, Postal postulated a PRO-like entity which he called *doom*: "Observe that one can naturally think of deletion governed by coreference as equivalent to the existence of some general pronoun, call it Doom, which accidentally has the null phonological shape" (Postal 1970: 458)

2.1.4 Subject control and object control

Subject control is dealt with in Publications III and IV, while object control is the theme of Publications I and II. Examples of subject control are given in (3a–b) below, and examples of object control in (4a–b). (All examples are taken from the original publication data and include the original corpus identifier tags.)

(3) a. The mayor promised to write back to the committee with the exact figures. (20-07-23 GB)
b. I'm a parent, and I still fear sitting next to babies and toddlers on planes as much as anyone. (16-12-06 NZ)

(4) a. The law prevents Jan from suing the taxi company. (CNZNE, *Sunday News*, 1996)
b. The intenseness of the situation pushes Amanda into reflecting on painful childhood memories. (CNZNE, *Daily News*, 1996)

Looking first at the constituent structure of the subject control patterns of (3a–b), the necessary sections are repeated below, using brackets for a clear representation of the syntactic structures.

(5) a. $[[\text{The mayor}_i]_{NP}\ [\text{promised}]_{Verb1}\ [[PRO_i]_{NP}\ [\text{to}]_{Infl}\ [[\text{write}]_{Verb2}$ back to the committee$]_{VP}]_{S2}]_{S1}$
b. $[[I_i]_{NP}\ [\text{fear}]_{Verb1}\ [[PRO_i]_{NP}\ [[\text{sitting}]_{Verb2}$ next to babies$]_{VP}]_{S2}]_{S1}$

In each case, PRO is marked as coreferent to the matrix NP and these are both therefore subject control structures. (5a) contains a *to* infinitive clause as the complement, which is the only option in terms of non-finite complements with the matrix verb *promise*, and in (5b) *fear* selects an *-ing* clause complement, which, in the case of this matrix verb, is interchangeable with the *to* infinitive. It is worth making the point that the *to* in the *to* infinitive of sentences like (3a) is assumed here to be a grammatical clause marker, and not a preposition (see Duffley 2000 for the prepositional analysis of *to*) and is generated under the INFL node (Chomsky 1981: 18–20) (or T node in more recent work, AUX node in older terminology).

Turning to the structural representation of the object control examples of (4a–b), the bracketed versions are given in (6a–b) below.

(6) a. [[The law]$_{NP}$ [prevents]$_{Verb1}$ [Jan$_i$]$_{NP}$ [[from]$_{Prep}$ [[PRO$_i$]$_{NP}$ [[suing]$_{Verb2}$ the taxi company]$_{VP}$]$_{S2}$]$_{PP}$]$_{S1}$

b. [[The intenseness of the situation]$_{NP}$ [pushes]$_{Verb1}$ [Amanda$_i$]$_{NP}$ [[into]$_{Prep}$ [[PRO$_i$]$_{NP}$ [[reflecting]$_{Verb2}$ on painful childhood memories]$_{VP}$]$_{S2}$]$_{PP}$]$_{S1}$

In contrast to the subject control structures, it can be seen in (6a–b) that an NP appears directly after the matrix verb, and that this NP object is coreferent with PRO.[10] Both sentences also incorporate a prepositional phrase as the third argument of the matrix verb, and within that phrase the *-ing* clause acts as complement to the preposition.[11] The pattern in (4a/6a) has been termed the *from -ing* pattern (Rudanko 2002), and the pattern in (4b/6b) the *into -ing* pattern (*ibid*). One major difference between the two complement types is that the preposition is omissible in the case of the *from -ing* pattern,[12] while Standard English does not allow the preposition to be omitted from the *into -ing* pattern.

2.1.5 Subject control and Subject-to-Subject Raising

The dichotomy between subject control and subject-to-subject raising is the topic of Publication III, with a focus on the predicates *promise* and *threaten*. These predicates are well-known to lend themselves to both control and raising structures, and as such

---

[10] Note that the subject control verb *promise* also allows an NP in the slot directly after the matrix verb (*The mayor promised his wife to write back to the committee*) while remaining a subject control structure. This aspect of *promise* has long been discussed for the counterexample it presents to Rosenbaum's (1967) Minimal Distance Principle.

[11] I use the label *preposition* for the *from* in the *from -ing* patterns. It has been argued, however, that in structures with the *prevent* class of verbs the (optional) *from* is not a preposition in all cases. Aarts (1990) suggests that verbs of this class may have two subcategorization frames available to them, and the choice depends on the animacy of the NP object. Inanimate NPs may trigger selection of a two-argument frame and a raising-to-object structure, and the *from* in that case should be seen as a complementizer rather than a preposition. Animate NPs on the other hand, trigger selection of a three-argument frame and a control structure, and within this structure *from* is a preposition. And, although he does not discuss the control properties of the verb structures in his work on these verbs, Mair (2002; 2009b) uses the label 'preposition/complementizer', which seems to acknowledge Aarts' argument. In any case, further discussion of that issue falls outside the scope of the present work, and the example given in (4a/6a) shows an animate NP object as an argument of the matrix verb, and *from* here is a preposition in an object control structure.

[12] This statement requires some qualification: AmE typically does not omit *from* from the pattern with the majority of verbs of the *prevent* class. In other varieties, while generally omissible in an active construction (*The law prevents Jan suing the taxi company*), it can only be omitted "very marginally with *prevent* and *stop*" (Huddleston & Pullum 2002: 1238) when the structure is passivized (?*Jan is prevented suing the taxi company*).

often exhibit semantic ambiguity. The standard structural differences between subject control and subject-to-subject raising are outlined in the publication and need not be reproduced here, but the fact that many ambiguous examples were found in the corpus data during the course of the investigation for Publication III makes it worthwhile to look into the basis for their classification in the publication in more detail.

The corpus data available on *promise* and *threaten* tends to turn up many ambiguous examples due to their dual roles as control and raising verbs, and lying at the less ambiguous end of the scale are sentences such as *Your daughter promised to be a pretty girl, Betteredge, and she has kept her promise*. Taken from a separate study (Rickman and Rudanko 2023), this example sentence was published in the year 1868,[13] when the grammaticalization process that resulted in the availability of *promise* and *threaten* as raising predicates in addition to their original control uses had been underway for some time (Traugott 1993; 1996; Rickman and Rudanko 2023). In the example, the subject of the first main clause, although [+HUMAN], and therefore a fairly prototypical NP subject for *promise* as a control verb, needs to be interpreted as not having been assigned a theta-role by the matrix predicate, but instead as having been theta-marked by the lower verb *be* and subsequently raised into the empty matrix subject slot. The example shows that the speakers/writers of the period were making good use of the polysemy offered by verbs such as *promise* for the purposes of creative and colourful writing. It is argued in Rickman and Rudanko (2023) that this very use may have been one of the driving forces behind the relatively quick development of *promise* and *threaten* as raising verbs during the early modern English period.

While the example given above does not seem overly confusing or difficult to interpret, there nevertheless remains scope for the less likely interpretation, whereby the daughter is seen as the Agent of the matrix verb and that she actually gave a verbal assurance of one day becoming pretty – and in cases such as this it is this ambiguity that provides the humour. Such ambiguity characterises many of the corpus tokens gathered for Publication III, and among the more challenging to classify are the types shown in examples (7a–b), both taken from the Publication III data.

[13] The data in Rickman and Rudanko (2023) is from CLMET3.0 - the Corpus of Late Modern English Texts, version 3.0 (Diller et al. 2011)

(7) a. the first gig I'd ever been to that employed Yondr, a technology that promises to provide "phone free spaces" for gigs, concerts and meetings. (17-07-07 NZ)

b. As Reviewed's favorite dry shampoo, Dove Refresh + Care Dry Shampoo promises to revive flat or unwashed hair while adding shine. (20-11-25 US)

This class of tokens calls for a more systematic method of classification. The subject in this class is a [–HUMAN] NP of the sort which is very typical of a raising structure, i.e., it cannot be an Agent to a commissive verb such as *promise* in its control use. However, it typically denotes a commercial product, service, television series or movie, or something along similar lines, and as such there is a potential link to a [+HUMAN] entity to fulfil the Agent role needed for a control reading. The lower clause expresses a commitment that could easily be seen as having been made by an Agent; for example, the examples above could be rephrased with a clear Agent subject with a natural result: *We promise to provide phone free spaces* / *We promise to revive flat or unwashed hair*. As discussed in the publication, the acceptability of adverb insertion was used as one indicator of a control reading (*a technology that explicitly promises to provide "phone free spaces"*) but the adverb insertion test also needed reinforcing with further tests.

In such tokens, evidence needed for classification as control is taken from the level of detail in the lower clause. In (7a), the quoted language provides the evidence, as it seems to have been taken directly from the written or spoken language of the linked Agent, i.e., the person/people representing the company, firm, manufacturer etc. of the product. In a similar way, in (7b) the detail in the content of the lower clause can be taken as evidence. While not actually quoted, the phrase *while adding shine* was judged as being too detailed to have been the assessment of an outside observer, and the token is therefore seen as a case of control.

In the absence of any of the pieces of evidence outlined above, the default classification for these types of tokens was raising. The corpus data presented several such examples which failed to exhibit the necessary control-like characteristics, and two such examples are given in (8a–b).

(8) a. … with an F1-esq feet up driving position, plenty of carbon fibre and excellent build quality, the car promises to feel like a properly premium track toy. (18-06-27 NZ)

b. Sea of Thieves promises to deliver high seas hilarity as you take on fantastical creatures and other swarthy players … (16-06-13 GB)

These examples seem to less naturally accept an adverb such as *explicitly* (*the car explicitly promises to feel like a properly premium track toy / Sea of Thieves explicitly promises to deliver high seas hilarity*); they less naturally convert to [+HUMAN] subject NPs (*we promise that our car will feel like a properly premium track toy /we promise to deliver high seas hilarity*); and they do not contain quoted material or overly detailed information in the lower clause. Instead, they seem more likely to be the assessment of an outside observer, and therefore cases of raising. It should be borne in mind though that the outside observer in cases of raising is typically a reviewer or journalist familiar with the field or the product, so the lower clause is likely to be somewhat detailed in any case – a fact that further complicates the picture.

The discussion in Publication III notes that of the two verbs, *promise* is the only one to show such a high degree of ambiguity. The corpus data on *threaten* contained far fewer borderline tokens, but there were several with the type of [–HUMAN] NP discussed above. These, however, were all much more clearly linked to a [+HUMAN] Agent, as NPs denoting things such as letters, emails, directives etc, in which the threat can only have come from such an Agent. Two examples of this type are given in (9a–b).

(9) a. The letters threatened to contaminate infant formula if New Zealand did not stop using 1080 (15-12-03 NZ)
b. The emails also threatened to expose the company for "supporting her opinions" whenever she engaged in such "slander and libel." (18-02-06 US)

Note also the quoted phrases included in (9b), which make it very easy to ascribe the verbal act to a [+HUMAN] Agent. The different levels of ambiguity present among the *promise* and *threaten* data is a further noteworthy aspect of the study, and one which deserves separate treatment in future work.

2.2 World Englishes

The main focus of this dissertation is on a native-speaker variety of English, with a secondary focus on two other native varieties. Despite the lack of a wider consideration of different English dialects, it will nevertheless be instructive to establish the place of NZE within the network of World Englishes. This section aims to do that, with a brief overview of prominent World English models, discussions on the topic of influence among Englishes, and the much-debated notion of 'colonial lag'.

2.2.1 Models of World Englishes

The 1980s saw a rise in interest in the Englishes of the world (see section 2.5 below for further discussion), and it was during this period that efforts were made to categorise the varieties and understand the relationships between them. Three important models of World English were postulated towards the end of this decade: Kachru's Three Circles model of World Englishes (Kachru 1985); McArthur's Circle of World English (McArthur 1987), and Görlach's Circle model of English (Görlach 1990). Of the three, Kachru's model is the one that has received the most recognition, and the labels "Inner", "Outer" and "Expanding" Circles are still widely used today as a convenient way of referring to the different forms of English used in different areas of the world (see e.g. Melchers et al. 2019, whose descriptive framework is based on the three circles model).

Kachru's model thus proposes a division of the world's nations into three areas, represented as circles, labelled Inner, Outer and Expanding. In Inner Circle countries, English is spoken as a first language by the majority of the population, is used as the language of politics, education and business, and while it may not be the only official language – or even an official language at all[14] – it is the main one. These are the countries to which English spread through settlement processes, often at the expense of a pre-existing language and culture. As Figure 1 shows, New Zealand is an Inner Circle country, along with Britain and the USA.

The Outer Circle contains those countries where English is used alongside at least one other language, with a significant proportion of the population bi/multilingual.

[14] In the USA for example, "[t]here is no "official" language at the federal level for the United States" although "[s]ome individual states list English as their official language." (https://www.usa.gov). Similarly English is not given official legal status in the UK (Melchers et al. 2019: 43).

The English language in these Outer circle countries "has acquired an important status in the language policies of most of such multilingual nations" (Kachru 1985 12–13). These nations are predominantly former exploitation colonies, colonised and used for various commercial or military purposes by Inner Circle powers, who introduced English in the process, and the language became an integral part of the countries' linguistic ecology.

The Expanding Circle comprises those areas of the world in which English is secondary to at least one native language, and often has little or no official role in government or education. These countries tend to have few, if any, historical or political ties to the Inner Circle countries, as "English was not brought to these parts

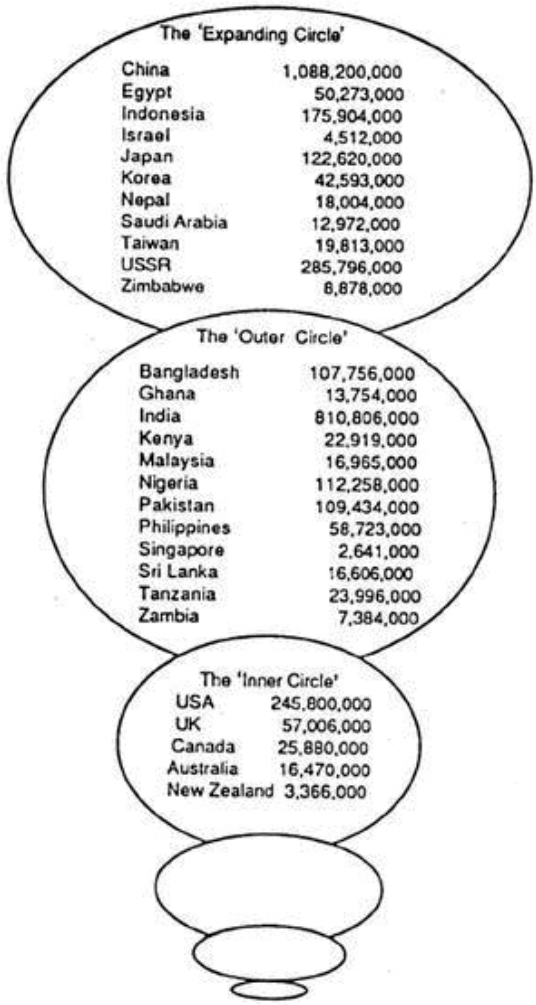

Figure 1. Kachru's Three Circles model of World Englishes (Kachru 1990: 179)

of the world by colonial activity in its usual sense" (Nelson et al. 2020: xxvii). The language has nonetheless made itself indispensable in various sectors of Expanding Circle nations, and is widely learned as a foreign language by an increasing proportion of the population. In today's rapidly changing economic and technological climate, it is clear that English is an essential tool for people in Expanding Circle countries to interact with the global community.

The division of the world into these three groups therefore reflects geography as well as the historical foreign policies of the world's most powerful nations. The Inner Circle is the result of the Second; the Outer Circle the Third; and the Expanding Circle the Fourth Diaspora of English (Nelson et al. 2020: xxvii).

An alternative model, proposed in the last decade, is Christian Mair's World System of Standard Englishes (Mair 2013). Based on an existing theory of Language Systems, which aims to provide a way of categorising languages to account for the way in which the speakers of the world's approximately 6,000 languages interact across linguistic barriers (de Swaan 2010), Mair's model is intended to reflect the hierarchy among the world's varieties of English in a more accurate way than the Three Circles model can. Under this model, AmE is at the "hub", the one "hyper-central" variety, ranked as such due to its demographic weight, or number of speakers, and institutional support, a factor measured in terms of "established orthography … official status and use in prestigious domains of communication" (Mair 2013: 258). The hub variety therefore has global influence (and within de Swaan's world languages model it is English that occupies the hub).

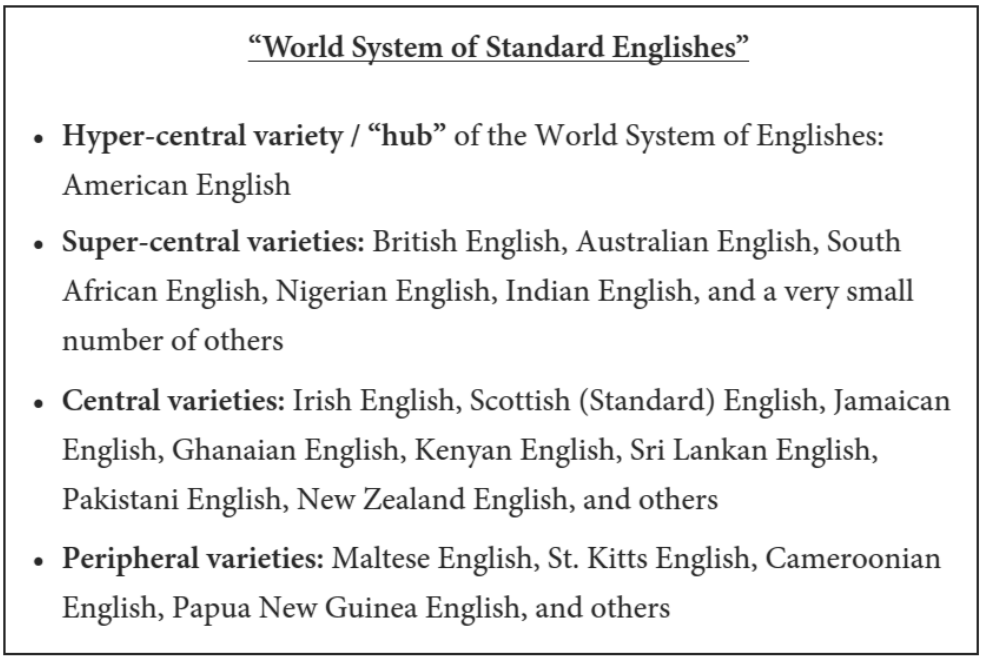

Figure 2. The World System of Standard Englishes (Mair 2013: 261)

The "super central" varieties, such as British, Australian, and Indian Englishes, have significant demographic weight and institutional support, and some influence beyond their borders but somewhat less than the hub variety. Below this level are the "central" varieties, where NZE is located, alongside Irish English, Scottish Standard English, and several traditionally non-native varieties such as Ghanaian and Pakistani Englishes. The central varieties are characterised by more modest demographic weight,[15] and are defined as languages which "are firmly entrenched in their own home bases but have limited impact on the development of varieties beyond this" (Mair 2022: 415). The lowest level contains the numerous "peripheral varieties", spoken by fewer people and more likely to diverge significantly from the standard varieties in many ways.

Mair's model therefore places the three Inner Circle varieties that form the basis for the present dissertation on three different levels of the model, a grouping which brings out the differences between them in terms of size and global relevance.

[15] In December 2022 New Zealand had a population of 5,151,600 (www.stats.govt.nz), most of whom can be assumed to be native speakers of New Zealand English.

2.2.2 Dynamics and influence among World Englishes

Section 2.2.1 above noted that a recent model of world Englishes (Mair 2013) placed AmE on the highest tier of a four-tiered model, the position occupied by the variety with the most speakers, the strongest institutional support, and the greatest transnational importance and global influence. It is likely that, had the model been proposed fifty years ago (and, perhaps, disregarding the population factor, which even 50 years ago was far greater in the United States), the position at the top could have gone to BrE. It is the purpose of the present section to discuss the power dynamics among World Englishes, and how the influence of the main varieties may affect NZE.

It is hard to argue against the view that since World War II, Britain's position as a world leader in terms of political, economic and military power has been overshadowed by the United States. Tombs (2014) writes:

> When Britain emerged as a significant force, after the War of the Spanish Succession in 1713, it was the smallest and yet most global of the world's half dozen or so most powerful states … It occupies a similar position three centuries later. The change in the world has not been the decline of Britain, but the post-1941 rise of America, which in wealth and military power outdistanced not only Britain, but every other state. (Tombs 2014: 768; see also Schneider 2007: 291–292)

The United States inevitably exerts a linguistic influence over developing second-language varieties of English as well as native "inner circle" varieties, and one explanation for this, as noted by Kachru (1986: 144), is simply that "language and power go together". Leech et al. (2009: 22) note that "[t]he latter part of the twentieth century, the era in focus here, seems to have been a period shaped … by a pervasive and growing influence of American usage on all varieties of English." Too much should not be made of this, however, and it is significant that Leech et al. (2009) contains a section entitled *Americanization?* (*ibid.* 252), with the term appended with a question mark, as does an earlier volume (Mair 2006: 193)[16]. It is pointed out in Mair (2006) that, while there are clear cases of other varieties adopting the AmE linguistic model, the overall picture that we gain when taking all factors into account is that of each variety making its own way in roughly the same direction as all the others, sometimes crossing paths and sometimes taking a slightly different route. As Mair puts it "we need to realize that many instances in which British (or Australian, or Irish, etc.) usage seems to follow American practice do not necessarily represent

[16] As does a sub-section in Hundt's (2017) chapter on Australian/New Zealand English.

direct American influence. Rather, they show all varieties of English developing along the same lines and toward the same putative end-point, but at slightly different speeds" (Mair 2006: 193–194).

Mair (*ibid.* 194) goes on to point out that it is the lexicon of a variety that is most susceptible to American influence, the grammar less so, and the pronunciation remaining almost impervious to external pressure. In fact, it is well known that American words have long been working their way into the lexical inventories of other varieties, and this knowledge was reinforced in a recent study focusing on vocabulary and spelling preferences in 30 countries (Gonçalves et al. 2018). The results of this study showed that a clear shift towards American patterns can be discerned in many areas of the world. Countries with stronger historical and cultural ties to Britain, however, such as the former colonies Australia, New Zealand and South Africa, tend to retain British norms more clearly.

Hay et al. (2008: 76), citing a 1989 study by Donn Bayard, note that New Zealanders have become more accepting of some American lexical items, with the American term replacing the British one in several cases over recent years.[17] Also, it bears pointing out that in Hay et al's (2008) book *New Zealand English*, the only chapter to have a subsection entitled *American influence* is the chapter dealing with vocabulary.[18]

2.2.3 The 'American Invasion' of New Zealand

To return to the latter period of the twentieth century, when a "pervasive and growing influence of American usage" (Leech et al. 2009: 22) becomes evident, if we are to identify a specific date which marked the introduction of American cultural influence in New Zealand, many would point to June 1942 when, in response to the

[17] Green and Bayard (2000) note a difference in the rates of acceptance of newer American terms over British ones in different parts of New Zealand, and Vine (1999) obtained similar results, although her study focused on one part of the country. And, more recently, Macalister has stated that "it is not necessarily the case that an existing word is being replaced by an American synonym. Speakers often have access to both terms for a referent, and allow context to determine which to use." (Macalister, fc).

[18] It is not only American English to have established a presence in the New Zealand vocabulary. New Zealand and its closest neighbour Australia have a long history of sharing words (see Hay et al. 2008: 74–75), and it is well-known that New Zealand English has a high percentage of words borrowed from Māori (Macalister 2006b; Davies and Maclagan 2006).

threat from Japan during World War II, the US began to station troops in New Zealand in large numbers:

> It had been only six months since the Japanese attack on Pearl harbour, and four since the fall of Singapore, but US was pouring vast resources into the Pacific to repulse the enemy. Over the next three years, approximately 150,000 American servicemen would pass through New Zealand, with up to 50,000 GIs camped in the country at a time … this was a cultural as well as military invasion, with little resistance offered. (Bourke 2010: 117)

The term "American invasion" has been used more than once in reference to the US troops' arrival (see also Bateman 1997), and the troops brought more than just weapons with them: during the years that they were there, they introduced American music, dance, food and, undoubtedly, vocabulary, to a hitherto relatively conservative New Zealand.

The 1970s brought further, more permanent exposure to American culture via television. Donn Bayard notes that "NAm media influence began to increase massively in 1974, when New Zealand acquired its second television channel" (Bayard 2000: 298), and the launch of a third, private channel in 1989, "can be said to mark the almost total Americanisation of New Zealand television in terms of number of commercials, style and frequency of trailers, and of course the shows themselves" (*ibid*). Bell (1988: 327) noted that in the late 1980s, "more than forty percent of television time is taken by American programmes (with twenty percent British)".

As New Zealand television and movie screens were awash with the sound of AmE, it can be said that the relevance and influence of Britain and BrE was concurrently waning among some sectors of society,[19] but it is not accurate to say that it was being *replaced* by AmE. This was a time when, according to Schneider (2007), New Zealand was entering a period of cultural and linguistic independence,

[19] Schneider (2007: 131) points to Britain's entry into the European Economic Community in 1973 as the time when New Zealanders began a psychological disconnection from Britain. In joining the EEC, Britain deprived New Zealand of a lucrative export market, which contributed to a period of economic hardship and caused the nation to start focusing on restructuring its own identity and independence. Schneider calls this New Zealand's 'Event X', an "exceptional, quasi-catastrophic political event which ultimately causes the identity alignment of STL [settler] strand speakers to switch from a self-association with the former mother country, however distant, to a truly independent identity," (*ibid.* 48). See also Belich (2001) on the 'decolonisation' of New Zealand following Britain's 1973 entry to the EEC. It should be added that, in describing the 'American invasion' of the 1940s here, I am not contesting Schneider's observation that the 1973 event outlined above was a turning point in the development of NZE. The arrival of the American troops did bring with it a certain amount of new cultural influence from a different part of the Anglophone world, but as far as I know it is not regarded as a particularly relevant event in the development of NZE or the identity construction of the nation.

the final stage in the typical development of a post-colonial variety. Rather than looking towards a replacement model however, a nation at this stage of development, according to Schneider, re-evaluates its own identity and takes pride in its independence. It is true that the American cultural and linguistic influence was pervasive but, as Vine stated in the late 1990s:

> There has definitely been an orientation away from Britain in recent times, Britain no longer being seen as 'Home' by the majority of New Zealanders. This does not mean, however, that New Zealanders are instead orientating themselves towards America. Rather, as noted in a recent TV programme on Kiwiana, they are turning more towards themselves. (Vine 1999: 20)

And in turning more towards themselves, they began to appreciate the part of their identity that makes them unique: the following section looks at the effect that the Māori language has had on NZE.

2.2.4 Influence within New Zealand: Māori

It was briefly noted in the Introduction above that the Māori language has had an impact on the NZE lexicon, and it is believed that this impact is greater than normal for an indigenous language in a post-colonial setting (Trye et al. 2020). Thanks to the Māori Renaissance of the late twentieth century (Hay et al. 2008: 8; Benton 2015),[20] NZE is experiencing a new wave of Māori borrowings. This is manifesting itself in

[20] The Māori Renaissance can be described as a renewed appreciation of New Zealand's indigenous culture following a period of decline. The decline is often seen as the result of the shift of a large number of Māori families from a rural to an urban environment and lifestyle following World War II. The shift was not backed up with sufficient institutional support for these families, however, who found themselves disconnected from their traditional way of life. King (2003: 477) describes the plight of the Māori language and culture during this period as follows:

> The language was in a relatively healthy state in the early 1930s. By the 1970s it was in serious danger of extinction as elderly native speakers died and were not replaced by younger ones. The policy of not speaking Māori in schools … had done some damage to the transmission of the culture but not nearly as much as that caused by the later breakdown of family and tribal links in the post-war years.

From the mid-1970s onwards, thanks to a number of factors, including national identity-related issues (Schneider 2003; 2007), increased media focus on the conditions Māori were facing, and the setting up in 1975 of the Waitangi Tribunal (a committee to oversee alleged breaches of the Treaty of Waitangi), efforts were made to revitalise the language and culture. The Māori language was given official language status in 1987.

an increase in Māori words appearing in newspaper articles, often without glossing (Davies and Maclagan 2006), and the increase has also been attested in other domains such as parliamentary debates and the *School Journal* [21] (Macalister 2006b). Research has also noted a rise in hybrid compounding as a word-formation technique (Degani and Onskyo 2010; Trye et al. 2020) and a more regular and comfortable use of Māori borrowings in everyday situations by non-Māori New Zealanders (Gordon 2005). In contrast to the Māori lexical items of the earlier main period of borrowing, 1769–1860 (Hay et al. 2008: 68), which were predominantly place-names and items denoting flora and fauna, those characteristic of the current wave are more related to Māori material and social culture (Macalister 2006b).

As for pronunciation, it has been noted that spoken NZE is becoming both faster and more syllable-timed in its rhythm in comparison to other English varieties (Nokes and Hay 2012), and the influence of the Māori language and/or Māori English[22] was thought to be a possible factor in this change (see discussion in Szakay 2006). A recent and rapid change in several aspects of the phonemic system of Māori[23] has also been noted (Harlow et al. 2009) and research strongly suggests that contact with English is responsible for this, but there is little evidence for the external influence of Māori on English pronunciation. Rather, Nokes and Hay (2012) suggest language-internal causes for the shift in spoken NZE rhythm, with a link to the well-known chain shift affecting the NZE short front vowels (see Bauer 1979; Gordon et al. 2004; Maclagan and Hay 2007).

The question most relevant to the present dissertation, however, is whether Māori has had, or is having, an influence on the grammar of NZE. Cases of structural transfer in language contact situations are numerous and well documented (Filppula

---

[21] The *School Journal* is an educational publication for New Zealand schoolchildren, established in 1907 and published several times a year, containing "New Zealand content to motivate, excite, and engage students across the curriculum", i.e., short stories, poetry, plays, non-fiction writing, current events, art, etc. (gazette.education.govt.nz).

[22] Māori English is an ethnic variety of New Zealand English, spoken mainly by Māori but also by non-Māori, often to signal solidarity with Māori people and their culture. It is thought to have developed following the mid-20th century urbanisation of many Māori, alongside their subsequent adoption of English and the decline in Māori language proficiency, discussed in footnote 20 above. In 2008 Māori English was thought to be the fastest-growing variety of New Zealand English (Maclagan et al. 2008: 668).

[23] The changes in question are "the abandonment in very large part of the range of allophones used for /f/; increasing aspiration of the three voiceless stops; decreasing distinction in both quantity and quality between the phonemically distinct long and short vowels; the shifts, apparently towards merger, of some of the diphthongs studied to date." (Harlow et al. 2009: 146)

et al. 2008; Heine and Kuteva 2005; Matras 2009; Thomason and Kaufman 1988; Winford 2003), and, in theory, grammatical transfer could have taken place at some level between Māori and NZE in New Zealand. According to Gordon et al. (2004), however, it did not:

> Māori had certainly already been in New Zealand some centuries at the time of the first European colonisation. A good number of Europeans became more or less fluent in Māori … Thus, while it is possible in principle that Māori could have had some impact on the pronunciation or grammar of New Zealand English, this does not appear to have happened. Māori did contribute to New Zealand English, but only lexically. (Gordon et al. 2004: 69–70)

Trudgill expresses a similar opinion, noting that "[c]ontact with indigenous languages seems to manifest itself mostly in terms of the acquisition of loanwords from these languages (Trudgill 2004: 4; see also Degani 2012: 14–15).

Thus, while its lexical influence on NZE is considerable and appears to be continuing, there is no evidence to suggest that the Māori language is having any effect on any aspect of NZE grammar.

2.2.5 Colonial lag

Colonial lag is the focus of research question 3, and the topic is also taken up in Publication I with some background provided there. Rather than repeat what is written in Publication I, however, the purpose of the present section is to briefly recap the main points and, since Publication I was written several years ago and the topic has attracted a fair amount of attention since then, I will survey some of the research that has been published on colonial lag more recently.

The term itself was coined in the mid-twentieth century (Marckwardt 1967 [1958]) to label a phenomenon that was earlier noted in a linguistic context in the latter part of the nineteenth century (Ellis 1869–89: 19; cited in Hundt 2009: 13):

> [T]here is a kind of arrest of development, the language of the emigrants remains for a long time at the stage in which it was at when emigration took place, and alters more slowly than the mother tongue, and in a different direction.

It is typically applied to a colonial language setting to describe the retention of features in the younger variety that have since either fallen out of use or progressed on to further stages of development in the parent variety (see also Görlach 1987 for a well-known account). It was recognised early on though, that the "arrest of

development" of a feature is only one possible scenario, and that "[c]olonial languages, like all dialects, exhibit both conservative and innovative traits" (Bryant 1907: 290). The latter type of scenario described by Bryant over a century ago would then be a "colonial innovation".

Dissatisfaction with the concept has certainly been expressed, perhaps most clearly by Mair (2009a: 13–14), who stated that "[o]nce again, the "colonial lag" has not provided an over-arching explanatory framework for developments in World Englishes but has been exposed as the myth it probably is". At around the same time, Marianne Hundt pointed out that "the dichotomy of 'colonial lag' and 'colonial innovation' … implies a far too simplistic view of the much more complex patterns and processes of language change" (2009: 14), and suggested an expansion of the framework to encompass a total of six possible scenarios. In doing so, Hundt replaced the terms 'colonial' and 'lag' with 'extraterritorial' and 'conservatism', respectively (with *ETE* in the below list denoting *Extraterritorial English*).

(a) Extraterritorial conservatism: older forms of home country usage are retained in the ETE
(b) Extraterritorial innovation
(c) Truly divergent patterns develop in both varieties
(d) Parallel developments in both varieties
(e) The revival of an older form in either variety
(f) Kick-down developments: what starts out as conservatism gathers momentum to overtake home country usage, and becomes innovation (adapted from Hundt 2009: 32–33)

In a study encompassing a range of variables in Australian English, Collins (2015: 39) agrees that "the oft-invoked notions of 'colonial lag'… and 'colonial innovation' … need supplementation by other patterns of change" and adopts Hundt's framework as a starting point but supplements 'lag' and 'innovation' with the five possibilities of "'revival', 'survival', 'divergence', 'parallel change', and 'overtake'" (*ibid.* 17), giving a total of seven possible scenarios. More recent studies that have provided support for colonial lag – although as one aspect of a more nuanced framework, and not in the "overarching explanatory framework" sense as rejected by Mair – are, among others, Schreier (2019), in a study on /h/-insertion in Tristan da Cunha English, and Fuchs (2022), in work on the Englishes of India and South Asia.

An important piece of relatively early work on colonial lag with a direct connection to NZE is Trudgill (1999; see also Trudgill 2004: 31*ff*), where he connects the concept of a one-generation developmental delay with the formation of a new dialect in a colonial setting such as New Zealand in the mid-ninteenth century. It is shown that, due to the fact that the children of the new colony lack a common peer-group dialect, they end up instead adopting a range of often very mixed speech patterns from the older generation. The development of the dialect is therefore delayed by around one generation, and, according to Trudgill, it is accurate to call this a case of colonial lag. He is, however, careful to restrict himself to this particular context, stating that "there is at least one sense in which "colonial lag" […] is, or at least in certain situations can be, a demonstrable linguistic reality" (Trudgill 1999: 227).

To conclude this section, a recent observation by Hundt on the relevance of the colonial lag framework to investigations into New Zealand (and Australian) English noted that "there is ample scope for further research on differential grammatical change in AusE and NZE and comparison with available evidence for British and American English. The likely picture to emerge from such studies is one that goes beyond the simple dichotomy of colonial lag and innovation" (Hundt 2017: 302). Colonial lag will be taken up again in the discussion of research question 3, in the results section below.

2.3 A brief historical background sketch

This section provides the necessary background information on the history of New Zealand and the dialect of NZE.

2.3.1 History of New Zealand

Human settlement of New Zealand, or, "the last habitable area of the globe to be colonised" (Bayard 1994: 108), is thought to have begun as recently as the early decades of the fourteenth century (Holdaway et al. 2019; Walter et al. 2017). The first arrivals were migrants from various island groups in Eastern Polynesia, most likely those of the Cook, Austral, and Society Islands (Walter et al. 2017: 369). The south-westerly push from Eastern Polynesia to New Zealand in the fourteenth century was the last phase of a three-thousand-year era of sea voyaging that spread

humans from Asia eastward across the Pacific Ocean as far as Hawaii in the north, Easter Island in the east, and New Zealand in the south: these three points are shown in Figure 3 below. The eventual end of the age of long-distance canoe voyaging isolated the early settlers of New Zealand from their islands of origin, and over the following centuries Polynesian culture evolved into what we now know as Māori culture, and the language of the settlers evolved into the Māori language.

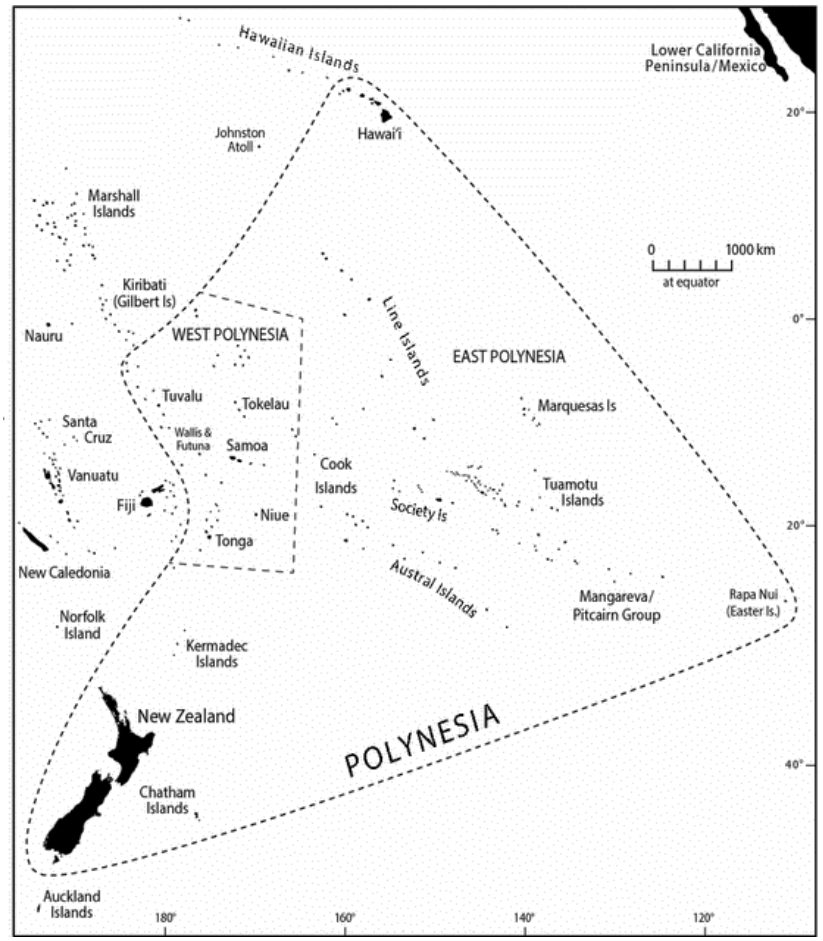

Figure 3. Map of Polynesia showing the likely origins of the first migrants to New Zealand: the Cook, Austral and Society Island groups to the north-east of New Zealand, and the outermost areas of settlement of the Pacific region.

Europeans first arrived some three and a half centuries after the Polynesians, in December 1642. These were Dutch explorers under the leadership of Abel Tasman,

and they were representatives of the Dutch East India Company, seeking to expand the commercial opportunities of the company in the uncharted area of the Southern Pacific, then known to Europe only as the myth of the great unknown continent Terra Australis. Due to a conflict with local Māori, the Dutch did not make landfall, and returned to Europe in January 1643 after spending only a short time in New Zealand waters (Belich 1996: 119–120).[24]

The next Europeans to arrive, led by the Englishman James Cook, were more successful. Cook's visit could well have been as short as Tasman's, but he was fortunate enough to have a Tahitian with him who was able to act as a translator between the English and the Māori. The Tahitian managed to facilitate discussion between Cook and the local Māori, and friendly relations were established. There were sporadic visits from other explorers in the years after Cook's first visit, and Europeans gradually began to find their way into the country to live among the Māori in the period from 1790 to 1840. These early European settlers were mainly former convicts (freed or escaped) from the Australian penal colonies that were established from 1788 onwards, missionaries, deserting seamen, and workers in New Zealand's newly established whaling, sealing, timber and flax industries. By 1840, there are thought to have been around 2,000 Europeans living in New Zealand on a permanent basis (McKinnon 1997).

[24] The reason for the conflict was a misunderstanding of cultural protocols between the Europeans and a local Māori tribe, which led to a battle at sea and the deaths of men on both sides. As emissaries from Europe did not return for another 127 years, Belich (1996: 120) has noted that the actions of the Māori tribe may amount to the "cheapest and most effective resistance to European expansion in all history."

Figure 4. Map of New Zealand, showing the main cities. Tasman arrived in Golden Bay (which he dubbed Murderer's Bay at the time), in the north-western corner of the South Island, while Cook made landfall in Gisborne, on the east coast of the North Island.

The year 1840 saw the signing of the Treaty of Waitangi, an agreement between the Māori and the British Crown under which New Zealand became a British colony. Schemes to attract settlers to the colony were put into effect by a commercial enterprise known as The New Zealand Company, and the discovery of gold in the 1860s, along with the ongoing New Zealand Wars (Belich 1986), which required the relocation to New Zealand of large numbers of British troops, also brought in new people, both permanent and temporary. The combination of these factors, together with natural growth within New Zealand among the settler population, resulted in a rapid rise in the numbers of non-Māori citizens; only 40 years after the signing of the Treaty, their number had risen to half a million (Belich 1996: 278).

The origins of the early non-Māori New Zealanders were largely various places in the British Isles, with a handful of English-speaking immigrants from North America and Australia. In addition, small numbers of people also came from non-English speaking areas, such as Germany, Scandinavia and China. Table 1 below shows the origins of the settlers born outside NZ as of 1881.[25]

Table 1. Origins and numbers of the NZ population born outside NZ as of 1881. (Source: www.stats.govt.nz)

Birthplace	Number	Percent
England	119,224	44.7
Scotland	52,753	19.8
Ireland	49,363	18.5
Australia	17,277	6.5
Wales	1,963	0.7
Other British possessions*	4,014	1.5
China	5,033	1.9
Germany	4,819	1.8
Denmark, Sweden, Norway	4,734	1.8
USA	841	0.3

*This category is not included in the figures from www.stats.govt.nz, but is taken from Gordon et al. (2004: 45)

It is important to take account of the origins of the settlers during the period of rapid settlement, because, as will be shown below, the period between 1840 and 1880 is seen as the period during which the New Zealand dialect formed. The theory of new-dialect formation (Trudgill 1986) is today the most widely-accepted explanation

[25] Details of the actual places of birth of the immigrants are not available, and the numbers in Table 1 reflect the situation in as precise terms as are available. Gordon et al. (2004: 37) write:

> [t]here is a serious lack of historical data on the origins of the early immigrants … in 1972 the New Zealand department of statistics destroyed all the nineteeth-century returned questionnaires that had been collected since the beginning of comprehensive census-taking in 1857, and which reported the original birthplaces of much of the New Zealand population.

However, since this dissertation is not focused on the pronunciation of New Zealand English, where the regional pronunciations of the early European settlers are crucial, the lack of detail is less important. (See Gordon et al. 2004: 36*ff* for discussion.)

of the origins of NZE, and it relies on the ratios of the varieties involved in the dialect mixture process. This theory is discussed, alongside other earlier theories, in the following section.

The term 'colony' was converted to 'dominion' in September 1907, although it did not bring with it any real change in the political relationship between New Zealand and Britain, and the same nominal upgrade to dominion status was applied to other "colonies with responsible government" (McIntyre 1999: 194) within the British empire around that time. The Balfour Declaration of 1926[26] used the term 'commonwealth' in reference to the group of nations under British leadership, and it emphasised the equality among the Dominions, and the increasing independence of the respective governments. Five years later, the ideals of the Declaration took shape in the 1931 Statute of Westminster, after which point "the Dominions were as independent as they wished to be, but in relation to Britain and each other they retained a common allegiance to the Crown and were freely associated as members of the British Commonwealth of Nations" (McIntyre 1999: 195). The proffered independence, however, was to take effect upon ratification of the statute, an act which was left up to the respective Dominions. For various reasons New Zealand chose not to accept full independence for another 16 years, putting it off until 1947.[27]

Today, New Zealand is, of course, still a member of the Commonwealth of Nations, and one of the smaller group of 15 Commonwealth realms to retain the British monarch as head of state, while maintaining full political independence.

2.3.2 History of New Zealand English

Since the early days, general observers and language scholars alike have proposed theories to explain the origins of NZE. Early theories – and, in some cases, fairly

[26] Not to be confused with the Balfour Declaration of 1917, which was an expression of British support for the "establishment in Palestine of a national home for the Jewish people" (as stated in the original document).

[27] Reluctance to take on full independence and thereby lose an aspect of the British identity that New Zealanders were very proud of at the time, may have been strengthened "by the length of time that New Zealand had been a British colony, by a continuing inflow of predominantly British immigrants, by the heightened feelings of imperialism – or, at the very least, that 'double patriotism' engendered by World War I – and by Britain's continuing role as receiver of New Zealand's exports and provider of its imports" (King 2003: 367). On this last point, there is an interesting connection with the point made in footnote 19, where it was noted that Britain's 1973 entry into the European common market, which effectively ended this lucrative export trade deal, was a factor in New Zealand's reassessment of its relationship with Britain, and a milestone in the creation of a new national identity (see Schneider 2003; 2007).

recent ones – have cited factors such as general laziness, poor teaching, excessive flies and dust, and the widespread use of false teeth to explain why New Zealanders speak the way they do (Gordon et al. 2004: 68–69). Of the more professional theories, though, two main perspectives emerged. One was monogenetic, supporting the idea of a single source for NZE (and, in general, other colonial varieties). The other was polygenetic, claiming that colonial varieties were the result of the mixture of several input varieties. Of the former perspective, its similarity to Australian English (AusE) and south-eastern England English in terms of pronunciation and vocabulary led many scholars to suggest that NZE was a transplanted variety of either Australian, Cockney, or south-eastern English. In 1994, after having considered the evidence for the various options, Bauer decided that "the hypothesis that New Zealand English is derived from Australian English is the one which explains most about the linguistic situation in New Zealand" (Bauer 1994b: 428).

Lass, in 1990, took account of multiple dialects coming into contact in the formation of a new variety, but part (i) of Lass' Law of Swamping states that "[i]n cases of mixed input to an ETE [Extraterritorial English], whatever the original demography, the output is (a) southern, and (b) more eastern than western." (Lass 1990: 269). Lass therefore accepted that a variety such as NZE has multiple source dialects but that this nevertheless has no effect on the resulting dialect, with all input features being swamped by the features of the south-east of England.

Peter Trudgill, in the landmark 1986 book *Dialects in Contact*, put forward a theory that took the polygenetic approach, clearly outlining for the first time the stages of a process in which several different dialects come together and mix, with the subsequent levelling out of surplus forms, resulting in one new dialect which is not exactly the same as any of the dialects which went into the mixture. Today "there seems to be a general consensus in the literature that the normal position is that colonial varieties are the consequence, at least in part, of dialect mixture" (Trudgill 2004: 11). With many recent studies based upon this theoretical framework (e.g. Kerswill 2020), it seems that the consensus has not shifted in the almost two decades since Trudgill made that statement.

New-dialect formation theory gained traction in the late 1990s, when linguists at New Zealand's Canterbury University learned of the existence of a large repository of audio recordings of interviews with elderly New Zealanders. The recordings were made by a team of technicians from the New Zealand Broadcasting Service travelling around the country in a mobile recording unit (a truck equipped with recording facilities) in 1946–48. The purpose was to collect and then later broadcast on the radio, for entertainment and education, the stories and reminiscences of the early

pioneers. The speakers on the recordings are the earliest generations of European children born in New Zealand during the period of heavy settlement following 1840 (though, Trudgill (2004: 25) notes that there are no speakers among those on the recordings who were born *in* the 1840s). The recordings, supplemented with recordings of later generations, make a unique corpus of spoken NZE that has given a great deal of support to new-dialect formation theory. The corpus and its related research project is known as the Origins of New Zealand English (ONZE) Project.[28]

Trudgill's new-dialect formation theory encompasses six key processes, which operate over three distinct periods. The periods correspond roughly to one generation each, and the processes are mixing, levelling, unmarking, interdialect development, reallocation, and focussing (Trudgill 2004: 84–89).

The period 1840–1880 is seen as the time during which the dialect formation processes took place in New Zealand[29] (Trudgill et al. 2000: 300), making NZE one of the youngest of the native, Inner Circle varieties of English. Of the other main Southern Hemisphere varieties, Trudgill et al. (2000: 302) state that "Australian English was formed in the period 1800-1840 … South African English in 1820-1860" (notes omitted).

As regards its applicability to the present dissertation, it should be noted that new-dialect formation, under Trudgill's formulation, is largely a theory concerned with the phonological aspects of a new dialect. Trudgill's work in this connection, and the work of the ONZE project team, is therefore heavily weighted towards the New Zealand accent. New-dialect formation theory is described here only to give some general background to the dominant theoretical framework that applies to dialects of the type under discussion here.[30]

---

[28] See www.canterbury.ac.nz/nzilbb/research/onze/ for details of the ONZE project, and the more detailed discussion in Gordon et al. (2004).

[29] It will be recalled that Europeans, mainly from the British Isles, were already living in New Zealand in the decades leading up to 1840 – as many as 2,000 in 1840 – but Trudgill discounts these speakers as having made little contribution towards the processes required for the formation of the new dialect, stating that "the English language arrived as a significant force in New Zealand, with large groups of immigrants only in the period from 1840 onwards. The crucial period for the formation of New Zealand English was thus between 1840 … and 1890" (2004: 24). It is true that many of the earliest pre-1840 settlers tended to assimilate themselves into Māori society, and were known as 'Pakeha Māori', and probably spoke more Māori than English.

[30] The theory is not necessarily restricted to phonology, however, and it has been applied to other areas. For example, Dollinger (2008) adapts the framework to the study of Canadian English modals.

2.4 Previous research on NZE

This section sets out to establish the place of the present dissertation within the existing literature. I first provide some general background to the study of NZE and a discussion of the primary foci of earlier research, before turning to a literature review of the more specific field of NZE grammar and syntax, and finishing off with a few words of justification for the present dissertation.

The late twentieth century saw a rise in scholarly interest in the non-central,[31] more recently developed, varieties of English. At the start of the 2000s, Schneider pointed to the early 1980s as the turning point, commenting that

> [t]he beginnings of the study of new varieties of English as a serious topic of linguistic research and a new subdiscipline of English linguistics can be dated to the early 1980s, with the publication of some groundbreaking books … and the launching of scholarly journals devoted to this topic … Prior to that time, no more than a handful of books on some of the major new varieties of English had been published, for example on English in Australia and New Zealand. (Schneider 2003: 234)

In the context of NZE, Gordon et al. (2004: 7) agree, noting that "the 1980s marked a considerable change in the study of New Zealand English with both overseas interest … and local research". To date, NZE has enjoyed thorough coverage of many of its areas. Starting in 1988, the *New Zealand English Journal* published bibliographies of writings on NZE at six-year intervals, and in 2000, Tony Deverson noted that "[i]n the past six years the number of items listed has increased by more than half, from 275 to 426, striking testimony to the continued growth and maturity of New Zealand English studies at the end of the twentieth century" (Deverson 2000: 18). In the 2006 update to the bibliography, John Macalister wrote that "[t]he growth has been just as spectacular this time – the list now contains 585 items – and once again it illustrates the amount of work being done in the description of New Zealand English and with the use of data from New Zealand English" (Macalister 2006a: 48). (It can be added that the first of the series of bibliographies, in 1988, listed 159 publications, with around half of those appearing in the period 1980–1987 (Gordon et al. 2004: 7).)

[31] The terms 'central' and 'non-central' are used here rather loosely, and are not to be confused with the terms as used in the categorisations in Mair's (2013) world system. The 'central' varieties here refer to British and American English, both of which have global influence, are generally seen as the 'reference varieties' among Englishes (Schneider 2007: 251), and have long been the subjects of a wide variety of research.

Topics of a sociolinguistic nature have been well researched, and work has covered areas such as attitudes to the dialect, gender and social class differences, Māori English, and the origins of the dialect. Of the areas more central to linguistics, it is clear from the 2006 bibliography that phonology was the field that had attracted the most research attention, and around that time, Hay et al. (2008: 47) observed that "[m]uch more research has been done on the ways that New Zealanders pronounce words than on the ways in which they organise them into sentences." Although there has been no update to the bibliography since 2006 (Macalister, pc.), a look through the available literature makes it clear that phonology has continued to dominate NZE research in the years since then. Already a popular topic in the late twentieth century, the discovery of the ONZE recordings in the 1990s served only to place the New Zealand accent more firmly on the world stage, encouraging a new phase of work in that area.

NZE lexis has also proven to be a rich area for research. As noted in section 2.2.4 above, Māori language contact led to a flow of borrowings that began almost as soon as Captain Cook first stepped ashore in 1769. The flow abated to a trickle for a period during the late nineteenth and early twentieth centuries but gained a new lease of life with the reassessment and renewed appreciation of the Māori language and culture in the latter decades of the twentieth century – see section 2.2.4. It is often pointed out that, since many of the accent-related features of NZE are shared by other varieties,[32] it is the vocabulary that makes it unique among the world's dialects of English (see Macalister fc. for a recent account of aspects of the development of the New Zealand lexicon).

The area of grammar has received attention, but, as noted above, due to the pre-1980s assumption that there was little of interest to be found, there was little scholarly inquiry into it. Unlike the accent, "the rules for standard written English in New Zealand are based on the syntax of standard British English with very few exceptions" (Gordon 2012: 323). Laurie Bauer is often seen as the pioneer of the study of NZE grammar, with his key publications of the late 1980s paving the way for subsequent work (Bauer 1987; 1988; 1989a; 1989b; 1989c). One important example of that work is Marianne Hundt's (1998) corpus-based study, which examined a wide range of morphosyntactic variables in comparison to their usage in three other varieties of English. Although now 25 years old, it remains the definitive work on the topic.

[32] Well-known NZE phonological features such as the NEAR-SQUARE merger (Gordon and Maclagan 2001), the raised short front vowels, centralised KIT, and, most certainly, /l/ vocalisation, can be heard elsewhere in the English-speaking world (Wells 1982).

Since then, reasonably detailed chapters on aspects of the grammar have appeared in thematic volumes (Quinn 2000; Hay et al. 2008; Hundt et al. 2008), with at least one volume dedicated to the study of grammatical issues in NZE and AusE (Peters et al. 2009). NZE is also occasionally included alongside several other varieties in larger-scale, article-length comparative studies of standard grammatical features such as the subjunctive and the dative alternation (see, for example, Hundt 2018; Szmrecsanyi et al. 2017).

There are thought to be very few grammatical structures that are used in New Zealand and nowhere else (Gordon 2012: 328). Rather, the difference lies for the most part in the frequencies with which NZE speakers use the available structures (Hay et al. 2008: 63). Included in the list of (standard and non-standard) grammatical features known to be relevant to, and in some cases favoured in, NZE are: variation in the use of past tense and past participle forms of verbs such as *ring*, *sing*; variation between the regular and irregular past tense and past participle forms of verbs such as *burn*, *dream*, *smell*; use of the auxiliary *have*; agreement patterns with some collective nouns; transitive use of the verb *farewell*, as well as non-transitive use of the verb *screen* (see Quinn 2000; Hundt 1998; Hay et al. 2008; Hay and Schreier 2004 for comprehensive discussions of these issues and more).

Thus, the ways in which general grammatical and morphological features behave in NZE are reasonably well researched and understood. As regards the area of complementation, however, it can be said that the number of previous studies is relatively modest. The cases of *farewell* and *screen*, mentioned above, (along with the often-cited examples of verbs such as *protest, appeal*) certainly fall within the scope of complementation, as they involve the selection of the phrasal or clausal elements by the head. Alongside these well-known examples, studies have focused on other complementation-related features such as verbal coda deletion with *as far as* constructions (Britain 2000), complementation pattern frequencies of verbs such as *appeal, protest, progress, battle, meet, consult, congratulate, fill, write* (Bauer 2001), complementation pattern frequencies of *help, prevent, stop, begin* and *start* (Mair 2009b), the zero-complementizer/*that* variation (Kearns 2009), and the effect of animacy on the dative alternation with *give* (Bresnan and Hay 2008). Many of these studies included a cross-varietal perspective, offering a comparison of NZE with other varieties.

The present dissertation seeks to build on the existing body of research in this area with the four publications, each of which examines a different verbal predicate or set of predicates, and a specific pattern and its permissible set of matrix verbs. The justification behind the project lies in the fact that NZE is a relatively young

variety, the result of colonial expansion and the transplantation of English around the world, and a recent mixture of native English dialects with the addition of a good amount of lexical input from the typologically unrelated language of Māori. Schneider (2007: 86), in his postulation of the five-phase stabilisation period of post-colonial English varieties, states the following:

> It should not come as a surprise, therefore, that grammatical innovations … typically start out where the regular meets the chaotic, i.e. at the intersection of grammar and lexis. Therefore, many of the characteristic innovations of [post-colonial Englishes] can be located at this boundary; they concern the co-occurrence potential of certain words with other words or specific structures. A classic example is the complementation patterns which verbs and also adjectives typically enter: in new varieties, in the process of structural nativization, verbs begin to allow and later prefer new structures to complement them and build a complete sentence. Well-known examples include the New Zealand use of *to farewell somebody* … the variability of the prepositions introducing complements of the adjective *different* (known to vary regionally between *from, than* and *to*)

According to Schneider's model, NZE is several decades past the structural nativization stage (Schneider 2007: 129*ff*), but it has been pointed out that the cross-varietal divergence of grammatical patterns is a slow process and is often detectable only during its process in shifting frequencies of usage (Mair 2002: 108–9). It has also been noted that "non-categorical (i.e. probabilistic) grammatical preferences can change significantly in a period of as little as 30 years" (Hundt and Szmrecsanyi 2012: 244), so new complement patterns do not exactly spring up overnight, and ground-breaking discoveries are not characteristic of this type of research, but change can certainly be detected within one generation. In support of this point, a recent study of agreement patterns with collective nouns showed that the preferences of some nouns had changed significantly in NZE over a relatively short period of time (Rickman 2018). By extension, it is reasonable to expect that the complementation preferences of verbs may also manifest some observable change within a similar time frame. Because relatively little systematic, large-scale research has been done in the area of complementation in NZE, the present dissertation aims to highlight changes in progress through observation of patterns of usage.

To close this chapter, the following quote brings together many of the points made in this section, and neatly summarises the place of NZE within the world's dialects of English as it stood in the opening years of the twenty-first century:

> New Zealand English is the dark horse of world English regional dialectology. It has long been neglected, mentioned only in passing as part of a treatment of Australian English … or assumed by outsiders to be identical with it in all salient

respects. During the 1980s, however, this state of affairs began to change, with several studies focusing directly on the variety, and taking into account the unique features of the New Zealand sociolinguistic situation. The results of this interest suggest that there is a great deal that the study of New Zealand English can contribute to our understanding of linguistic variation and change, and – more excitingly – that some of its most distinctive developments have yet to take place. (Crystal 2003: 354, reference omitted)

3 DATA AND METHODOLOGY

This chapter discusses aspects of the corpus data and methodological approaches used in the four publications.

3.1 Corpus linguistics and newspaper corpora

The four publications that comprise this dissertation are based almost exclusively on the empirical evidence gleaned from newspaper corpora. Given the amount of work in recent decades that takes this type of approach, the use of corpora as evidence in linguistic analysis these days requires little justification. Despite the fact that the syntactic perspective adopted here is of a tradition that is well-known for not using corpus evidence, I can see the value in the application of authentic language data to questions of language variation and change, and in principle I agree with Meyer (2002: 4), who believes that "it is a mistake to assume that the analysis of corpora has nothing to offer to generative theory in particular or to theorizing about language in general." A great deal has changed in the years since corpus-based methodology was just emerging and beginning to polarize attitudes in the western linguistic world (see Tognini-Bonelli 2001: 47*ff* for a summary, and Leech 1968 for an early discussion of the different types of evidence), and it is my feeling that little space needs to be taken up here in defending the use of corpus methods.

As for the reliance on newspaper data, however, a few words need to be said. It has now become widely accepted as one of the main genres lending itself to corpus compilation, with one of the most important and influential studies on NZE morphosyntax to date based heavily on newspaper corpus data (Hundt 1998). On the creation of a large newspaper corpus of Norwegian, its compilers commented that:

> Web newspapers lend themselves to use in a corpus for a number of reasons, such as their ease of access, their large and regular text production, their wide range of topics, their systematicity and coherence in text categorisation, their coverage of local, regional, national and global events … Unlike many other sections of the World Wide Web, newspapers contain journalistic text which is professionally written and edited.

> This entails a certain level of quality and standardisation in terms of spelling, language use, genre conventions and other features. (Andersen and Hofland 2012: 3)

The point made by Andersen and Hofland regarding ease of access is an important one, and the widespread availability of data in electronic form – undoubtedly even more widespread now than a decade ago when that was written – means that enormous amounts of data can be obtained with relatively little effort, and, crucially, cost, a factor which often makes it the only real option for studies taking a large-scale approach.

This cheap and easy access does come with certain limitations, however, the main ones being the lack of representativeness inherent in the use of only one text type, and the subsequent restrictions on generalizability of results. In the late 1990s, Hundt noted that "[t]he restriction in terms of text types - i.e. to journalistic prose - poses a more obvious problem as far as possible generalizations of the results are concerned. Journalistic prose is a heavily edited type of language." (Hundt 1998: 140). Admittedly then, there are some drawbacks to be kept in mind when working with news data, but in terms of its connection with the language on the street, the newspaper genre has been described as an "agile" one (Hundt and Mair 1998). This is due to its tendency to reflect the linguistic innovations of the community more quickly and readily than other written genres, such as academic prose – an "uptight" genre at the opposite end of the spectrum. The value of newspaper data as a resource for corpus analysis therefore seems clear.

A further fact to be borne in mind – a fact that is particularly relevant when comparing language varieties – is that there is often an element of the unknown in the authorship of newspaper articles. Journalists often write articles that are shared among different subsidiaries of a larger media corporation and altered only slightly for publication in another paper, and, naturally in today's world, people simply move around the world. New Zealand has been shown to have a high emigration rate among the native-born population (Carey 2019), with the diaspora more than offset by the number of immigrants arriving in the country (www.stats.govt.nz). This means that, while the chances are very good, there is still no real guarantee that an article published in a New Zealand newspaper was written by a native speaker of NZE. Clearly, the same applies to most countries, but the issue would seem to be most salient in English-speaking countries, and those that are seen as attractive destinations – and New Zealand has certainly always been one of those.

3.2 Two methodological approaches

Two main methodological approaches to corpus data extraction in complementation research can be identified: the head-based approach, and the pattern-based approach (Kaunisto and Rudanko 2019; Ruohonen and Rudanko 2021). The head-based approach is associated with research that seeks to focus on one complement-taking item – the head – and study the variation in the types of complements that the head can select. Conversely, the pattern-based approach is appropriate for work that focuses on one complement type – which can be termed the 'pattern' – and seeks to find and study the range of predicates that are available to select it. The publications in the present dissertation has employed both methodologies. Publications I, III and IV are within the head-based approach. Each study identifies one or more verbal matrix predicates and finds and analyses the non-finite complements that result from the corpus searches. Publication II falls within the pattern-based method; it takes as its starting point the transitive *into -ing* pattern and catalogues and analyses all the matrix verbs that are available to select it.

In terms of devising an effective search string for each of the above approaches, a string for the head-based approach would need to specify the head item itself, and the pattern-based approach would leave the head slot unspecified, and spell out the details of the pattern itself. Details on the corpus search procedures for each study of this dissertation are of course given in each publication, and a good demonstration of the pattern-based approach is to be found in Kim and Davies (2015).

3.3 The Corpus of New Zealand Newspaper English (CNZNE)

The drawbacks of using newspaper data in corpus design were discussed above, and the main reason for my reliance on it is simply that it was the only real option for compiling a sufficiently large, diachronic corpus of NZE. Other corpora available at the time included ICE-NZ, from the International Corpus of English family (https://www.ice-corpora.uzh.ch/en.html), and the Wellington Corpus of Written New Zealand English (Bauer 1994a). These corpora are of course both excellent for certain projects but, at one million words each, they were too small for the present project. While a more balanced and representative corpus design would have been ideal and would have allowed for a wider generalizability of results, the data was simply not available.

The early stage of the present research project was dedicated to the compilation of the corpus, with the main requirements being size – it needed to be big enough to provide data on less commonly used predicates – and as wide a diachronic perspective as possible. The answer was found in the archives of the set of newspapers published under the Fairfax Media group, which, at the time, was available for a fee through the commercial archiving agency The Knowledge Basket (www.knowledge-basket.co.nz).[33] The digitized archive of Fairfax (and other media groups') material reaches back to 1995. This data was used to create a 100 million word diachronic corpus to be used for the publications of the present dissertation. A detailed description of the corpus compilation process is given in Publication I but a brief description is also in order here. The corpus is diachronic, with the space of 15 years between the two sections of the corpus. The first section, at approximately 42.5 million words, contains data from 12 newspapers published during the period 1995–1998; the second section, at approximately 58.5 million words, contains data from 10 newspapers published during the period 2010–2012.[34] Efforts were made to select papers from different parts of the country, and to offer a similar balance between broadsheet and tabloid as is found in the press section of the British National Corpus. With a similar makeup and time-frame, the first section of the CNZNE is comparable to the BNC's press section.

The corpus is POS tagged, using CLAWS4 (Garside and Smith 1997), and AntConc concordancing software was used to carry out the searches of the CNZNE (https://www.laurenceanthony.net/software/antconc/). The data was then handled in Microsoft Excel.

3.4 The News on the Web Corpus (NOW)

Between the writing of the second and the third publications, in May 2016, the News on the Web (NOW) corpus was released and became another viable source of data for the present dissertation (Davies 2017). The NOW corpus is updated regularly and allows comparison between twenty varieties of English. Importantly for this project, NZE, BrE and AmE are all very well represented in NOW. As noted by Davies (2017), corpora tend to go stale rather quickly, and the value of producing

[33] I am grateful to the GlobE consortium and Tampere University English department, who jointly covered the costs of our one-year license to the archive.

[34] The more recent articles of this dissertation were written 9–10 years after the last period of the corpus, but my access to the Fairfax archive was short-term and after 2013 the corpus could not be updated. This contributed to the decision to use data from the NOW corpus for the later articles.

research based on very recent data was weighed against the value of continued use of the CNZNE data, and the former was seen as more important. As stated in footnote 34, the CNZNE could not be updated due to the short-term access. NOW contains data from 2010 to the present day and so it lacks the slightly deeper diachronic perspective of the CNZNE, but because of its obvious advantages, NOW was used as the data source for Publications III and IV.[35] The english-corpora.org online interface was used for all searches involving the NOW data, and the data was then handled in Microsoft Excel.

3.5 Other corpora

Publication I is the only study to use data from corpora other than the two discussed above. The additional corpora used in Publication I are the British National Corpus (BNC), the Corpus of Contemporary American English (COCA), and the Corpus of Late Modern English Texts, version 3.0 (CLMET3.0).

[35] Like the CNZNE, NOW is a corpus of newspaper language, but NOW draws its data from a much wider range of newspapers, as well as magazines and other online publications. It contains all of the newspapers that the CNZNE does, plus material from over 150 additional sources, meaning that it cannot simply be used as an extension of the CNZNE. A complete list of sources can be found at https://www.english-corpora.org/now/

4 RESULTS AND DISCUSSION

This section addresses the research questions that were listed in section 1.2. The questions aim to combine the results of the publications to provide wider discussion of selected issues.

4.1 Research question 1

"Does NZE show significantly different complementation patterns of the selected predicates to British and American Englishes?"

Research question 1 investigates the difference in the use of the selected complement patterns between NZE, and BrE and AmE. One of the overall aims of the present dissertation is to shed some light on the ways in which the grammar of native-speaker Englishes may be changing – perhaps converging in some areas or perhaps diverging – and uncovering differences at the level of complementation patterns is the justification for this research question.

"Significantly different" here means statistical significance. The intersection of statistics and linguistics is a notoriously difficult area (Wallis 2020), and the more complex statistical tests – both the reasons for their application and the interpretation of the results – are not well understood by many linguists, the present author included. The emphasis on the importance of statistical foundations for the claims made in linguistic research today, however, makes it almost mandatory to include some kind of mathematical justification in a publication. The latter two publications of this series do include some basic statistical information, but the former two do not, so in order to answer the present research question comprehensively, where statistical tests were not applied in the original publication, they will be applied here and the results reported in this section.

I restrict the statistical analyses to simple two-tailed binomial tests using Fisher's exact test. In cases where the score produced by the test is greater than the significance threshold of 5%, and the result therefore deemed significant, it means that the variables are likely to be dependent on one another, and not the result of

pure chance. Fisher's test produces a p-value directly and rather than reproducing the precise value here, I have rounded the p-values up the nearest decimal place and reported them in the form $p < X$.

The publications that include a cross-varietal aspect are Publications I and III, and the results of these two studies are discussed below.

4.1.1 Verbs of prevention

Publication I analyses two variants of one complementation pattern which is most commonly associated with verbs of prevention, and the study uses 14 matrix verbs of this type. The data are taken from three corpora: the (then) newly compiled CNZNE (42.6 m. + 58.5 m. words), the BNC newspaper sub-section (9.4 m. words), and the COCA news sub-section (20.4 m. words).

The pattern is V + NP + (*from*) + *-ing*, and in terms of its control properties, it is one of object control. The preposition *from* is, structurally speaking, an optional element in the pattern with most of these verbs when used in the active voice.[36] AmE is known to strongly favour the inclusion of *from*, while BrE is known to allow, but not require, its omission. Previous research (e.g. Mair 2009b) as well as the results of Publication I show that NZE usage is broadly in line with BrE in terms of omission of *from*. Examples (10a–d) below, taken from the NZE corpus, show the pattern with the two prototypical verbs of prevention, *prevent* and *stop*. (10a) and (10c) include *from*, and (10b) and (10d) omit it.

(10) a. Australian authorities have intervened to prevent a female stalker from contacting Dannii Minogue. (CNZNE, *Manawatu Standard*, 2011)
b. Modesty prevents me commenting on my own intelligence. (CNZNE, *Southland Times*, 2012)
c. This not only contributes flavour and texture, but it stops the bread from going soggy with the meat juices. (CNZNE, *Dominion Post*, 2011)
d. I fiddle with the TV controls for half an hour to try to stop everyone looking stretched sideways, then give up. (CNZNE, *Sunday Star Times*, 2012)

[36] *Dissuade* and *restrain* seem to be exceptions, with no *from*-less examples found in the data used for the study. A search of the NOW corpus, however, suggests that *dissuade* allows omission of *from* to some extent, while *restrain* appears to completely disallow it.

The main focus of Publication I, therefore, is on the presence or omission of *from*, and the 14 verbs are: *ban*, *bar*, *block*, *deter*, *discourage*, *dissuade*, *forbid*, *hinder*, *prevent*, *prohibit*, *restrain*, *save*, *spare* and *stop*.

As noted above, no statistical data is given in Publication I, and the relevant figures are therefore supplied here. As also noted above, the test is Fisher's exact test.[37] The cutoff point for statistical significance was set at the standard level of 5% ($p < 0.05$) (Wallis 2020: 35).

Table 2 (a reproduction of Table 3 from Publication I) gives the data on the frequencies of the pattern with and without *from* in the three varieties in the period of the mid-1990s.

Table 2. Prevention verbs in NZE, BrE and AmE

	cnzne	**bnc**	**coca**
ban NP *-ing*	0.3 (13)	0.1 (1)	-
ban NP *from -ing*	2.6 (110)	5.1 (44)	1.5 (30)
bar NP *-ing*	0.1 (6)	-	0.0 (1)
bar NP *from -ing*	1.2 (51)	2.3 (22)	5.4 (110)
block NP *-ing*	0.2 (10)	0.2 (2)	-
block NP *from -ing*	0.3 (14)	0.1 (1)	1.3 (26)
deter NP *-ing*	0.3 (11)	0.1 (1)	-
deter NP *from -ing*	2.5 (107)	2.4 (23)	0.8 (17)
discourage NP *-ing*	0.4 (18)	0.1 (1)	0.0 (1)
discourage NP *from -ing*	2.1 (90)	2.1 (20)	4.6 (94)
dissuade NP -ing	-	-	-
dissuade NP *from -ing*	0.8 (36)	0.6 (6)	0.7 (15)
forbid NP *-ing*	0.2 (8)	-	0.0 (1)
forbid NP *from -ing*	0.5 (21)	1.1 (10)	1.9 (38)
hinder NP *-ing*	0.1 (3)	0.2 (2)	-
hinder NP *from -ing*	0.0 (1)	0.1 (1)	0.2 (5)
prevent NP *-ing*	21.2 (904)	29.5 (278)	0.3 (2)
prevent NP *from -ing*	19.5 (832)	23.5 (221)	30.5 (194)

[37] Statistical data was obtained using the calculator available at http://corpora.lancs.ac.uk/sigtest/.

prohibit NP -ing	0.4 (19)	-	-
prohibit NP *from -ing*	0.9 (39)	0.7 (7)	7.9 (161)
restrain NP -ing	-	-	-
restrain NP *from -ing*	0.4 (16)	0.6 (6)	0.3 (6)
save NP *-ing*	0.9 (37)	1.1 (10)	-
save NP *from -ing*	2.2 (92)	3.4 (32)	3.0 (10)
spare NP *-ing*	0.0 (1)	-	-
spare NP *from -ing*	0.1 (4)	0.2 (2)	0.2 (4)
stop NP *-ing*	31.1 (1325)	39.9 (376)	0.5 (1)
stop NP *from -ing*	13.5 (575)	8.0 (75)	13.6 (28)

Numbers outside brackets represent normalised frequencies per one million words; bracketed numbers represent raw frequencies.

Looking first at the comparison of the NZE and AmE data, the significance test shows significant results between the two varieties in the following eight verbs:

bar ($p < 0.01$); *block* ($p < 0.0001$); *discourage* ($p < 0.0001$); *forbid* ($p < 0.01$); *hinder* ($p < 0.05$); *prevent* ($p < 0.0001$); *prohibit* ($p < 0.0001$); *stop* ($p < 0.0001$)

Given the well-known tendency for AmE users to include the preposition in the structure, these differences are to be expected, and the only surprise is that there are not significant differences between NZE and AmE for all of the verbs – e.g. *ban*, *deter*, *save* all have reasonably high numbers in three of the four cells, but the differences are non-significant.

The frequencies between NZE and BrE, on the other hand, are for the most part non-significant. The exception is the verb *stop*, where a highly significant difference was identified ($p < 0.0001$). The difference is due to the NZE data containing a greater percentage of the pattern with *from* than BrE. In terms of percentages, 30.3% of the NZE data comprises the pattern with *from*, while the percentage in BrE is a much lower 16.6%. The *prevent* frequencies in the NZE and BrE datasets are, by contrast, non-significant, with 47.9% of the NZE data comprising the pattern with *from*, as compared to 44.3% of the BrE data. (In AmE, for the sake of comparison, the pattern contains *from* in 99% and 97% of the cases with *prevent* and *stop*, respectively.)

Publication I then goes on to investigate diachronic change in NZE in 12 of the 14 verbs[38] across the approximately 15-year timespan between CNZNE subsections (1995-98 to 2010-12). Table 3 (a reproduction of Table 4 from Publication I) shows the frequencies in NZE in the use of the pattern with the two main verbs *prevent* and *stop* across this time period.

Table 3. *Prevent* and *stop* in NZE (1995–98 and 2010–12)

	1995–98	**2010–12**
prevent NP *-ing*	21.2 (904)	15.8 (925)
prevent NP *from -ing*	19.5 (832)	20.3 (1190)
stop NP *-ing*	31.1 (1325)	27.1 (1585)
stop NP *from -ing*	13.5 (575)	16.2 (948)

Numbers outside brackets represent normalised frequencies per one million words; bracketed numbers represent raw frequencies.

The observed frequencies in Table 3 point to a drop in the usage of the *NP -ing* complement and a concomitant increase in the *NP from -ing* complement across the short time period. Fisher's test shows a high degree of significance in both cases: *prevent* ($p < 0.0001$); *stop* ($p < 0.0001$). Four other verbs out of the remaining 10 also showed significance in a shift in the same direction, although to a lesser degree: *bar* ($p < 0.01$); *block* ($p < 0.01$); *discourage* ($p < 0.01$); *prohibit* ($p < 0.01$).

The data available at the time the study was carried out did not allow any comparable measure of change in BrE or AmE usage, but a separate study on *prevent* using different corpora carried out at around the same time shows that BrE appears to be further decreasing its use of the version with *from* (Collins 2015: 33), and moving further away from AmE. With this information unavailable to me at the time, my suggestion was that the NZE shift could be seen as one pattern of change within the various possible scenarios under the 'colonial lag' heading, one which sees the change in NZE driven by twentieth century American global influence.

[38] As was stated in footnote 36, *dissuade* and *restrain* were not found with the *from*-less complement in any of the corpora. They were therefore excluded from further analysis in Publication I.

4.1.2 *Promise* and *threaten*

Publication III is a study of the semantically related matrix verbs *promise* and *threaten* selecting a *to* infinitive complement clause. In terms of argument structure, both verbs are known to function as matrix predicates in both subject control and subject-to-subject raising structures, the subject control type being the older, more established use, the raising type having developed via grammaticalization and its use fully attested by the eighteenth century (Traugott 1996). The study uses data from the NOW corpus to give quantitative and qualitative accounts of the verbs' usage in both structures in NZE, BrE and AmE. Examples of each verb in both syntactic roles are given in (11a–d), with subject control in (11a–b) and raising in (11c–d).

(11) a. The company promises to refund customers if a ticket is rejected at an event. (19-02-04 NZ)
b. He's also been left frustrated by Customs after they threatened to put West Coast Brewery into liquidation over unpaid debts. (21-03-18 NZ)
c. This show promises to be hugely entertaining for both young and old. (17-05-31 NZ)
d. … ongoing wet conditions have threatened to give weeds a head start as spring growth starts to kick off. (17-11-26 NZ)

The study includes basic statistical information on differences between the three varieties. The test used in the paper is the Chi-squared test, and, although it does not change the level of significance, the numbers were processed again here using Fisher's exact test, since that test was used in section 4.1.1.

The usage aspect applicable to the present research question concerns the difference in the ratios of control to raising structures.[39] The results for the matrix verb *promise* showed a significant difference between the NZE and BrE uses ($p < 0.05$). In terms of percentages, in the NZE dataset the subject control structure occupies 72% of the total, and in BrE 64.5%. Differences between the ways NZE and AmE speakers use the verbs, however, are not statistically significant; they are, in fact, from a quantitative point of view, very similar: in AmE the subject control structure occupies 73%, as compared to NZE 72%.

[39] The results of this study are presented in terms of ratios due to the methodology, described in more detail in the publication, whereby the datasets were comprised of a random selection of 500 tokens for each variant in each regional variety. Only the percentages are given in this section as a general illustration, and the reader is referred to the publication for the more detailed results.

The results for the matrix verb *threaten*, on the other hand, showed no significant differences between any of the three varieties. From a percentage point of view, the similarity was striking: the control structure occupies 62.2% in NZE, 62.3% in AmE, and 61.8% in BrE.

4.1.3 Summary of research question 1

The combined results of Publications I and III clearly show that there are significant quantitative differences in the ways that both of the reference varieties and NZE use non-finite complements with some groups of verbs. The major differences in the 1990s data between NZE and AmE in the use of *from* with *prevent*-type verbs was expected and came as no surprise. On the other hand, the major difference in that same time period between NZE and BrE in the case of *stop*, along with the documented shift in NZE away from the *from*-less variant over the following 15-year period with *stop*, *prevent* and several other verbs of this class, is a strong suggestion that NZE is not following BrE in its increased use of the *from*-less variant – an increase which is noted in Collins (2015); see also the data in Mair (2019). Without comparable diachronic data on the reference varieties at the time of writing though, Publication I was only able to provide conjecture and it was suggested that NZE could be starting to favour the more explicit pattern under the influence of AmE. This suggestion is in fact in line with the results of work on *prevent* in Australian English by Collins (2015), which was published shortly after the work on Publication I was finalised, but with smaller datasets comprising fiction and news material covering the nineteenth and twentieth centuries. Collins found that "AusE diverges from BrE, its steadily decreasing endorsement of the *from*-less version – possibly influenced by the American dispreference for this variant – contrasting with the continuing British endorsement of it" (*ibid* 33).

Seen alongside the prevention verbs, the results from the study on *promise* add further weight to the hypothesis that NZE and BrE may be parting company in some areas of verb complementation. The fact that NZE and AmE are using the control and raising structures with *promise* in such similar proportions suggests more of a parallel scenario in those two varieties, rather than a shift towards AmE, as was seen in the prevention verbs. The significantly higher use made of the raising pattern in BrE makes this another area for future observation.

4.2 Research question 2

Is there any observable innovation in NZE complementation?

4.2.1 The transitive *into -ing* pattern

Publication II provides the clearest example of NZE innovation in the area of complementation. The innovation in question is not part of the complement itself in this case, but should rather be seen as lexico-grammatical innovation at the head of the pattern.

The paper is a study of the transitive *into -ing* pattern in NZE based on data from the CNZNE alongside results from previous studies. It is the only paper in the series to take a pattern-based approach, with the aim of identifying and cataloguing all the matrix verbs in the corpus that are available in NZE to select that pattern. The pattern is V + NP + *into* + *-ing*, and in terms of control properties, it is an object control structure. The pattern is shown in (12a–b) below, with two typical matrix verbs, *fool* and *force*.

(12) a. The emails were designed to fool people into believing they had come from Kiwibank, the National Bank and ASB. (CNZNE, *Daily News*, 2010)
b. It probably forced us into making a few decisions we probably shouldn't have made. (CNZNE, *The Press*, 2012)

The study makes a comparison of the findings emerging from the NZE data to those of earlier studies (Bridgeman et al. 1965; Davies 2012; Francis et al. 1996; Kim and Davies 2015; Rudanko 2005), the results indicating that NZE shows lexico-grammatical innovation in 19 of the matrix verbs used in the pattern. Three of the 19 innovative NZE verbs are classed as "radically innovative", to use Rudanko's (2015: 81*ff*) classification, i.e., they have not been documented in previous surveys, nor are they listed as verbs in the *Oxford English Dictionary* (*OED*). The 19 innovative verbs uncovered by the study are *buoy, bustle, court, fast-track, force-feed, fuel, hard-talk, heavy, irk, link, mobilise, monster, peer pressure, pressure-cook, re-energise, reinvigorate, stir up, turbo-shock* and *wrongfoot*. Of this group, the three radically innovative ones are *hard-talk*, *peer pressure* and *turbo-shock*, all of which are from the semantic class of verbs influencing causation "by means of force or pressure", under Rudanko's (2005) typology, a class that has been prominent as the source of innovative verbs (*ibid.*).

With 19 previously undocumented matrix verbs coming to light in this study, it can be said that NZE does indeed show innovation. It has to be added, however, that the *into -ing* pattern is such a well-known vehicle for lexico-grammatical innovation that it would have been unusual to find no new matrix verbs in the NZE data. Indeed, the study set out with the hypothesis "that some kind of innovative usage will be found".

The pattern has seen a steady rise in frequency in recent decades and the number of different matrix verbs occurring in the pattern is continually increasing, as many recent studies have attested. The Construction Grammar approach (Goldberg 1995), which is the framework often adopted for studies of this pattern as it is the one which helps to explain the most in the case of such a productive complement type, places an emphasis on the semantic contribution of the pattern itself – its constructional meaning. As this constructional meaning strengthens over time, it frees up the restrictions on the types of matrix verbs that can occur in the pattern, giving speakers more freedom to use it in increasingly creative ways (for a recent discussion see Flach 2021; and for a control-based perspective, see Landau 2021).

In such cases it is likely that many innovative verbs used in this environment are simply one-off nonce formations, appearing in print one day thanks to a journalist's flash of creative inspiration, gone the next and never to surface again. This may be the case with the NZE verb *turbo-shock*, which does not seem to be found in this pattern (or any other, for that matter) in a recent search of the NOW corpus.[40] Searches of NOW also reveal some of the other innovative NZE verbs appearing in the *into -ing* pattern in other varieties, e.g. *buoy* is used in a similar way in a handful of tokens in Irish English and AmE; *court* is found in AusE and BrE; *wrongfoot* is found in AusE and AmE; *monster* appears in the British section of NOW, and *heavy*, listed as a verb in the *New Zealand Oxford Dictionary,*[41] also shows up in tokens in the Australian section of NOW. Crucially though, the searches carried out for this additional information did not turn up any examples of any of these verbs predating the NZE examples cited in Publication II, so it seems that it is still reasonable to call them NZE innovations.

The pattern was, until relatively recently, generally portrayed in the literature as having an inherently negative semantic prosody and was notable for hosting matrix verbs of negative causation (Wierzbicka 1998: 125). Studies have since gone on to uncover *into -ing* matrix verbs with a neutral or even positive semantic orientation

[40] The additional searches were carried out in August 2023.

[41] *Heavy* was also used as a verb in Old and Middle English but is now marked as obsolete (*OED* sv. *heavy*), while the *NZOD* lists it as verb in current NZE.

(e.g. Rudanko 2006b; Davies and Kim 2019), and lower predicates encoding positive actions and events. Rudanko (2006b) suggests that BrE may be leading the way ahead in the use of manner-neutral matrix verbs, while Davies and Kim (2019), with the benefit of data from the mega-corpora of more recent years, find that neutral verbs have been associated with the pattern in AmE since the early 1800s, but they have seen "little or no increase" (*ibid* 43) over time in that variety.[42] In the case of matrix verbs with a positive semantic prosody, on the other hand, Davies and Kim note an increase in frequency, with examples such as *charm, jolly, enchant, urge, ease, motivate* and *romance* (*ibid* 44–46).

As for the semantic orientation of the innovative NZE verbs, a good many of them are unambiguously negative, e.g. *bustle, force-feed, hard-talk, heavy, irk, monster, peer pressure, pressure-cook* and *wrongfoot*, but they are by no means all of that type. Some can be seen as effecting causation in a neutral or, arguably, even a positive manner. This is illustrated in (13a–e), with the verbs *buoy, fuel*, *mobilise*, *re-energise* and *reinvigorate*.

(13) a. It was an effort which buoyed rider David Walsh into thinking he could win his second Wellington Cup. (CNZNE, *Evening Post* 1996)
b. … the story goes that the artist fell in love with a younger woman, but she did not fall in love with him. That fuelled him into painting his study of the female body with a grotesque beast lying on her stomach. (CNZNE, *Nelson Mail* 2010)
c. … it was a matter of mobilising people into putting their concerns in writing. (CNZNE, *Southland Times* 1997)
d. They are designed to revitalise and re-energise mature jobless people into wanting to get back into the work-force. (CNZNE, *Press* 2011)
e. I want us to reinvigorate our families into taking up their full responsibility for their own. (CNZNE, *Dominion Post* 2011)

Some of the verbs are synonymous with one another to some degree: *fuel, re-energise* and *reinvigorate*, with *reinvigorate* even forming part of the *OED* gloss of *re-energise*, and all of them are to some extent logical semantic extensions of the idea of *X provides Y with energy to perform a task*. It therefore seems possible to say that these verbs are more than simply neutral in their semantic prosody, but are in fact positive.

[42] Flach (2021) suggests that the apparently conflicting findings of Rudanko and Davies and Kim are the result of approaching the analysis on the basis of verb *types* versus *tokens*. She notes that "both approaches tap into the same phenomenon, they just look at two sides of the coin." (Flach 2021: 266).

The study carried out by Davies and Kim (2019) established that, while neutral verbs have been used in the pattern in AmE since at least the eighteenth century, verbs with a positive semantic prosody have not. They state that "it appears that there was always the possibility of extending this [use of the pattern with verbs other than negative ones - PR] to explicitly positive verbs, such as *motivate*, *enchant*, and *love*, but it is only in the last few decades that such an extension has actually taken place" (*ibid* 46).

The innovative, positive NZE usage certainly extends back across the last few decades – to the mid-1990s, in fact – and therefore it seems possible to suggest that NZE was extending the pattern to positive verbs at around the same time as AmE was. It will be recalled that the latter decades of the twentieth century were notable for the increase in American media influence on New Zealand television and radio, after all. The types of positive NZE verbs, however, are not the same as the positive AmE verbs cited by Davies and Kim, and NZE might therefore be seen as innovative with respect to the use of the *X provides Y with energy to perform a task* verb group.

4.2.2 *Promise* and *threaten*

Publication III provided results which, for the most part, saw NZE usage patterning either more like AmE (*promise*) or in remarkably similar ways to both AmE and BrE (*threaten*). One area where there may be a suggestion of NZE innovation, however, was seen with the raising predicate *threaten* when it is used in the progressive aspect. In terms of overall percentages of raising tokens which were progressives, NZE was only slightly ahead of AmE, with both being noticeably, but statistically-speaking non-significantly, ahead of BrE. In terms of the subject types of those progressives, however, the NZE data was shown to contain more animate subjects than the other two varieties. Examples are given in (14a–b).

(14) a. A hundred horsemen on mountain bikes are threatening to swoop down Te Mata like a wolf on the fold. (18-05-13 NZ)
b. Their talisman Joe Root scored 25 but took a stunning catch to send back Dasun Shanaka who, partnering Mathews, was threatening to take the game away from England. (16-03-26 NZ)

Previous research (Hundt 1998; Hundt and Szmrecsanyi 2012) indicates that the extension of the progressive construction in general to *in*animate subject types is an area in which NZE was at one stage innovative. In the narrower context of raising, however, it is thought that raising structures "tend not to occur with progressives, at least with animate subjects" (Traugott 1996: 194). The results given in Publication III, therefore, which show that the NZE news data has a relatively high number of progressives with the raising verb *threaten*, a good proportion of which (35.7%) have animate subjects, may be taken as an indication of further NZE innovation in the area of progressives.

4.2.3 Summary of research question 2

The results taken from Publications II and III show that there are grounds to claim that NZE is showing innovative tendencies. Lexico-grammatical innovation with the transitive *into -ing* pattern is, in general, very common, but here it is suggested that NZE speakers may have been among the earliest to use positively-oriented verbs in the pattern, and furthermore, they may be using a specific semantic sub-group of those verbs.

Recent research has not yet gone in-depth in the case of the raising verb *threaten*, especially not in NZE, but Publication III raises the possibility that there is more to be uncovered here. It has already been established that NZE was once an innovator with inanimate subjects in progressive structures, but it may be that the dialect is also leading the way ahead with animate subjects with raising verb progressives. The data gathered here on *threaten* will need to be compared with that of other raising verbs in order to offer a more complete picture.

4.3 Research question 3

How can the present research contribute to our current understanding of "colonial lag/innovation"?

To repeat a quote given in section 2.2.5, it was recently noted that "there is ample scope for further research on differential grammatical change in AusE and NZE and comparison with available evidence for British and American English. The likely picture to emerge from such studies is one that goes beyond the simple dichotomy

of colonial lag and innovation" (Hundt 2017: 302). Collins (2015: 39), too, believes that "the oft-invoked notions of 'colonial lag' […] and 'colonial innovation' […] need supplementation by other patterns of change". More recently, in a study on a recessive pronunciation feature in the isolated speech community of Tristan da Cunha, Schreier (2019: 59) provides clear support for colonial lag, but points out that such analyses can only operate at the level of individual features, and it is not possible to claim that the entire dialect is archaic on the basis of the discovery of a case of an individual archaic feature.

Thus, since the early days of discussion of the topic, scholars have recognised 1) the need for a refined typology of differential language change scenarios, the number of which in recent years has grown to at least six; and 2) the fact that individual features within a dialect are subject to complex pressures and rarely develop in unison, resulting in archaisms existing alongside innovations. While the present research cannot claim to do much to further the development of either of these views, several of the results of the present research are very much in line with both of them.

Publication I suggests that verbs of prevention may be shifting their preferences in different ways. Among other suggestions made in the publication, the two main verbs, *prevent* and *stop* initially showed signs of lag, but may now be in the process of following the AmE lead towards a more explicit prepositional variant. *Bar* was deemed a possible candidate for innovation; while *forbid* and *prohibit* were suggested as possible cases of colonial lag. It was emphasised in the publication, however, that the corpus data which would be needed to confirm these hypotheses was not available, and conclusive claims could not be made.

Publication III also provides some evidence of a scenario under which NZE is closely aligned with AmE in its division of *promise* between control and raising uses, with statistically significant differences between both NZE and AmE on the one hand and BrE on the other. Based on data taken from the genres of fiction, non-fiction, drama and letters spanning the Late Modern English period (i.e. CLMET3.0 data), Rickman and Rudanko (2023) showed that the use of *promise* as a raising verb in BrE steadily increased across the LModE period to the point where it occupied 13.4% of all relevant usage by 1920 (Rickman and Rudanko 2023: 268). Publication III, based on data from the NOW corpus, shows that the raising use of *promise* occupies 34.6% of all relevant usage in BrE today. Clearly, the two corpora are not comprised of the same text types, and data from more closely matching sources would be needed to prove it conclusively, but it nevertheless seems possible to state that the raising use of *promise* in BrE seems to have continued to increase throughout

the period between 1920 and the present day. The history of NZE and AmE usage, on the other hand, is less certain due to the lack of LModE data, but a likely hypothesis is that both varieties are also increasing their use of raising *promise* but levels are currently somewhat lower than BrE. The scenario could therefore be a possible case of colonial lag.

Somewhat more conclusive, though, is the case of colonial innovation seen in the NZE (and possibly AmE) use of the progressive with the raising predicate *threaten*, details of which were described in 4.2.2 above.

Taken together, the applicable results serve to reinforce the opinion, which now seems hard to dispute, that the mosaic of global English diffusion and change that still today tends to go by the archaic, but catchy, title of 'colonial lag' is complicated to say the least. In order to achieve conclusive results, the researcher needs to have access to sufficient amounts of corpus data of the right type, covering the right time periods – criteria which the data for the present research does not always manage to fulfill.

4.4 Research question 4

What insights are gained from the application of the Choice Principle to the verb *fear* in NZE?

Publication IV is an investigation into aspects of the two main non-finite complements of the matrix verb *fear* in NZE, using data from the NOW corpus. The two complement types are the *to* infinitive and the gerund, and in terms of its control properties, the pattern is one of subject control. Examples are given in (15a–b).

(15) a. The indomitable Queenslander has never feared to ask the hard question … (21-04-02 NZ)
b. A Mount Maunganui resident says he fears losing his flatmates and not being able to pay his mortgage. (21-11-25 NZ)

The main theoretical goal of the paper is to provide new testing grounds for the hypothesis known as the Choice Principle (Rudanko 2010; 2014; 2017; Rickman and Rudanko 2018; Ruohonen and Rudanko 2021). The Choice Principle is based on the observation that when a *to* infinitive and a gerundial clause co-exist as complement choices of certain control matrix verbs or adjectives, the *to* infinitive complement

tends to be selected when the lower clause subject[43] exercises agentivity/choice over the event encoded by the lower predicate, and the gerund tends to be selected when the event encoded by the lower predicate requires no agentivity/choice from its subject. The two contexts are labelled [+Choice] and [–Choice], respectively. The principle aligns itself with the theory of semantic roles (Fillmore 1968; Dowty 1991): the subject in a [+Choice] context is an Agent, and the subject in a [–Choice] context is not, often being a Patient or Undergoer instead. Typical examples of the two different contexts, with the appropriate complement as predicted by the Choice Principle, were given in (15a–b) above, where the lower clause predicate *ask* in (15a) requires an Agent, and the lower predicate *lose* in (15b) requires an Undergoer.

The principle joins the toolkit of syntactic and semantic generalisations that have been postulated over the years to help explain the variation between competing complement types (for other generalisations see Allerton 1988; Rohdenburg 1996; 2003; Vosberg 2003a; 2003b).

Earlier observers have noted a contrast in the agency of the subjects of *to* infinitives and gerunds (e.g. Wood 1956; Wierzbicka 1988: 33*ff*; Huddleston and Pullum 2002: 1243), but the Choice Principle and the related research represent the first attempt to home in on the contrast and to systematically test it using authentic English data. It has been applied to complement pairs of verbal and adjectival matrix heads in mainly BrE and AmE, in both present-day English and the English of recent centuries. Publication IV, however, is the first time it has been applied to NZE data, to the matrix verb *fear,* and to a pair of complements involving a non-prepositional, or bare gerund. While some matrix verbs have been previously tested, earlier studies have tended to focus more on matrix adjectives such as *afraid, scared, accustomed*, and *prone*, among others. The present work, therefore, represents fresh research in a number of ways.

Findings of the study highlight a significant tendency ($p < 0.001$) for the *to* infinitive to be selected in [+Choice] cases and the *-ing* clause in [–Choice] cases. This finding is in line with those of previous studies in this area, and it provides further empirical evidence for the link between agency/volition and the *to* infinitive. Furthermore, and probably more importantly, it shows that, with respect to the parameters of the Choice Principle, the bare gerund behaves in very similar ways to the previously-tested prepositional gerunds (i.e. *of -ing*, *to -ing*), and that the verb *fear*

[43] The subject of the lower clause is covert, but co-referential with the matrix clause subject.

behaves in similar ways to the semantically related adjectives *afraid*, *scared* and *terrified*.[44]

The study is intended as an investigation into the use of *fear* in NZE only, and as such it does not include data from any other variety, but it is reasonable to assume that NZE uses *fear* in broadly similar ways as other English varieties do. This assumption is based on the results of another recent study on the verb (Duffley and Fisher 2021), which provides information on AmE using data from Corpus of Contemporary American English (COCA). Aside from some discrepancies in overall frequencies, which are likely to be the result of the different methodological approaches of the two studies, the overall distribution of the two complement types is very similar.

A follow-up study providing a close comparison of several varieties would be the next step in determining more widespread tendencies, and Publication IV makes a valuable contribution in providing a starting point for future research of this type.

[44] That is, as regards the semantics of the Choice Principle. As regards the quantitative complement selection aspect, the *to* infinitive is greatly outnumbered by the bare gerund with *fear* in NZE – and in AmE; see Duffley and Fisher (2021).

5 CONCLUSION

The final chapter of this Introduction comprises sections dealing with a discussion of the strengths and weaknesses of the project as a whole, an outline of future avenues of research, and my concluding observations.

5.1 Strengths and weaknesses

The four publications analyse several different aspects of non-finite verb complementation, and in doing so they provide snapshots of various areas of the inner workings of the verb complementation system of NZE. The areas covered are *from* and *from*-less gerundial complements with the large group of prevention verbs, the wide range of matrix verbs available for the *into -ing* pattern, verbs that straddle the syntactic boundary between subject control and subject-to-subject raising, and the effect of subject agentivity on the selection of competing non-finite complements. The range of topics covered can therefore be seen as one of this project's strengths.

A further strength of the present project is the corpus data used in Publications I and II. The corpus was compiled specifically for this project, and it is unique in that it is a diachronic corpus of NZE press material covering the period 1995–2012. While the Fairfax archives were always there (but not freely available), a dedicated corpus of this type did not exist at the time of its compilation. The time period covered by the corpus can be seen as an important one for NZE, as it is during this stage that we may see evidence of "the gradual emergence of new dialects and sociolects within the new variety" (Schneider 2014: 12). Therefore the data used in the first two publications is applied to this type of research for the first time, and it offers new insights into NZE at an important stage of its development.

The corpus used for the latter two publications (the NOW corpus) is different, but nonetheless it is still made up of material of the press genre, and the more up-to-date data of the latter publications should also be viewed as a positive aspect of this project. Furthermore, it needs to be stated that the reliance of the entire project on press material is something of a drawback, but it can always be argued that the

genre reflects changes in the language of society relatively quickly, thereby giving a more reliable reflection of actual language use. Also, it bears repeating that data of this type is generally easier to acquire in large amounts than that of the fiction, academic, and spoken genres, and as such it is widely used today.

A further potential weakness is the inconsistency among the four publications in terms of the inclusion of BrE and AmE. This is a result of the time that was taken to produce the studies, and changes of plan during that time. As a result, we get a rather patchy picture of the relationship between the three varieties, and it means that further research is going to be needed to provide the missing information.

Nevertheless, there is a clear consistency in thematic approach among the four studies, and a wide range of topics are addressed, with results that make a significant contribution to the literature.

5.2 Future directions

Like any major project, besides answering many of the questions it set out to answer, this work has also left several unanswered, and raised various issues that now need to be taken up in future work.

The incomplete nature of the data for Publication I at the time meant that there were no corresponding BrE/AmE datasets to match the latter period of the CNZNE, making it impossible to go on anything more than educated guesswork in the question of change in BrE/AmE and the relationship with the very clear changes that were evident in NZE with several of the verbs of prevention. With the data available today in the form of the NOW corpus, the gaps can be filled, to give a clearer picture of the situation among the large group of verbs in three dialects across the now almost 30-year time period.

Similarly, with more than a decade having passed since the latest data was collected from the CNZNE on the *into -ing* construction, the work can be revisited, the data supplemented with more up-to-date data from NOW, and NZE results compared with the recent work that has been done on the pattern in other varieties. Considering the highly productive nature of the pattern, it is very likely that new, innovative verbs will be found, and there is every reason to expect that NZE speakers would not shy away from using Māori borrowings in the pattern.[45]

[45] Admittedly though, in terms of Māori borrowings in NZE, verbs are vastly outnumbered by nouns (Macalister 2006b).

Further work is also needed on verbs participating in both control and raising constructions, with more information needed on the rate of the grammaticalisation of *promise* in BrE over the course of the last century, and comparative information on NZE and AmE. In addition, the striking similarity in the frequencies of *threaten* in each of its uses in the three varieties invites further work with more data, and with different types of data. A further natural extension of the study would take account of the uses of *promise* and *threaten* in other varieties of English. This is of interest not only for the field of complementation, but to scholars of grammaticalisation, and language contact.

The NZE data on *fear* provided by Publication IV can form the basis of a follow-up study on the verb in other varieties. This would bolster the position of the Choice Principle by giving a better understanding of its predictive power in complement selection.

Finally, as of 2025, work on updating the CNZNE can begin. The aim is to produce a corpus of NZE newspaper English spanning a 30-year period. The data could be taken either directly from comparable sources in the NOW corpus, or downloaded from the relevant newspapers' own websites and cleaned up using freely-available software.

5.3 Concluding remarks

The present research was undertaken with the aim of providing new insights into the area of non-finite verb complementation in NZE, and it began with the assumption, based on the results of earlier work, that there would be discernible, statistical differences in some areas between NZE and the other varieties represented in some of the articles. This assumption was borne out in the results, with conclusive evidence allowing the statement that the native speakers of NZE are indeed using the grammar on their own terms in many of the areas investigated in the present study.

Echoing the sentiments of earlier research, it should be added here that these differences are not categorical, and are "mostly a matter of degree" (Hundt 1998: 140). They merely reflect the preferences NZE speakers show when selecting from the "one common underlying system of options" Mair (2007: 98) that are available within the boundaries of Standard English grammar. As was stated above, it seems that there are very few, if any, grammatical structures that are used only in NZE and nowhere else. This apparent lack of categorical grammatical variation is in sharp

contrast to the ongoing addition to NZE of lexical borrowings from Māori, to the point where tourists may even need to "use a dictionary just to read the newspaper" (Hay et al. 2008: 71), and the phonological idiosyncracies that are the cause of frequent misunderstandings in spoken exchanges for New Zealanders when they go abroad. Indeed, it has been shown time and again that, in terms of grammar, the native varieties have a great deal in common.

Historically speaking, research into the grammar of NZE has been underway for only a very short time. Its early years roughly coincided with the takeoff of corpus linguistics, and during the whole period, the latter few decades especially, the field has developed in ways which have significantly extended the researcher's scope, and today we have far greater insight into synchronic and diachronic language use than ever before. Due to its age and its position in the Anglosphere, NZE is an ideal subject for the study of shifting grammatical preferences under changing global influence. It is my hope that work of this kind will continue, and will continue to produce results that enhance our understanding of the ways in which the grammars of young dialects develop.

REFERENCES

Corpora

The Corpus of New Zealand Newspaper English

The News on the Web Corpus (https://www.english-corpora.org/now/)

The British National Corpus (http://www.natcorp.ox.ac.uk/)

The Corpus of Contemporary American English (https://www.english-corpora.org/coca/)

The Corpus of Late Modern English, version 3.0 (https://perswww.kuleuven.be/~u00444 28/ clmet3_0.htm)

Bibliography

Aarts, Bas. 1990. Prevent-type verbs in a GB-framework. *UCL Working Papers in Linguistics* 2: 147–164.

Algeo, John. 2006. *British or American English?: A Handbook of Word and Grammar Patterns*. Cambridge University Press.

Allerton, David J. 1988. 'Infinitivitis' in English, in Josef Klegraf and Dietrich Nehls (eds.), *Essays on the English Language and Applied Linguistics on the Occasion of Gerhard Nickel's 60th Birthday*. Heidelberg: Groos, 11–23.

Andersen, Gisle and Knut Hofland. 2012. Building a large corpus based on newspapers from the web, in Gisle Anderson (ed.) *Exploring newspaper language*. John Benjamins Publishing Company, 1–28.

Bauer, Laurie. 1979. The second great vowel shift? *Journal of the International Phonetic Association* 9: 57–66.

Bauer, Laurie. 1987. Approaching the grammar of New Zealand English. *New Zealand English Newsletter* 1: 12–15.

Bauer, Laurie. 1988. Number agreement with collective nouns in New Zealand English. *Australian Journal of Linguistics* 8 (2): 247–259.

Bauer, Laurie. 1989a. Irregularity in past non-finite verb-forms and a note on the New Zealand weekend. *New Zealand English Newsletter* 3: 13–16.

Bauer, Laurie. 1989b. Marginal modals in New Zealand English. *Te Reo* 32: 3–16.

Bauer, Laurie. 1989c. The verb have in New Zealand English. *English World-Wide* 10(1): 69–83.

Bauer, Laurie. 1994a. Introducing the Wellington Corpus of Written New Zealand English. *Te Reo* 37: 21–28.

Bauer, Laurie. 1994b. English in New Zealand, in Robert Burchfield (ed.), *The Cambridge History of the English Language, Volume V: English in Britain and Overseas: Origins and Development*. Cambridge: Cambridge University Press. 382–429.

Bauer, Laurie. 2001. Some verb complements in New Zealand English. *New Zealand English Journal* 15: 29–34.

Bayard, Donn. 1994. Review of *The Origins of the First New Zealanders*, *Te Reo* 37: 108–110.

Bayard, Donn. 2000. The cultural cringe revisited: changes through time and Kiwi attitudes towards accents, in Allan Bell and Koenraad Kuiper (eds.), *New Zealand English*. John Benjamins Publishing Company. 297–324.

Belich James. 1986. *The New Zealand Wars and the Victorian Interpretation of Racial Conflict*. Auckland: Auckland University Press.

Belich, James. 1996. *Making Peoples. A History of the New Zealanders from Polynesian Settlement to the End of the Nineteenth Century*. Harmondsworth: Allen Lane.

Belich, James. 2001. *Paradise Reforged: A History of the New Zealanders from the 1880s to the Year 2000*. Auckland: Allen Lane/Penguin Press.

Bell, Allan. 1988. The British base and the American connection in New Zealand media English. *American Speech* 63(4): 326–344.

Benton, Richard A. 2015. Perfecting the partnership: revitalising the Māori language in New Zealand education and society 1987–2014. *Language, Culture and Curriculum* 28(2): 99–112.

Bourke, Chris. 2010. *Blue Smoke: The Lost Dawn of New Zealand Popular Music 1918-1964*. Auckland: Auckland University Press.

Bresnan, Joan and Jennifer Hay. 2008. Gradient grammar: An effect of animacy on the syntax of *give* in New Zealand and American English. *Lingua* 118(2): 245–259.

Bridgeman, Lorraine, Dale Dillinger, Constance Higgins, David Seaman and Floyd Shank. 1965. *More classes of verbs in English*. Bloomington: Indiana University Linguistics Club.

Britain, David. 2000. As far as analysing grammatical variation and change in New Zealand English with very few tokens is concerned, in Allan Bell and Koenraad Kuiper (eds.), *New Zealand English*. John Benjamins Publishing Company. 198–220.

Bryant, Frank Egbert. 1907. VIII.—On the Conservatism of Language in a New Country. *PMLA* 22(2): 277–290.

Carey, David. 2019. Improving well-being in New Zealand through migration. *Economics Department Working Papers* No. 1566.

Carnie, Andrew. 2013. *Syntax: A Generative Introduction*, 3rd edition. John Wiley and Sons.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Massachusetts: MIT Press.

Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin and Use*. Westport, CT: Praeger.

Collins, Peter. 2015. Diachronic variation in the grammar of Australian English: corpus-based explorations. In Peter Collins (ed.), *Grammatical Change in English World-Wide*. John Benjamins Publishing Company. 15–42.

Crystal, David. 2003. *The Cambridge Encyclopedia of the English Language*. Cambridge: Cambridge University Press.

Davies, Caroline and Margaret MacLagan. 2006. Māori words – read all about it: Testing the presence of 13 Māori words in four New Zealand newspapers from 1997 to 2004. *Te Reo* 49: 73–99.

Davies, Mark A. 2012. Some methodological issues related to corpus-based investigations of recent syntactic changes in English. In Terttu Nevalainen and Elizabeth Traugott (eds.), *The Oxford Handbook of the History of English*. Oxford: Oxford University Press. 157–174.

Davies, Mark A. 2017. The new 4.3 billion word NOW corpus, with 4–5 million words of data added every day. *The 9th International Corpus Linguistics Conference*. Vol. 2017.

Davies, Mark A. and Jong-Bok Kim. 2019. Historical shifts with the into-causative construction in American English. *Linguistics: An Interdisciplinary Journal of the Language Sciences* 57(1): 29–58.

Degani, Marta. 2012. Language contact in New Zealand: A focus on English lexical borrowings in Māori. *Academic Journal of Modern Philology* 1: 13–24.

Degani, Marta and Alexander Onysko. 2010. Hybrid compounding in New Zealand English. *World Englishes* 29(2): 209–233.

De Smet, Hendrik. 2013. *Spreading Patterns: Diffusional Change in the English System of Complementation.* Oxford: Oxford University Press.

De Swaan, Abran. 2010. Language systems, in Nikolas Coupland (ed.), *The Handbook of Language and Globalization*. Malden, MA: Blackwell. 56–76.

Deverson, Tony. 2000. A revised bibliography of writings on New Zealand English. *New Zealand English Journal* 14: 18–33.

Diller, Hans-Jürgen, Hendrik De Smet and Jukka Tyrkkö. 2011. A European database of descriptors of English electronic texts. *The European English Messenger* 19: 21–35.

Dollinger, Stefan. 2008. *New-Dialect Formation in Canada: Evidence from the English Modal Auxiliaries*. Amsterdam: John Benjamins.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67: 547–619.

Duffley, Patrick. 2000. Gerund versus infinitive as complement of transitive verbs in English: The problems of "tense" and "control". *Journal of English Linguistics* 28(3): 221–248.

Duffley, Patrick. 2014. *Reclaiming Control as a Semantic and Pragmatic Phenomenon*. John Benjamins Publishing Company.

Duffley, Patrick and Ryan Fisher. 2021. *To*-infinitive and gerund-participle clauses with the verbs *dread* and *fear*. *Studia Linguistica* 75(1): 72–96.

Fanego, Teresa. 1996. The development of gerunds as objects of subject-control verbs in English (1400-1760). *Diachronica* 13: 29–62.

Fanego, Teresa. 2004. On reanalysis and actualization in syntactic change: The rise and development of Englis verbal gerunds. *Diachronica* 21(1): 5–55.

Fanego, Teresa. 2007. Drift and the development of sentential complements in British and American English from 1700 to the present day, in Javier Pérez-Guerra, Dolores González-Álvarez, Jorge L. Bueno-Alonso and Esperanza Rama-Martínez (eds.), *Of Varying Language and Opposing Creed: New Insights into Late Modern English*. Bern: Peter Lang. 161–235.

Fillmore, Charles. 1968. The Case for Case, in Emmon Bach and Robert T. Harms (eds.) *Universals in Linguistic Theory*. New York: Holt, Rinehart and Winston.

Filppula, Markku, Juhani Klemola and Heli Paulasto. 2008. *English and Celtic in Contact*. London: Routledge.

Flach, Susanne. 2021. From movement into action to manner of causation: changes in argument mapping in the into-causative. *Linguistics* 59(1): 247–283.

Francis, Gill, Susan Hunston and Elizabeth Manning (eds.). 1996. *Collins Cobuild Grammar Patterns 1: Verbs*. London: Harper Collins.

Fuchs, Robert. 2022. Colonial lag or feature retention in postcolonial varieties of English: the negative scalar conjunction "and that too" in South Asian Englishes and beyond, in Paula Rautionaho, Hanna Parviainen, Mark Kaunisto and Arja Nurmi (eds.), *Social and Regional Variation in World Englishes: Local and Global Perspectives*. New York and London: Routledge. 123–148.

Garside, Roger and Nicholas Smith. 1997. A hybrid grammatical tagger: CLAWS4, in Roger Garside, Geoffrey Leech, Anthony McEnery (eds.), *Corpus Annotation: Linguistic Information from Computer Text Corpora*. Longman, London. 102–121.

Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Goldberg, Adele. 2003. Constructions: A New Theoretical Approach to Language. *Trends in Cognitive Sciences* 7(5): 219–224.

Goldberg, Adele. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Gonçalves Bruno, Lucía Loureiro-Porto, José J Ramasco and David Sánchez. 2018. Mapping the Americanization of English in space and time. PLOS ONE 13(5): e0197741.

Gordon, Elizabeth. 2005. Macmillan Brown Lecture (2nd in series, September 23), Christchurch:http://www.radionz.co.nz/nr/programmes/the_macmillan_brown_lectures

Gordon, Elizabeth. 2012. Standard New Zealand English, in Raymond Hickey (ed.), *Standards of English: Codified Varieties Around the World.* Cambridge: Cambridge University Press. 318–335.

Gordon, Elizabeth, Lyle Campbell, Jennifer Hay, Margaret Maclagan, Andrea Sudbury, and Peter Trudgill. 2004. *New Zealand English: Its Origins and Evolution*. Cambridge: Cambridge University Press.

Gordon, Elizabeth and Margaret Maclagan. 2001. Capturing a sound change: a real time study over 15 years of the NEAR/SQUARE diphthong merger in New Zealand English. *Australian Journal of Linguistics* 21(2): 215–238.

Görlach, Manfred. 1987. Colonial lag? The alleged conservative character of American English and other 'colonial' varieties. *English World-Wide* 8: 41–60.

Görlach, Manfred. 1990. *Studies in the History of the English Language*. Heidelberg: Winter.

Green, James and Donn Bayard. 2000. Zed to Zee Revisited: Notes from Dunedin on the Americanisation of the NZE lexicon through time. *Te Reo* 43: 63–70.

Harlow, Ray, Peter Keegan, Jeanette King, Margaret Maclagan and Catherine Watson. 2009. The changing sound of the Māori languages, in James N. Stanford and Dennis R. Preston (eds.), *Quantitative Sociolinguistic Studies of Indigenous Minority Languages*. Amsterdam: John Benjamins. 129–152.

Hay, Jennifer, Margaret Maclagan and Elizabeth Gordon. 2008. *New Zealand English.* Edinburgh: Edinburgh University Press.

Hay, Jennifer and Daniel Schreier. 2004. Reversing the trajectory of language change: subject–verb agreement with BE in New Zealand English. *Language Variation and Change* 16(3), 209–235.

Heine, Bernd and Tania Kuteva. 2005. *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.

Hoffmann, Sebastian. 2018. I would like to request for your attention: On the diachrony of prepositional verbs in Singapore English, in Mark Kaunisto, Mikko Höglund and Paul Rickman (eds.), *Changing Structures: Studies in Constructions and Complementation.* John Benjamins Publishing Company. 171–196.

Holdaway, Simon J., Joshua Emmitt, Louise Furey, Alex Jorgensen, Gerard O'Regan, Rebecca Phillipps, Matthew Prebble, Roderick Wallace and Thegn N. Ladefoged. 2019. Māori settlement of New Zealand: The Anthropocene as a process. *Archaeology in Oceania* 54(1): 17–34.

Huddleston, Rodney and Geoffrey Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Hundt, Marianne. 1998. *New Zealand English Grammar: Fact or Fiction?* Amsterdam/Philadephia: John Benjamins.

Hundt, Marianne. 2009. Colonial lag, colonial innovation, or simply language change?, in Günter Rohdenburg and Julia Schlüter (eds.), *One Language, Two Grammars*. Cambridge: Cambridge University Press. 13–37.

Hundt, Marianne. 2017. Australian/New Zealand English, in Alexander Bergs and Laurel Brinton (eds.), *Varieties of English*. De Gruyter, Inc. 289–310.

Hundt, Marianne. 2018. It is time that this (should) be studied across a broader range of Englishes: a global trip around mandative subjunctives, in Sandra C. Deshors (ed.), *Modeling World Englishes. Assessing the Interplay of Emancipation and Globalization of ESL Varieties*. Amsterdam: John Benjamins Publishing. 217–244.

Hundt, Marianne, Jennifer Hay and Elizabeth Gordon. 2008. New Zealand English: Morphosyntax, in Kate Burridge and Bernd Kortmann (eds.), *Varieties of English. 3, The Pacific and Australasia.* Berlin: Mouton de Gruyter. 305–340.

Hundt, Marianne and Christian Mair. 1998. 'Agile' and 'uptight' genres. The corpus-based approach to language change in progress. *International Journal of Corpus Linguistics* 4(2): 221–242.

Hundt, Marianne and Benedikt Szmrecsanyi. 2012. Animacy in early New Zealand English. *English World-Wide* 33(3): 241–263.

Jespersen, Otto. 1992 [1924]. *The Philosophy of Grammar.* Chicago: The University of Chicago Press.

Jespersen, Otto. 1961 [1940]. *A Modern English Grammar on Historical Principles. Part V: Syntax* (Vol. IV). London: Allen and Unwin.

Kachru, Braj B. 1985. Standards, codification and sociolinguistic realism: The English language in the outer circle, in Randolph Quirk and Henry G. Widdowson (eds.), *English in the World: Teaching and Learning the Language and Literatures.* Cambridge: Cambridge University Press. 11–30.

Kachru, Braj B. 1986. *The Alchemy of English: the Spread, Functions and Models of Non-native Englishes*. Oxford: Pergamon Press.

Kachru, Braj B. 1990. World Englishes and applied linguistics. *World Englishes* 9(1): 3–20.

Kaunisto, Mark and Juhani Rudanko. 2019. *Variation in Non-Finite Constructions in English: Trends Affecting Infinitives and Gerunds*. Springer International Publishing.

Kearns, Kate. 2009. Zero complementizer, syntactic context, and regional variety, in Pam Peters, Peter Collins and Adam Smith (eds.), *Comparative Studies in Australian and New Zealand English: Grammar and Beyond*. John Benjamins Publishing Company. 243–262.

Kerswill, Paul. 2020. Contact and new varieties, in Raymond Hickey (ed.), *The Handbook of Language Contact*. John Wiley and Sons, Incorporated. 241–260.

Kim, Jong-Bok and Mark A. Davies. 2015. The into-causative construction in English: a construction-based perspective. *English Language and Linguistics* 20: 55–83.

King, Michael. 2003. *The Penguin History of New Zealand*. London: Penguin Books.

Koster, Jan and Robert May. 1982. On the Constituency of Infinitives. *Language* 58(1): 116–143.

Landau, Idan. 2013. *Control in Generative Grammar: A Research Companion*. Cambridge University Press.

Landau, Idan. 2021. Duality of control in gerundive complements of P. *Journal of Linguistics* 57: 1–31.

Lass, Roger. 1990. Where do extraterritorial Englishes come from? Dialect, input, and recodification in transported Englishes, in Sylvia Adamson, Vivian Law, Nigel Vincent, and Susan Wright (eds.), *Papers from the 5th International Conference on English Historical Linguistics*. Philadelphia: John Benjamins. 245–280.

Leech, Geoffrey. 1968. Some assumptions in the metatheory of linguistics, *Linguistics* 39: 87–102.

Leech, Geoffrey, Marianne Hundt, Christian Mair and Nicholas Smith. 2009. *Change in Contemporary English: A Grammatical Study*. Cambridge: Cambridge University Press.

Los, Bettelou. 2015. *A Historical Syntax of English*. Edinburgh University Press.

Macalister, John. 2006a. A note on the newly revised bibliography of writings on New Zealand English. *New Zealand English Journal* 20: 48–50.

Macalister, John. 2006b. The Māori presence in the New Zealand English lexicon, 1850–2000. *English World-Wide* 27(1): 1–24.

Macalister, John. fc. NZE lexis, in Kingsley Bolton (ed.), *The Wiley Blackwell Encyclopedia of World Englishes*. Wiley-Blackwell.

Maclagan, Margaret and Jennifer Hay. 2007. Getting fed up with our feet: contrast maintenance and the New Zealand English "short" front vowel shift. *Language Variation and Change* 19(1): 1–25.

Maclagan, Margaret, Jeanette King and Gail Gillon. 2008. Māori English. *Clinical Linguistics and Phonetics* 22(8): 658–670.

Mair, Christian. 2002. Three changing patterns of verb complementation in Late Modern English: A real-time study based on matching text corpora. *English Language and Linguistics* 6(1): 105–131.

Mair, Christian. 2006. *Twentieth-Century English: History, Variation, and Standardization*. Cambridge: Cambridge University Press.

Mair, Christian. 2007. British English/American English grammar: Convergence in writing – divergence in speech? *Anglia* 125: 84–100.

Mair, Christian. 2009a. Corpus linguistics meets sociolinguistics: The role of corpus evidence in the study of sociolinguistic variation and change. *Language and Computers* 69(1): 7–32.

Mair, Christian. 2009b. Infinitival and gerundial complements, in Pam Peters, Peter Collins and Adam Smith (eds.), *Comparative Studies in Australian and New Zealand English: Grammar and beyond*. John Benjamins Publishing Company. 263–276.

Mair, Christian. 2013. The world system of Englishes. *English World-Wide* 34(3): 253–278.

Mair, Christian. 2019. American English: No written standard before the twentieth century?, in Nuria Yáñez-Bouza, Emma Moore, Linda van Bergen, and William Hollman (eds.),

Categories, Constructions and Change in English Syntax. Cambridge: Cambridge University Press. 336–363.

Mair, Christian. 2022. Migration, media, and the emergence of pidgin- and creole-based informal epicentres. *World Englishes* 41(3): 414–428.

Marckwardt, Albert. 1967 [1958]. *American English*. New York: Oxford University Press.

Matras, Yaron. 2009. *Language Contact*. Cambridge: Cambridge University Press.

McArthur, Tom. 1987. The English Languages? *English Today* 11: 9–11.

McIntyre, W. David. 1999. The strange death of dominion status. *The Journal of Imperial and Commonwealth History*, 27(2): 193–212.

McKinnon, Malcolm (ed.). 1997. *New Zealand Historical Atlas*. Bateman: Albany.Melchers, Gunnell, Philip Shaw and Peter Sundkvist. 2019. *World Englishes* (3rd ed.). Routledge.

Meyer, Charles F. 2002. *English Corpus Linguistics: An Introduction*. Cambridge: Cambridge University Press.

Mitchell, Bruce and Fred C. Robinson. 2012. *A Guide to Old English* (8th ed.) John Wiley and Sons.

Mukherjee, Joybrato and Sebastian Hoffmann. 2006. Describing verb-complementational profiles of new Englishes: a pilot study of Indian English. *English World-Wide* 27(2): 147–173.

Mustanoja, Tauno F. 1960. *A Middle English Syntax. Part I: Parts of speech*. Helsinki: Société Néophilologique.

Nelson, Cecil L., Zoya G. Proshina and Daniel R. Davis (eds.). 2020. *The Handbook of World Englishes*. John Wiley & Sons, Incorporated.

New Zealand Oxford Dictionary. 2005. Tony Deverson and Graeme Kennedy (eds.). Australia and New Zealand: Oxford University Press.

Nokes, Jacqui and Jennifer Hay. 2012 Acoustic correlates of rhythm in New Zealand English: A diachronic study. *Language Variation and Change* 24(1): 1–31.

Noonan, Michael. 1985. Complementation, in Timothy Shopen (ed.), *Language Typology and Syntactic Description. Volume 2: Complex Constructions*. Cambridge: Cambridge University Press. 42–140.

Olavarría de Ersson, Eugenia and Philip Shaw. 2003. Verb complementation patterns in Indian Standard English. *English World-Wide* 24: 137–61.

Oxford English Dictionary. 2nd ed. 1989. OED online. Available from http//:www.oed.com. [Accessed January–October 2023] Oxford: Oxford University Press.

Peters, Pam, Peter Collins and Adam Smith (eds.). 2009. *Comparative Studies in Australian and New Zealand English: Grammar and Beyond*. John Benjamins Publishing Company.

Postal, Paul. 1970. On coreferential complement subject deletion. *Linguistic Inquiry* 1(4): 439–500.

Poutsma, H. 1929 [1905]. *A Grammar of Late Modern English*. Part 1, The Sentence. Second edition. Groningen: P. Noordhoff.

Quinn, Heidi. 2000. Variation in New Zealand English syntax and morphology, in Allan Bell and Koenraad Kuiper (eds.), *New Zealand English*. John Benjamins Publishing Company. 173–197.

Ransom, Evelyn. 1986. *Complementation: Its Meanings and Forms*. Amsterdam and Philadelphia: John Benjamins.

Rickman, Paul. 2018. Collective nouns and number agreement in New Zealand English. *New Zealand Studies in Applied Linguistics* 24(1): 40–55.

Rickman, Paul and Juhani Rudanko. 2018. *Corpus-Based Studies on Non-finite Complements in Recent English*. London: Palgrave Macmillan.

Rickman, Paul and Juhani Rudanko. 2023. Straddling a syntactic divide: tracking sentential complements of *promise* and *threaten* in CLMET3.0, in Hendrik de Smet, Peter Petré and Benedikt Szmrecsanyi, (eds.), *Context, Intent and Variation in Grammaticalization*. de Gruyter. 253–276.

Riemsdijk, Henk van and Edwin Williams. 1986. *Introduction to the Theory of Grammar*. Cambridge: MIT Press.

Rohdenburg, Günter. 1996. Cognitive complexity and increased grammatical explicitness in English. *Cognitive Linguistics* 7(2): 149–182.

Rohdenburg, Günter. 2003. Cognitive complexity and *horror aequi* as factors determining the use of interrogative clause linkers in English, in Günter Rohdenburg and Britta Mondorf (eds.), *Determinants of Grammatical Variation in English*. Berlin: De Gruyter. 205–250.

Rohdenburg, Günter. 2006. The Role of Functional Constraints in the Evolution of the English Complementation System. In Christine Dalton-Puffer, Dieter Kastovsky, Nicholas Ritt, and Herbert Schendl (eds.), *Syntax, Style and Grammatical Norms: English from 1500–2000*, 143–166. Bern: Peter Lang.

Rohdenburg, Günter. 2009. Nominal complements, in Günter Rohdenburg and Julia Schlüter (eds.), *One Language, Two Grammars?: Differences Between British and American English*. Cambridge University Press. 194–211.

Rosenbaum, Peter. 1967. *The Grammar of English Predicate Complement Constructions*. MIT Press.

Rudanko, Juhani. 1996. *Prepositions and Complement Clauses: A Syntactic and Semantic Study of Verbs Governing Prepositions and Complement Clauses in Present-Day English.* State University of New York Press.

Rudanko, Juhani. 1998. *Change and Continuity in the English Language: Studies on Complementation over the Past Three Hundred Years*. University Press of America.

Rudanko, Juhani. 2002. *Complements and Constructions*. Lanham MD: University Press of America.

Rudanko, Juhani. 2005. Lexico-grammatical innovation in current British and American English: A case study on the transitive *into -ing* pattern with evidence from the Bank of English corpus. *Studia Neophilologica* 77: 171–187.

Rudanko, Juhani. 2006a. Watching English grammar change: a case study on complement selection in British and American English. *English Language & Linguistics* 10(1): 31–48.

Rudanko, Juhani. 2006b. Emergent alternation in complement selection: the spread of the transitive 'into -ing' construction in British and American English. *Journal of English Linguistics* 34: 312–331.

Rudanko, Juhani. 2010. Explaining grammatical variation and change: a case study of complementation in American English over three decades. *Journal of English Linguistics* 38(1): 4–24.

Rudanko, Juhani. 2012. Exploring aspects of the Great Complement Shift, with evidence from the TIME corpus and COCA. In Terttu Nevalainen and Elizabeth Closs Traugott (eds.), *The Oxford Handbook of the History of English.* Oxford: Oxford University Press. 222–232.

Rudanko, Juhani. 2014. A new angle on infinitival and of -ing complements of afraid, with evidence from the TIME Corpus. In Kristin Davidse, Caroline Gentens, Lobke Ghesquière and Lieven Vandelanotte (eds.), *Corpus Interrogation and Grammatical Patterns*. Amsterdam: John Benjamins. 223–238.

Rudanko, Juhani. 2015. *Linking Form and Meaning: Studies on Selected Control Patterns in Recent English*. Basingstoke, UK: Palgrave Macmillan.

Rudanko, Juhani. 2017. *Infinitives and Gerunds in Recent English: Studies on Non-Finite Complements with Data from Large Corpora*. London: Palgrave Macmillan Springer.

Rudanko, Juhani and Lea Luodes. 2005 *Complementation in British and American English: Corpus-based Studies on Prepositions and Complement Clauses in British and American English*. University Press of America.

Ruohonen, Juho and Rudanko, Juhani. 2021. *Infinitival vs Gerundial Complementation with* afraid*,* accustomed*, and* prone*: Multivariate Corpus Studies*. Springer International Publishing.

Schneider, Edgar. 2003. The dynamics of new Englishes: from identity construction to dialect birth. *Language* 79(2): 233–281.

Schneider, Edgar. 2007. *Postcolonial English: Varieties Around the World.* Cambridge: Cambridge University Press.

Schneider, Edgar. 2014. New reflections on the evolutionary dynamics of world Englishes. *World Englishes* 33(1): 9–32.

Schreier, Daniel. 2019. /h/insertion as a 'camouflage archaism'? Dialect contact, colonial lag and the feature pool in South Atlantic English. *Diachronica* 36(1): 37–65.

Szakay, Anita. 2006. Rhythm and pitch as markers of ethnicity in New Zealand English, in Paul Warren and Catherine Watson (eds.), *Proceedings of the 11th Australian International Conference on Speech Science and Technology*. 421–426.

Szmrecsanyi, Benedikt, Jason Grafmiller, Joan Bresnan, Anette Rosenbach, Sali Tagliamonte and Simon Todd. 2017. Spoken syntax in a comparative perspective: The dative and genitive alternation in varieties of English. *Glossa: a journal of general linguistics* 2(1): 1–27.

Tajima, Matsuji. 1983. *The Syntactic Development of the Gerund in Middle English*. Doctoral dissertation, University of Ottawa.

Thomason, Sarah G. and Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.

Tognini-Bonelli, Elena. 2001. *Corpus Linguistics at Work*. John Benjamins Publishing Company.

Tombs, Robert. 2014. *The English and their History*. Allen Lane.

Traugott, Elizabeth C. 1993. The conflict promises/threatens to escalate into war, in Joshua Guenter, Barbara Kaiser and Cheryl Zoll (eds.), *Proceedings of the Nineteenth Annual Meeting of the Berkeley Linguistics Society General Session and Parasession on Semantic Typology and Semantic Universals*. 348–358.

Traugott, Elizabeth C. 1996. Subjectification and the development of epistemic meaning: the case of *promise* and *threaten*, in Toril Swan and Olaf Westvik (eds.), *Modality in Germanic Languages: Historical and Comparative Perspectives*. Berlin: de Gruyter Mouton. 185–210.

Trudgill, Peter. 1986. *Dialects in Contact.* Oxford: Blackwell.

Trudgill, Peter. 1999. A window on the past: "Colonial lag" and New Zealand evidence for the phonology of nineteenth-century English. *American Speech* 74(3): 227–239.

Trudgill, Peter. 2004. *New-Dialect Formation: The Inevitability of Colonial Englishes.* Edinburgh: Edinburgh University Press.

Trudgill, Peter, Elizabeth Gordon, Gillian Lewis and Margaret MacLagan. 2000. Determinism in new-dialect formation and the genesis of New Zealand English. *Journal of Linguistics* 36(2): 299–318.

Trudgill, Peter and Jean Hannah. 1982. *International English: A Guide to Varieties of Standard English*. London: Edward Arnold.

Trye, David, Andeea S. Calude, Felipe Bravo-Marquez and Te Taka Keegan. 2020. Hybrid hashtags: #Youknowyoureakiwiwhen your tweet contains Māori and English. *Frontiers in Artificial Intelligence* 3(15).

Vine, Bernadette. 1999. Americanisms in the New Zealand English lexicon. *World Englishes* 18(1): 13–22.

Vosberg, Uwe. 2003a. Cognitive complexity and the establishment of *-ing* constructions with retrospective verbs in modern English, in Marina Dossena and Charles Jones (eds.), *Insights into Late Modern English*. Berlin: Mouton de Gruyter. 197–220.

Vosberg, Uwe. 2003b. The role of extractions and *horror aequi* in the evolution of *-ing* complements in modern English, in Günter Rohdenburg and Britta Mondorf (eds.), *Determinants of Grammatical Variation in English*. Berlin: Mouton de Gruyter. 305–327.

Vosberg, Uwe. 2006. *Die Grosse Komplementverschiebing*. Tübingen: Narr.

Vosberg, Uwe. 2009. Non-finite complements, in Günter Rohdenburg and Julia Schlüter (eds.), *One Language, Two Grammars? Differences between British and American English*. Cambridge: Cambridge University Press. 212–227.

Wallis, Sean. 2020. *Statistics in Corpus Linguistics Research: A New Approach*. Routledge.

Walter, Richard, Halley Buckley, Chris Jacomb and Elizabeth Matisoo-Smith. 2017. Mass migration and the Polynesian settlement of New Zealand. *Journal of World Prehistory,* 30(4): 351–376.

Wells, John. 1982. *Accents of English*. Cambridge: Cambridge University Press.

Wierzbicka, Anna. 1988. *The Semantics of Grammar*. Philadelphia: John Benjamins.

Wierzbicka, Anna. 1998. The semantics of English causative constructions in a universal-typological perspective. In Michael Tomasello (ed.), *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure, Volume I.* Mahwah, NJ: Lawrence Erlbaum. 113–153.

Winford, Donald. 2003. *An Introduction to Contact Linguistics*. Blackwell Publishing.

Wood, Frederick T. 1956. Gerund vs. infinitive. *English Language Teaching* 11: 11–16.

Websites

http://corpora.lancs.ac.uk/sigtest/ [accessed August–November 2023]

https://gazette.education.govt.nz/ [accessed May 2023]

https://www.canterbury.ac.nz/nzilbb/research/onze/ [accessed February 2023]

https://www.english-corpora.org/now/ [accessed February 2023]

https://www.ice-corpora.uzh.ch/en.html [accessed January 2023]

https://www.knowledge-basket.co.nz [accessed 2011–12 and January 2023]

https://www.laurenceanthony.net/software/antconc/ [accessed February 2023]

https://www.stats.govt.nz/ [accessed January–February 2023]

https://www.usa.gov/ [accessed March 2023]

https://www.wgtn.ac.nz/lals/resources/corpora-default [accessed May 2023]

PUBLICATIONS

Publication I Rickman, Paul. 2017. Aspects of verb complementation in New Zealand newspaper English. In Markku Filppula, Juhani Klemola, Anna Mauranen and Svetlana Vetchinnikova (eds.), *Changing English: Global and Local Perspectives*, 169–190. Berlin, Boston: Mouton de Gruyter.

Publication II Rickman, Paul and Mark Kaunisto. 2018. Aspects of the use of the transitive *into -ing* pattern in New Zealand English. In Mark Kaunisto, Mikko Höglund and Paul Rickman (eds.), *Changing Structures: Studies in Constructions and Complementation*, 71–88. Amsterdam: John Benjamins.

Publication III Rickman, Paul and Juhani Rudanko. 2023. Facing two ways syntactically: on the grammar and use of *promise* and *threaten* in three regional varieties. *English Studies* 104(2), 365–382.

Publication IV Rickman, Paul and Juhani Rudanko. 2024. The Choice Principle and the matrix verb *fear* in recent New Zealand English. *Studia Neophilologica* (Published online: 7 Feb 2024).

PUBLICATION I

Aspects of verb complementation in New Zealand newspaper English

Paul Rickman

Changing English: Global and Local Perspectives, 169–190
https://doi.org/10.1515/9783110429657

Publication reprinted with the permission of the copyright holders.

Aspects of Verb Complementation in New Zealand Newspaper English[1]

Paul Rickman

University of Tampere

Abstract

The present paper describes a new diachronic corpus of New Zealand newspaper English compiled at the University of Tampere. Using material from the Fairfax archives, the corpus covers selected periods from the last two decades, and comprises approximately 100 million words. Using data from the corpus in addition to pre-existing data on British and American English, the paper then explores aspects of the complementation patterns of the verbs of prevention. The study aims to highlight selected aspects of NZE usage in comparison to BrE usage, with the older postcolonial variety AmE used as a point of reference. The study goes on to discuss the direction of change discernible in the development of NZE, with reference to the concept of 'colonial lag'. It is argued that while certain changes may be classified as developmental lags, examples of other types of diachronic change, including innovation, are also evident.

Keywords: New Zealand English (NZE), British English (BrE), American English AmE), complementation, verbs of prevention, colonial lag, diachronic change, newspaper language, Corpus of New Zealand Newspaper English (CNZNE), British National Corpus (BNC), Corpus of Contemporary American English (COCA), Corpus of Late Modern English Texts (CLMET3.0)

[1] This research was made possible by funding from the Academy of Finland, under the GlobE consortium. The author would like to thank all members of the consortium, the audience at ChangE Helsinki 2013 and an anonymous reviewer for helpful comments on the paper.

1. Introduction

The main Southern Hemisphere Englishes – New Zealand, Australian and South African – have provided fertile ground for linguistic analysis, and what is emerging is a picture of young nations that are comfortable with their identities – a large part of that identity being the language that they take pride in calling their own. In each of these three varieties, English has taken on its own unique flavour, having absorbed features of the pre-existing native language(s), and, in the case of NZE at least, the old and the new now exist side by side. In New Zealand, both English and Māori have official status, and the lexical presence and influence of the latter in New Zealand today is clear and undeniable, evident in place names, terms for flora and fauna, and, perhaps to a lesser extent, more general vocabulary items (see e.g. Davies & Maclagan 2006; Macalister 2006).

The view that New Zealand English (NZE) is a separate and unique variety of English is not new, with commentary and studies on its various phonological and lexical features dating back many years (for a summary see Gordon & Abell 1990). However the view that NZE contains unique grammatical features was only established in more recent decades. Regarding genuine syntactic New Zealandisms, it was recently noted that "there are actually relatively few syntactic features that are wholly unique to New Zealand" (Hay et al. 2008: 47), and some 20 years prior to that, it was noted that any differences that exist are not radical, but are manifest in "matters of degree rather than in categorical distinctions" (Bauer 1989b: 82). In the closing years of the twentieth century, Hundt (1998) presented a convincing argument supporting the view that several aspects of NZE grammar do serve to set it apart from the other main varieties, and it has been argued that "even if, in terms of grammar, usage in New Zealand is found to agree closely with the standards of the United States and Britain, that does not mean that it makes no sense to speak of New Zealand English morphosyntax" (Hundt et al. 2008: 305).

The present paper aims to contribute to the literature on NZE by using new NZE data alongside BrE and AmE data to address the question of 'colonial lag' and other manifestations of postcolonial linguistic development and differentiation. The paper is structured as follows: Section 2 describes a new corpus compiled at the University of Tampere in order to support further studies on NZE; Sections 3 presents a study of the sentential complements of the *prevent* type verbs, with evidence from NZE, BrE, and AmE corpora. A diachronic approach is taken, allowing categorisation of the apparent changes underway in NZE within the framework of the six-way typology of diachronic change proposed in Hundt (2009). Section 4 summarises the colonial lag question, introduces the typological framework, and provides a discussion of results. Here I provide evidence in support of the idea that the linguistic 'lag' scenario is only a small part of a more complex system of changes taking place in NZE. Section 5 provides concluding remarks.

2. The New Zealand Newspaper Corpus[2]

Basing research on corpora consisting entirely of newspaper material requires little justification these days; a look at some recent work (Macalister 2001; Rohdenburg 2002; Mukherjee & Hoffmann 2006; Davies & Maclagan 2006; Calude & James 2011) shows that a carefully compiled corpus of media language can provide a solid foundation for linguistic analysis, provided the usual restrictions regarding generalisability of results are kept in mind.

The currently available set of NZE corpora are well compiled and useful for a wide range of research goals, but they are nevertheless small by today's standards, and therefore unsuitable for certain uses. In a recent study on constructions in four varieties of English, Mair found aspects of his investigation restricted by the small size of the corpora used, commenting at one point that his results are "bedeviled by the very low figures in some cells" (Mair 2009: 271). The corpora were the Brown/LOB quartet for written AmE and BrE, the Wellington Corpus of Written New Zealand English (WWC) and the Australian Corpus of English (ACE) for written NZE and AusE. The latter two are based on the Brown/LOB template, with 1980s/90s material, allowing comparison to Frown/FLOB. For spoken English, Mair used the 600,000 word spoken sections of the ICE corpora, which, as noted above, were not large enough to give a full picture of the patterns associated with some predicates. In an earlier study Hundt et al. comment that "the data on *dare* from the one-million-word corpora are too meagre to verify any hypotheses on diachronic change or regional variation" (Hundt et al. 2008: 319), and some years prior to that Bauer noted the sparse data available for *farewell, screen,* and *appeal* in the same group of corpora (Bauer 2001).

With size thus being one of the main prerequisites, I set about compiling a corpus of NZE newspaper text (The Corpus of New Zealand Newspaper English – CNZNE) using material from the archives of Fairfax Media.[3] Fairfax is one of the main media groups in Australasia, controlling publications that span the length of New Zealand and include major metropolitan newspapers from three of the four main centres – Auckland, Wellington, and Christchurch – as well as several provincial papers. Care was taken to ensure that papers from as many different regions of the country as possible were included, since, while it is generally acknowledged that NZE is more homogeneous than BrE and AmE, there is still known to be some, albeit minor, regional variation (Turner 1966: 163 ff.; Bauer & Bauer 2000; Hay et al. 2008: 95 ff.; Calude & James 2011).

[2] Prof. Dr. Sebastian Hoffmann gave valuable and much needed assistance in the compilation of the corpus in connection with the data retrieval and processing, for which I am extremely grateful. All Perl scripts used in the process were written by him.

[3] Fairfax Media New Zealand kindly granted permission for their material to be republished in my work.

The aim was to compile a diachronic corpus, the earlier sub-section of which would parallel the British National Corpus (BNC)[4] newspaper sub-section. The composition of the BNC sub-corpus, according to the numbers retrieved from the BNCweb (CQP-Edition) interface, is given in Table 1.

Table 1. BNC newspaper sub-section word counts

Arts	593,014
Commerce	850,071
Editorial	102,718
Miscellaneous	1,040,943
Report	3,403,683
Science	121,199
Social	1,234,095
Sports	1,333,385
Tabloid	733,066
TOTAL	**9,412,174**

While BNC news material covers the period 1985–93, the Fairfax archives contain electronic material going back only as far as 1995, which was deemed close enough to the BNC time frame to allow valid comparison.[5] From the period 1995–98, one-year samples (which included all available material) were taken from a selection of 10 broadsheet newspapers (nine daily and one weekly), and two-year samples were taken from two weekly tabloid newspapers. Two-year tabloid samples were taken in order to try and increase the amount of this type of material, which still remains somewhat lower in comparison to broadsheet material than the ratio found in the BNC.

One- and two-year samples (which, again, included all available material) of the same papers, where possible, were taken from the 2010–12 period, forming the second section of the corpus. The 2010–12 sub-section contains material from only 10 papers, due to the fact that in 2002 two of the newspapers merged, and in 2006 one of the tabloids ceased its contribution to the archives. Table 2 provides the details of both sub-sections of the NZE corpus.

[4] A genre-balanced corpus comprising 100 million words of British English (90 % written, 10 % spoken) covering the years 1960–1993.

[5] Details of the BNC newspaper material show that the majority of the data, which is mainly taken from the larger metropolitan papers *The Independent, The Guardian,* and *The Daily Telegraph*, is from the period Oct 1989-Apr 1992.

Table 2. Newspapers, periods sampled, and word counts in the CNZNE

NEWSPAPER	PERIOD SAMPLED	WORD TOKENS
SECTION 1: 1995–98		
Daily News*	Aug 1996–Jul 1997	2,676,265
Dominion**	Jan 1995–Dec 1995	7,725,459
Evening Post**	Jan 1996–Dec 1996	6,874,558
Evening Standard*	Jun 1996–May 1997	2,158,689
Nelson Mail	Jul 1997–Jun 1998	2,221,658
Press	Jan 1997–Dec 1997	3,710,614
Southland Times	Mar 1997–Feb 1998	4,223,612
Sunday News (tabloid)	Dec 1995–Nov 1997	3,626,271
Sunday Star Times	Dec 1995–Nov 1996	4,096,501
Timaru Herald	Jul 1996–Jun 1997	1,422,032
Truth (tabloid)***	Dec 1995–Nov 1997	808,631
Waikato Times	May 1996–Apr 1997	3,049,242
TOTAL		**42,593,532**
SECTION 2: 2010–12		
Dominion Post	Jan 2011–Dec 2011	10,142,874
Manawatu Standard	Jan 2011–Dec 2011	5,022,767
Nelson Mail	Jan 2010–Dec 2010	4,668,709
Press	Jan 2012–Dec 2012	10,980,886
Southland Times	Jan 2012–Dec 2012	5,940,329
Sunday News (tabloid)	Jan 2010–Dec 2011	3,435,614
Sunday Star Times	Jan 2012–Dec 2012	3,583,489
Taranaki Daily News	Jan 2010–Dec 2010	4,809,591
Timaru Herald	Jan 2011–Dec 2011	4,134,794
Waikato Times	Jan 2010–Dec 2010	5,775,395
TOTAL		**58,494,448**

* name of paper changed - *Daily News* Section 1 = *Taranaki Daily News* Section2
- *Evening Standard* Section 1 = *Manawatu Standard* Section 2

** merged to become *Dominion Post* in Section 2

*** not found in the archive after 2006

The sample periods of the 2010–12 sub-section are noticeably more uniform than those of the earlier sub-section. This is due to most of the newspapers having begun electronic archiving at different times in the mid-1990s – not always at the beginning of the calendar year – and the date from which each paper first appears in the electronic archive was taken as the starting point for each paper's sample period. When compiling the

2010–12 sub-section on the other hand, the priority was to maximise the gap between the two sub-sections, while spreading the sample years out over a three year period in an attempt to minimise or avoid the duplicate problem (discussed below).

The Fairfax archives provide relatively clean material consisting of the full text of each article, title, name of author if available, and a set of tags denoting section, sub-section, topic and sub-topic. As well as texts from all the usual newspaper sub-genres, the archives naturally also include all normal newspaper service information. Macalister (2001: 37) mentions this type of text, defining it as "lists – sports results, tv programmes, share prices, weather forecasts, and so on". He chose to exclude it from his work, and I have done the same, on the grounds that it does not contribute much towards a clear picture of the aspects of NZE under investigation. (1) below is an example of service information text, and (2) is an abridged article from the sports sub-genre.

(1) BOX OFFICE
1 (new) Transformers: Dark of the Moon 2 (1) Cars 2 3 (2) Bridesmaids 4 (4) Bad Teacher 5 (3) Green Lantern 6 (8) My Afternoons with Margueritte 7 (5) X-Men: First Class 8 (6) The Hangover Part II … (Dominion Post, 2011)

(2) The Canterbury Crusaders will name their team tomorrow. … The players from outside the region are: Mark Weedon (North Harbour), Norman Maxwell (Northland), Pat Lam (North Harbour), Andy Miller (Bay of Plenty) and Graham Dempster (South Canterbury) … (Sunday News, 1995)

In order to deal with service information text as efficiently and objectively as possible, a Perl script was used to identify and remove articles whose content exceeds predefined thresholds for the number of words consisting of numerical characters, or beginning with a capital letter (*word* defined here as any character or series of characters separated on either side by a white space). The thresholds were set, based on consideration of the material, at 15 per cent for numericals and 45 per cent for capitals. The full text of example (1) exceeds both numerical and capital thresholds, at 19.66 per cent and 56.84 per cent respectively; while (2) exceeds only the capital threshold, at 49.06 per cent (articles exceeding either or both thresholds were removed). Service information-type text accounted for approximately seven million words in the original set of texts in the 1995–98 section, and around 10 million in the 2010–12 section.[6]

[6] Another issue, concerning the presence of duplicate articles in the material, surfaced when test concordances were being carried out with an early version the corpus. It seems that sometime between the mid-1990s and 2012, publications within the Fairfax group began sharing articles to a much greater extent than they had done previously, i.e. the same article is published in more than one Fairfax newspaper, often with slight paraphrasing, or sometimes with none. This results in excessive numbers of duplicate, triplicate and even quadruplicate tokens showing up in the search results. Indeed, duplicates are found in other, more high-profile corpora to a certain extent, but I wanted to eliminate them from the CNZNE as much as possible. The chosen solution was to run all 2010–12 data sharing the same sample year through the

The corpus has been tagged for part-of-speech, using the CLAWS4 tagger (Garside & Smith 1997) – the same software that was used to tag the BNC. At this stage it is not possible to carry out genre-specific searches of the corpus.

To summarise, the CNZNE was designed and compiled for the purpose of researching grammatical patterns in NZE that have thus far been beyond the reach of existing corpora. Comparison with BNC – and indeed COCA (1995–99) – newspaper data is possible. The CNZNE period covers the greater part of the last two decades; a period in which enormous advances in information technology and global communications have taken place, and for this reason alone investigations into change and variation in postcolonial Englishes over this period are justified, and can be expected to produce worthwhile results.

3. Verbs of prevention: background and preliminary results

3.1 Background

In this part of the paper CNZNE data is used alongside BrE and AmE data to investigate the variable use of the complementiser/preposition[7] *from* with verbs of prevention in NZE, BrE, and AmE. The pattern in question is V + NP + (*from*) + *-ing*, as illustrated in examples (3–7) below with the two prototypical verbs of this class, *prevent* and *stop*.

(3) … to drag their careless owners out of range of exploding barbecues or *prevent their foolish children from gulping* down bottles of bleach. (CNZNE, Evening Post, 1996)

(4) Now only bad luck or selectorial whims look capable of *preventing the pair representing* New Zealand at next year's Games in Dublin. (CNZNE, Sunday Star Times, 1997)

(5) Antony was struck with polio as a young boy. This *prevented his pursuing* the musical career he almost certainly would have chosen. (CNZNE, Sunday Star Times, 1997)[8]

---

plagiarism detection software Wcopyfind, which resulted in the detection of around 13,000 articles that were more or less copies of other articles. These were removed, reducing the word count of the 2010–12 section to 58,494,448. The 1995–98 section did not require this step, as the duplicate issue here is minimal.

[7] *From* in prevent-type patterns has been analysed as a complementiser (see e.g. Rosenbaum 1967: 89ff; Aarts 1990), and the label *complementiser/preposition* is used in Mair's work. In the present study however, while recognising the implication of these earlier works, for the sake of simplicity I adopt a permissive attitude towards terminology and use the label *preposition.*

[8] (5) represents the pattern that is commonly referred to in the literature as *poss -ing*: a possessive/genitive NP preceding the gerund. The *poss -ing* pattern in this context is no longer as common as it once was, nowadays being mainly restricted to formal registers, and structures with pronominal NPs or NPs with

(6) … appears to have lost the house he mortgaged to fund a case aimed at *stopping the creationists from* "*peddling* their poppycock to schoolchildren". (CNZNE, Daily News, 1997)

(7) His response was neither clever nor funny but that didn't *stop him flying* to Ohio to deliver his line and make himself look a total jerk. (CNZNE, Southland Times, 1997)

Variable preposition use with these verbs in some varieties of English is an area that is well covered in the literature, the earlier work documenting both quantitative and qualitative aspects of the variation. The present study takes a quantitative approach, using new data to assess the recent trends in NZE. The new NZE data will also shed light on the less frequently used verbs *ban, bar, block, deter, discourage, dissuade, forbid, hinder, prohibit, restrain, save* and *spare*,[9] the details of which have been difficult to assess with the smaller corpora.

Earlier work, based on the Brown/LOB quartet and the WWC (Mair 2009), has found that regarding *prevent*, NZE resembles BrE in making use of both the *from* and *from*-less patterns in varying degrees, while AmE speakers, on the other hand, favour the former, more explicit version almost exclusively. The general situation with *stop* was found to be similar to that of *prevent*, though use of the *from*-less variant with *stop* was somewhat more advanced than *prevent* in BrE. There has been a pronounced shift in BrE speaker preference over the past century in the case of both of these verbs, and it has been noted that "changes in the complementation of *prevent* represent one of the rare instances in which the grammars of the British and American standard have definitely moved apart in the course of the twentieth century" (Mair 2006: 130). Mair confirms that over the course of approximately one century the *from*-less pattern with *prevent*, having been at one time also a possibility in AmE, was all but eradicated from that variety, and gained ground with some speed in BrE.

It has also been noted that "a similar trend seems to be affecting the entire semantic class of what could be called verbs of prevention" (Mair 2006: 132). This observation both justifies research into the wider group of verbs, and provides grounds for some of the assumptions that it will be necessary to make in the present research. These

---

personal reference (Quirk et al. 1985: §16.42), and its presence in the corpora used for this study was indeed found to be minimal. It has also been pointed out that *prevent* and *stop* differ in respect to their compatibility with the *poss -ing* pattern, with *stop* being less likely to license it (Huddleston and Pullum 2002: 1238; Heyvaert et al. 2005: 83–84). More to the point, *poss -ing* does not accommodate *from* with *prevent*, nor any other verbs in this group (Rudanko 1989: 118). All tokens of this type are therefore excluded from the analysis section below.

[9] This group of 12 of the lesser-used verbs of prevention/discouragement was selected from the larger group as being those which may promise at least some degree of variation with *from*. While some small benefits might have arisen from the inclusion of more verbs, the analysis of every verb from e.g. Huddleston and Pullum's (2002, 657) list was seen as casting the net unnecessarily wide.

assumptions are necessary due to the unfortunate fact that this study does not have access to BrE newspaper data from the 2010–12 period to match the recent CNZNE material, and it will therefore not be possible to provide evidence of BrE movement over the past two decades. Mair's view lends weight to the assumption that the BrE shift in favour of the *from*-less variant is widespread – affecting the whole semantic class – and ongoing.

Much of the earlier discussion in the literature surrounding this class of verbs is of a qualitative nature, and the question of factors motivating the variable preposition usage has been approached from several angles (Poutsma 1904: 649; Jespersen 1940: 148; Rohdenburg 2002; Rudanko 2002; Herbst et al. 2004; Dixon 2005; Mair 2009; Sellgren 2010). The subject of much of this earlier discussion lies outside the scope of the present study, but one previously noted structural factor bears mentioning here: Rohdenburg's (1996: 151) Complexity Principle states that:

> In the case of more or less explicit grammatical options the more explicit one(s) will tend to be favored in cognitively more complex environments.

The principle suggests that in the present context, sentences hosting one or more features that contribute to increased cognitive complexity (passives and other types of structural discontinuity, complex NPs etc.) would tend to include the preposition, in order to ease the processing burden. This is indeed the case when the matrix clause is in the passive voice, where, with the majority of verbs in this group, there is little option but to include it in Present-day English[10] (Huddleston & Pullum 2002: 1238). The following tokens, from NZE and BrE illustrate *prevent*, *stop*, and *ban* in the passive.

(8) Every day deaf New Zealanders are *prevented from* doing something that hearing New Zealanders take for granted. (CNZNE, Southland Times, 2012)
(9) The accident report also said the company should be *stopped from* operating out of Porirua Harbour … (CNZNE, Evening Post, 1996)
(10) Kim Basinger is refusing to open an Argentinian disco named after her unless blondes are *banned from* standing near her in a photo session. (BNC, CEK 2986)

Due to this lack of variation, passive tokens are omitted from the results in the data analysis section below. One verb, however, in an interesting departure from the behaviour of the rest of the class, does allow *from* omission with the passive in all three varieties:

(11) Wellington supporters were *spared* having to watch the clinical dissection of their team as it happened. (CNZNE, Dominion, 1995)

[10] Huddleston and Pullum (*ibid.*) note that the *from*-less passive is possible "very marginally with *prevent* and *stop* (?*She was prevented/stopped writing to us*)".

(12) Since it's breakfast, you are mercifully *spared* ordering salad, because salad comes with a metronomically-recited choice of dressing … (BNC, AL3 1518)
(13) With the machine I was *spared* having to relay dissertations. (COCA, NEWS, 1998)

From-less *spare* passives are indeed found in the corpus data, but in very low numbers.[11] For the sake of consistency therefore, they will be omitted along with the rest of the passives.

3.2 Data and methodology

The data used for this section are taken from both sub-sections of the CNZNE (42.6 m. + 58.5 m. words), the BNC newspaper sub-section (9.4 m. words), and the Corpus of Contemporary American English (COCA)[12] (1995–1999) news sub-section (20.4 m. words). As is well known, AmE shows little variation in the use of *from* with *prevent* and *stop* (Mair 2009), and so AmE data is used in this study only as a point of reference. With this in mind, the large numbers of COCA tokens for some verbs were thinned down to more manageable random samples for analysis.[13] With the main focus being NZE and BrE, all available data for these varieties has been used.

In addition, it will be helpful at times to look beyond the boundaries of the corpora mentioned above, and therefore brief reference is made to evidence from the entire BNC and COCA, as well as the latest version of the Corpus of Late Modern English Texts (CLMET3.0)[14]. This extra material, while not matching the genre of the main data set under investigation here, is employed only to provide clues as to the behaviour of certain verbs – in the case of CLMET3.0 – in years past.

The CNZNE and CLMET3.0 were processed with AntConc 3.2.4w concordancing software. The BNC and COCA were accessed through the BNCweb (CQP-Edition) interface and the BYU interface respectively. With corpus word counts being unbalanced, normalised frequencies are given alongside raw numbers in all tables and charts below.

[11] CNZNE 1995–98: two tokens; CNZNE 2010–12: two tokens; BNC: two tokens; COCA: one token.
[12] A genre-balanced corpus comprising 450 million words of American English, covering the period 1990–2012, compiled by Mark Davies at Brigham Young University.
[13] A search for the lemma [prevent] returns 1608 hits. From these, 500 were selected at random and analysed, which resulted in the raw numbers given in Table 3. Normalised frequencies were calculated according to the reduced size of the total available data set (31.09 % of all hits). The same procedure was followed for [stop].[v*] (500 from a total of 4968) and [save].[v*] (500 from a total of 3103). All other verbs produced more manageable numbers of hits.
[14] CLMET3.0 contains around 34.4 m. words of genre-balanced BrE text from the period 1710–1920, arranged in three sub-sections (1710–1780, 1780–1850, 1850–1920). Compiled by Hendrik De Smet, Hans-Jürgen Diller and Jukka Tyrkkö (https://perswww.kuleuven.be/~u0044428/).

3.3 Results

At this point I introduce the synchronic (1990s) data from the three varieties, with Table 3 below containing both the normalised frequencies and the raw frequencies for all predicates in the CNZNE 1995–98, the BNC and COCA. The table lists the two prototypical members of the group, *prevent* and *stop*, first, then the remaining 12 in alphabetical order.

Table 3. Prevention verbs in NZE, BrE and AmE

	CNZNE 1995–98	BNC	COCA
prevent + NP + *-ing*	21.2 (904)	29.5 (278)	0.3 (2)
prevent + NP + *from* + *-ing*	19.5 (832)	23.5 (221)	30.5 (194)
stop + NP + *-ing*	31.1 (1325)	39.9 (376)	0.5 (1)
stop + NP + *from* + *-ing*	13.5 (575)	8.0 (75)	13.6 (28)
ban + NP + *-ing*	0.3 (13)	0.1 (1)	-
ban + NP + *from* + *-ing*	2.6 (110)	5.1 (44)	1.5 (30)
bar + NP + *-ing*	0.1 (6)	-	0.0 (1)
bar + NP + *from* + *-ing*	1.2 (51)	2.3 (22)	5.4 (110)
block + NP + *-ing*	0.2 (10)	0.2 (2)	-
block + NP + *from* + *-ing*	0.3 (14)	0.1 (1)	1.3 (26)
deter + NP + *-ing*	0.3 (11)	0.1 (1)	-
deter + NP + *from* + *-ing*	2.5 (107)	2.4 (23)	0.8 (17)
discourage + NP + *-ing*	0.4 (18)	0.1 (1)	0.0 (1)
discourage + NP + *from* + *-ing*	2.1 (90)	2.1 (20)	4.6 (94)
dissuade + NP + *-ing*	-	-	-
dissuade + NP + *from* + *-ing*	0.8 (36)	0.6 (6)	0.7 (15)
forbid + NP + *-ing*	0.2 (8)	-	0.0 (1)
forbid + NP + *from* + *-ing*	0.5 (21)	1.1 (10)	1.9 (38)

hinder + NP + *-ing*	0.1 (3)	0.2 (2)	-
hinder + NP + *from* + *-ing*	0.0 (1)	0.1 (1)	0.2 (5)
prohibit + NP + *-ing*	0.4 (19)	-	-
prohibit + NP + *from* + *-ing*	0.9 (39)	0.7 (7)	7.9 (161)
restrain + NP + *-ing*	-	-	-
restrain + NP+ *from* + *-ing*	0.4 (16)	0.6 (6)	0.3 (6)
save + NP + *-ing*	0.9 (37)	1.1 (10)	-
save + NP + *from* + *-ing*	2.2 (92)	3.4 (32)	3.0 (10)
spare + NP + *-ing*	0.0 (1)	-	-
spare + NP + *from* + *-ing*	0.1 (4)	0.2 (2)	0.2 (4)

Numbers outside brackets represent normalised frequencies per one million words; bracketed numbers represent raw figures.
COCA normalised frequencies for *prevent*, *stop* and *save* were calculated based on adjusted corpus numbers. See footnote 12 for further details.

As mentioned earlier, all examples of the *poss -ing* pattern and matrix clause passives are excluded from these figures, as there is no possibility of prepositional variation. These omissions, coupled with the genre bias introduced by the use of newspaper corpora, mean that the results given here are not directly comparable with results of earlier work. Despite this, on a general level the profile for NZE that emerges here appears to agree with that suggested in earlier work; NZE exhibits a broadly BrE profile. The avoidance of the *from*-less pattern with all verbs of the group in AmE is also clear.

No tokens of the *from*-less complement were found with *dissuade* and *restrain* in the corpora, despite speculation that with all but one (*keep*) of the verbs in Huddleston and Pullum's (2002: 657) list, "*from*-less variants are possible in theory, and – on the strength of the *prevent* model – expected in British English." (Mair 2006: 132). An extended search covering the entire BNC also failed to turn up any relevant tokens. *Dissuade* and *restrain* will therefore receive no further attention in the analysis section below.

Section 4 below provides the main analysis of the results, where the more recent NZE data is placed alongside the synchronic data. The issue of colonial lag is also raised in Section 4, since the discussion here aims to classify all instances of change according to this framework.

4. Discussion

This section opens with a discussion of colonial lag, the conceptual background against which the subsequent data analysis will be conducted. In the analysis section, the verbs, rather than being treated separately, are treated in groups according to similarity of changes apparent.

4.1 Colonial lag

A number of recent studies on variation among varieties of English make reference to the concept of 'colonial lag'. The term itself was coined in the mid-twentieth century (Marckwardt 1958: 80) to label a phenomenon that had been the subject of comment since the late nineteenth century, and is typically used to refer to older, now archaic features of the original mother country language which are still used in the language of a colony. The term 'colonial innovation' similarly refers to cases of linguistic innovation in the language of the colony. The concept and terminology have been critically discussed and further refined in recent years (see e.g. Görlach 1987; Trudgill 1999 and 2004; Bauer 2002; Dollinger 2006; Hundt 2009). The discussion I focus on here is that of Hundt (2009: 14), who notes that "the dichotomy of 'colonial lag' and 'colonial innovation' – especially when it is applied to features of post-colonial English […] – implies a far too simplistic view of the much more complex patterns and processes of language change". Hundt goes on to restrict colonial lag to one of a wider set of six possible scenarios in the development of grammatical features of postcolonial Englishes. The term 'colonial' English is replaced with 'extraterritorial' English (ETE) (following Lass 1990, cited in Hundt 2009: 15), and 'lag' is superseded by 'conservatism' to avoid any negative and misleading connotations. The six scenarios outlined by Hundt are:

(a) Extraterritorial conservatism: older forms of home country usage are retained in the ETE
(b) Extraterritorial innovation
(c) Truly divergent patterns develop in both varieties
(d) Parallel developments in both varieties
(e) The revival of an older form in either variety
(f) Kick-down developments: what starts out as conservatism gathers momentum to overtake home country usage, and becomes innovation

Hundt's case study concerns BrE and AmE, and indeed she is able to categorise all divergent features she addresses according to the above typology. In the case of the two main varieties of English, there appears to be little need to allow for an additional scenario, one which might prove relevant when dealing with less dominant varieties. In

this regard I suggest a seventh possibility – scenario (g): change under the influence of another main variety. This would be a case in which a variety (e.g. NZE) that has traditionally been seen as being grammatically closer to one main variety (BrE) than the other (AmE), begins to exhibit grammatical tendencies more associated with AmE.[15] With the addition of this extra scenario, the seven-way typology can be applied to the data, and will be referenced throughout the discussion below.

On a final note, it is important, particularly in the context of NZE, to consider Trudgill's view, that a degree of lag is likely to be present, at least for a short time, in the language of a certain type of founder society.[16] All in all there appears to have been a complex blend of forces at work on the emerging NZE dialect during the lifetimes of the first generations of New Zealanders, and straightforward explanations of variation and change are not always going to be possible.

[15] It must be pointed out that in the absence of the comparable 2010–12 BrE data, any potential scenario (g) classifications – or indeed any classifications that require knowledge of current BrE trends – can only really be speculative at this point, since there is no way of knowing whether it is in fact AmE that is providing the model for NZE.

[16] In Trudgill's view, a lag can occur when "there is no common peer-group dialect for children to acquire in first-generation colonial situations involving dialect mixture" (Trudgill 2004: 34). The children therefore reach adulthood speaking more like their parents than is normally the case, and the language of the colony is thus prevented from following the natural paths of development by the space of around one generation. The early New Zealand settlers were just this type of mixed-dialect founder group; a good proportion originated from various parts of the British Isles and spoke a wide variety of dialects, and there was very little or no pre-established English-speaking population in most of the areas in which they settled. Figures denoting the origins of the early settlers, according to census data from the year 1881, are: England 49 %; Scotland 22 %; Ireland 20 %; Australia 7 %; Wales 1 %; North America 1 %. (from McKinnon 1997, cited in Trudgill et al. 2000).

4.2 Prevent and stop

Table 4 presents the data on the two main verbs from the 1995–98 and 2010–12 sections of the CNZNE.

Table 4. *Prevent* and *stop* in NZE (1995–98 and 2010–12)

	CNZNE 1995–98	CNZNE 2010–12
prevent + NP + *-ing*	21.2 (904)	15.8 (925)
prevent + NP + *from* + *-ing*	19.5 (832)	20.3 (1190)
stop + NP + *-ing*	31.1 (1325)	27.1 (1585)
stop + NP + *from* + *-ing*	13.5 (575)	16.2 (948)

Numbers outside brackets represent normalised frequencies per one million words; bracketed numbers represent raw figures.

The situation in NZE with the two main verbs of this group is somewhat surprising, with the data suggesting change in NZE over the CNZNE period in favour of the more explicit pattern with both verbs – a reversal of the same type of shift that has taken place in BrE over the twentieth century. Note that due to the absence of sufficient diachronic data no such movement has actually been documented in NZE, but it is reasonable to assume that a shift very similar to that which BrE has experienced has taken place.

Concerning NZE developments from the perspective of colonial lag, the twentieth-century restructuring of the patterns of sentential complements for *prevent* and *stop* could, up until the late 1900s, have been classed as a lag, since it could be seen as though NZE was following and trailing behind the BrE lead. It now appears that NZE is in the process of reverting to the more explicit pattern, which could arguably be seen as a result of AmE influence. Keeping in mind the important fact that comparable 2010–12 BrE data is unavailable, and thus going on the assumption that the BrE change continues in favour of the *from*-less option, I suggest scenario (g): change under the influence of another main variety, as a classification for *prevent* and *stop*, pending the availability of BrE data.

4.3 Forbid, hinder, prohibit, and save

Table 5 gives the data on these verbs from the 2010–12 section of the CNZNE, alongside the earlier data of Table 3 for comparison.

Table 5. *Forbid*, *hinder*, *prohibit* and *save* in NZE (1995–98 and 2010–12)

	CNZNE 1995–98	CNZNE 2010–12
forbid + NP + *-ing*	0.2 (8)	0.0 (2)
forbid + NP + *from* + *-ing*	0.5 (21)	0.4 (22)
hinder + NP + *-ing*	0.1 (3)	0.1 (3)
hinder + NP + *from* + *-ing*	0.0 (1)	0.1 (4)
prohibit + NP + *-ing*	0.4 (19)	0.2 (10)
prohibit + NP + *from* + *-ing*	0.9 (39)	1.2 (70)
save + NP + *-ing*	0.8 (37)	0.7 (39)
save + NP + *from* + *-ing*	1.9 (92)	2.5 (146)

Numbers outside brackets represent normalised frequencies per one million words; bracketed numbers represent raw figures.

These four verbs are grouped together by virtue of their being the only four – outside of the two main verbs – that are found in CLMET3.0 (in low numbers) with the *from*-less complement. The following tokens illustrate this earlier BrE usage.

(14) … but should inimical circumstances *forbid* me closing with your kind offer … (CLMET 2)
(15) We went to bed rather fatigued, but not so much so as to *hinder* us getting up this morning to mount Skiddaw. (CLMET 2)
(16) Paul *prohibits* all Christians, in every age, celebrating it at all. (CLMET 3)
(17) It'll *save* you having to walk from the Maypole, there and back again. (CLMET 3)

Two of these – *forbid* and *prohibit* – are not found governing *from*-less complements in the BNC data; an absence which may indicate a rise and subsequent fall of that pattern in BrE, and present a counterargument to the assumption that the BrE move toward the *from*-less complement is taking place among all verbs of this class. *Forbid* and *prohibit* are found with *from*-less complements in the CNZNE however, with numbers declining

across the period of the last two decades. The fortunes of the *from*-less complement in NZE could thus be seen as tracing the same course as BrE, but lagging behind by at least one generation. Scenario (a): ETE conservatism (colonial lag) therefore seems an appropriate classification of the progress of *forbid* and *prohibit*.

In the case of *save*, diachronic NZE shows only a slight drop in the *from*-less complement between sub-corpora. The drop being so slight, it could easily be a case of random fluctuation, and thus relatively unremarkable. *Save* does, however, stand out from the rest of the group, in recording the highest use of the *from*-less complement behind the two main verbs *prevent* and *stop* in both NZE and BrE, as well as having been noticed (Mair 2006) to be acting like the prototypical prevention verbs in BrE, showing a clear predilection for the *from*-less pattern. Furthermore, the NZE rise in the more explicit pattern is something to be noted; it may also be instructive to take account of the raw frequencies in comparison to the normalised data in this regard. For the above reasons I would suggest that *save* is likely to be following the same path as *prevent* and *stop* in NZE, that is, an apparent return to the more explicit complement pattern, possibly under AmE influence: scenario (g). As with *prevent* and *stop* in the previous section, this classification is speculative, pending the missing BrE data.

Hinder maintains similar numbers in both halves of the CNZNE, and aside from not being attested in the *from*-less form in AmE, it shows no major discrepancy in numbers across varieties. As noted above, it is attested in CLMET3.0 – where the more explicit complement far outweighs the less explicit – and is still evident, with a total of only three tokens, in the BNC, where the more explicit complement happens to be in the minority. It could be assumed to be following the course of *prevent* and *stop* in BrE, but with numbers overall being too low to reveal much, *hinder* will not be categorised here.

4.4 Bar and spare

Bar and *spare* are both found with the *from*-less complement in the CNZNE data but not the BNC data. In total, NZE yields six *from*-less tokens of *bar*, and one of *spare*, all of which are in the 1995–98 sub-section. Table 6 presents the relevant data.

Table 6. *Bar* and *spare* in NZE (1995–98 and 2010–12)

	CNZNE 1995–98	CNZNE 2010–12
bar + NP + *-ing*	0.1 (6)	-
bar + NP + *from* + *-ing*	1.2 (51)	1.0 (59)
spare + NP + *-ing*	0.0 (1)	-
spare + NP + *from* + *-ing*	0.1 (4)	0.0 (2)

Numbers outside brackets represent normalised frequencies per one million words; bracketed numbers represent raw figures.

Expanding the search to take in the entire BNC and COCA reveals that the *spare* + NP + *-ing* pattern is in fact also to be found in both BrE and AmE, as attested by the following tokens.

(18) I have fetched in enough buckets of coal this afternoon to *spare* you having to carry them in tomorrow. (BNC, fiction, HGE 1179)
(19) His slouch hat *spared* him being cut by the mirror shards raining down around him. (COCA, FIC, 2008)

With AmE generally not conducive to the *from*-less complement, the fact that there are several other examples of this type in COCA – as well as the issue of *from*-less passivisation in NZE, BrE, and AmE (§3.1) – separates *spare* from the other verbs here, and makes it a worthwhile topic for further research. In the meantime, it will not be classified here, due to low numbers and incomplete data.

Bar + NP + *-ing*, also the subject of an expanded search of the BNC and COCA, is found only in NZE – with a fairly convincing six tokens – and in AmE. One token, from NZE, is given below.

(20) Wellington lawyer Tony Ford said the law *barred* relatives claiming exemplary damages after a death. (Evening Post 1996)

The *bar* + NP + *-ing* pattern could be a candidate for scenario (b): ETE innovation. A complication, however, is presented by the fact that *bar* + NP + *-ing* is found only in the 1995–98 section, and not the later section of the CNZNE. Future research will have to determine whether the case for innovation can be upheld.

4.5 Ban, block, deter and discourage

The remaining four verbs are all found with the *from*-less complement in the modern data. Table 7 below provides the relevant numbers.

Table 7. *Ban*, *block*, *deter* and *discourage* in NZE (1995–98 and 2010–12)

	CNZNE 1995–98	CNZNE 2010–12
ban + NP + *-ing*	0.3 (13)	0.2 (10)
ban + NP + *from* + *-ing*	2.4 (110)	2.8 (162)
block + NP + *-ing*	0.2 (10)	0.1 (5)
block + NP + *from* + *-ing*	0.3 (14)	0.7 (43)
deter + NP + *-ing*	0.3 (11)	0.2 (14)
deter + NP + *from* + *-ing*	2.5 (107)	2.1 (124)
discourage + NP + *-ing*	0.3 (18)	0.1 (8)
discourage + NP + *from* + *-ing*	2.1 (90)	2.1 (124)

Numbers outside brackets represent normalised frequencies per one million words; bracketed numbers represent raw figures.

Absence of the *from*-less complement from CLMET3.0 would indicate that the use of this complement with these verbs is a twentieth-century addition to both BrE and NZE. The *from*-less pattern with *ban*, *block*, *deter*, and *discourage* in NZE, while being equal to or – in most cases – higher than that of BrE in the 1990s, shows a decline (at least with *ban*, *block*, and *discourage*) over the period between the NZE sub-corpora. This appears to be a situation where a pattern has emerged at roughly the same time in both varieties – although admittedly it is not possible to discern the progression timeline in each variety from the current data – and is now beginning to drop out of use in at least one of those varieties (comment cannot be made on recent BrE developments). Classification under Hundt's type (d) change: parallel developments in both varieties, would seem appropriate.

5. Conclusion

Bigger is not always better, but when focus is on the lesser-used predicates in a language, corpus size becomes all-important. Various earlier studies mention target predicates, then, due to the lack of suitably large corpora, fail to provide conclusive results. To address this, a diachronic newspaper corpus was compiled and put to the test in a case study on the complementation of 14 *prevent*-type verbs. The results show that the creation of the corpus was justified (though admittedly two of chosen verbs were still not well represented), and the literature on NZE will benefit from it.

Certain aspects of the results of the case study go some small way towards strengthening the justification for spending more time on investigating New Zealand grammar, as it sits in relation to the main varieties of English. The diachronic overview of the last two decades of NZE newspaper language given here allows qualification of the comment that, in regard to this area of the grammar, "no rapid recent "Americanization" of usage can be observed" (Mair 2009: 263): it does appear that, in the case of *prevent* and *stop* – and at least one of the minor verbs from the group – NZE usage is beginning to drift away from the BrE model, possibly under the influence of AmE, though it would be rash to apply the term "Americanization" at this stage. This was not found to be the case for all verbs of this semantic group; in fact one important finding of this study was that while some – *prevent*, *stop* etc. – appear to be drifting in one direction, many are still apparently on the same path as their BrE counterparts. Thus the assumption that all verbs of prevention are following the same trend – movement in the BrE direction – cannot be upheld. It is important to reiterate the caveat that, until these results are cross-checked against the relevant BrE data, they must remain speculative at this point. The present study, therefore, opens up possibilities for future research, and provides justification for the creation of a comparable corpus of BrE newspaper data.

The present study supports the view, expressed in much of the literature on extraterritorial Englishes, that while there often is a type of lag evident in one area of the grammar of an extraterritorial variety, belated development is only a small part in the more complex system of change taking place in the different varieties of English worldwide. Hundt suggests six possible scenarios for extraterritorial English change, I have added one more to that list, and four of these scenarios were applied to 10 of the verbs analysed here.

References

Primary sources

BNC – The British National Corpus (http://www.natcorp.ox.ac.uk/)

COCA – The Corpus of Contemporary American English (http://corpus.byu.edu/coca/)

CLMET3.0 – The Corpus of Late Modern English Texts, version 3.0 (https://perswww.kuleuven.be/~u0044428/clmet3_0.htm)

CNZNE – The New Zealand English Newspaper Corpus

Secondary sources

Aarts, Bas. 1990. *Prevent*-type verbs in a GB-framework. *UCL Working Papers in Linguistics* 2: 147–164.

Bauer, Laurie. 1989b. The verb *have* in New Zealand English. *English World-Wide* 10(1): 69–83.

Bauer, Laurie. 2001. Some verb complements in New Zealand English. *New Zealand English Journal* 15: 29–34.

Bauer, Laurie. 2002. *An Introduction to International Varieties of English.* Edinburgh: Edinburgh University Press.

Bauer, Laurie & Winifred Bauer. 2000. Nova Zelandia est omnis divisa in partes tres. *New Zealand English Journal* 14: 7–17.

Calude, Andreea & Paul James. 2011. A diachronic corpus of New Zealand newspapers. *New Zealand English Journal* 25: 35–52.

Davies, Carolyn & Margaret Maclagan. 2006. Maori words – read all about it: testing the presence of 13 Maori words in four New Zealand newspapers from 1997 to 2004. *Te Reo* 49: 73–99.

Dixon, R. M. W. 2005. *A Semantic Approach to English Grammar.* Oxford: Oxford University Press.

Dollinger, Stefan. 2006. The modal auxiliaries *have to* and *must* in the Corpus of Early Ontario English: Gradient change and colonial lag. *Canadian Journal of Linguistics* 51(2): 287–308.

Garside, Roger & Nicholas Smith. 1997. A hybrid grammatical tagger: CLAWS4. In Roger Garside, Geoffrey Leech & Tony McEnery. (eds.) *Corpus Annotation: Linguistic Information from Computer Text Corpora*. Longman, London. 102–121.

Gordon, Elizabeth & Marcia Abell. 1990. 'This objectionable colonial dialect': Historical and contemporary attitudes to New Zealand speech. In Allan Bell & Janet Holmes (eds.), *New Zealand Ways of Speaking English*. Clevedon: Multilingual Matters Ltd. 21–48.

Görlach, Manfred. 1987. Colonial lag? The alleged conservative character of American English and other 'colonial' varieties. *English World-Wide* 8: 41–60.

Hay, Jennifer, Margaret Maclagan, & Elizabeth Gordon. 2008. *New Zealand English.* Cambridge: Cambridge University Press.

Herbst, Thomas, David Heath, Ian F. Roe, Dieter Götz & Michael Klotz. 2004. *A Valency Dictionary of English.* Berlin: Mouton de Gruyter.

Heyvaert, Lisbet, Hella Rogiers & Nadine Vermeylen. 2005. Pronominal Determiners in Gerundive Nominalization: A ''Case'' Study. *English Studies* 86(1): 71–88.

Huddleston, Rodney & Geoffrey Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Hundt, Marianne. 1998. *New Zealand English Grammar – Fact or Fiction?* Amsterdam: John Benjamins.

Hundt, Marianne, Jennifer Hay & Elizabeth Gordon. 2008. New Zealand English: morphosyntax. In Kate Burridge & Bernd Kortmann (eds.) *Varieties of English: The Pacific and Australasia*. Berlin: Mouton de Gruyter. 305–340.

Hundt, Marianne. 2009. Colonial lag, colonial innovation or simply language change? In Günter Rohdenburg & Julia Schlüter (eds.), *One Language, Two Grammars?* Cambridge: Cambridge University Press. 13–37.

Jespersen, Otto. 1940. *A Modern English Grammar on Historic Principles.* 7 vols. Copenhagen: Munksgaard.

King, Michael. 2003. *The Penguin History of New Zealand*. Auckland and London: Penguin Books.

Macalister, John. 2001. Introducing a New Zealand newspaper corpus. *New Zealand English Journal* 15: 35–41.

Macalister, John. 2006. The Mäori presence in the New Zealand English lexicon, 1850–2000. *English World-Wide* 27: 1–24.

Mair, Christian. 2006. *Twentieth Century English*. Cambridge: Cambridge University Press.

Mair, Christian. 2009. Infinitival and gerundial complements. In Pam Peters, Peter Collins & Adam Smith (eds.), *Comparative Studies in Australian and New Zealand English*. Amsterdam: John Benjamins. 263–276.

Marckwardt, Albert. 1958. *American English.* New York: Oxford University Press.

Mukherjee, Joybrato & Sebastian Hoffmann. 2006. Describing verb-complementational profiles of New Englishes: A pilot study of Indian English. *English World-Wide* 27(2): 147–173.

Poutsma, Hendrik. 1904. *A Grammar of Late Modern English: Part 1 - The Sentence*. Groningen: P. Noordhoff.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language.* London: Longman.

Rohdenburg, Günter. 1996. Cognitive complexity and increased grammatical explicitness in English. *Cognitive Linguistics* 7(2): 149–182.

Rohdenburg, Günter. 2002. Processing complexity and the variable use of prepositions in English. In Hubert Cuyckens & Günter Radden (eds), *Perspectives on Prepositions.* Max Niemeyer: Tübingen. 79–100.

Rosenbaum, Peter. 1967. *The Grammar of English Predicate Complement Constructions.* Cambridge: MIT Press.

Rudanko, Juhani. 1989. *Complementation and Case Grammar: A Syntactic and Semantic Study of Selected Patterns of Complementation in Present-day English.* Albany, New York: State University of New York Press.

Rudanko, Juhani. 2002. *Complements and Constructions: Corpus-based Studies on Sentential Complements in English in Recent Centuries.* Lanham MD: University Press of America.

Sellgren, Elina. 2010. *Prevent* and the battle of the *–ing* clauses. In Ursula Lenker, Judith Huber & Robert Mailhammer (eds.), *English Historical Linguistics 2008: Volume 1: The history of English verbal and nominal constructions*. Amsterdam: John Benjamins. 45–62.

Trudgill, Peter. 1999. A window on the past: "Colonial lag" and New Zealand evidence for the phonology of nineteenth-century English. *American Speech* 74(3): 227–239.

Trudgill, Peter. 2004. *New Dialect Formation.* Edinburgh: Edinburgh University Press.

Trudgill, Peter, Elizabeth Gordon, Gillian Lewis & Margaret MacLagan. 2000. Determinism in new-dialect formation and the genesis of New Zealand English. *Journal of Linguistics* 36(2): 299–318.

Turner, G. W. 1966. *The English Language in Australia and New Zealand.* London: Longman.

PUBLICATION II

Aspects of the use of the transitive *into -ing* pattern in New Zealand English

Paul Rickman & Mark Kaunisto

Changing Structures: Studies in Constructions and Complementation, 71–88
https://doi.org/10.1075/slcs.195

Publication reprinted with the permission of the copyright holders.

Aspects of the Use of the Transitive *into -ing* Pattern in New Zealand English

Paul Rickman & Mark Kaunisto

University of Tampere

Abstract

This chapter investigates the use of the transitive into -ing pattern in one of the world's youngest dialects of English, New Zealand English (NZE). We draw on evidence collected from a diachronic NZE corpus of newspaper English, which comprises 100 million words from the years 1995–98 and 2010–12. We categorise the matrix verbs according to semantic orientation, and provide comment on any change evident between the two time-frames of the corpus, and we identify several previously undocumented matrix verbs in this pattern. Comparison is drawn to American English in an attempt to determine similarities in the use of this surprisingly productive pattern.

Keywords: New Zealand English; complementation; transitive into -ing pattern; corpus linguistics

1. Introduction

The transitive *into -ing* pattern, exemplified by the sentence *Philip tricked Phyllis into paying for dinner*, has been given a thorough treatment in recent years (see, e.g. Davies 2012; Kim and Davies 2015; Rudanko 2000; 2002; 2005; 2006; 2011; 2015a; 2015b; Rudanko and Luodes 2005; Wulff et al. 2007).[1] This surge in attention parallels a rise in the general frequency of the pattern in the main varieties of English; something that has been convincingly demonstrated in the literature (see, e.g. Rudanko 2005, and Davies 2012 for figures). Despite all the recent coverage, the usage of the pattern in the

[1] As regards terminology, here we use 'transitive *into -ing*', as coined by Juhani Rudanko.

young dialect of New Zealand English remains, to the best of our knowledge, unexplored. Should the best of our knowledge prove faulty, however, the diachronic approach afforded by a corpus created specifically for studies in NZE requiring large amounts of data, such as the present chapter, will at least offer important new insights into the use of the transitive *into -ing* pattern in this lesser studied variety of English.

With this study we aim to address the questions of how the frequency in NZE generally compares to other dialects; whether there is any apparent change in the qualitative or quantitative uses of the pattern in NZE over the time frame investigated, and whether NZE shows any innovative aspects of its use. With regard to the last point, the expectation that some kind of innovative usage will be found almost goes without saying, as each of the southern hemisphere Englishes is well known for lexical idiosyncrasies (Turner 1966: 112ff.; Bauer 1994; Hay et al. 2008: 65ff.), and NZE speakers are almost certain to add some local flavour to such a productive pattern.

This chapter is structured as follows: Section 2 introduces the pattern, covers the necessary theoretical aspects of transitive *into -ing*, and discusses the relevant points that have arisen from earlier research. Section 3 presents the corpus used in this study and methodological issues; Section 4 comprises the results and discussion, and Section 5 provides the summary and concluding remarks.

2. Background and earlier research

As earlier research has noted, a compelling property of the pattern is its apparent ability to provide fertile ground for new and innovative language use, and the array of matrix verbs documented with the transitive *into -ing* pattern is striking, to say the least. The following example, taken from the Corpus of New Zealand Newspaper English (CNZNE), illustrates the pattern.

(1) They force the roading authorities into providing expensive new motorways. (CNZNE *Dominion* 1995)

The pattern manifests certain syntactic and semantic properties that require discussion at this point. Syntactically, it should be noted that in sentence (1), the matrix verb *force* selects three arguments: the higher subject NP, the higher object NP, and the lower clause introduced by the preposition *into*. The lower clause contains a null subject NP, an understood argument, which, in an expanded syntactic representation, would be represented by the null element PRO. The postulation of a null subject, while not uncontroversial, is in accordance with much of the literature on argument structure today, and is found in recent mainstream influential grammars of English, as this except from Huddleston & Pullum (2002: 65) shows: "[m]ost non-finite clauses have no overt subject, but the interpretation of the clause requires that an understood subject be retrieved from the linguistic or non-linguistic context" (see also Brinton & Brinton

2010: 275). PRO obtains its reference from the object of the higher clause, *the roading authorities,* and the sentence, therefore, is one of object control.

Verbs occurring in this pattern have been found to co-occur in some other patterns (Rudanko 2006; Kim and Davies 2015), the main two being the V + NP + *to*-infinitive pattern, which hosts three-argument verbs such as *force, persuade* and *pressure,* as in *They force the roading authorities to provide expensive new motorways*; and the V + NP pattern, which hosts two-argument verbs such as *fool, manoeuvre* and *tempt*, as in *They fool the roading authorities*. The third and less central pattern likewise hosts two-argument verbs, but in this case the second argument is not of the same semantic type as the previous pattern. The difference between the semantic roles of the NP object permitted with verbs like *fool/tease/tempt* versus those permitted with verbs like *talk*, is shown by comparing the awkward sounding (?)*they talked the roading authorities* with the more acceptable *they talked business*. Verbs such as *talk* form the third "very limited set of verbs" (Rudanko 2006: 315).

Semantically, it is important to note that the arguments of sentence (1) are assigned the semantic roles of Agent, Patient, and Goal, for the higher subject, higher object, and lower clause, respectively. Examples of superficially similar patterns are plentiful in the data, such as sentence (2); these do not fit the bill semantically or syntactically: the semantic role of the higher object has to be seen as Instrument, rather than Patient; and the understood subject of the lower clause obtains its reference from the subject of the higher clause, making this a case of subject, rather than object control. Tokens of this type fall outside the scope of the present study.

(2) She had put her life into building the bush reserve. (CNZNE *Timaru Herald* 1997)

A connection between the transitive *into -ing* pattern and the Construction Grammar framework (Goldberg 1995) has been postulated in the literature; the pattern in this context is seen as a subtype of the caused motion construction. This helps to explain why so many unusual matrix verbs are so readily used in this pattern, and how speakers are able to understand the meaning without difficulty (for the original proposal of this, see Rudanko 2000: 81ff.; and for a recent detailed argument for the caused motion subtype, see Kim and Davies 2016). While we recognise the significance of the Construction Grammar approach, it will not be a focus of this analysis.

The history of the pattern has been traced with the help of large historical corpora. Using the Corpus of Historical American English (COHA), Davies (2012) provides a timeline for the development in AmE, and suggests that the pattern began life as an analogy of the *into* NP pattern, with the strongly nominal verb *being* first entering into that pattern, as in *somehow I inadvertently willed it into being*, and providing a foothold for other verbs. Davies argues that "[r]ather than have the construction created "ex nihilo", it apparently started where [V NP into V -ing] would be least noticeable—where [into V-ing] could also be analysed as a noun, as with the pre-existing [into N]. And then once the [into V-ing] construction was firmly "established" in about the 1850s, the percentage of *being* decreased markedly" (Davies 2012: 166). Rudanko (2015a) traces the development in BrE using the Corpus of Late Modern English Texts

(CLMET) version 3.0, and finds numerous examples of the fully fledged transitive *into -ing* appearing as early as the mid 18th century, along with a few of the *into being* examples, so it seems clear that the pattern was already firmly established in BrE prior to the 19th century. This supports earlier work by Rudanko on the history of the pattern in BrE (e.g. Rudanko 2000).

It has also been observed that verbs occurring in the transitive *into -ing* pattern express negative causation, and that the construction has a negative semantic prosody. Wierzbicka (1998: 125) stated that "the set of main verbs that can be used in this construction is quite limited and gives a clear clue to this construction's meaning. Thus, one can not only "talk" someone "into" doing something, but also "trick" them, "manoeuver" them, or "push" them. On the other hand, one cannot "encourage" or "induce" someone "into" doing anything. […] the *into* construction takes verbs that either imply, or are at least compatible with, the idea of manipulation (tricking, manoeuvring, and the like)" (see also Hunston and Francis 2000: 102 on this point). On the other hand, Rudanko (2006), citing the influence of the Great Complement Shift (Rohdenburg 2006) – a major feature of which sees *-ing* complements spreading at the expense of *to*-infinitive complements – points out a number of verbs of unflavoured, rather than negative, causation that are found in the pattern. The argument cites the verbs *impel, induce, influence, lead, motivate, prompt* and *stimulate*, all of which are also compatible with the *to*-infinitive complement (Kim and Davies 2015: 25f. also address this issue).

Verbs found to occur in the pattern have been categorised under various different semantic groups by different authors. One such grouping, comprising seven categories, is given below, with example verbs provided in italics.

Deceiving or tricking – *deceive, trick, bamboozle, con*
Forcing or pressuring – *force, pressure, arm-twist, coerce*
Bullying or frightening – *bully, frighten, harass, panic*
Verbal persuasion – *cajole, inveigle, persuade, sweet-talk*
Luring or tempting – *lure, tempt, entice*
Coaxing or stimulating – *coax, stimulate, prod, stir*
Other – *annoy, bounce, hijack, rush* (adapted from Rudanko 2000: 83)

With regard to the semantics of the matrix verb, it has been suggested that there is variation to be found between BrE and AmE in the type of verb used most frequently with the transitive *into -ing*. Wulff et al. (2007) suggest that BrE speakers tend to favour verbs denoting physical force, such as *force, bully* and *push*, while AmE speakers lean towards more orally persuasive verbs, such as *talk, persuade* and *coerce*. This suggestion, based on the results of statistical analyses of newspaper data, has since been reinforced by Kim and Davies (2016), on the basis of more stratified corpora.

We close this section with a brief outline of the target dialect of this study.[2] As far as native speaker varieties of English are concerned, New Zealand English is one of the world's youngest, with the formation period 1840–1880 (Trudgill et al. 2000).[3] It was during this time that settlers, mainly European, the largest single group of whom embarked in various parts of England, arrived in New Zealand in significant numbers. The large influx is due mainly to settlement schemes, and was additionally helped along by the discovery of gold. Prior to this, New Zealand was populated by fewer than 100,000 Māori and little more than several hundred Europeans.[4] From 1840 then, new dialect formation processes began to take place (Gordon et al. 2004; Trudgill 2004), and by the early 20th century a distinctive new dialect was emerging, and various, mainly accent-related, aspects of it were being commented on, often unfavourably, in newspapers and other forums (for a good summary of early attitudes to NZE, see Gordon and Abell 1990).

The phonological and phonetic features of NZE have received much attention since those early years, but it has only been in recent decades that NZE has been seriously examined with the aim of documenting any distinguishing grammatical features, and it is on this area that the present chapter seeks to shed further light. The nature of the pattern also inevitably takes us across the borders of the fields of vocabulary and word-formation – areas in which NZE creativity, to repeat a point made in the introduction, is known to excel.

The following section continues by outlining the practical and methodological aspects of the study, and discussing the corpus that provided the data.

3. Data and methods

Our data come from the recently compiled Corpus of New Zealand Newspaper English (CNZNE), a 100-million-word, part of speech tagged corpus comprising two sub-sections, 1995–98 (42.6 million words), and 2010–12 (58.5 million words).[5] Material from a total of 13 metropolitan and provincial newspapers contributes to the total word count. The 1995–98 sub-section is structured so as to allow comparison to the British National Corpus (BNC) newspaper sub-section (9,412,174 words). A detailed account

[2] The present chapter represents part of a larger research project being carried out by the first author into various aspects of complementation in NZE. The aim is to document areas in which NZE differs from other varieties of English in this respect, and to discern to what extent, if at all, the phenomenon of colonial lag (Marckwardt 1958: 80; Görlach 1987; Hundt 2009) is relevant.

[3] Compare Australian English (1800–1840), and South African English (1820–1860) (Trudgill et al. 2000: 302)

[4] Precise figures are not available for the Māori population, but Belich (1996:178) estimates the number to have been around 86,000 in 1769, the year when Europeans first made landfall, and around 70,000 in 1840. As for Europeans, the same source (*ibid.*:198) states that they numbered around 300 in 1830, and 2,000 in 1840.

[5] The CNZNE was compiled by the first author of the present paper. At the time of writing, it is not freely available.

of the compilation process of the CNZNE can be found in Rickman (2017), but a few details here are in order.

The CNZNE was compiled using material from the archives of Fairfax Media, accessed via The Knowledge Basket news and information archive service.[6] Fairfax is one of the largest media conglomerates in Australasia, and, within New Zealand, it controls publications that span the length of the country. In compiling the corpus, care was taken to ensure that major metropolitan newspapers from as many of the four main centres as possible were included (New Zealand's four main centres[7] are traditionally understood as being Auckland, the capital Wellington, Christchurch and Dunedin), as well as provincial papers from as many different regions of the country as possible.[8] The obvious reason for this is that while it is generally acknowledged that NZE is a great deal more homogeneous than BrE and AmE, there is still known to be some regional variation (Turner 1966: 163ff.; Bauer & Bauer 2000; Hay et al. 2008: 95ff.; Calude & James 2011).

As well as texts from all the usual newspaper sub-genres, material obtained from the Fairfax archive also includes service information. Macalister (2001: 37) discusses service information in newspapers, defining it as "lists – sports results, tv programmes, share prices, weather forecasts, and so on". He chose to exclude it from his corpora, and the same has been done with the CNZNE, on the grounds that it does not help all that much in giving a clear picture of the aspects of NZE under investigation in the research project that the corpus was compiled for.

Naturally, corpus size is an all-important factor with a study of this type, and the entire 100 million word corpus was used. By today's standards, however, even 100 million words, for some research purposes, lies towards the more modest end of the scale, so in order to ensure that we were getting all that the corpus had to offer, we opted to use the simple search string [into *ing], using the concordancing software AntConc 3.2.4w, and then manually sort the relevant tokens from the roughly 5,600 total hits. The main reason for this all-inclusive approach is that, as is well known, no corpus tagging software can provide 100 per cent accurate results, and it was noticed during the preliminary search phase of this project that a search string designed to find words tagged as verbs followed by 0-4 intervening words, followed by into v_*ing, for example, was not going to achieve perfect recall.

Furthermore, it was noted by Rudanko (2015b: 81ff.) as part of his search for what he terms "radically innovative" matrix verbs, that words not traditionally recognised as performing a verbal function (i.e. words not listed as verbs in the OED,

[6] www.knowledge-basket.co.nz

[7] Many thanks to an anonymous reviewer, who pointed out that the comparatively smaller cities of Hamilton and Tauranga are now seen by Radio New Zealand and Television New Zealand as main centres, as they have been included in weather forecasts of the main centres for at least a decade now. Thus the 'four main centres' notion may be a thing of the past. It is still in use, however, as can be seen on the pages of the Government-managed online encyclopedia of New Zealand, Te Ara <http://www.teara.govt.nz>, and we continue to use the term here.

[8] Rickman (2017) also addresses the issue of multiple versions of a single article that can surface among the various publications under single ownership, like the publications under Fairfax, and the problems this causes in corpus compilation.

according to Rudanko's criteria) may nonetheless function as matrix verbs in the transitive *into -ing* pattern. Therefore, a search string aimed at identifying verbs may well miss all such radically innovative examples. With our aim of identifying innovative NZE usage, it is of course vital to retrieve all such tokens.

The coordination of matrix verbs, as in *parents have to push and cajole their children into doing their homework*, is reasonably common in the data, and needs to be resolved in a systematic way. We chose not to record both verbs occurring in such structures, but rather to follow Rudanko (2005: 175) in recording only the rightmost verb, in this case *cajole*, due to its proximity to the complement. Only in the case of established verbal units, such as *name and shame*, which occurs once in our data, were the verbs not separated; these were recorded as one verb.

Although the connection between the preposition and the *-ing* clause is tight, it is nonetheless possible to find material inserted here, and, in addition to the primary search string given above, additional searches were run to find such tokens. Examples of the results are given below.

(3) a. …I find myself forced into not panning Jon Bon Jovi's latest effort. (CNZNE *Sunday News* 1997)
b. …the Palestinians will be pounded into basically giving up on their rights and leaving. (CNZNE *Dominion Post* 2011)

The most common word – and it is only ever a single word in our data – occurring in this position is *not*, which is found eight times in the entire corpus. Other words in this slot include: *again, almost, basically, illegally, regularly*. In total 21 tokens of this type were retrieved.

One aim of this analysis is to identify new matrix verbs in the target pattern, and to this end we cross-checked our results against those of a selection of prominent sources on the same topic – Bridgeman et al. (1965), Francis et al. (1996), Kim and Davies (2016),[9] and Rudanko (2005). Any verbs not documented in any of these sources were further cross-checked against the relevant entries in the *Oxford English Dictionary* (OED), and the *New Zealand Oxford Dictionary* (NZOD). If the examples supplied in these two reference works did not include the verb used in the transitive *into -ing* pattern, the verb was then deemed eligible to be put forward as an example of innovative usage. The results of this process are presented in the following section, along with other results.

[9] We would like to thank Mark Davies for giving us access to the data that he and Jong-Bok Kim collected for their 2016 study, since only a small percentage of the large number of verbs they identified were listed in the actual paper.

4. Results

This section presents the main observations on the uses of the transitive *into -ing* pattern in the two parts of the CNZNE. First we will pay attention to the overall frequencies in the use of the pattern and the most common matrix verbs used with the pattern, followed by a section observing verbs found in the data which have been previously unattested as selecting the complement pattern.

4.1 Type and token frequencies of the transitive into -ing pattern in the CNZNE

On the whole, the transitive *into -ing* pattern was quite frequent in the corpus data examined, with a total of 1,485 relevant tokens found. The token and type frequencies of the matrix verbs are given in Table 1.

Table 1. The frequencies of transitive verbs selecting *into -ing* complements in two sections of the Corpus of New Zealand Newspaper English (pmw = per million words).

CNZNE 1995-98 (42.6 million words)	**CNZNE 2010-12 (58.5 million words)**
635 tokens (14.9 pmw)	850 tokens (14.5 pmw)
134 different matrix verbs	145 different matrix verbs
190 different matrix verbs in total	89 matrix verbs common to both sets

As can be seen in Table 1, the normalized token frequencies of the pattern in the two periods were fairly similar, amounting to 14.9 instances per million words in 1995–98, and 14.5 per million words in the 2010–12 data. The type frequencies of the relevant transitive matrix verbs in the two periods were likewise close to one another, with 134 verbs found in the earlier period, and 145 verbs found in the latter. There was some variation as regards the sets of the matrix verbs with *into -ing* complements in that 45 verbs were unique to the 1995–98 period, 56 verbs were unique to the 2010–12 period, and 89 verbs were found in both sets.

Considering the normalized frequencies of the matrix verbs in the CNZNE data, some interesting observations can be made on them with regard to the findings in Davies (2012) and Rudanko (2015b). Firstly, in his study on the diachronic developments in the use of the pattern in COHA, Davies (2012) noted that in 2000–09 the normalized frequency of the pattern in the AmE data was 13.9 instances per million

words, which is remarkably close to the frequency in the NZE data. Another point perhaps worth observing has to do with the slight decrease in the frequencies of the pattern in the two sets of NZE data. Although the drop in the normalized frequencies (14.9 to 14.5 instances per million words) may not be striking as such, in his study of the pattern in the TIME corpus, Rudanko (2015b: 82–83) also perceives a decline in the use of the pattern from the 1990s to the 2000s, with the 1970s being the peak period in its popularity.[10] However, examination of data with a longer time span might be needed to conclude whether there is indeed a declining trend in the frequency of the transitive *into -ing* pattern in NZE.

The most commonly occurring matrix verbs taking an *into -ing* complement were largely the same in the two sets of data. In fact, the top 15 verbs were the same in both periods (namely *bully, coax, coerce, con, draw, dupe, fool, force, lure, mislead, pressure, provoke, push, talk,* and *trick*), with only slight differences in the order of frequency of the verbs in the two periods. Since the difference in time between the two periods studied is roughly that of half a generation, the similarity among the most prominent matrix verbs could be expected.

As regards the semantic characteristics of the verbs, the same major categories can be perceived as in the earlier studies of the pattern. There are instances where the verb involves different forms of pressure; for example, intimidation or surprise (e.g. *badger, blackmail, browbeat, jolt, kickstart, panic, shock*), or the use of physical force (e.g. *beat, bludgeon, bulldoze, stampede, thrust*). Persuasion (usually verbal) is another prominent semantic group (e.g. *cajole, convince, court, flatter, seduce, smooth talk, woo*), as well as deception (e.g. *bamboozle, confuse, deceive, hoodwink*). A less prominent semantic group in the data included words involving the idea of guidance or inspiration (e.g. *coach, encourage, enthuse, inspire, guide, lead, steer, support, usher*), which also serve as examples of verbs that carry a positive rather than negative sense.

4.2 Innovative usages

The productivity of the pattern is evident also on the basis of the findings in the CNZNE data, as altogether 19 innovative usages were found in the material, that is, 19 matrix verbs which have not been previously observed as having been used in the transitive *into -ing* pattern. In their study, Kim & Davies (2016) identified 335 innovative usages of the pattern in present-day AmE and BrE corpora (the Corpus of Contemporary American English [COCA], BNC, and the US and UK sections of the Corpus of Global Web-based English [GloWbE]) containing in total as many as 1.32 billion words. From

[10] Interestingly enough, the normalized frequencies of the transitive *into -ing* pattern in the last couple of decades of the TIME Corpus are notably higher than those found in COHA and the CNZNE, as the frequencies in the TIME Corpus were 29.5 and 23.1 instances per million words in the 1990s and the 2000s, respectively (Rudanko 2015b: 82). It must be noted here that the three corpora are different as regards the text types or registers represented, and it is possible that the different frequencies observed in the TIME Corpus are reflective of the use of the pattern in magazine texts.

that perspective, finding innovative usages of the pattern in the NZE data is further testimony to how productive the pattern is today.

A particularly interesting case among the 19 innovative matrix verbs is the verb *monster*, the single instance of which is presented in example (4):

(4) … Mr Brownlee – who spent the week beavering away with officials while Mr Key was in Australia – had been monstered into bringing forward an announcement by the Beehive's ninth floor … (CNZNE *Dominion* 2011)

The use of *monster* as a verb is seen only in some varieties of English, and the OED and the NZOD list the verbal use with the regional label "Australian". This verb can thus be regarded as a genuinely new entry among the verbs selecting the *into -ing* pattern, arising from the AusE and NZE varieties. Another verb in this respect is *heavy*, which is listed in the OED as a verb, but not with argument patterns that would match that of the transitive *into -ing* pattern. The NZOD does include the word as a transitive verb, but no example is given of the verb with an *into -ing* complement. The example of the verb in the CNZNE is given in example (5).

(5) VEHICLE franchise war has been declared in Taranaki, with a major car manufacturer heavying its New Plymouth dealership into relinquishing an opposition franchise. (CNZNE *Daily News* 1997)

Of the 19 innovative matrix verbs found in the CNZNE data, three verbs comply with Rudanko's (2015b) definition of a "radically innovative" verb, i.e. they were not listed in the OED as verbs: *hard-talk, peer pressure,* and *turbo-shock*, illustrated in examples (6a)–(c):

(6) a. In a statement at James' hearing he said he was peer pressured into taking cocaine on a night out with friends and said it was a mistake. (CNZNE *Waikato Times* 2010)
b. The original version was popular Down Under but Sabbath apparently had to be hard-talked into playing it live because they didn't like having to use a piano and Mellotron (a type of keyboard) on stage. (CNZNE *Waikato Times* 2010)
c. … the experience of being hammered for pretty much eight months in a row seems to have turbo-shocked ministers into pulling up their socks in recent weeks. (CNZNE *Press* 2012)

The verbs *peer pressure, hard-talk,* and *turbo-shock* can also be regarded as innovative with respect to their structural property of consisting of two parts. In the case of *peer pressure* and *hard-talk*, the verbal use is probably a result of conversion or zero-derivation from a corresponding compound noun or an adjective + noun compound, while in *turbo-shock* the combining form *turbo* is attached to the base verb *shock*. In fact, morphological complexity is found in half of the innovations found in the corpus,

as verbs involving compounding, conversion, prefixation, or the use as a phrasal verb are seen in seven other innovations of the pattern, namely *fast-track, force-feed, pressure-cook, stir up, re-energise, reinvigorate,* and *wrongfoot*. Here are illustrations of these verbs:

(7) a. Fast-tracking young graduates into teaching will result in disaster for many. (CNZNE *Dominion Post* 2011)
b. A European political elite had force-fed nations into accepting the euro, and despite multi-lateralism, it was still a world of nation states, he said. (CNZNE *Southland Times* 2012)
c. ARE horses getting their jumping tickets too easily these days? Are horses being pressure-cooked into jumping? A leading trainer posed these questions this week. (CNZNE *Evening Post* 1996)
d. Competition is a wonderful thing for stirring up local bodies into realising on assets. (CNZNE *Evening Standard* 1996)
e … they are designed to revitalise and re-energise mature jobless people into wanting to get back into the work-force. (CNZNE *Press* 2011)
f. I want us to reinvigorate our families into taking up their full responsibility for their own. (CNZNE *Dominion Post* 2011)
g. So Bakugan are swamped by Gogos, who cede to Mighty Beans, whose time in the sun lasts just long enough to wrongfoot a grandparent into buying the present.[11] (CNZNE *Dominion Post* 2011)

The structural complexity seen in connection with innovative verbs was also noted by Rudanko (2015b: 85), and it may be that the innovativeness in word-forms correlates with the colourful and figurative usage characteristic of the language in magazines and newspapers.

The CNZNE data contained a further seven innovative matrix verbs with *into -ing* complements. These were the verbs *buoy, bustle, court, fuel, irk, link, mobilise,* with examples as follows:

(8) a. It was an effort which buoyed rider David Walsh into thinking he could win his second Wellington Cup … (CNZNE *Evening Post* 1996)
b. Canterbury looked a more focused unit early in the second half. Immediately from the kickoff, it bustled Southland into knocking the ball dead in goal and from the resulting drop out, hooker Malcolm Aldridge scored in the corner. (CNZNE *Southland Times* 1997)
c. But the immigrants themselves can hardly be blamed for that. They are courted into coming here as part of our immigration policy. (CNZNE *Sunday News* 1996)

[11] *Bakugan, Gogos* and *Mighty Beans* are brands of children's toys, and the point here is that such toys enjoy only a brief period of popularity.

d. … the story goes that the artist fell in love with a younger woman, but she did not fall in love with him. That fuelled him into painting his study of the female body with a grotesque beast lying on her stomach. (CNZNE *Nelson Mail* 2010)
e. A cluster of real estate signs at the eastern entrance to Kurow has irked a resident into complaining to the community board representative. (CNZNE *Timaru Herald* 2011)
f. Girls' clothing is more aligned with fashion trends, linking girls early into being objects of the gaze and judgements of others. (CNZNE *Dominion Post* 2011)
g. … it was a matter of mobilising people into putting their concerns in writing. (CNZNE *Southland Times* 1997)

In some instances, the entries of the verbs in the OED and the NZOD may have included examples with *into* + NP or *to*-infinitive complements. The use of the verbs with these complements has become established before the *into -ing* pattern. Indeed, some of the instances of *into -ing* patterns in (8a)–(g) could be rephrased by using the other types of complements, some of them rather easily (e.g., *an effort which buoyed rider David Walsh to think that he could win his second Wellington cup; they are courted to come here; it was a matter of mobilising people to put their concerns in writing*). The V + NP + *to*-infinitive pattern, as noted in Section 1, is one of the main patterns with which many transitive *into -ing* verbs are also compatible.[12]

5. Summary and conclusion

This paper has investigated what could be seen as one of the most productive complement patterns in current English, with the aim of documenting its use and status in NZE. The data were drawn from a diachronic corpus of newspaper English with a time frame covering the better part of the last two decades. It was found that the overall frequency of the pattern does not change much over the half generation covered by the corpus, although the later section shows slightly lower frequency than the earlier. Further research would be needed to show whether this might be linked to Rudanko's (2015b) suggestion of a decrease in the use of the transitive *into -ing* pattern, as shown by data from the TIME corpus. The frequencies observed here are in line with those produced by Davies (2012), with data from COHA.

[12] We agree with an anonymous reviewer, who pointed out that it would be highly relevant to compare the frequencies of the V + NP + *to*-infinitive pattern and the V + NP + *into -ing* pattern in NZE, among verbs that select both. While this falls outside the scope of the present study, we can refer to an observation by Juhani Rudanko on the verb *force* in AmE; it was noted that the results of two separate searches in COCA point to an "overwhelming predominance of the infinitival pattern in comparison to the *into -ing* pattern" (2015: 72). A complete appraisal of the situation in NZE will need to wait for future research.

The data uncovered 19 matrix verbs that were previously undocumented in the transitive *into -ing* pattern, and at least two of these, *monster* and *heavy*, are attributed to local NZE lexical inventiveness. Clearly we are not claiming that the appearance of these verbs in this pattern in the CNZNE material represents the very first usage, only that these are possibly the first recorded examples in a study of this type. Among many of the innovative verbs, various word formation processes were seen to be actively used, and from this point of view, complex and creative matrix verbs such as *turbo-shock* will no doubt continue to appear, as the combination possibilities are almost endless.

The present analysis is based on newspaper language, and admittedly, the restriction to just one text type is not ideal. It should be remembered, though, that a large number of important studies have been conducted on the basis of newspaper material, and the genre does have its advantages, not least the wealth of different subject matters covered and the fact that it reflects current trends in language use. Provided the different types of newspapers – e.g. elite, popular, etc. – are represented in a corpus, valid inferences can be drawn from the results.

On a final note, it was something of a surprise that no influence of te reo Māori was evident among the matrix verbs in our material. We were pleased, however, to at least come across this example of the *into* NP pattern in the newspaper *The Press* from earlier years:

(9) Why, only ten years after a city's Jubilee, should the people be hakaed into another outburst, while a blank interval of forty years intervenes before the centenary celebrations? (*Press* 1928)

We have seen that the *into* NP pattern has a connection with the development of the transitive *into -ing* pattern, and, despite not having surfaced in our corpus data, it is quite possible that *haka* and/or other similar words from te reo Māori are being used in the *into -ing* pattern somewhere in New Zealand. Given the well-known negative attitudes towards te reo Māori during the early part of the 20th century, it seems that the author of this example was ahead of his/her time, but that *haka* did not succeed in making it into the ranks of verbs used with the *into -ing* pattern.

Acknowledgements

The authors would like to thank the audience of the 2015 Symposium on Complementation, held at the University of Tampere in honour of Professor Juhani Rudanko. Thanks must also go to our two anonymous reviewers, whose suggestions have helped to improve the present work. We would also like to take this opportunity to acknowledge the scholarly achievements of Professor Rudanko, without which the present paper would not have been written.

References

Bauer, Laurie. 1994. English in New Zealand. In *The Cambridge History of the English Language*, Vol V: *English in Britain and Overseas, Origins and Developments*, Robert W. Burchfield (ed.), 382–429. Cambridge: CUP.

Bauer, Laurie & Bauer, Winifred. 2000. Nova Zelandia est omnis divisa in partes tres. *New Zealand English Journal* 14: 7–17.

Belich, James. 1996. *Making Peoples: A History of the New Zealanders from Polynesian Settlement to the End of the Nineteenth Century*. Harmondsworth: Allen Lane, The Penguin Press.

Bridgeman, Lorraine, Dillinger, Dale, Higgins, Constance, Seaman, David & Shank, Floyd. 1965. *More classes of verbs in English*. Bloomington IN: Indiana University Linguistics Club.

Brinton, Laurel J. & Brinton, Donna M. 2010. *The Linguistic Structure of Modern English*. Amsterdam: John Benjamins. doi: 10.1075/z.156

Calude, Andreea & James, Paul. 2011. A diachronic corpus of New Zealand newspapers. *New Zealand English Journal* 25: 35–52.

Davies, Mark A. 2012. Some methodological issues related to corpus-based investigations of recent syntactic changes in English. In *The Oxford Handbook of the History of English*, Terttu Nevalainen & Elizabeth Traugott (eds), 157–174. Oxford: OUP.

Deverson, Tony & Kennedy, Graeme (eds.). 2005. *The New Zealand Oxford Dictionary*. Victoria: OUP.

Francis, Gill, Hunston, Susan & Manning, Elizabeth. 1996. *Collins Cobuild Grammar Patterns* 1: *Verbs*. London: HarperCollins.

Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Gordon, Elizabeth, Campbell, Lyle, Hay, Jennifer, MacLagan, Margaret, Sudbury, Andrea & Trudgill, Peter. 2004. *New Zealand English: Its Origins and Evolution*. Cambridge: CUP. doi: 10.1017/CBO9780511486678

Gordon, Elizabeth & Abell, Marcia. 1990. This objectionable colonial dialect: historical and contemporary attitudes to New Zealand speech. In *New Zealand Ways of Speaking English*, Allan Bell & Janet Holmes (eds), 21–48. Clevedon: Multilingual Matters.

Görlach, Manfred. 1987. Colonial lag? The alleged conservative character of American English and other 'colonial' varieties. *English World-Wide* 8: 41–60. doi: 10.1075/eww.8.1.05gor

Hay, Jennifer, Maclagan, Margaret, & Gordon, Elizabeth. 2008. *New Zealand English.* Cambridge: CUP. doi: 10.3366/edinburgh/9780748625291.001.0001

Huddleston, Rodney & Pullum, Geoffrey. 2002. *The Cambridge Grammar of the English Language*. Cambridge: CUP.

Hundt, Marianne. 2009. Colonial lag, colonial innovation or simply language change? In *One Language Two Grammars?* Günter Rohdenburg & Julia Schlüter (eds), 13–37. Cambridge: CUP. doi: 10.1017/CBO9780511551970.002

Hunston, Susan, & Gill Francis. 2000. *Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English*. [Studies in Corpus Linguistics 4]. Amsterdam: John Benjamins. doi: 10.1075/scl.4

Kim, Jong Bok & Davies, Mark A. 2015. The into-causative construction in English: a construction-based perspective. *English Language and Linguistics* 20: 55–83.

Macalister, John. 2001. Introducing a New Zealand newspaper corpus. *New Zealand English Journal* 15: 35–41.

Marckwardt, Albert. 1958. *American English*. Oxford: OUP.

Oxford English Dictionary, online. <http://www.oed.com> (February 2016).

Rickman, Paul. 2017. Aspects of verb complementation in New Zealand Newspaper English. In *Changing English: Global and Local Perspectives*, Markku Filppula, Juhani Klemola, Anna Mauranen & Svetlana Vetchinnikova (eds), 169–190. Berlin: Mouton de Gruyter.

Rohdenburg, Günter. 2006. The role of functional constraints in the evolution of the English complementation system. In *Syntax, Style and Grammatical Norms*, Christiane Dalton-Puffer, Dieter Kastovsky, Nikolaus Ritt & Herbert Schendl (eds), 143–166 Bern: Peter Lang.

Rudanko, Juhani. 2000. *Corpora and Complementation: Tracing Sentential Complementation Patterns of Nouns, Adjectives and Verbs over the Last Three Centuries*. Lanham, MD: University Press of America.

Rudanko, Juhani. 2002. *Complements and Constructions*. Lanham, MD: University Press of America.

Rudanko, Juhani. 2005. Lexico-grammatical innovation in current British and American English: A case study on the transitive into -ing pattern with evidence from the bank of English corpus. *Studia Neophilologica* 77, 171–87. doi: 10.1080/00393270500384791

Rudanko, Juhani. 2006. 'Emergent alternation in complement selection: the spread of the transitive 'into -ing' construction in British and American English' *Journal of English Linguistics* 34: 312–331. doi: 10.1177/0075424206293600

Rudanko, Juhani. 2011. *Changes in Complementation in British and American English: Corpus- based Studies on Non-finite Complements in Recent English*. Houndmills: Palgrave Macmillan. doi: 10.1057/9780230305199

Rudanko, Juhani. 2015a. Wheedled me into lending him my best hunter: Comparing the emergence of the transitive *into -ing* pattern in British and American English. In *Perspectives on Complementation: Structure, Variation and Boundaries*, Mikko Höglund, Paul Rickman, Juhani Rudanko & Jukka Havu (eds), 128–140. Houndmills: Palgrave Macmillan.

Rudanko, Juhani. 2015b. *Linking Form and Meaning: Studies on Selected Control Patterns in Recent English*. Houndmills: Palgrave Macmillan.

Rudanko, Juhani & Lea Luodes. 2005. *Complementation in British and American English.* Lanham, MD: University Press of America.

The Corpus of New Zealand Newspaper English (CNZNE).

Trudgill, Peter. 2004. *New-Dialect Formation: The Inevitability of Colonial Englishes*. Oxford: OUP.

Trudgill, Peter, Gordon, Elizabeth, Lewis, Gillian & Maclagan, Margaret. 2000. 'Determinism in new-dialect formation and the genesis of New Zealand English' *Journal of Linguistics* 36: 299–318. doi: 10.1017/S0022226700008161

Turner, G. W. 1966. *The English Language in Australia and New Zealand*. London: Longman.

Wierzbicka, Anna. 1998. The semantics of English causative constructions in a universal typological perspective. In *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure,* Vol. I, Michael Tomasello (ed.), 113–153. Mahwah NJ: Lawrence Erlbaum Associates.

Wulff, Stefanie, Stefanowitsch, Anatol, & Gries, Stefan T. 2007. 'Brutal Brits and persuasive Americans' in Radden, Günter, Klaus-Michael Köpcke, Thomas Berg and Peter Siemund (eds.) *Aspects of Meaning Construction*. Amsterdam: John Benjamins, 265–282.

PUBLICATION III

Facing two ways syntactically: on the grammar and use of *promise* and *threaten* in three regional varieties

Paul Rickman & Juhani Rudanko

English Studies 104(2), 365–382
https://doi.org/10.1080/0013838X.2023.2179205

Publication reprinted with the permission of the copyright holders.

ENGLISH STUDIES
2023, VOL. 104, NO. 2, 365–382
https://doi.org/10.1080/0013838X.2023.2179205

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Facing Two Ways Syntactically: On the Grammar and Use of *Promise* and *Threaten* in Three Regional Varieties

Paul Rickman and Juhani Rudanko

Department of English, Tampere University, Tampere, Finland

ABSTRACT
The predicates *promise* and *threaten* with a *to* infinitive complement can be used in two main senses, with each sense representing a different argument structure: subject control, and subject to subject raising. Recent work has shown both uses to have been available for several centuries, with the raising variant emerging later than the control. The present investigation aims to provide insight into the usage patterns of the two predicates in the Englishes of New Zealand, Britain, and North America. The former is compared to the two latter varieties to help further our understanding of the complex system of the grammar of English predicate complementation. The data is examined from quantitative and qualitative perspectives, and the results indicate not only significant variation in usage among the varieties, but also high levels of correlation in some areas. New findings emerge from the study, including on the notion of stance, offering fresh avenues of research into the nature of the complementation patterns examined and their use in three core varieties of English.

ARTICLE HISTORY
Received 16 November 2022
Accepted 27 January 2023

KEYWORDS
Promise; *threaten*; subject control; subject to subject raising; New Zealand English; spin

1. Introduction

One of the fundamental insights in the area of sentential complementation to emerge from the generative-transformational model of linguistic analysis concerns the basic dichotomy of two types of *to* infinitive complements in English. To illustrate the dichotomy, consider sentences (1a–b).

(1) a. John seems to be a millionaire.

b. John wants to be a millionaire.

The matrix verb of (1a) is *seem*, and that of (1b) is *want*, with each selecting a *to* infinitive complement. In each case the sequence of overt constituents is: "NP-Verb-*to*-*be* NP". However, in the generative model their derivations and the resulting structures are sharply different. In (1a) the matrix subject is moved from the lower subject position into the matrix subject position by subject to subject raising, which is a type of NP

CONTACT Paul Rickman paul.rickman@tuni.fi

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

Movement rule, leaving a coindexed NP trace behind. The matrix subject, the NP *John*, receives its theta role in its original subject position in the lower clause. A convenient way to describe the nature of the matrix verb in (1a) is to say that *seem* is a subject to subject raising verb.

As regards (1b), there is no movement into the higher subject position. Instead, the higher subject is generated in the matrix subject position by phrase structure rules. The *to* infinitive complement is still sentential, for instance because of the need to satisfy the theta grid of the lower verb. It has its own subject, which is implicit or covert, and represented by the symbol PRO. PRO is a pronominal NP that is not pronounced. The complement structure construction is one of subject control, and the matrix subject controls the reference of PRO and imposes coreference between the two NPs in (1b). In this case, the matrix subject receives its theta role from the higher predicate, with PRO getting its theta role from the lower predicate. Regarding the nature of the matrix verb in (1b), it can be said that *want* in (1b) is a subject control verb.

The model explains for instance why a sentence of the type of (2a), which is from the News on the Web (NOW) Corpus, is completely natural, whereas a sentence of the type of (2b), freely invented using (2a) as a model, is unlikely.

(2) a. The latest news seems to be encouraging. (21-12-30 US)

b. *The latest news wants to be encouraging.

In (2b) the higher predicate, *want*, again assigns a theta role to its subject, and it would prefer a human (or at least an animate) subject. By contrast, in (2a) *seem* does not assign a theta role to the NP in the higher subject position, which explains why even an abstract NP, raised into that position from the lower clause, by subject to subject raising, is fine in that position, as in (2a).

The number of matrix verbs of the subject control type is relatively large in English, while the number of matrix verbs of the subject to subject raising type is relatively small. This conclusion is grounded in the list in Alexander and Kunz.[1] Their list is illustrated with the example *My father WANTED to see the world*, but it comprises both types of verbs. The number of subject to subject raising verbs is relatively small in their list, while verbs of the subject control type is much larger (see also the lists in Rudanko).[2] At the same time, the disparity in the numbers of verb types does not mean that they are infrequent at the level of usage, since for instance *seem* is very frequent with *to* infinitive complements.

The list of verbs in Alexander and Kunz is also of interest for another reason. When the verbs listed are considered from the point of view of the subject control versus subject to subject raising dichotomy, it is observed that for the vast majority of the verbs it is true to say that with *to* infinitive complements they subcategorise either for subject control or for subject to subject raising. In other words, verbs that are compatible with either structure are very rare in the list. When the verbs listed by Rudanko in the subject control pattern with *to* infinitives are considered, it is likewise observed that very few of them are triggers for subject to subject raising. The purpose of this study is to investigate two matrix verbs that are exceptional in this respect in that they can be found both in

[1]Alexander and Kunz, *Some Classes of Verbs in English*, *Vol. 1*, 75.
[2]Rudanko, *Complementation and Case Grammar*, 21–3.

subject control and subject to subject raising structures. The verbs in question are *promise* and *threaten*, the latter the "evil twin" of *promise*, as Culicover and Jackendoff[3] label *threaten*.

Initial illustrations of these verbs are given in (3a–b) and (4a–b), from the Great Britain part of the NOW Corpus.

(3) a. The mayor promised to write back to the committee … (20-07-23 GB)

b. This promises to be quite an exciting biking year … (10-01-14 GB)

(4) a. The accused […] threatened to post clips on social media if the victim went to the police. (16-12-07 GB)

b. Ultra-low rates threaten to cause serious distortions; (15-12-13 GB)

The sentences in (3a) and (4a) illustrate subject control constructions, with the matrix verbs assigning theta roles to their subjects, and there is no relevant movement rule applying. For their part, the sentences in (3b) and (4b) illustrate subject to subject raising constructions. In them the higher subject is generated in the lower subject position and then raised into the higher subject position, with a coreferential NP trace remaining behind.

The difference in the interpretations of the control and raising variants is also reflected in a difference of the senses of the matrix verbs in the two constructions. To consider *promise* first, the sense of the verb in (3a) is the one represented under sense 1 of the verb in the *Oxford English Dictionary (OED)*. Part of this sense is "to undertake or commit oneself to do or refrain from (a specified thing or act)". On the other hand, for sentence (3b), where raising is applicable, the sense of the verb is that of sense 5 in the *OED*. Part of this sense is "to give strong or reasonable grounds for expecting (future achievements or good results)". An analogous difference can be discerned for *threaten*. In the case of the subject control construction, as in (4a), the sense of *threaten* is the one represented by sense 3.a in the *OED* "to hold out or offer (some injury) by way of a threat; to declare one's intention of inflicting". Also relevant to the subject control construction is the *OED* sense 7, "In weakened use: to express an intention *to do* something, not necessarily evil", with the example "*he threatened to give me money*". As for the subject to subject raising construction, illustrated by (4b), sense 4.b in the *OED* is appropriate, which is "to appear likely *to do* some evil". It is possible to say that the control constructions involve meanings of the matrix verbs that are more lexical and more specific, and that the subject to subject raising constructions involve meanings of the matrix verbs that are more bleached and more grammaticalised.[4] In this sense the raising constructions have a degree of epistemic modality about them and, as Traugott notes,[5] they are similar to semi-auxiliaries, such as *be going to*. The *OED* senses of raising *promise* and *threaten* reproduced above include the verb phrases *to give* …

[3]Culicover and Jackendoff, *Simpler Syntax*, 437.

[4]Traugott, "The Conflict Promises/Threatens to Escalate into War", and "Subjectification and the Development of Epistemic Meaning" for discussion of *promise* and *threaten* in this context; Verhagen, "Subjectification, Syntax, and Communication"; Cornillie, "On the Pace of Syntactic Elaboration" and "Syntactic Complexity in Standard Average European"; Kissine, "Metaphorical Projection, Subjectification and English Speech Act Verbs". See Hopper and Traugott, *Grammaticalization*, for background on this topic.

[5]Traugott, "Subjectification and the Development of Epistemic Meaning," 190.

grounds for expecting and *to appear likely to* … , which strongly suggest that the evaluation of the likelihood of the enactment of the action denoted by the lower verb rests upon the speaker. Traugott[6] provides what is probably the most detailed discussion of the development of the epistemic uses of *promise* and *threaten* within the concept of subjectification. This widely-discussed process, which, in Traugott's conception, places an emphasis on "meanings that express speaker attitude or viewpoint",[7] is important to note here, because in the classification of corpus data much can rest on the extent to which the evaluation of the lower clause proposition can be attributed to a subjective viewpoint.

Earlier work has shown that both the more lexical and the more grammaticalised senses of the matrix verbs have existed in English for several centuries.[8] It also emerges from this body of earlier work that the proportions of the two senses and of the two types of syntactic structures associated with them have sometimes shifted in the course of recent centuries.[9] Regarding earlier studies, Traugott's 1993 study is an important investigation of the evolution of the two types of constructions with the two verbs in British English from Middle English onwards, and the present authors recently examined the patterns in question, again in British English, up to the 1920s on the basis of a systematic study of the Corpus of Late Modern English Texts, version 3.0.[10] However, as far as we are aware, the incidence of the properties of the two types of constructions have not been explored in very recent English on the basis of large electronic corpora. It is one of the aims of the present study to fill this gap in the literature. In view of their earlier work on the patterns in question in BrE up to the 1920s, the present authors naturally chose BrE as one of the varieties to be investigated in this article. This study also investigates the two types of constructions in very recent New Zealand English and American English. The choice of the former is made because the area of sentential complementation has sometimes been neglected in the study of that regional variety,[11] and is further motivated by the historical link between the two varieties of NZE and BrE.[12] AmE is chosen in order to represent another major source of potential influence for a smaller English variety. The three varieties are all inner-circle varieties, facilitating the comparison of the uses of the variants and their properties from that perspective.

A comparison of regional varieties invites discussion of language change under the influence of the dominant varieties of AmE and BrE, or, in the case of the varieties under discussion here, what has been termed "colonial lag".[13] The influence of AmE on other Englishes in general has been the subject of research in recent years,[14] and

[6]Ibid.
[7]Traugott, "(Inter)Subjectivity and (Inter)Subjectification," 60.
[8]Traugott, "The Conflict Promises/Threatens to Escalate into War" and "Subjectification and the Development of Epistemic Meaning"; Cornillie, "On the Pace of Syntactic Elaboration"; Rickman and Rudanko, "Straddling a Syntactic Divide."
[9]Cornillie, "On the Pace of Syntactic Elaboration"; Rickman and Rudanko, "Straddling a Syntactic Divide."
[10]Traugott, "The Conflict Promises/Threatens to Escalate into War"; Rickman and Rudanko, "Straddling a Syntactic Divide"; on the corpus in question, see Diller, De Smet, and Tyrkkö, "A European Database of Descriptors of English Electronic Texts."
[11]The present study also forms part of a larger project of research on predicate complementation in NZE currently being conducted by the first author.
[12]Gordon et al., *New Zealand English*, for in-depth discussion of the development of NZE and its relationship with BrE, though from a largely phonological perspective.
[13]Görlach, "Colonial Lag?"; Hundt, "Colonial Lag, Colonial Innovation or Simply Language Change?"
[14]Gonçalves et al., "Mapping the Americanization of English in Space and Time", offers a large-scale perspective on the influence of AmE.

the availability of appropriate corpora has no doubt helped to make this research possible. The present study is synchronic, however, and thus we will not inquire into changes of complementation tendencies with *promise* and *threaten*. Nonetheless, a picture of the current state of affairs can be pieced together, which will be useful in providing a foothold for future work.

2. Data and Methodology

The NOW Corpus was selected as the source of data for the present study because it offers a large body of data of very recent English and because it is possible to collect information from it on a country-by-country basis. The NOW Corpus currently comprises 15.3 billion words of data collected from online newspapers and other publications from 20 countries around the English-speaking world. The majority of the data are from American and British-based sources, but NZE is also very well represented, making it possible to obtain large amounts of data on *promise* and *threaten* in all three varieties.

The search strings used were "[promise]_v* to" and "[threaten]_v* to". These strings were designed to find the verbal uses of *promise* and *threaten* followed directly by *to*. From each variety, a random set of 500 tokens produced by each string was taken using the built-in random sample function of the www.english-corpora.org interface. This retrieves a non-reduplicating set of random tokens from the entire period covered by the corpus, which is 2010 to the present day. (At the time the data were collected, the most recent material available in NOW was from the end of 2021.) This approach was chosen over simply taking the most recent 500 available tokens in order to minimise the presence of duplicate tokens in the datasets. Duplicate tokens were still found in the *promise* datasets of all three varieties, however, where, in each case, a sentence was reproduced in different publications several times, resulting in the removal of as many as 62 tokens in the case of the US *promise* dataset, and somewhat fewer in the NZ and GB datasets.[15]

The vast majority of tokens retrieved are relevant to the present investigation, but there are some that are not. They include tokens where *to*, rather than introducing a *to* infinitive complement, is instead the head of a prepositional phrase in a different complement pattern, as in (5).

(5) … the money promised to him by his aunt in the UAE was still nowhere to be seen. (16-07-28 NZ)

The rare instances of the type shown in (6a–b) were also excluded.

(6) a. More stock is promised to soon arrive at Canterbury's Superdrug. (20-03-04 GB)

b. We call upon all nations threatened to join us in that endeavour. (19-06-16 NZ)

Sentence (6a) does involve a sentential complement of *promise* and the surface subject of *promise,* the NP *more stock,* does originate in the lower clause. However, the raising rule it has undergone is subject to object raising, after which it is moved by the NP movement

[15]We henceforth refer to the datasets of the three regional varieties using the abbreviations GB, US and NZ, and to the varieties themselves as BrE, AmE and NZE.

rule of Passivisation into the subject position. Subject to subject raising is not involved in its derivation, and the pattern, apparently quite rare with *promise*, can be set aside in the present investigation. Sentence (6b) also involves a *to* infinitive clause, but the construction is of the type "NP call [upon NP] [to Verb …]", with the *to* infinitive sentence being a complement of *call*.

Also set aside in the present investigation are tokens of the type in (7a–b).

(7) a. No school is being threatened to close. We're not doing any of that. (19-09-10 US)

b. … five years after the affair she was threatened to keep quiet by a man she did not recognize in a Las Vegas parking lot. (18-12-03 NZ)

Although the *to* infinitive clause in (7a–b) is a sentential complement of *threaten*, there are structural differences between this pattern and the subject to subject raising and subject control types discussed so far which rule it out of consideration in the present study. The clearest difference is the fact that, like (6a), the higher subject, the NPs *no school* and *she*, are the surface subjects of passive clauses and were therefore moved into that position via passivisation, instead of being generated in that position and receiving their theta roles as the original subjects of *threaten*. It seems clear that this type of subject originated as the object of *threaten*, with the theta role of Patient, and left behind an NP trace as a result of the passivisation process, which acts as the controller of PRO.[16] This is therefore neither a case of subject control nor subject to subject raising. In addition to the examples in (7a–b) there was only one other such example in the dataset, found in the US data, which appears to be a reproduced token stemming from the same news item as (7b).

The search string was intended to retrieve verbs but part-of-speech taggers are not infallible and other word classes also turned up in the datasets: nouns in the case of *promise*, and adjectives in the case of *threaten*, as shown in (8a–b).

(8) a. … fulfilled the French president's campaign promises to pull no punches with Russia … (17-05-30 NZ)

b. None were as divisive or as threatening to the royal family's standing as those surrounding race. (21-03-10 US)

The datasets of both verbs were manually scanned for irrelevant tokens. The total numbers of the remaining relevant tokens are given in Section 3.

3. Results and Discussion

This section presents the results of the corpus data analysis, dealing firstly with the results for *promise*, followed by the results for *threaten*, and a discussion of the data taken as a whole.

3.1. Promise in NOW GB, US and NZ

Table 1 gives the numbers of *promise* tokens categorised as subject control and subject to subject raising in the three varieties.

[16]Rickman and Rudanko, "Straddling a Syntactic Divide" discuss this.

Table 1. Subject control and subject to subject raising uses of *promise* in NOW GB, US and NZ.

	GB	US	NZ
Control	308 (65.4%)	303 (73%)	329 (72%)
Raising	163 (34.6%)	112 (27%)	128 (28%)
Totals	471	415	457

The results show that *promise* is used more frequently as a control predicate than a raising predicate in all three varieties. Furthermore, in our datasets, there is some clear variation in the ratio of control to raising tokens in each variety. In BrE the share of the total number of instances of *promise* that is occupied by the raising variant is slightly over one-third, and as such it is higher than it is in either of the other varieties. AmE and NZE are at comparable levels, with the raising variant accounting for 27% and 28%, respectively, of all instances of *promise*. The difference between the BrE and the AmE uses of *promise* is statistically significant at the 0.5% level (chi-squared = 5.63, $p < 0.05$). The difference between the BrE and the NZE results is also statistically significant at the 0.5% level (chi-squared = 4.39, $p < 0.05$). The difference between the AmE and NZE uses of *promise*, however, is not statistically significant, indicating that NZE and AmE are more closely aligned to one another than they are to BrE in terms of the ratio of control to raising.

(9a–c) give examples of *promise* in its subject control use, from the datasets of each of the three regional varieties.

(9) a. As always happens at such times, politicians promised to go to the ends of the earth to find the perpetrators. (18-08-11 GB)

b. That includes 120W HyperCharge technology that promises to juice your phone full in less than 20 minutes. (21-09-16 US)

c. … it's the story of four teenage friends who promised to be at each other's weddings no matter what. (18-04-12 NZ)

These are, at least in the case of (9a) and (9c), standard, uncontroversial examples. The subjects of *promise* in these two tokens are both animate, human entities and as such are "able to perform an illocutionary act",[17] and make suitable subjects for *promise* in its performative (subject control) meaning. These are thus prototypical subject control constructions. The subject in (9b), on the other hand, is not human, nor animate, but this example – and several others in the datasets like it – was classed by the present investigators as subject control. The commitment to undertake the act encoded in the lower clause is interpreted as having been made by the creators and/or marketers of the 120W Hypercharge technology, thus a link between a human agent and the inanimate grammatical subject seems evident, even if such tokens do not represent prototypical subject control constructions.[18] Clues lending support to this interpretation can often be found in the complement clause in these types of tokens, i.e., the detail in the

[17]Traugott, "Subjectification and the Development of Epistemic Meaning," 186.

[18]That they are not prototypical subject control constructions is shown by their failure to pass the *eager* test, offered for instance by Soames and Perlmutter (*Syntactic Argumentation and the Structure of English*, 102–6) in a standard discussion of the difference between subject control and subject to subject raising. Thus ??*120W HyperCharge technology that is eager to juice your phone full in less than 20 minutes*, modeled on (9b), seems very unlikely.

prepositional phrase *in less than 20 minutes* seems less likely to be the evaluation of an outside observer – as it would have to be under a raising interpretation – and more likely to be the repetition of assurances given by the creators of the technology. The compatibility with a suitable manner adverb was used as a test to determine the classification of a number of the borderline tokens that turned up in the datasets. For example, the insertion of *explicitly* into the frame of (9b) is shown below, in (9′).

(9′) That includes 120W HyperCharge technology that *explicity* promises to juice your phone full in less than 20 minutes.

In all such examples, the insertion of the adverb had to give a plausible-sounding result in order for the token to be classed as subject control.

Turning to the subject to subject raising use of *promise*, (10a–d) give examples.

(10) a. This year's route has all the same ingredients, so promises to be a great event. (16-02-25 GB)

b. There's little in the way of a shiny floor here, but there promises to be plenty of entertainment … (20-06-06 GB)

c. It has been raining all night, and it promises to rain all day, possibly for the rest of the month. (21-01-23 US)

d. He promises to give an outstanding lecture on Antarctic science, and how science can help us … (17-10-03 NZ)

The examples in (10a–c) are the prototypical raising types. The subject of (10a), *this year's route*, is [-Animate], and, in particular, the dummy subject *there* and weather *it* of (10b–c), are raising subjects par excellence.[19] Alongside such prototypical examples, however, it can also be seen that [+Human] subjects, as in (10d), are also possible with raising *promise*. Rickman and Rudanko[20] observed in their dataset from Early Modern English that in such cases often "the lower verb describes a human quality that one does not usually have control over, such as physical and mental attributes", as in *the girl promises to be a good athlete*.

As well as a few of these types, another type is found in the data, one which may be described as the expectation of a certain type of performance, based on the reputation of the referent of the NP subject. (10d) was an example of this, and we supplement it with the illustrations in (11a–c).

(11) a. … they promise to mesmerise you with their voices and invisible instruments … (12-05-31 GB)

b. Brutal honesty is Thompson's reputation, and the 2016 NBA champion with Cleveland promises to be more expressive with his teammates in being complimentary and critical. (21-02-21 US)

c. Event founder Mark Tupuhi puts in an appearance with his band Runt and as ever promises to be interesting. (16-09-11 NZ)

[19]Davies and Dubinsky, *The Grammar of Raising and Control*, 7–8.
[20]Rickman and Rudanko, "Straddling a Syntactic Divide."

The majority of such types in our datasets describe entertainment or sporting events, and the subjects are the performers or athletes from whom a good performance is expected. Judgements on this type of example are likely to be a matter of interpretation dependent on context: it is theoretically possible, in an appropriate context, to interpret these as examples of subject control, but to the present investigators it seems unlikely that e.g., the lecturer of (10d) made a promise to give an outstanding talk, or the musician of (11c) made a pre-concert pledge to be interesting. Instead, the sentences express the speakers' (or writers') views of what is likely to be the case and involve subject to subject raising. The [+Human] subject type may have been among the first of the epistemic uses of *promise* with a non-finite complement to emerge in the eighteenth century,[21] and it seems in Present-day English to have found a certain semantic area with which it tends to be linked.

Turning to the complement clause with *promise* raising structures, the predominant lower verb by far in all three varieties is *be*. As noted by Traugott,[22] this was the earliest verb to emerge with the non-finite complement clause as the epistemic use of *promise* developed in the eighteenth century. It was initially restricted to descriptions of inchoative events, i.e., *the girl promises to be a good athlete*, and, although the inchoative type is infrequent in the NOW datasets, *be* is clearly the most common lower verb in raising structures today. The GB dataset has *be* in 89 (or 54.6%) out of all 163 examples; US has 41 (36.6%) out of 112; and NZ has 84 (65.6%) out of 128. A wide variety of other lower verbs are of course possible, as evidenced in (10c–d), the more frequent include *deliver*, *bring*, *make*, and *change*. (Lower clause verbs in *promise* control structures, on the other hand, are much more uniform in their distribution, with *be* generally as frequent as activity verbs such as *do*, *make*, *take*, *return*.)

One further pattern, an off-shoot of the raising structure of *promise*, turned up in the data with noticeable frequency and deserves attention here. Examples of the "what promises to be X" pattern are given in (12a–d).

(12) a. … on Sunday Pittodrie will be the scene of what promises to be a full-blooded collision between Aberdeen and Rangers … (19-03-02 GB)

b. We're thrilled to have her here with 2degrees for what promises to be a stunning free show. (16-09-04 NZ)

c. … nationwide protests over racial injustice and what promises to be the most divisive presidential election in living memory. (20-07-12 US)

d. … and what promises to be the first fun-to-drive hybrid, the Honda CR-Z (10-01-12 GB)

The pattern appears to be a useful discourse strategy, providing focus for the following NP, and it is worth noting that the *what promises to be* clause is structurally entirely unnecessary, as the removal of those four words shows in (12′a–b).

(12′) a. Pittodrie will be the scene of […] a full-blooded collision between Aberdeen and Rangers

[21]Traugott, "The Conflict Promises/Threatens to Escalate into War," 352–3; Traugott, "Subjectification and the Development of Epistemic Meaning," 188–9.
[22]Traugott, "Subjectification and the Development of Epistemic Meaning," 188.

b. We're thrilled to have her here with 2degrees for […] a stunning free show.

It is most commonly used in the context of a much-anticipated sporting event or concert performance, as seen in (12a–b), and occasionally some other type of event or product (12c–d) that deserves to be heralded with the type of extravagant and superlative NP modifiers seen in most of these examples. The past tense *what promised to be* is possible but it is far less common than the present simple, and is found only once in our data; this pattern is used predominantly with a future orientation. There is little noteworthy variation in the context of use across the three varieties in our data. Frequencies of usage though, are a little different: the pattern comprises 26 (or 16%) of the 163 GB raising examples, 12 (10.7%) of the 112 US, and 15 (11.7%) out of the 128 NZ tokens. These ratios are largely in line with the general raising usage frequencies for each variety seen in Table 1, with BrE somewhat higher than the other two varieties.[23]

3.2. Threaten in NOW GB, US and NZ

Table 2 provides the numbers of control and subject to subject raising tokens in the *threaten* data in the three varieties.

Like *promise*, *threaten* is clearly used more frequently as a control predicate than a raising predicate. Unlike *promise* though, the ratio of raising to control uses with *threaten* appears to be much more uniform across the three regional varieties. AmE and NZE are almost exactly the same, at 37.7% and 37.8% raising, respectively, and BrE is only a small margin ahead, with 38.2% raising. Beginning with subject control, (13a–c) below give examples.

(13) a. The court heard in September 2015 she threatened to bite the nose off a security guard at the RVI … (19-12-12 GB)

b. The NFL is threatening to force teams to forfeit games if they have COVID-19 outbreaks among unvaccinated players. (21-07-22 US)

c. The letters threatened to contaminate infant formula if New Zealand did not stop using 1080. (15-12-03 NZ)

Subjects in control structures with *threaten* are most typically animate and human, as in (13a), or collective bodies such as governments and associations, as in (13b). These are prototypical subject control constructions. For its part, the inanimate subject control type that is fairly frequent with *promise* (*120W HyperCharge technology that promises to …*) is much less commonly found with *threaten*. One of the few examples is given in (13c). Here, however, the human agent – the writer of the letters – is unmistakable.

[23]Given the frequency of *promise* in this pattern, it is of interest to look into the pattern with *threaten*. Our dataset contains only two examples of *what threatens/threatened to be NP*, and these were found in the NZ and US data. Both are from the sports/entertainment context, and, in contrast to *promise*, both are in the past tense. They are given in (a–b) below.

a. It was an incredible finish but it should not completely gloss over what threatened to be a major controversy with France once again accused of manipulating the Head Injury Assessment protocols … (18-02-04 NZ)
b. … thrillingly meshing physical production, lighting, projection, sound and hydraulics – lifts what threatened to be a movie retread into a live entertainment triumph. (21-09-13 US)

The finding suggests that the pattern is used almost exclusively with *promise*, compared to *threaten*.

Table 2. Subject control and subject to subject raising uses of *threaten* in NOW GB, US and NZ.

	GB	US	NZ
Control	307 (61.8%)	302 (62.3%)	308 (62.2%)
Raising	190 (38.2%)	183 (37.7%)	187 (37.8%)
Totals	497	485	495

A handful of similar NP subjects denoting e.g., messages, emails, cards, executive orders, are found in the data.

Moving on to *threaten* in its raising role, (14a–c) give examples.

(14) a. … several major blazes threatened to merge into a single dangerous fire front that could reach outlying districts of Sydney. (13-10-20 GB)

b. Even the film's style, characterized by conspicuous crosscutting and stuttering camera pans, is loose-lipped, threatening to let cats out of bags. (19-01-24 US)

c. Heads threatened to roll unless they could find an immediate response and, to coach Michael Maguire's credit, he did just that. (16-09-03 NZ)

The examples in (14a–c) all have the prototypical inanimate raising subject, and (14c) is particularly noteworthy in this respect, with the subject idiom chunk from the idiom *heads will roll*. The compatibility of raising predicates with subject idiom chunks – only possible with raising predicates – is among the standard set of diagnostics for distinguishing cases of control from raising.[24]

[+Human] subjects are found with the raising predicate *threaten*, as they are with raising *promise*, and, like *promise*, they appear to be used most often in sporting/entertainment contexts. (15a–c) give examples.

(15) a. In Duloc we meet pint-sized baddie Lord Farquaad, who threatens to steal the show with his camp comedy and incredible timing. (15-12-14 GB)

b. Wylie was threatening to pull off a sweep with a 22-19 cushion in the third, but the Lady Cats hung tough and won the next six points to force a fourth set. (19-08-06 US)

c. … the free-swinging 21-year-old threatened to cause a huge boilover, only for Kerber to regroup in the nick of time. (17-01-17 NZ)

Recalling the way in which raising *promise* with a [+ Human] subject can be used to express an evaluation of a performance based on the subject's reputation, it can be seen that *threaten* can be used in a comparable way. The difference though, is that the description is often that of a surprisingly good performance, one contrary to expectations, the loser coming up from behind to almost "steal the show". A comparison of this type of use in the GB, US and NZ data shows its frequency to be relatively high in GB and NZ, and in US somewhat lower.[25] Thus, all three varieties use raising *promise* and *threaten* with [+ Human] subjects in similar ways to describe topics related to sports/entertainment, and in the case of *threaten* this may be more widespread in BrE and NZE than it is in AmE.

[24]Postal, *On Raising*.

[25]GB: 20 (or 10.5%) out of 190 raising tokens; US: 7 (3.8%) out of 183; NZ: 22 (11.8%) out of 187. The frequencies of the comparable *promise* version: GB: 8 (4.9%) out of 163 raising tokens; US: 4 (3.6%) out of 112; NZ: 5 (3.9%) out of 128.

There are also examples of the more standard type of raising use, as shown in (16).

(16) Rapidly growing numbers of New Zealanders living with dementia threaten to overwhelm our health system unless government acts quickly … (20-11-11 NZ)

A further noteworthy feature of raising *threaten* in the NOW data is its appearance in the progressive, as shown in (15b). While quite acceptable with control uses of *threaten*, the progressive is thought to be possible but somewhat limited when *threaten* is used as a raising verb.[26] (As noted by Traugott, raising verbs in general are not normally compatible with the progressive, particularly with animate subjects). Our data support the observation of this particular property of *threaten*, and indicate that raising *threaten* is in fact used in the progressive form with what might be called reasonable frequency, as shown in Table 3. As for *promise*, Traugott notes a split between the two verbs in this area, with *promising* "restricted to adjectival and gerund constructions".[27] This is supported by the NOW data, with no *promise* raising tokens and only one *promise* control token in the progressive.

Examples of raising *threaten* in the progressive are given in (17a–c).

(17) a. … the 17 million visitors they get annually are threatening to swamp the city. (17-11-01 GB)

b. … they released water from dams that were threatening to overflow. (21-07-23 US)

c. Storms are threatening to spoil Guy Fawkes parties across the north of the North Island. (13-11-05 NZ)

NZE is known as a variety that favours the use of the progressive, or at least did so in the late twentieth century,[28] and this may go some way towards explaining the slightly higher figure for that variety. AmE, however, is not far behind NZE in this respect, with BrE noticeably further behind. In terms of subject animacy, which is a variable known to be a significant factor affecting not only raising verbs but also progressives in general,[29] the data show that six of the 18 GB *threaten* progressives contain animate subjects, as do 10 of the 28 NZ progressive tokens, while only three of the 26 US tokens have animate subjects.[30] Such differences point to a need for further investigation of the relationship between the progressive aspect and raising predicates in different regional varieties of English.

As for complement clauses with *threaten*, in its raising use the two verbs found to be the most frequent in all three varieties are *derail* and *undermine*, with others, such as *break*, *destroy*, and *overwhelm* also somewhat common but more variety-specific. (18a–c) give examples.

[26]Traugott, "Subjectification and the Development of Epistemic Meaning," 194.
[27]Ibid., 204.
[28]Hundt, *New Zealand English Grammar*, 75–7; Collins, "The Progressive Aspect in World Englishes"; Collins, "The Progressive."
[29]Hundt and Szmrecsanyi, "Animacy in Early New Zealand English."
[30]Like Hundt and Szmrecsanyi, "Animacy in Early New Zealand English", we include collective nouns – *the All Blacks, the group* etc. – among the animate types. These in fact account for a good many of the animate types in the datasets of all three varieties.

Table 3. Use of the progressive in raising *threaten* in NOW GB, US and NZ.

	GB	US	NZ
Threaten	18 (9.5%)	26 (14.2%)	28 (15.0%)

Note: Brackets = percentages of the total number of raising tokens.

(18) a. An increase in Saudi oil flowing to Europe threatens to undermine Russia's principle market. (15-11-18 GB)

b. As the Celtics navigated a tumultuous regular season that threatened to derail their title aspirations … (19-05-03 US)

c. … cleavers (aka sticky weed) will be climbing up and over many plants, threatening to overwhelm them if left to their own devices. (17-10-12 NZ)

The frequencies of *derail* and *undermine* are very similar in the three varieties: *derail* – GB: 9, US: 10, NZ: 9; *undermine* – GB: 9, US: 9, NZ: 8.[31] With a strong collocation with lower predications headed by verbs such as these, it is obvious that *threaten* is a verb that entails a negative semantic prosody.[32] Indeed, it has been described as the "evil twin"[33] of *promise*, and, as noted above, the *OED* sense most applicable to raising *threaten* is "to appear likely *to do* some evil". With such descriptions, it is safe enough to say that wherever *threaten* is involved, bad things are rarely far away. Nevertheless, *OED* sense 7, also quoted above, is "to express an intention to do something, not necessarily evil", but even so, positive or even neutral events encoded by the lower predications of *threaten* are unlikely to be frequent. This expectation is indeed borne out by the subject control section of the data, with only a handful of candidates for the weakened "not necessarily evil" sense emerging. Two of these are given below.

(19) a. Kate had threatened to do a naked cartwheel if she won – something no one wanted to see. This was the Great British Bake Off after all … (17-10-31 GB)

b. The developer of a Mosgiel subdivision has won the right to name his own streets and pathways, but only after threatening to use hard-to-pronounce Scottish village names instead. (13-04-26 NZ)

The semantic prosodies of the lower predicates themselves in (19a–b) can be seen as neutral or, at best "not necessarily evil", but the selection of the higher predicate *threaten* lends them the somewhat negative implication that the intended act is undesirable. Still, these examples show that *threaten* can be used in a light-hearted, non-serious way, consistent with the *OED* sense 7. It also seems possible to say that the degree of commitment to carry out the act expressed in the lower clause is weaker than it would be had *promise*

[31]These figures prompted a further search of the NOW corpus as a whole in order to gain a better understanding of the collocational strength of the *threaten to derail* combination. The results of a search using the string "VERB to derail" shows that not only is *derail* the preferred lower verb of the raising verb *threaten*, but *threaten* is the most common matrix verb selecting *derail* in NZ, GB and US datasets in the NOW corpus. The NZ and GB results were manageable enough to allow us to use all tokens returned by the search string (274 and 822 tokens, respectively), while a sample of 1000 tokens was taken from the total 3028 US results. The three varieties show a striking similarity in the collocational strength of *threaten* and *derail*: the two verbs were found together in 50.7% of the sampled US tokens, in 51.3% of the NZ tokens, and in 52.4% of the GB tokens.

[32]Louw, "Irony in the Text or Insincerity in the Writer?"; Partington, "Evaluative Prosody" for comments on semantic prosody in individual lexical items.

[33]Culicover and Jackendoff, *Simpler Syntax*, 437.

been the matrix verb instead, e.g., *Kate promised to do a naked cartwheel if she won* seems to entail a stronger commitment on the part of the subject.

The choice of matrix verb can also be useful in expressing stance or "spin" in political contexts.[34] In terms of stance, Gray and Biber[35] list the nouns *promise* and *threat* as stance nouns with a *to*-complement clause, and it seems clear that this can be extended to cover the verbs as well. Our focus here is on "spin", however, which is a term typically associated with US politics. It may be described as "the intentional slanting of ambiguous political events and situations to promote an interpretation favorable to one's own side".[36] A political agent, whether a politician or a newspaper editor committed to a particular party, may engage in positive spin by highlighting the advantages of his or her political agenda and the assumptions that go with it. It is also possible to talk of negative spin ("Frequently in *to put a positive (negative, etc). spin on.* colloquial (chiefly U.S. Politics",[37])). In the latter case a political agent generally seeks to attack, and to find fault with, the competing political agenda of his or her enemies. In texts that are meant to argue for a certain point of view, where reality can be constructed in a particular way in support of that point of view, the choice of the higher predicate is an important one. In the context of the two verbs *promise* and *threaten,* the former is the clear choice for a writer seeking a positive spin, and the latter verb for a negative spin.

The influence that the verb choice can have on the framing of a speaker's message is seen clearly in (20a–b).

(20) a. Montgomery threatened to fight the lawsuit all the way to get Voice Tech Corp's patents invalidated. (20-02-19 GB)

b. In two tweets, Trump threatened to bring in the National Guard to control the situation. (20-05-29 US)

Here, both the referents of the higher NP subjects and the events described in the lower clauses are invariably affected by the negative connotations of *threaten*, and are presented to the reader as undesirable. Replacing *threaten* with *promise*, however, would lend a positive spin to the story, and would help to encode a positive assessment of the referent of the higher subjects, and the events of the lower clauses. We would be more inclined to view the idea of *fighting the lawsuit* and *bringing in the National Guard* as good things. Consider the example of *promise* in (21).

(21) Since his days on the campaign trail, President Donald Trump has promised to roll back environmental regulations, boost the use of coal and pull out of the Paris climate agreement … (20-02-01 US)

Even though many people would no doubt think that they are undesirable and detrimental, the events described in the lower clauses – as well as the subject NP referent – are given a decidedly positive spin, situated in the benevolent shade of *promise*, the good twin of *threaten*. Such devices can be powerful tools of persuasion, but they can also

[34]Gray and Biber, "Stance Markers"; Rudanko, "Representations of the Baltimore Riots of July 1812."
[35]Gray and Biber, "Stance Markers," 248.
[36]Schaefer and Birkland, *Encyclopedia of Media and Politics*, 272; Press, *Spin This!*, for various forms of spin and their uses in fairly recent American politics; Guriev and Treisman, *Spin Dictators*, for a world-wide perspective.
[37]*OED,* s.v. *spin*.

be used to simply raise reader interest in a story or to add humour, as (19a–b) appear to do.

Turning to the raising tokens, however, more obviously neutral or positive lower clause propositions are more likely to be found. They appear to be limited to the [+Human] subjects in descriptions of sports/entertainment, as in (22a–c) below.

(22) a. … made it feel a bit like one minute France were threatening to win, the next the game was all over … (18-06-09 NZ)

b. Kuate threatened to become a cult hero with his stunning goal against Morton … (17-05-25 GB)

c. As the song's energy begins to rise, she threatens to get up, but maintains her post. (21-12-03 US)

The lower predications here, *win, become a cult hero* and *get up* are undoubtedly positive/neutral events, but due to the choice of the matrix verb, the hypothetical events of the lower clause can be framed as events that are undesirable. This is clearly the case in (22a–b) but (22c) appears to be an exception, with *threaten* being used in a truly weakened sense. The example in question continues as follows, in an undeniably positive review of the performance: *Then … she can't control the emotions of the song and rises up to face the crowd and belt the incredible chorus as the song peaks and closes.* In light of this example, it seems that the weakened *OED* sense, viewed as a control construction in Section 1 above and seemingly presented as such in the *OED*, can also be associated with the raising use of *threaten*.

4. Conclusion

The verb pair *promise* and *threaten* are rare cases among the English system of predicates selecting non-finite complements, in straddling the syntactic divide between the subject control and subject to subject raising categories. Few other verbs lend themselves to use in quite such a variety of ways, with a colourful and arresting range of subject types and contexts. This study set out to shed light on the uses and distribution of these verbs in three regional varieties of English, based on the evidence of very recent newspaper data. Both prototypical and less prototypical cases of each type have been discussed, with the implication being that a gradient exists between the two syntactic categories, and the establishment of clear boundaries is not always a straightforward matter. Indeed, it is often the ambiguity stemming from the lack of clear boundaries that makes these verbs so intriguing and worth investigating.

In terms of the findings to emerge from this study, we have shown that there are significant differences between the three regional varieties in terms of how *promise* is used, with more scope for the raising variant in BrE, and less in both NZE and AmE, the latter two varieties appearing to be relatively similar in this respect. A synchronic snapshot is all we have given here, but the finding invites further diachronic work to uncover any evidence of regional influence. We have also shown that raising *promise*, in all three varieties, lends itself well to descriptions of entertainment and sporting events when used with a [+Human] subject. The sports/entertainment domain is also seen to be a

significant factor behind the "what promises to be X" pattern, a structurally superfluous but pragmatically useful device for providing focus for an upcoming NP.

On the other hand, the data on *threaten* has shown some unexpected similarities among the three varieties, with raising and control ratios running almost in parallel in the three datasets. As well as the correlation in the overall frequencies, there appear to be clear similarities in the collocational strength between *threaten* and certain groups of lower predicates. Support was found in our data for previous observations on the compatibility of *threaten* with the progressive, with progressive forms reaching their highest levels in NZE. In short, the findings on *threaten* all deserve to be investigated in more detail in future work.

Taking the notion of semantic prosody as a starting point, the study has pointed to the use of the two verbs in the language of politics and of advocacy more generally. Political agents typically promote particular agendas involving sets of assumptions about the nature and status of controversial issues and each of the verbs can be used by a speaker as a vehicle of spin in an attempt to frame the perceptions of such issues, either positively (*promise*) or negatively (*threaten*), with a view to influencing voting behaviour. The present article illustrates such uses, providing a stimulus for further work combining grammatical analysis with the study of pragmatic function.

Going beyond the datasets examined here, the present article also invites follow-up studies on the uses of the two verbs and their complementation patterns in other regional varieties. Later work can for instance focus on the patterns in question in non-core or non-native varieties of English. In investigations of those varieties, attention would obviously need to be paid to the potential influence of the native languages of the speakers in question on their use of the patterns, but the analytic framework presented here may still serve as an inspiration for, and as a suitable theoretical underpinning of, such follow-up studies.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Paul Rickman http://orcid.org/0000-0003-3407-0245
Juhani Rudanko http://orcid.org/0000-0002-8692-7784

Bibliography

Data

The NOW Corpus = *News on the Web*. Compiled by Mark Davies. [cited August 2022]. Available from https://www.english-corpora.org/now/.

Secondary sources

Alexander, Dee, W. J. Kunz, and Fred Walter Householder. *Some Classes of Verbs in English. Vol. 1*. Linguistics Research Project. Indiana University, 1964.

Collins, Peter. "The Progressive Aspect in World Englishes: A Corpus-Based Study." *Australian Journal of Linguistics* 28, no. 2 (2008): 225–49.

——— "The Progressive." In *Comparative Studies in Australian and New Zealand English: Grammar and Beyond*, edited by Pam Peters, Peter Collins, and Adrian Smith, 115–24. Amsterdam: John Benjamins Publishing Company, 2009.

Cornillie, Bert. "On the Pace of Syntactic Elaboration from Latin Calques. Evidence from Spanish, Dutch and English." *Belgian Journal of Linguistics* 33 (2019): 81–106.

——— "Syntactic Complexity in Standard Average European: Language Contact and Discourse Traditions in the Domain of Communicative Distance." In *Manual of Discourse Traditions in Romance*, edited by Esme Winter-Froemel and Álvaro S. Octavio de Toledo y Huerta, 721–38. Berlin: De Gruyter, 2022.

Culicover, Peter, and Ray Jackendoff. *Simpler Syntax*. Oxford: Oxford University Press, 2005.

Davies, Mark. "The NOW Corpus (News on the Web): 15.3 Billion Words, 2010–Present." Available from www.english-corpora.org/now/.

Davies, William, and Stanley Dubinsky. *The Grammar of Raising and Control*. Malden, MA: Blackwell Publishing, 2004.

Diller, Hans, Hendrik De Smet, and Jukka Tyrkkö. "A European Database of Descriptors of English Electronic Texts." *The European English Messenger* 19 (2011): 21–35.

Gordon, Elizabeth, Lyle Campbell, Jennifer Hay, Margaret Maclagan, Andrea Sudbury, and Peter Trudgill. *New Zealand English: Its Origins and Evolution*. Cambridge: Cambridge University Press, 2004.

Gonçalves, Bruno, Lucía Loureiro-Porto, José J. Ramasco, and David Sánchez. "Mapping the Americanization of English in Space and Time." *PLoS One* 13, no. 5 (2018): e0197741.

Gray, Bethany, and Douglas Biber. "Stance Markers." In *Corpus Pragmatics: A Handbook*, edited by Karin Aijmer and Christoph Rühlemann, 219–48. Cambridge: Cambridge University Press, 2014.

Guriev, Sergei, and Daniel Treisman. *Spin Dictators: The Changing Face of Tyranny in the 21st Century*. Princeton, NJ: Princeton University Press, 2022.

Görlach, Manfred. "Colonial Lag? The Alleged Conservative Character of American English and Other 'Colonial' Varieties." *English World-Wide* 8 (1987): 41–60.

Hopper, Paul J., and Elizabeth C. Traugott. *Grammaticalization*. Cambridge: Cambridge University Press, 1993.

Hundt, Marianne. *New Zealand English Grammar – Fact or Fiction? A Corpus-Based Study in Morphosyntactic Variation*. Amsterdam: Benjamins, 1998.

——— "Colonial Lag, Colonial Innovation or Simply Language Change?" In *One Language, Two Grammars?*, edited by Günter Rohdenburg and Julia Schlüter, 13–37. Cambridge: Cambridge University Press, 2009.

Hundt, Marianne, and Benedikt Szmrecsanyi. "Animacy in Early New Zealand English." *English World-Wide* 33, no. 3 (2012): 241–63.

Kissine, Mikhail. "Metaphorical Projection, Subjectification and English Speech Act Verbs." *Folia Linguistica* 44, no. 2 (2010): 339–70.

Louw, Bill. "Irony in the Text or Insincerity in the Writer? The Diagnostic Potential of Semantic Prosodies." In *Text and Technology: In Honour of John Sinclair*, edited by Mona Baker, Gill Francis, and Elena Tognini-Bonelli, 157–76. Amsterdam: John Benjamins, 1993.

Oxford English Dictionary. 2nd ed. OED Online. Oxford: Oxford University Press, 1989 [cited August 2022]. Available from http//:www.oed.com.

Partington, Alan. "Evaluative Prosody." In *Corpus Pragmatics: A Handbook*, edited by Karin Aijmer and Christoph Rühlemann, 279–303. Cambridge: Cambridge University Press, 2014.

Postal, Paul. *On Raising: One Rule of English Grammar and Its Theoretical Implications*. Cambridge: MIT Press, 1974.

Press, Bill. *Spin This! All the Ways We Don't Tell the Truth*. New York: Simon and Schuster, 2001.

Rickman, Paul, and Juhani Rudanko. "Straddling a Syntactic Divide: Tracking Sentential Complements of *Promise* and *Threaten* in CLMET3.0." In Press.

Rudanko, Juhani. *Complementation and Case Grammar: A Syntactic and Semantic Study of Selected Patterns of Complementation in Present-Day English*. Albany: State University of New York Press, 1989.

——— "Representations of the Baltimore Riots of July 1812: Political Spin in the Early American Republic." In *Discourses of Disorder: Riots, Strikes and Protests in the Media*, edited by Christopher Hart and Darren Kelsey, 21–37. Edinburgh: Edinburgh University Press, 2019.

Soames, Scott, and David M. Perlmutter. *Syntactic Argumentation and the Structure of English*. Berkeley: University of California Press, 1979.

Traugott, Elizabeth C. "The Conflict Promises/Threatens to Escalate into War." In *Proceedings of the Nineteenth Annual Meeting of the Berkeley Linguistics Society General Session and Parasession on Semantic Typology and Semantic Universals*, edited by Joshua Guenter, Barbara Kaiser, and Cheryl Zoll, 348–58. Berkeley, CA: Berkeley Linguistics Society, 1993.

——— "Subjectification and the Development of Epistemic Meaning: The Case of *Promise* and *Threaten*." In *Modality in Germanic Languages: Historical and Comparative Perspectives*, edited by Toril Swan and Olaf Westvik, 185–210. Berlin: de Gruyter Mouton, 1996.

——— "(Inter)Subjectivity and (Inter)Subjectification: A Reassessment." In *Subjectification, Intersubjectification and Grammaticalisation*, edited by Kristin Davidse, Lieven Vandelotte, and Hubert Cuyckens, 29–71. Berlin: de Gruyter Mouton, 2010.

Verhagen, Arie. "Subjectification, Syntax, and Communication." In *Subjectivity and Subjectivisation in Language*, edited by Dieter Stein and Susan Wright, 103–28. Cambridge: Cambridge University Press, 1995.

PUBLICATION IV

The Choice Principle and the matrix verb *fear* in recent New Zealand English

Paul Rickman & Juhani Rudanko

Studia Neophilologica, published online: 7 Feb. 2024
https://doi.org/10.1080/00393274.2023.2292025

Publication reprinted with the permission of the copyright holders.

STUDIA NEOPHILOLOGICA
https://doi.org/10.1080/00393274.2023.2292025

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

The Choice Principle and the matrix verb *fear* in recent New Zealand English

Paul Rickman and Juhani Rudanko

ABSTRACT
The present study investigates aspects of the non-finite complementation of the subject control verb *fear* using recent data from the New Zealand section of the NOW corpus. The matrix verb *fear* selects both infinitival and gerundial complements and a first objective of this article is to inquire into their incidence in the corpus selected. Another, more theoretical objective, is to test the applicability of the Choice Principle to a new set of data and a pair of competing complements not previously analysed from this angle. In earlier work, the Principle has been shown to be applicable to contexts of prepositional gerunds, and this article indicates that its scope can also be extended to contexts of bare gerunds. The article also places an emphasis on examining constructions containing a complement other than the one predicted by the Choice Principle, and the authors argue that in many such cases the principle affords a useful window for examining the difference in meaning between the two constructions.

ARTICLE HISTORY
Received 21 March 2023
Accepted 3 December 2023

KEYWORDS
NOW corpus; New Zealand English; *fear*; Choice Principle; *to* infinitive; *-ing* clause

1. Introduction

Consider the sentences in (1a, b), both from the New Zealand part of the NOW corpus. (For the NOW corpus, see Davies 2008–.)

(1) a. They [. . .] often fear to take the first step. (18-09-30 NZ)
b. Companies [. . .] fear revealing too much. (21-02-18 NZ)

In both of (1a) and (1b) there is one level of sentential embedding in that the matrix verb *fear* is a head that syntactically selects a non-finite complement. In (1a) the non-finite complement is a *to* infinitive construction, and in (1b) it is a non-prepositional, or bare, gerundial complement. The verb *fear* thus selects two different types of non-finite complements.[1] The general purpose of this article is to shed light on certain semantic properties of the two types. The more precise objectives of this article are defined below.

It is an assumption made by major traditional grammarians (for instance, see Jespersen 1961 [1940]: 140) and by many current syntacticians that both these types of constructions are sentential, with their own understood subjects (see, for instance, Chomsky 1981: 20–26; Chomsky 1986: 119–132). Such understood subjects may also be called covert or

CONTACT Paul Rickman paul.rickman@tuni.fi Department of English, FI-33014 Tampere University, Finland
[1] *Fear* also selects other complement types, e.g. the finite *that* clause – cf. (8b) below – but the focus of the present article is on the variation between the two non-finite types.

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

implicit. The postulation of understood subjects is not accepted by all linguists, but it is motivated for instance by the consideration that such an understood subject makes it possible to represent the argument structure of the lower verb in a simple fashion, to saturate its theta grid. The lower verbs (and predicates) are different, of course, but in both (1a) and (1b) the theta role of the lower (implicit) subject is that of Agent. A further similarity shared by the two types of complement constructions is that in both (1a) and (1b) the matrix verb *fear* assigns a theta role to the higher subject, with that theta role being an Experiencer. The postulation of an understood subject therefore also makes it possible to represent the difference in theta roles assigned by the lower and higher predicates in an intuitively clear fashion.

The fact that the matrix verb *fear* assigns a theta role to its subject argument in the absence of the movement that is associated with raising structures (see e.g. Davies and Dubinsky 2004; Carnie 2013) in both (1a) and (1b), means that both (1a) and (1b) are control constructions. More precisely, they are subject control constructions, since in each case the higher subject controlling the reference of the understood subject receives a theta role from the matrix verb in its original position. In other words, neither construction involves subject to subject raising, which is a rule that is also sometimes involved in the derivation of sentential complements with covert subjects. Following Chomsky (1986: 119–132) the symbol PRO is used to represent understood subjects in control constructions, representing an abstract pronominal NP that is not pronounced. With these analytic assumptions in place, the sentences in (1a) and (1b) may be assigned the basic bracketed structures given in (1a´) and (1b´).

(1) a.´ $[[\text{they}]_{NP1}$ often $[[\text{fear}]_{Verb1}$ $[[PRO]_{NP2}$ $[\text{to}]_{Aux}$ $[[\text{take}]_{Verb2}$ the first step$]_{VP}]_{S2}]_{VP}]_{S1}$
b.´ $[[\text{companies}]_{NP1}$ $[[\text{fear}]_{Verb1}$ $[[[PRO]_{NP2}$ $[\text{revealing}]_{Verb2}$ too much$]_{S2}]_{NP}]_{VP}]_{S1}$

The representation in (1a´) also incorporates the assumption that infinitival *to* is analyzed as an Aux in present-day English (see Warner 1993: 64 for justification). As for the representation in (1b´), it incorporates the traditional insight that a gerundial clause can be viewed as a nominal clause. (For a comprehensive discussion of the various ways in which gerundial clauses are at the nominal end of sentential complements, see Ross 2004.)

There are many matrix verbs and adjectives in English selecting subject control constructions that only permit one of the two types of complement illustrated in (1a, b). However, there are also verbs and adjectives that can select both types. As shown by (1a, b), *fear* is a verb of this latter type. An initial, descriptive objective of this study is therefore to shed light on the incidence of the two types of complements of the matrix verb *fear* in a sample of recent New Zealand English (henceforth NZE). NZE was selected as a native variety of English that has often been neglected in studies of English sentential complementation. It is also a suitable selection as a follow-up to a recent study on *fear* (Duffley & Fisher 2021), which drew on current American English for its authentic data.

Beyond the descriptive objective, this article seeks to shed light on the two types of complement selected by the matrix verb *fear* by using a semantic principle. It is observed that in the *Oxford English Dictionary* both variants are accommodated under sense 5.b of the verb, without being differentiated: 'with *infinitive* (*vbl. n*., etc.) as object: To hesitate (*to* do something) through fear of the consequences.' (The sense in question is under the general sense of II 'To feel fear; to regard with fear,' of the verb.) One illustration of each construction is given in the *OED*. Both are from the 1790s, and they are reproduced in (2a, b).

(2) a. Dorothée..feared to obey. (1794, Radcliffe, *Myst. of Udolpho*)
b. You feared disturbing our tranquillity. (1799, tr. D. Diderot, *Nat. Son*)

That the two variants are placed side by side in the *OED* testifies to their semantic similarity and to the difficulty of teasing them apart from the point of view of their meanings. However, the similarity in meaning is also an invitation to probe the semantic ground occupied by the two forms. Such work is in the spirit of what is often termed Bolinger's Principle. This says that a 'difference in syntactic form always spells a difference in meaning' (Bolinger 1968: 127). In the spirit of Bolinger's Principle, it has been proposed in very recent work that the analysis of semantic roles, or theta roles, may provide an angle to examine the uses and meanings of *to* infinitive and gerundial complements, and the alternation between them on the basis of what has been labelled the Choice Principle (Rickman & Rudanko 2018; Ruohonen & Rudanko 2021). A more precise definition of the principle is given in section 3, but its essence is that the presence of an agentive lower subject favors a *to* infinitive complement, whereas a non-agentive lower subject favors a gerundial complement.

In earlier work, the Choice Principle has been found to be applicable to the comparison of *to* infinitive complements and gerundial complements introduced by prepositions selected by the same head (see Rudanko 2011; 2012; 2014; Ruohonen and Rudanko 2019; 2021). In the present instance the gerundial complement is what may be called a bare gerundial complement, since there is no preposition introducing the complement. A major research task of this article, significant from a theoretical point of view, is therefore to inquire into the question whether the Choice Principle also applies to the comparison of *to* infinitives and gerundial complements that are not introduced by a preposition.

A further major research task of the present article is to compare the meanings and usages of the two types of non-finite complements in the context of the Choice Principle. The goal in this part of the article is to identify the shades of meanings that inhere in each of the two variants, which has been the theme of some other recent work, including Duffley and Fisher (2021). The Choice Principle expresses a tendency that makes predictions about expected usages, but it also provides a principled basis for comparing the expected variants with unexpected variants. The comparison of the two variants in such cases then provides a new window to identifying nuances of meaning that are inherent in each variant, irrespective of how frequently or how rarely each variant is used in a particular dataset or text type. The nuances of meaning represent communicative resources that are at the disposal of speakers. The qualitative research task in question, lying at the intersection of syntax and semantics, is undertaken in the spirit of Bolinger's Principle. It is this research task that the early part of the article leads to and where the present authors place an emphasis in the present article.

2. Data and methodology

The corpus selected as the source of data is the News on the Web (NOW) corpus. It consists of large amounts of recent material from online newspapers and magazines starting from 2010 and is updated regularly. It contains texts from 20 countries where

English is an official language or spoken natively by a segment of the population. For this study, the NZE datasets were taken from the entire period of 2010 to 17 December 2021. As of August 2023, this section of NOW stood at just over 665 million words.[2]

NZE is now approximately 170 years old (Gordon et al. 2004), making it the world's youngest major native variety of English. Its phonological and lexical properties are by now well understood and continue to be well researched, but there remains a good deal to be explored in the area of grammatical variation (Quinn 2000).

For the purposes of this study, the present authors chose not to make use of the part-of -speech tagging on the word *fear* itself, to avoid tagging errors, and instead adopted two search strings designed to retrieve relevant instances of *fear* even if erroneously tagged. The search strings used were '[fear] to' for *to* infinitive complements, and '[fear] _v?g' for gerundial complements (the _v?g tag was used to retrieve *-ing* forms of verbs after *fear*). At the time the searches were carried out, the '[fear] to' search string retrieved 388 tokens and the '[fear] _v?g' string 977 tokens. Clearly irrelevant tokens, for instance, tokens where *fear* is obviously a noun, were removed manually, and in this way, the samples were reduced to 165 tokens for *to* infinitives and 639 for gerundial complements. However, some tokens in these sets still need to be set aside, even if the reason for excluding them may be slightly less obvious. One type of construction involving a *to* infinitive complement selected by *fear,* but irrelevant to the study of subject control, may be illustrated with the sentences in (3a, b).

(3) a. Australian Prime Minister Scott Morrison said three Australians were feared to be among the confirmed fatalities, with 13 among the injured. (19-12-27 NZ)
b. Foreign interest in residential property was feared to be most acute in Auckland, [. . .] (19-02-07 NZ)

Constructions of the type of (3a, b) are excluded because they do not involve subject control. It is also worth noting that the alternation between the two types of complements, observed in (1a, b) and in (2a, b), does not extend to sentences of the type of (3a, b). Thus, for instance, a sentence such as **Three Australians were feared being among the confirmed fatalities*, modeled on (3a), is ill formed.

With respect to the analysis of sentences of the type of (3a, b) it is pertinent to note that the subjects in question in them – *three Australians* in (3a) and *Foreign interest in residential property* in (3b) – are not generated in the matrix subject positions of their sentences. Instead, the subjects in these are generated and assigned their theta roles in the lower sentence, then moved by subject to object raising into the object position in the higher sentences and then in the matrix sentence, they are moved by NP Movement (passivization) into the matrix subject position. In such raising constructions the lower subject is covert, but it is not PRO; instead, it is an NP trace. Sentences of this type often feature NP Movement in the higher sentence, as is likewise the case with some of the other types of matrix verbs that trigger subject to object raising (see Postal 1974). However, while passivization often occurs in the higher sentence, tokens can be found in other parts of the NOW corpus where it has not applied, as for instance in (4).

[2]NOW is primarily made up of newspaper/magazine material, so the caveat that applies to any study drawing conclusions from data representing only one genre must apply here too.

(4) Others feared it [the drafting of the constitution to completion] to be another grab for power. (12-12-06 US)

As far as *to* infinitive complements of *fear* are concerned, there is also another type of construction that deserves attention at this point. Consider (5) from the New Zealand part of the NOW corpus.

(5) These guys go where angels fear to tread and that's not an exaggeration. (18-01-08 NZ)

The construction harks back to a well-known line in a poem by Alexander Pope (1709), which runs in part 'Fools rush in where Angels fear to tread.' There is some variation in the construction in the present corpus, as for instance in (6a, b).

(6) a. [. . .] four wheels are constantly in contact with the ground over terrain where other vehicles would fear to tread. (20-10-30 NZ)
b. Where US presidents have feared to tread I can imagine our Winston happily popping over to Pyongyang. (17-11-18 NZ)

The construction, as used in (5) and (6a, b), does involve subject control, but it has entrenched itself as a fixed combination and it is therefore set aside in the present article, along with such variants as play on the construction, including those in (6a, b).[3]

Among the tokens retrieved with the '[fear] to' search string, there is also the token in (7).

(7) There are definitely many parents who fear to admitting that parenthood is not like the adverts would have us believe. (16-02-12 NZ)

Sentence (7) contains an instance of what may called a *to -ing* complement. It has been observed in the literature that that type of complement, which does involve subject control in (7), has been spreading (Denison 1998; Rudanko 1998; 2006), but caution needs to be exercised when coming across a single token of what appears to be a new construction, and the present authors were unable to find other tokens of the pattern in the NOW corpus. Under these circumstances, it would be too bold to regard the finding as representing an established pattern.

As regards gerundial complements, there are also exclusions. Illustrations are given in (8a, b).

(8) a. Kids should be able to do these things without this fear hanging over their heads. (21-09-08 NZ)
b. But he feared doing it now would see it bogged down in political posturing [. . .] (19-08-21 NZ)

[3]The status of the *to* infinitive as the established non-finite complement in constructions of the type of the patterns in (5) and (6a, b) is confirmed when the numbers of *to* infinitive and gerundial tokens retrieved by the search strings used in the NZE part of the NOW corpus are considered. The numbers of tokens are 43 for *to* infinitives as compared to 0 for gerundial complements. This relatively clear finding motivated the decision of the present authors to exclude the *to* infinitive complement containing the verb *tread* in constructions of the type of (5) and (6a, b) from further analysis. For their part, Duffley and Fisher (2021: 77, footnote 4, and 86: example 59) did not exclude them and therefore there is a difference in perspective between the two approaches. (A search of the entire NOW corpus turns up no instances of *fear treading* that could be seen as being of the same type as (5) and (6a, b), so the lack of these in the NZE data is not due to a peculiarity of that variety.)

In (8a) the word *fear* is a noun. As for (8b), it illustrates *fear* selecting a *that* clause complement, with the complementizer omitted, which obviously also is to be excluded in this study.

When the exclusions of the types of tokens described are carried out and when obvious duplicates are omitted, there remain 35 *to* infinitive complements and 548 *-ing* complements of *fear* in the database. The totals of the two types of complement are therefore still very unequal in size, with the gerundial complement being the more frequent by a very long way.

Given the main research questions of this study, described at the end of section 1, it is also appropriate at this point to exclude sentences where complement choice has been influenced by one of two syntactic factors. The first of these is the Extraction Principle. The principle says that the infinitive is favored 'in environments where a complement of the subordinate clause is extracted [. . .] from its original position and crosses clause boundaries' (Vosberg 2003a; see also Vosberg 2006). The data turned up as many as 10 *to* infinitive tokens showing the extraction of material from the complement clause. By contrast, there are only eight extractions among the hugely larger set – 548 – of *-ing* complements. There is therefore a clear disparity in the proportions of extractions, with extraction contexts favoring *to* infinitives, which is as predicted by the Extraction Principle. Illustrations of two such tokens are provided in (9a, b).

(9) a. [. . .] there will be plenty of driveways and ramps you'll fear to point the Aventador toward. (21-10-04 NZ)
b. They're striving to turn Twickenham into a fortress and make it the place everyone fears to play at, [. . .] (13-11-14 NZ)

The application of the Fisher exact test confirms the bias towards *to* infinitives when extractions have been made, with the results significant at the 0.01% level ($p < 0.0001$ ($df = 1$), Phi coefficient = 0.3209). Given that the status of the Extraction Principle is by now well established in the literature on sentential complementation (Vosberg 2003a; 2003b; 2006; Rohdenburg 2006; 2016; Ruohonen and Rudanko 2019; 2021) and that it is confirmed by the present data, the 10 *to* infinitive extraction tokens are disregarded in the further analysis.

The second factor known to exert an influence on complement choice that we take into account here is the *horror aequi* constraint (Rohdenburg 2003; Vosberg 2003a). This constraint is based on 'the widespread (and presumably universal) tendency to avoid the use of formally (near-) identical and (near-)adjacent (non-coordinate) grammatical elements or structures' (Rohdenburg 2003: 236). In other words, from the dataset of *to* infinitive complements, we separate and exclude any tokens in which the matrix verb is an *-ing* form, and from the gerundial complement dataset we exclude all tokens in which the matrix verb is in the *to* infinitive form, as the selection of these complements is likely to have been based on the need to avoid identical consecutive structures. Tokens showing violations of the constraint, on the other hand, are not accounted for by the principle and they are left in the datasets. Nine tokens showing the effects of *horror aequi* were found in the relatively small *to* infinitive dataset, and 30 in the larger dataset of gerunds. One example of each is given in (10a, b).

(10) a. Wilkshire stopped short of calling on the major Australian banks to join Kiwibank's embargo on third tier lenders, fearing to be accused of organising a cartel-style operation. (18-06-23 NZ)
b. Also, we're a business and we don't want our members to fear coming back to the club. (20-10-29 NZ)

The Fisher exact test was then applied to the numbers, with the results significant at the 0.01% level ($p < 0.0001$ ($df = 1$), Phi coefficient = 0.1996). This indicates that the horror aequi constraint has a significant effect on complement selection in the present data.

After the exclusion of tokens showing the influence of the Extraction Principle and *horror aequi*, we are left with 16 *to* infinitive tokens, and 518 gerundial tokens.

3. The Choice Principle and the matrix verb *fear*

The Choice Principle has been defined as in (11).

(11) In the case of infinitival and gerundial complement options at a time of considerable variation between the two patterns, the infinitive tends to be associated with [+Choice] contexts and the gerund with [–Choice] contexts. (Rudanko 2017; Ruohonen & Rudanko 2021)

The two types of context are then defined on the basis of theta theory: a context is [+Choice] if the understood subordinate subject has the theta role of Agent. When this is not the case, the context is [–Choice].

As seen in (11), the definition is stated as applying to all gerundial clauses, whether or not they are introduced by a preposition. However, as noted above, earlier work on the principle has concerned gerundial clauses introduced by a preposition, and this study of *fear* is undertaken in part in order to find out if the principle can also be relevant to gerundial complements that are not preceded by a preposition and if the general formulation of the principle can stand as stated.

The Choice Principle is thus predicated on the notion of Agent. Defining the notion of Agent, and the associated notion of agency or agentivity, is far from being a trivial task, and it is probably impossible to offer a definition that would satisfy every linguist. However, the general outline of what constitutes an Agent seems fairly clear. Valuable work on this notion was done in the 1960s, for instance, by Jeffrey Gruber, whose approach can be an appropriate point of departure here. He was concerned with the notion of an agentive verb, and made the statement in (12) on the notion.

(12) [A]n agentive verb is one whose subject refers to an animate object which is thought of as the willful source or agent of the activity described in the sentence. (Gruber 1967: 943)

In later work it has come to be recognized that when assigning the label of Agent, or indeed any semantic role label, to a subject that is generated in that position, it is often helpful to consider not only the verb of the sentence, but also the larger predicate of the sentence (see Marantz 1984: 25–26; Chomsky 1986: 59–60). Further, it has come to be realized that it is helpful to conceive of the notion of Agent as a cluster of features. Such

an approach to the concept was advocated by Lakoff (1977). His discussion was focused on what he termed 'prototypical agent-patient sentences' and he proposed a list of as many as 14 properties (Lakoff 1977: 244) as characterizing such sentences. In current work on the Agent role in the context of the Choice Principle three features found in Lakoff's list have often been used (Rickman & Rudanko 2018; Ruohonen & Rudanko 2021). These three are listed in (13), slightly edited to make the formulations gender neutral. The numbers represent the numbers that the features had in Lakoff's list.

(13) 4. the agent's action is volitional.
5. the agent is in control of what he [or she] does.
6. the agent is primarily responsible for what happens (his [or her] action and the resulting change).

Regarding the nature of volitionality, it is recalled how Gruber used the phrase 'willful source or agent of the activity described' in his approach, and it is helpful to flesh out volitionality as 'volitional involvement in the event or state,' which formulation goes back to Dowty (1991: 572). As regards control, it is reasonable to accept that an Agent has at least some degree of control over the action or event described by the verb phrase of the sentence in question. (For discussion of control, see also Berman 1970.)

To gain a sense of the nature of the notion of Agent adopted for this study, consider the examples in (14a, b) and (15a, b), from the corpus under consideration. Those in (14a, b) are examples of *to* infinitive complements, and those in (15a, b) examples of gerundial complements.

(14) a. They have a lot of good qualities, but often fear to take the first step. (18-09-30 NZ)
b. Unfortunately, backed by politicians who fear to be seen as "soft on crime," the police continue to ignore this wise advice. (16-08-09 NZ)

(15) a. She never feared fighting for what she believed in [. . .] (17-08-13 NZ)
b. [. . .] members of the crew were aware of the situation but took no action because they feared losing their jobs. (18-02-14 NZ)

The examples in (14a, b) and (15a, b) illustrate the application of the concept of the Agent adopted to authentic data. In (14a) and (15a) the covert subjects of the lower clauses dependent on the matrix verb *fear* are Agents, because the referents of the noun phrases in question are conceptualized or encoded as being volitionally involved in the actions or events described by their predicates, *to take the first step* and *fighting for what she believed in*. The referents of these understood noun phrases are also conceptualized as being in control of the actions or events and as being responsible for them. The context is therefore [+Choice] in these cases.

By contrast, in (14b) and (15b) the referents of the lower subjects are not Agents. The covert subject of (14b) represents the object of the verb *see*, with the Patient or Undergoer theta role, which has been moved from its original direct object position into the subject position by NP Movement (Passivization). In the surface sentence the predicate of the lower sentence has the form *be seen as soft on crime*, but the semantic role of the covert subject is assigned before the NP in question is moved into the subject position. When the movement takes place, an NP trace remains in the original position, and is coindexed with the NP moved. The persons referred to are not depicted as being volitionally involved in

the event nor in control of it, rather the opposite. In (15b) the lower clause is not in the passive, but the predicate *losing their jobs* still encodes an event where something happens to the referents of the subject, not something that the persons in question are volitionally involved in. Nor are they in control of the event or responsible for it.

The three features coincide in a large number of cases, including (14a, b) and (15a, b). However, there may also sometimes be a tension between volitionality and control, on the one hand, and responsibility, on the other. For instance, consider sentence (16).

(16) I didn't feel comfortable moving around freely in the new space because I feared bashing into things. (16-06-01 NZ)

Sentence (16) encodes a situation where a person is not volitionally involved in bashing into things nor in control of (the act of) bashing into things. However, it can be held that the person is responsible for bashing into things, if it happens, because he or she is moving around too freely in the new space. In scenarios of this type, with lack of volitional involvement (and of control) of the referent of the subject in the event in question, the present authors do not regard the lower subject as an Agent, and the context is therefore [–Choice].

Another consideration that can be taken into account in a sentence in the active voice is whether a purpose clause can be added as a modifier of the complement clause. When the subject of the lower clause is an Agent, such a clause is generally possible. For instance, to build on sentence (14a), *They fear to take the first step in order to bring about better relations* is natural, but in a sentence in the active voice such an insertion seems less likely with a subject that is not an Agent.

The considerations mentioned are in general sufficient to make it possible to make a determination about the agentivity of a subject, but sometimes an ambiguity can occur. Perhaps the most famous example of such an ambiguity is Jackendoff's famous sentence *Max rolled down the hill*. The invented example may sound artificial, but Jackendoff makes the comment in (17) on his sentence.

(17) On one reading Max may be asleep not even aware of his motion. On the other reading he is rolling under his own volition; for this reading he must be an Agent. (Jackendoff 1972: 34)

In such cases the analyst needs to go beyond the sentence to make a determination about the agentivity or lack of it of the NP in question. With corpus data, it is normally possible to go beyond sentence boundaries.

There is an additional point of a theoretical nature to be added. This is that the Choice Principle is sensitive to the agentivity of the lower subject. Thus, it does not refer to the agentivity of the lower predicate. It should be borne in mind that in a passive sentence the derived subject is assigned its semantic role before being moved into the subject position. That theta role is Patient or Undergoer, and when the NP is moved into the subject position, it of course retains its theta role. However, it can be argued that a verb phrase can still be agentive even with passive morphology. The most famous example to illustrate this point is Roeper's (1987) sentence *The boat was sunk in order to collect insurance*. The subject *the boat* is not an Agent, but

the predicate can still be considered agentive since it is still compatible with a purpose clause, with the sentence remaining well formed.

4. Results and discussion

The Choice Principle, as defined in (11), was applied to the present data, and the results obtained are given in Table 1.

As noted above, the Choice Principle has proven to be a salient predictor of complement choice in certain environments (see also Rickman & Rudanko 2018) but is as yet untested with patterns of the type under consideration here. As a first insight into its application to the verb *fear* in NZE, the numbers in Table 1 show that there is a clear correlation between the [+Choice] context and the *to* infinitive, on the one hand, and the [–Choice] context and the *-ing* form, on the other. The Fisher exact test shows a significant result at the 0.1% level ($p < 0.001$, df = 1, Phi coefficient = 0.1495). The percentages given in Table 1 show that 8.4% of all [+Choice] tokens are *to* infinitives, while the *to* infinitive accounts for only 1% of all [–Choice] tokens.

Examples (14a, b) and (15a, b) above were given as initial illustrations of *to* infinitive and *-ing* complements in [+Choice] and [–Choice] contexts and further examples from the corpus data are given here.

(18) a. The indomitable Queenslander has never feared to ask the hard question [. . .] (21-04-02 NZ)
b. [. . .] said she sometimes fears to be identified as an Uber driver. (15-08-21 NZ)

(19) a. Gee said graduates in New Zealand feared taking on extra debt to own cars or houses [. . .] (17-05-15 NZ)
b. I'm a parent, and I still fear sitting next to babies and toddlers on planes as much as anyone. (16-12-06 NZ)

(18a) is a clear [+Choice] context, with the referent of the understood subject volitionally in control of, and responsible for, the action denoted by the lower verb *ask*. (18b), by contrast, is a clear [–Choice] example; the referent of the understood subject here having been assigned its theta role as the original object of the lower verb *identify* prior to passivization. The examples in (19a, b), from the much larger dataset of gerundial complements, show [+Choice] (19a), and [–Choice] (19b) contexts. (19b) is perhaps less obvious since *sit* is a fairly typical agentive predicate, but in this case the context makes it clear that we do not always have control over who sits next to us.

The passive construction has a strong connection to the non-volitional [–Choice] category, and requires some comment here. (20a, b) are examples of *to* infinitive and *-ing* clause passives (with (20a) a reproduction of (14b) above).

Table 1. Results of the application of the Choice Principle to the NOW NZ data.

	+Choice	–Choice	Totals
to infinitive	14 (8.4%)	2 (1%)	16
ing	151 (91.6%)	367 (99%)	518
Totals	165	369	534

(20) a. Unfortunately, backed by politicians who fear to be seen as 'soft on crime,' the police continue to ignore this wise advice. (16-08-09 NZ)
b. She is one of several worried Kiwis living overseas who fear being arrested as soon as they step foot in their homeland [. . .] (16-06-02 NZ)

As implied above, passives are not compatible with the [+Choice] context, with the surface subject being coreferential with the moved NP object of the lower verb, and as such typically having the Patient role. No examples of [+Choice] passives are found in the data. Of the [–Choice] side of the dataset, however, passives comprise both of the two [–Choice] infinitive tokens, and just under half (178) of the 367 [–Choice] gerundial tokens in the NOW data. This is not a striking result in itself, but the fact that it correlates well with what has been observed in another dataset in recent work on gerundial-infinitival complement variation is of some interest. It was noted in connection with a study on complement patterns selected by adjectival heads – in this case the semantically related adjective *afraid* – that 'passive complements, which are predominantly [–Choice], seem to attract the gerundial variant with the same intensity as active [–Choice] complements' (Ruohonen & Rudanko 2021).

Turning to the question of the interchangeability of the two complement types, in many cases we find that the competing complements are interchangeable, and, with Bolinger's Principle in mind, it is one of the aims of the present research to investigate this more thoroughly. We begin with the cases in which a lower verb occurs in the datasets in both its *to* infinitive and gerundial form. Illustrations are given in (21a, b) and (22a, b) with the verbs *report* and *be identified*, in [+Choice] and [–Choice] contexts, respectively (again, due to the low number of *to* infinitive tokens to choose from, (22a) is a repeat of (18b)).

(21) a. The case has been cited [. . .] as a prime example of why women fear to report cases of sexual abuse. (16-09-12 NZ)
b. [. . .] they had heard about Savile's predatory reputation but feared reporting their concerns to managers (16-01-21 NZ)

(22) a. [. . .] said she sometimes fears to be identified as an Uber driver. (15-08- 21 NZ)
b. Many 'disengage' because they fear being identified as indigenous [. . .] (20-03-03 NZ)

In addition to these two verbs, five others were found in both the datasets, albeit at very low frequencies. These verbs are as follows: *be seen as, ask, return, express*, and *break*. Table 2 lists all seven verbs shared by both datasets, with frequencies.

Following Bolinger, we assume that a difference in meaning is going to be discernible to some degree whenever one constituent replaces another (see Rudanko 2014 and

Table 2. Frequencies of verbs shared by both datasets.

	to infinitive	*-ing*
ask	1	3
be identified	1	2
be seen as	1	3
break	1	1
express	1	1
report	2	1
return	1	6

Rudanko 2015 for discussion). Duffley & Fisher (2021) discuss the case of *fear* with *to* infinitive and *-ing* complements from several angles, and the one most relevant to the present discussion concerns the notion of volition. The *to* infinitive is thought to convey an element of volition in the potential carrying out of the event encoded by the lower predicate; the volition arising from the Goal-like aspect inherent in the *to* infinitive construction (see Rickman & Rudanko 2018: 64, 70). On the other hand, the volitional aspect is absent from the gerundial construction. In that construction the complement is NP-like. It is recalled that only in (1b') is the complement represented as an NP. It is also possible to say that the gerundial complement designates 'that which is feared' (Duffley and Fisher 2021; see also Duffley 2000), and that it has the function of naming the source of fear (Rickman and Rudanko 2018).

This view finds support in the literature. Huddleston & Pullum (2002) compare the sentences [a] *They fear to go out at night* and [b] *they fear going out at night*, and conclude that '[w]ith *fear* infinitival [a] involves an element of volition/intentionality: [a] implicates that they don't/won't go out, while the gerund-participial [b] lacks this meaning and is comparable to an NP object' (Huddleston & Pullum 2002: 1243).

With this semantic opposition in mind, in (21a), *to report cases of sexual abuse* may be interpreted as an action that the referents of the NP *women* do not/will not carry out – but the choice is theirs – and in (21b), *reporting their concerns to managers* may be interpreted as an action that the referents of *they* are afraid of carrying out, and do so unwillingly, if at all. Note that in (21a, b) the lower verb *report* requires an Agent subject, and the examples are thus [+Choice], therefore the lack of volition which is contributed by the gerund in (21b) is not entirely in harmony with the agentive reading of *report*.

In (22a, b) the examples are [–Choice], the lower verb is passivized, and in (22a) the *to* infinitive contributes the underlying suggestion that the referent of *she* may play some part in potentially being identified as an Uber driver. This contrasts with the scenario described in (22b), however, where *their* identification as indigenous is beyond their control, with the *-ing* complement NP-like and expressing the source of fear (cf. Rickman & Rudanko 2018: 64).

While it can be argued that there is some small degree of semantic incompatibility discernible in (21b) and (22a) due to the semantics of the complement clause clashing with the dominant theta role reading triggered by each verb, these examples certainly seem natural enough. It is, however, not hard to find examples where the substitution of one complement for the other results in a less natural combination. Such cases from the present dataset are shown in (23a) and (24a), and the second member of each pair is an invented version created by substituting the complement.

(23) a. Anyone fearing falling should visit their doctor. (18-03-28 NZ)
b. ? Anyone fearing to fall should visit their doctor.

(24) a. Nevertheless, almost 50% of these people fear to unveil their sexual orientation or gender, to avoid discrimination and inequalities. (12-09-05 NZ)
b. ? almost 50% of these people fear unveiling their sexual orientation or gender

Here, the notions of volition, control, and responsibility lie much more clearly behind the unnaturalness of (23b) and (24b). (23a) illustrates a [–Choice] context with the *-ing* complement, while (23b), with the verb *fall* in the *to* infinitive form, leads the reader towards a clearly less acceptable [+Choice], agentive interpretation (such an interpretation might be suitable for a deliberate falling scenario, however, like sky-diving). A search of the entire NOW corpus for 'fear* to fall' shows that it is an infrequent combination, with very few relevant examples coming to light, and non-native use evident in some of them. The search string 'fear* falling', on the other hand, returns numerous examples comparable to (23a), confirming that the gerund is by far the more common option. This goes some way towards supporting the present authors' judgment that (23b) is not as acceptable as (23a). It is also worth noting that the structure of (23a) runs counter to the *horror aequi* constraint, and the desire to avoid using *fall* in its *to* infinitive form seems to have overridden the tendency to avoid consecutive gerunds in this case.

As for (24a), a search of the entire NOW corpus is less helpful than it was in the case of (23a, b) above, showing that (24a) is the sole combination of the verbs *fear* and *unveil*. However, it is observed that the lower verb *unveil* seems less conspicuous in the *to* infinitive form than it does with the substituted *-ing* form in (24b), where it suggests a Patient-like subject reading incongruous with the agentive meaning encoded by the verb *unveil*. The adjunct *to avoid discrimination and inequalities* goes with *fear* in the example, but the adjunct conveys purpose, which links well with the agentive and Goal-like meaning of the *to* infinitive. The sentence conveys the idea that if they do not unveil their sexual orientation, they may avoid discrimination. On the other hand, a gerundial complement would be more in line with lack of choice, as in *these people fear having to unveil their sexual orientation*, with the complement identifying the source of fear. An adjunct of purpose would be less likely in such a sentence.

This discussion suggests that certain lower verbs are more flexible than others in terms of their compatibility with both complement types; indeed, as noted by Duffley & Fisher, a great deal depends on 'contextual factors and the nature of the verb itself' (Duffley & Fisher 2021: 93). In the case of *fear*, we can say that lower verbs denoting random or inevitable events or processes that most of us do not look forward to or would want to avoid, and which most clearly assign a Patient theta role and resist an Agent reading, are the verbs that are found most commonly with the gerund – verbs such as *go bald, age, die, fall, become ill* etc. This finding is in line with the Choice Principle. At the other end of the spectrum are those verbs which are more open to the alternative Agent *or* Patient interpretations; such verbs can be used more or less interchangeably, perhaps with some semantic discordance or difference to a greater or lesser degree. It is, after all, often the aim of a writer to catch the reader's attention with unexpected combinations.

5. Conclusion

The present study has used recent data from the New Zealand section of the News on the Web corpus to investigate aspects of the sentential complementation of the matrix verb *fear*, a predicate which, until recently, has not attracted a great deal of attention in this

field. The verb *fear* selects both *to* infinitive and gerundial complements involving subject control, and one aim of the present study has been to investigate a well-defined body of corpus data systematically in order to obtain information on the incidence of the two variants in recent English. Our study indicates that the gerundial variant is clearly predominant in the body of data examined. The finding naturally invites further work in other varieties of current English, in particular, British, American, and Australian English, to give a comparison with the NZE data and help to provide a fuller picture of the relationship between the younger postcolonial varieties, the 'mother' variety of BrE, and the globally influential AmE.

From a theoretical perspective, the aim of the study has been to test the applicability of the Choice Principle to a new set of data and a pair of competing complements which have not previously been analysed from this particular angle. The Choice Principle has its roots in the theoretical framework of semantic roles, with the notions of agency, volition, and choice assumed to be contributing factors in complement selection, and here, with the application of the theory to a previously untested pair of complements and verbal predicate, we have offered a contribution to clarifying the scope of the Choice Principle. In earlier work, it has been shown to be applicable to contexts of prepositional gerunds, and this article indicates that its scope can also extend to contexts of bare gerunds. This finding naturally also invites further work on the complementation of other matrix verbs selecting *to* infinitive and bare gerundial complements, to determine the scope of the principle in such other contexts.

A further benefit of the Choice Principle pointed out in the article is that it draws attention to authentic examples where a *to* infinitive is encountered instead of the expected gerundial variant, or vice-versa. Studying such examples, the degree to which the complements are interchangeable, and the classes of lower verbs that allow both complement types is shown in the article to lead to a more nuanced analysis of the meaning of each construction. Further, this article brings to light a class of predicates, including *go bald* and *become ill*, that typically involve non-agentive or Patient-like interpretations of their subjects and such predicates are quite unlikely to be found in the infinitival form when selected by *fear*. It is suggested here that these findings receive an explanation from the agentive and Goal-like associations of the *to* infinitive, in line with the content of the Choice Principle.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Paul Rickman http://orcid.org/0000-0003-3407-0245
Juhani Rudanko http://orcid.org/0000-0002-8692-7784

References

Data

Davies, Mark. 2008. The NOW Corpus (News on the Web): 10.3 billion words, 2010–present. www.english-corpora.org/now/ [accessed December 2021].

Secondary sources

Berman, Arlene. 1970. Agent, experiencer, and controllability. In S. Kuno (ed.), *Mathematical linguistics and automatic translation* [Report NSF-24]. Cambridge, MA: Harvard University.

Bolinger, Dwight. 1968. Entailment and the meaning of structures. *Glossa* 2, 119–127.

Carnie, Andrew. 2013. *Syntax: A generative introduction*. 3rd ed. John Wiley and Sons.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 1986. *Knowledge of language: Its nature, origin and use*. New York: Praeger.

Davies, William & Stanley Dubinsky. 2004. *The grammar of raising and control*. Malden, MA: Blackwell Publishing.

Denison, David. 1998. Syntax. In Suzanne Romaine (ed.), *The Cambridge history of the English language, volume IV: 1776–1997*, 90–329. Cambridge: Cambridge University Press.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67, 547–619.

Duffley, Patrick. 2000. Gerund vs infinitive as complements of transitive verbs in English: The problems of 'tense' and 'control.' *Journal of English Linguistics* 28(3), 221–248.

Duffley, Patrick & Ryan Fisher. 2021. *To*-infinitive and gerund-participle clauses with the verbs *dread* and *fear*. *Studia Linguistica* 75(1), 72–96.

Gordon, Elizabeth, Lyle Campbell, Jennifer Hay, Margaret Maclagan, Andrea Sudbury & Peter Trudgill. 2004. *New Zealand English: Its origins and evolution*. Cambridge: Cambridge University Press.

Gruber, Jeffrey S. 1967. Look and see. *Language* 43(4), 937–947.

Huddleston, Rodney & Geoffrey Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: The MIT Press.

Jespersen, Otto. [1940] 1961. *A modern English grammar on historical principles. Part V: Syntax (Vol. IV)*. London: Allen and Unwin.

Lakoff, George. 1977. Linguistic gestalts. In A. B. Woodford, Samuel Fox & Sholamuth Philosoph (eds.), *Papers from the thirteenth regional meeting of the Chicago Linguistics Society*, 236–287. Chicago: Chicago Linguistics Society.

Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge, MA: The MIT Press.

Oxford English Dictionary, 2nd edn. 1989. OED online. Available from http://www.oed.com. [Accessed January 2022] Oxford: Oxford University Press.

Pope, Alexander. 1709. An essay on criticism. Available online at http://olympos.cz/Antika/Uceni/Sarkissian/Pope.pdf

Postal, Paul. 1974. *On raising: One rule of grammar and its theoretical implications*. Cambridge, MA: The MIT Press.

Quinn, Heidi. 2000. Variation in New Zealand English syntax and morphology. In Alan Bell & Koenraad Kuiper (eds.), *New Zealand English*, 173–197. Amsterdam/Philadelphia: John Benjamins.

Rickman, Paul & Juhani Rudanko. 2018. *Corpus-based studies on non-finite complements in recent English*. London: Palgrave Macmillan.

Roeper, Thomas. 1987. Implicit arguments and the head-complement relation. *Linguistic Inquiry* 18(2), 267–310.

Rohdenburg, Günter & Britta Mondorf (eds.). 2003. *Determinants of grammatical variation in English*. Berlin: De Gruyter.

Rohdenburg, Günter. 2003. Cognitive complexity and *horror aequi* as factors determining the use of interrogative clause linkers in English. In Rohdenburg & Mondorf (eds.), 205–250.

Rohdenburg, Günter. 2006. The role of functional constraints in the evolution of the English complementation system. In Christiane Dalton-Puffer, Dieter Kastovsky & Herbert Schendl (eds.), *Syntax, style and grammatical norms*, 143–166. Bern: Peter Lang.

Rohdenburg, Günter. 2016. Tracking two processing principles with respect to the extraction of elements out of complement clauses in English. *English Language and Linguistics* 20(3), 463–486.

Ross, John Robert. 2004. Nouniness. In Bas Aarts, David Denison, Evelien Keizer & Gergana Popova (eds.), *Fuzzy grammar: A reader*, 351–422. Oxford: Oxford University Press.

Rudanko, Juhani. 1998. *To* infinitive and *to -ing* complements: A look at some matrix verbs in Late Modern English and later. *English Studies* 79(4), 336–348.

Rudanko, Juhani. 2006. Watching English grammar change. *English Language and Linguistics* 10(1), 31–48.

Rudanko, Juhani. 2011. *Changes in complementation in British and American English: Corpus-based studies on non-finite complements in recent English*. London: Palgrave Macmillan.

Rudanko, Juhani. 2012. Exploring aspects of the Great Complement Shift, with evidence from the TIME corpus and COCA. In Terttu Nevalainen & Elizabeth Closs Traugott (eds.), *The Oxford handbook of the history of English*, 222–232. Oxford: Oxford University Press.

Rudanko, Juhani. 2014. A new angle on infinitival and *of -ing* complements of *afraid*, with evidence from the *TIME* Corpus. In Kristin Davidse, Caroline Gentens, Lobke Ghesquière & Lieven Vandelanotte (eds.), *Corpus interrogation and grammatical patterns*, 223–238. Amsterdam: John Benjamins.

Rudanko, Juhani. 2015. *Linking form and meaning: Studies on selected control patterns in recent English*. Houndmills, Basingstoke: Palgrave Macmillan.

Rudanko, Juhani. 2017. *Infinitives and gerunds in recent English: Studies on non-finite complements with data from large corpora*. London: Palgrave Macmillan Springer.

Ruohonen, Juho & Juhani Rudanko. 2019. Comparing explanatory principles of complement selection statistically: A case study based on Canadian English. *Studia Neophilologica* 91(3), 296–313.

Ruohonen, Juho & Juhani Rudanko. 2021. *Infinitival vs gerundial complementation with afraid, accustomed, and prone: Multivariate corpus studies*. Cham: Springer International Publishing.

Vosberg, Uwe. 2003a. The role of extractions and *horror aequi* in the evolution of *-ing* complements in Modern English. In Rohdenburg & Mondorf (eds.), 305–327.

Vosberg, Uwe. 2003b. Cognitive complexity and the establishment of *-ing* constructions with retrospective verbs in Modern English. In Marina Dossena & Charles Jones (eds.), *Insights into Late Modern English*, 197–220. Berlin: Mouton de Gruyter.

Vosberg, Uwe. 2006. *Die grosse Komplementverschiebung*. Tübingen: Narr.

Warner, Anthony. 1993. *English auxiliaries: Structure and history*. Cambridge: Cambridge University Press.