

**Drifting towards a new form of English – Reversing the
regularization of irregular verbs in New Zealand newspaper
English 1996–2012**

Lauri Sten
Tampere University
Faculty of Information Technology and Communication Sciences
Master's Programme in English Language and Literature
MA Thesis
May 2019

Tampereen yliopisto
Informaatioteknologian ja viestinnän tiedekunta
Englannin kielen ja kirjallisuuden maisteriopinnot

STEN, LAURI: Drifting towards a new form of English – Reversing the regularization of irregular verbs in New Zealand newspaper English 1996–2012

Pro gradu -tutkielma, 82 sivua + 1 liite

Toukokuu 2018

Tämä pro gradu -tutkielma käsittelee epäsäännöllisten verbien käyttöä uusiseelantilaisissa sanomalehtiteksteissä 1990- ja 2010-luvuilla. Tarkastelun kohteena ovat sellaiset verbit, joista on mahdollista käyttää sekä säännöllisiä että epäsäännöllisiä menneen ajan muotoja (imperfekti ja partisiippi). Uudenseelanninenglanti pohjautuu brittienglantiin, jossa epäsäännöllisten verbimuotojen käyttö on esimerkiksi amerikanenglantia yleisempää. Aiemmissä tutkimuksissa onkin havaittu uudenseelanninenglannin muistuttavan brittienglantia huomattavasti tässä suhteessa.

Tutkimuksessa analysoidaan 18:aa verbiä: *sneak, dive, knit, lean, dream, spoil, learn, burn, smell, spell, leap, hang, quit, speed, wed, light, prove* ja *get*. Verbit valikoituvat sen perusteella, että ne edustavat kattavasti erilaisia epäsäännöllisten verbien tyyppejä ja siksi, että niistä on tehty aiempaa tutkimusta britti- ja amerikanenglannissa. Verbien käyttöä tutkitaan imperfektissä sekä verbi- ja adjektiivimuotoisina partisiippeina. Tutkimusaineistona toimii uusiseelantilaisista sanomalehtiteksteistä koostuva Corpus of New Zealand Newspaper English, joka on diakronisen muutoksen tutkimiseksi jaettu kahteen alikorpukseen. Alikorpuksista toinen käsittää tekstejä vuodelta 1996 ja 1997 ja toinen vuosilta 2011 ja 2012. Epärelevanttien hakutulosten poistamisen jälkeen jokaisen tutkimukseen valitun verbin eri muotojen osuudet kaikissa kirjataan ylös kummassakin alikorpuksessa. Saatuja lukuja verrataan keskenään, jolloin selviää mahdollisten muutosten suunta ja suuruus.

Tutkimuksessa havaitaan uudenseelanninenglannin olevan siirtymässä hypoteesin vastaisesti kohti laajempaa epäsäännöllisten verbimuotojen käyttöä – tutkimuksen verbeistä puolella epäsäännöllisten muotojen osuus kasvoi tilastollisesti merkittävällä tavalla. Korpusteksteissä epäsäännöllisten muotojen käyttö lisääntyi enemmän partisiippi- kuin imperfektimuodoissa. Adjektiivina käytettyjen partisiippien osuuksissa ei kuitenkaan ole havaittavissa merkittäviä muutoksia alikorpusten välillä. Havaittuja muutoksia ei voi selittää britti- eikä amerikanenglannin vaikutuksella, sillä muutosta on tapahtunut sekä amerikanenglannin (muodot *snuck, proven* ja *gotten*) että brittienglannin (muodot *leant, spoilt, learnt, spelt, leapt* ja *sped*) suuntaan. Yhteistä muutoksille on siirtymä kohti epäsäännöllisempien muotojen suurempaa käyttöä.

Avainsanat: epäsäännölliset verbit, korpustutkimus, morfologia, uudenseelanninenglanti

Table of Contents

1. Introduction.....	1
2. Background.....	4
2.1. New Zealand English.....	4
2.1.1 Schneider's <i>Dynamic Model</i>	5
2.1.2 Colonial lag.....	9
2.1.3 Key features.....	11
2.2. Regular and irregular verbs.....	14
2.2.1 Grammars.....	14
2.2.2 <i>New Zealand English Grammar: Fact or Fiction?</i>	18
2.2.3 Regularization.....	21
3. Methodology and materials.....	25
3.1 Corpus linguistics.....	25
3.2 <i>Corpus of New Zealand Newspaper English</i>	27
3.3 Obtaining corpus data.....	29
3.4 Methodological issues.....	30
3.5 Statistical significance and the chi-squared test.....	32
4. Results.....	37
4.1 <i>Sneak</i>	37
4.2 <i>Dive</i>	40
4.3 <i>Knit</i>	42
4.4 <i>Lean</i>	43
4.5 <i>Dream</i>	45
4.6 <i>Spoil</i>	46
4.7 <i>Learn</i>	48
4.8 <i>Burn</i>	50
4.9 <i>Smell</i>	52
4.10 <i>Spell</i>	54
4.11 <i>Leap</i>	56
4.12 <i>Hang</i>	57
4.13 <i>Quit</i>	59
4.14 <i>Speed</i>	61
4.15 <i>Wed</i>	63
4.16 <i>Light</i>	65
4.17 <i>Prove</i>	67
4.18 <i>Get</i>	69
4.19 Discussion.....	72
5. Summary.....	76
References.....	80
Appendix.....	83

1 INTRODUCTION

Some verbs can be either regular or irregular in their past tense forms, with irregular forms being generally associated with British English and regular forms with American English. The two forms are typically identical in meaning, as is the case in the following example sentences:

- (1) It was only in his later years that he **learned** to speak Maori.
(ST_13_07_2011_59)
- (2) Originally from the Channel Islands, in New Zealand he **learnt** to speak and write fluent Maori. (TD_11_06_2011_28)

In this thesis, I will research verbs of this type in New Zealand English by using a corpus containing newspaper texts from the years 1996 to 2012. I aim to find out whether or not the preferred forms have changed between the 1990s and the 2010s, and if they have, what the direction of the change is.

As one of the youngest varieties of English (Hay 2009: 84), New Zealand English has not been studied as extensively as bigger, more established varieties such as British English and American English. With this thesis, I hope to add to the existing research on NZE, updating the results from previous research into the 2010s and observing diachronic change over a relatively short period, particularly from the point of view of verbal regularization. Verbs with varying degrees of irregularity are a fruitful research topic because their use is in most cases a matter of preference – when both variants are equally correct grammatically, the choice is made based on extralinguistic factors such as identity construction and prescription. Thus, the differences between different varieties of English are readily apparent in the form distributions of certain verbs and can be quantified in this way.

As elaborated in subsection 2.1.1, New Zealand English is moving towards a phase of linguistic development in which its grammar is codified (Hundt 1998: 1). Before this codification happens, the linguistic norms governing NZE are not specified: they are not exclusively British anymore, not Australian either (though some people bundle AusE and NZE together as ‘Antipodean English’ (Schneider 2007: 127)), and certainly not American. Researching irregular verbs may shed some light on the norms that are presently prevalent in New Zealand English. Diachronic change is of particular interest for this study, as studying the direction of the change is useful in updating the results of previous research to the present and can help predict future developments.

Previous studies (e.g. Hundt 1998, Bauer 1987) have found New Zealand English to greatly resemble British English in its use of irregular verb forms where American English would prefer the regular form, and even lagging behind BrE in regularization with some verbs (Hundt 1998: 135). Based on this, the logical hypothesis would be that New Zealand English takes after British English in its use of irregular verb forms and is currently undergoing regularization. In other words, the proportions of regular and irregular forms can be expected to be quite similar to those found in British English in previous literature. Irregular forms such as *learnt*, *dreamt*, and *leapt* would either lose ground to their regular counterparts (*learned*, *dreamed*, and *leaped*, respectively) or see very little change over time. Does this hypothesis hold up? In order to survey the state of verbs with varying degrees of irregularity in New Zealand English, I aim to provide answers to the following research questions:

1. Which of the studied verbs prefer regular forms and which irregular forms in New Zealand English corpus data?
2. Do regular and irregular forms have different distributions in past tense, past participle, and adjectival usage?

3. How have the proportions of the different forms changed over the years for each verb?

The structure of this thesis is as follows: Chapter 2 is divided into two sections, the first of which concerns New Zealand English and its historical development. The second section of the chapter is a survey of background literature on the verbs chosen for this study and regularization of irregular verbs. Next, the third chapter concerns materials and methodology. After an introduction to corpus linguistic methodology and justifying its use for this study, the corpus used in this study is introduced. The section on methodology mainly describes the process of obtaining the desired results from the corpus and the testing of statistical significance. In the fourth chapter, all 18 verbs examined in this thesis are presented in their own sections. There is also a further section for identifying the overall trends in the corpus data. Finally, the fifth chapter consists of a summary of the main research results in the form of explicit answers to the research questions, as well as some reflection on the possible reasons for the results and its implications for future research.

2 BACKGROUND

In this chapter I will outline the theoretical framework this thesis is based on. First, section 2.1 will shed some light on the development of New Zealand English from its origins as a British contact variety to its present status as a full-fledged postcolonial variety of English by using Schneider's Dynamic Model as an outline. A brief overview of the defining features of New Zealand English is also provided. Section 2.2 presents a concise overview on previous literature about verbs which can take both regular and irregular forms. The main points of focus are descriptions of the chosen verbs in three major English grammars and studies carried out by Biber et al. and Hundt, as well as regularization, a diachronic process affecting some irregular verbs.

2.1 New Zealand English

New Zealand English refers to a variety of English spoken primarily in New Zealand that is spoken by around 4 million people either as a first or a second language (Stats NZ 2013), meaning that almost everyone in the country speaks or at least understands it. It is one of the New Zealand's three official languages, the others being Māori and New Zealand Sign Language. The status of English is very strong in New Zealand – the language with the second highest number of speakers is Māori, which is spoken by around one fourth of the Māori population (Hay et al. 2008: 11), or just under 4 per cent of the overall population. The largest foreign languages are Samoan, Hindi, Mandarin, and Yue (Cantonese), all spoken by 1–2% of the population (Stats NZ 2013).

2.1.1 *Schneider's Dynamic Model*

Several linguists have presented models categorizing global varieties of English and explaining their unique traits. Edgar Schneider (2007) provides one such model, the Dynamic Model of the evolution of Postcolonial Englishes, which provides a framework for classifying different varieties of English and predicting their future developments. Schneider's model postulates that all postcolonial Englishes go through similar phases of development, numbered 1 to 5. The phases are as follows:

- Phase 1: Foundation
- Phase 2: Exonormative stabilization
- Phase 3: Nativization
- Phase 4: Endonormative stabilization
- Phase 5: Differentiation

Schneider argues that New Zealand English is in the early stages of Phase 5, making it an advanced variety. Its development is illustrated by the stages of Schneider's model in the following paragraphs.

In Phase 1 (foundation), the settlers (STL strand) transplant their language into a new, non-English-speaking country. The indigenous people (IDG strand) form the majority, and the limited contacts between the two strands are typically facilitated by bilingual members of the IDG strand (Schneider 2007: 33–6). New Zealand English entered Phase 1 upon the arrival of early settlers in the late 18th century. The first Europeans to set foot on New Zealand soil were the men of Captain James Cook, in 1769¹ (Hay et al. 2008: 3–4). Cook subsequently claimed the islands for the British crown

¹ A Dutch crew led by the explorer Abel Tasman reached New Zealand over 100 years earlier in 1642. The crew did not, however, make landfall. The country is named after the Dutch region of Zeeland (Hay et al. 2008: 4).

(*ibid.*). The contacts with the native Māori population were far from extensive, with Māori influence being limited mostly to toponyms (Schneider 2007: 127–8).

The signing of the Treaty of Waitangi between the British colonists and local Māori chiefs in 1840 effectively established British sovereignty over the country. Mass immigration from the British Isles followed, ushering NZE into Phase 2, exonormative stabilization (Schneider 2007: 128). The identities of both IDG and STL strands were still relatively unchanged – the settlers still viewed themselves as outposts of Britain (*ibid.*: 36–8). English, which was held to British standards, was made the language of government, law, and education (*ibid.*). The native Māori population, now a minority in their own country, was forced to adjust to the new situation and began speaking English in larger numbers (*ibid.*: 128–9). Lexical borrowing from Māori got more extensive, with names for local flora and fauna and concepts relating to Māori culture entering NZE during this stage (*ibid.*).

The first settlers from the British Isles were a heterogeneous mixture: according to an 1871 census, slightly more than a half (51%) came from England, with Scots and the Irish being represented by 27 and 22 per cent, respectively (Hay et al. 2008: 6). The largest non-British settler group was Australians, amounting to 6.5% of the total number of settlers (*ibid.*). Among these groups of people, a process of *koinéization* took place. In *koinéization*, or dialect leveling, the varying dialects of colonial settlers come into contact, with new generations speaking a new linguistic variety distinct from the ones spoken by their parents (Schneider 2007: 35). After the massive British influx following the signing of the Treaty of Waitangi, this is precisely what took place in New Zealand. During the latter half of the 19th century, New Zealand-born children grew up speaking a rather uniform English which contained features from English dialects spoken in various

regions of England, Scotland, and Ireland (ibid.: 128–9). The perceived similarity of NZE to Australian English may be explained with the analogous koinéization processes of the two nations: early British immigration to Australia, though taking place decades earlier, featured settlers from different areas of the British Isles in very similar proportions (Trudgill 2000: 158). The resulting koiné Englishes on both sides of the Tasman Sea, therefore, went on to resemble each other more so than other varieties of English, despite there being relatively small amounts of linguistic contact.

In Phase 3 (nativization), both IDG and STL strands realize that fundamental changes in society have happened – it is no longer possible to remain locked in old IDG or STL communities or identities (Schneider 2007: 40). Schneider places the onset of this phase in Britain granting New Zealand the status of Dominion in 1907 (ibid.: 129). Many countries gain political independence during this stage (ibid.: 130), as did New Zealand². The indigenous Māori population underwent a large-scale language shift to English (ibid.: 42, 130). It was during this phase that many of the defining features of NZE, to be discussed in 2.1.3, took form (ibid.: 130–1). A specific New Zealand accent formed (ibid.: 130) – people from all over the country began noticing (and complaining about) a perceived ‘colonial twang’ in the language of local children (Hay et al. 2008: 84). The English of New Zealanders was seen by some prescriptivists as “an inferior colonial and corrupted version of British English” (Hundt 1998: 2) – NZE was still not recognized as a legitimate variety of English. These attitudes were widespread and manifested themselves in a *complaint tradition*: letters of complaint about a decline in linguistic standards of children (Hay et al. 2008: 87). Among the numerous postcolonial Englishes, NZE has had a particularly vibrant and well-documented complaint tradition (ibid.).

² New Zealand has no single, fixed date for its independence. It is, however, universally agreed that New Zealand was an independent country by the time Phase 4 of NZE development began in 1973.

According to Schneider's model (2007: 48–9), Phase 4 in the development of a postcolonial English, endonormative stabilization, typically begins with an Event X – an incident after which it becomes abundantly clear to the residents of the colony that the colonial center of power (Britain in this case) cares considerably less about them than the other way around. The immediate reaction from the STL strand is to redefine their positions and (national) identities (ibid.). The entry of the United Kingdom into the European Economic Community (the predecessor of the present-day European Union), which made New Zealand lose its priority status as a trading partner of the UK, has been postulated by Schneider as this type of cataclysmic event (ibid.: 131). The identity of the STL strand is now that of the new nation, one that also includes the IDG strand (ibid.: 49–50). In New Zealand, this meant making Māori the co-official language and recognizing Māori heritage as a unique and important part of national identity (ibid.: 131). In a move away from the complaint tradition of Phase 3, local forms of English are more readily accepted (ibid.: 49–51). There is typically a remarkable degree of linguistic homogeneity, and heterogenous elements are often overlooked in the name of national unity (ibid.). Besides linguistic homogeneity, NZE demonstrates other Phase 4 hallmarks as well, including a literary tradition and codification by means of dictionaries and lists of defining grammatical features (ibid.: 132).

New Zealand can be said to have entered Phase 5 (differentiation) of the Dynamic Model in the 1990s, when signs of fragmentation into regional dialects began to appear (Schneider 2007: 53–4, 132). The image of a homogenous language no longer needs to be kept up because the nation is confident in its identity, and various peer-group memberships typically eclipse nationality as identity markers (ibid.: 52–3). For example, Māori English, a variety of NZE spoken amongst the indigenous population, is used to

mark Māori identity, though some researchers have found it to be “elusive” (ibid.: 133). It is to be noted that Schneider’s model cannot predict the course of change regarding a single feature such as irregular verbs. On one hand, one might expect further differentiation through rejection of British norms from a variety advancing deeper into Phase 5, but on the other hand, the identity construction of Phase 5 English speakers is not as dependent on the explicit rejection of Britishisms as it was in Phase 4.

2.1.2 *Colonial lag*

The concept of *colonial lag* has been utilized in explaining the allegedly conservative nature of postcolonial Englishes such as NZE. The term refers to a perceived delay in normal linguistic change lasting roughly one generation, or thirty years (Trudgill 1999: 227). This delay is a natural consequence of koinéization: there is most often no shared peer-group dialect among children to acquire in first-generation colonial situations (ibid.). This, of course, carries the assumption that children speak like other children instead of their parents, teachers or other adults – Trudgill argues that nearly all children, up to a certain age, accommodate their speech patterns totally or almost totally to match those of their peer group (ibid.: 227–8). In cases of koinéization, of course, there is no common dialect to accommodate to (ibid.: 228–30).

Trudgill’s research of colonial lag in NZE uses mid-20th century recordings of New Zealanders born in mid-to-late 19th century, providing a rare look into the language of people whose formative years took place during the koinéization process of NZE (Trudgill 1999, 229). The results speak heavily in favor of the colonial lag hypothesis:

people who had grown up in mixed-origin, non-isolated communities³, did show innovative combinations in their dialect mixtures, but they mostly turned out to be combinations of conservative features (ibid.: 229–231). This kind fossilization, of course, temporarily reverses the usual trajectory of linguistic change. As a result, Trudgill (ibid.) claims that the speech of New Zealanders he studied, born between 1850 and 1890, resembles that of Britons born 30 years earlier.

In her overview of linguistic developments in British and American English, Hundt (2009: 24–27) provides a critical look into utilizing the concept of colonial lag when discussing verb regularization. Because American English has advanced further in regularization of irregular verbs such as *burn*, *learn*, and *spell* than British English, ‘home lag’ seems to be a more appropriate term. Calling the situation home lag is not unproblematic or straightforward, either: AmE did initially lag behind BrE in regularization, having only been the most regularized variety since the 20th century. It is, however, to be noted that as some irregular forms such as *smelt* and *dreamt* are more recent innovations (ibid.: 27), labeling their use as a lag of any kind would not adequately represent the reality of linguistic change.

Hundt (2009: 34) concludes that the concept of colonial lag should not be used haphazardly to explain any and all differences between Englishes: it is perhaps best suited for discussing the early stages of colonization and drawing synchronic comparisons between an emerging colonial variety and BrE. After all, postcolonial Englishes are affected by several diachronic processes that can take various patterns: there are innovations like replacing the third person singular *-th* with *-s*, parallel developments (in

³ Some informants had grown up in isolated, rural communities and had therefore acquired the only variant of English available to them: the English, Scottish, or Irish English of their immigrant parents (Trudgill 1999: 230).

the concord of collective nouns, for instance), and resurrections of older features such as the form *gotten* in AmE (Hundt 2009: 32). Furthermore, the word *lag* is problematic because it assumes that linguistic change is linear and has a clear direction of change (ibid.: 34), which is far from being true. As Hundt (2009: 34) puts it: “Differential language change in BrE and AmE is not merely a case of ETE [extraterritorial English] conservatism or home lag. The reality is much more complex [---]”. This issue will be discussed in greater detail with regularization in subsection 2.2.3.

2.1.3 *Key features*

Despite long being considered a form of British English (Hundt 1998: 3–4), New Zealand English has a number of standout features which are rare or nonexistent in other Englishes. On the phonological level, most of the uniqueness stems from the vowel system. In NZE, the vowel /e/ has undergone such a radical raising that it is currently virtually indistinguishable from /i/ in all ways except vowel length (Hay et al. 2008: 24). Conversely, /ɪ/ has shifted from being a front central vowel, as in BrE or AusE, to a mid central vowel that has merged with /ə/ (ibid.: 23). Other notable traits of the NZE vowel system are the relative centrality of /a/ and /u/, both of which are back vowels in RP (ibid.: 22–4). On the other hand, the consonants of NZE are mostly very similar to other varieties of English (ibid.: 17). General NZE is non-rhotic, features virtually no /h/-dropping outside of unstressed grammatical words, and intervocalic /t/ and /d/ are rarely if ever realized as glottal stops [ʔ] (ibid.: 17–20). Postvocalic /l/ is losing its tongue tip contact and becoming vocalized, to the point of sounding like [ʊ] (ibid.: 35). This is a trait shared by numerous World Englishes, but Hay et al. (ibid.) claim that it has spread the furthest

in NZE. On the suprasegmental side, New Zealand accent has come to include a rising intonation (High Rising Terminal, or HRT) in non-interrogative sentences (ibid.: 27–29).

Regarding syntax, some verbs with separate simple past and past participle forms tend to be merged, with the participle form being used in simple past tense as well. Consider the following example from Hay et al. (2008: 49):

(3) I liked it. I only **done** it till fourth form though.

This past-tense merger is particularly noticeable in the speech patterns of young people in general, and young women in particular (Hay et al. 2008: 48–9). Prominence among these groups suggests a spreading feature – after all, women have been found to use more innovative, up-and-coming linguistic forms (Labov 1990: 206). Auxiliary *have* tends to be deleted in spoken language (Hay et al. 2008: 50–1), as in the following example (ibid.: 50):

(4) Cause **I been** through concussion and that was horrible.

Regarding collective nouns such as *team*, *government*, or *crowd*, NZE tends to go for plural agreement (“*The crowd are cheering*”, as opposed to “*The crowd is cheering*”) more than American English, but less than British English (Hay et al. 2008: 56). What is more, the proportions of singular and plural with collective nouns have been noted to change over time, with plural use gaining prominence (Rickman 2018). Rickman’s findings have some interesting implications for the present study: if similar development were to occur in verbs, the regularization of irregular verbs would actually reverse, since irregular verb forms are a feature associated with British English more than American English, with their use in New Zealand English occupying an intermediate position.

The singular *they*, typically used when the referent is of unknown or unspecified gender, is noted by Hay et al. (2008: 58–9) to register “very high” rates of usage in NZE.

In fact, the pronoun is gaining use even when talking about a specific person whose gender is known (ibid.). Another NZE pronoun not found in standard British English is the second-person plural *yous* which is typically used by children (ibid.: 60). The presence of *yous* in NZE (and AusE) is likely to be caused by Irish English influence (Burridge and Musgrave 2014: 31). Finally, Hay et al. (2008: 61–2) note that double comparison – the use of both *more* or *most* and *-er* or *-est* with an adjective – is becoming increasingly accepted in NZE.

The biggest standout feature of NZE vocabulary are loanwords from Māori language (Hay et al. 2008: 67–8). Lexical Māori influence can be divided into two periods, before 1860 and after 1970 (ibid.: 68), which correspond to Phase 2 and Phases 4–5 of Schneider’s Dynamic Model, respectively. Early Māori loans fall into three categories: flora and fauna (*kotuku* ‘white heron’), society and culture (*tapu* ‘sacred’), and place names (*Tauranga*) (ibid., 68–9). The more recent wave of borrowings from Māori is tied to a wider societal acceptance of Māori culture and customs, and includes even words with direct equivalents in English (ibid., 70–2). For example, NZE now uses the Māori loan *waka* instead of *canoe* and *iwi* instead of *tribe* (ibid.: 71).

Hay et al. note an expanding influence on NZE by American English. This influence is spreading via popular culture: American TV shows, movies, and music have all left their mark on the English spoken on the other side of the Pacific Ocean (Hay et al. 2008: 75–6). For example, NZE uses the AmE *stove* over the BrE *cooker*, and the AmE *truck* over the BrE *lorry*. With other words, such as *movie/film*, both usages are acceptable (ibid.: 76–7). The influence has been noted by Bayard (1989) as spreading: in other words, using American English vocabulary is gaining wider acceptance in New Zealand society. Out of the verb forms in this study, *snuck* and *gotten* in particular are associated with

AmE, and their potential spread might indicate a more widespread acceptance of grammatical forms associated with American English in NZE.

2.2 Regular and irregular verbs

Most verbs in the English language are regular⁴, having both their simple past and past participle forms end in *-ed* (Biber et al. 1999: 392). However, around 200 to 250 English-language verbs are irregular⁵, meaning that their past tense and past participle forms are not formed with the suffix *-ed* (ibid.: 394, Quirk et al. 1985: 104). Many of these verbs are in common use (ibid.) and therefore provide sufficient corpus data for a quantitative study. What is more, some irregular verbs can take a regular *-ed* ending as well as an irregular one (Biber et al. 1999: 396). The choice between the two possible past tense forms is governed by factors such as register (spoken or written, formal or informal), grammatical function (simple past or participle), and the verb itself (different verbs show different preferences) (ibid.). Notably, AmE has been found to favor regular forms to irregular ones in more verbs more so than BrE, and the general trajectory of change has been towards a wider use of regular forms (ibid.).

2.2.1 Grammars

Based on their patterns in past tense and participial forms, Biber et al. (1999: 394–6) divide English irregular verbs into seven classes:

- Class 1 verbs take a voiceless *-t* suffix (*send, sent; learn, learnt*)

⁴ In this thesis I will use the term ‘regular’ to describe verbs whose past tense forms have the ending *-ed* in both simple past and participle forms, and ‘irregular’ for all other verbs. The terms ‘weak verb’ and ‘strong verb’ are occasionally used in the literature to describe the same phenomenon. For further discussion of the problematics of terminology, see Anderwald 2009: 4–5.

⁵ Discounting derivations such as *undo* from *do*, the number of irregular verbs decreases drastically to around 70 (Peters 2009: 13).

- Class 2 verbs change their base vowel and take the same -t or -d suffix in both past tense and past participle forms (*keep, kept; sell, sold*)
- Class 3 verbs take the regular -ed suffix in past tense and -(e)n in past participle (*show, showed, shown*)
- Class 4 verbs have no suffix in simple past and the suffix -(e)n in past participle. The vowel changes either once or twice (*fall, fell, fallen; wear, wore, worn*)
- With Class 5 verbs, only the vowel changes between tenses (*come, came, come; hang, hung, hung*)
- Class 6 verbs are identical in all tenses (*hit, hit, hit; let, let, let*)
- Class 7 verbs have one or more completely unrelated form (*go, went, gone*)

Biber et al. (1999: 397) provide a corpus study of British and American English data on the past tense forms of sixteen verbs, ordered here from the most regular to the most irregular: *sneak, dive, knit, lean, dream, spoil, learn, burn, smell, spell, leap, hang, quit, speed, wed, and light*. All of the above verbs were chosen for the present study in order to represent different types of irregular verbs which can take the regular past tense suffix -ed as well. Out of the seven classes of irregular verbs, the selected verbs represent all classes except Class 3 and Class 4. In addition to the verbs on the aforementioned study by Biber et al., two more verbs were selected: *get* and *prove*. These two verbs break the pattern of the sixteen previous verbs: *get* has two possible irregular participle forms, of which *gotten* is considered to be an Americanism (Quirk et al. 1985: 116), while the usual “roles” of the participial forms of *prove* are swapped – *proven*, despite being irregular, is more common in AmE than in BrE (ibid.: 107). With the participial form *gotten*, *get* would be classified as a Class 4 verb, while *proven* represents Class 3, so all categories of irregular verbs can be said to be featured in this study. The results of Biber et al. confirm that American English has advanced further in the process of regularization than the more conservative British English: for all the verbs researched by Biber et al.

(1999: 396–8), past tense forms ending in *-ed* are more common in AmE than in BrE, or equally common.

Next, a more in-depth look is provided into the degrees of regularity that the verbs chosen for this study are said to exhibit according to previous literature. The results of the NZE corpus study are compared to these samples of British and American prescriptive tradition in the fourth chapter of this thesis. This overview is based on three major English language grammars: *Longman Grammar of Spoken and Written English* by Biber et al., *A Comprehensive Grammar of the English Language* by Quirk et al., and *The Cambridge Grammar of the English Language* by Huddleston and Pullum.

Sneak has been found to prefer the regular form in both British and American English (Biber et al. 1999: 397). The form *snuck* is considered to be “jocular”, and thus nonstandard, in tone by Huddleston and Pullum (2002: 1604). Biber et al. (1999: 398) single out *dive* as the only verb in their study to exhibit variation only in past tense – most commonly, both forms or only the participle vary. The preference for *dived* is overwhelming in both BrE and AmE (ibid.). Huddleston and Pullum (2002: 1604) and Quirk et al. (1985: 115) both note that *dove* is an American English innovation. *Knit*, too, is considered to be a primarily regular verb in all three featured grammars. Huddleston and Pullum even go as far as classifying *knit* as a regular verb with only a superficial resemblance to verbs like *hit* and *bid* (2002: 1601). The results from the corpus study by Biber et al. (1999: 397) also show *knit* as preferring the *-ed* ending, as do Quirk et al. (1985: 111). According to Biber et al. (1999: 397), *lean* shows an inclination for regular forms, though the preference is not as big in past tense. Quirk et al. (1985: 107) classify *leant* as a Britishism and *leaned* as an Americanism.

Dream shows preference for the regular ending, though the irregular *dreamt* is fairly common in BrE and not rare in AmE, either. (Quirk et al. 1985: 107; Biber et al. 1999: 397). The verb *spoil* shows a clear split between the majority-regular past tense and the more evenly distributed past participle, as well as considerable trans-Atlantic variation, with AmE preferring regular and BrE irregular forms (Biber et al. 1999: 397). Biber et al. (ibid.), once again, find *learn* to favor the regular ending, though by a larger margin in AmE than in BrE. Quirk et al. (1985: 105) note that adjectival use of *learned* has gained an additional related meaning, ‘scholarly’, not shared by *learnt*. *Burn* is observed by Biber et al. (1999: 396–7) to show a preference for the irregular form *burnt* in past tense and the regular form *burned* in past participle in news texts – the opposite of *spoil*. Again, British usage skews more towards the irregular than American usage (ibid.). *Smell* and *spell* show clear differences between BrE (chiefly irregular) and AmE (chiefly regular) usage (Biber et al. 1999: 397). Their distribution is very similar across the board (ibid.). *Leaped*, the regular form of *leap*, is marked by Quirk et al. (1985: 197) as a feature occurring especially in American English. Similarly, Biber et al. (1999: 397) observe that *leap* has a strong preference for the irregular form in British English and a preference for the regular form in American English.

The verb *hang* is unique among the verbs in this study in that it has a different meaning in BrE for its regular variant *hanged* ‘dead by hanging’ (Quirk et al. 1985: 112, Biber et al. 1999: 396). However, Huddleston and Pullum (2002: 1604) do note that *hung* is occasionally used in that sense as well. Perhaps due to the rather limited usage for *hanged*, *hung* is the more common variant in both British and American English (Biber et al. 1999: 397). *Quit* and *wed* are overwhelmingly irregular verbs in both BrE and AmE, though regular variants are still possible and technically correct (ibid.). The irregular *sped*

is the more common past tense form of *speed*, as stated by Biber et al. (1999: 397). According to Quirk et al. (1985: 112), the regular form *speeded* is used mainly to describe mechanisms and is required for the phrasal verb *speed up*. Huddleston and Pullum (2002: 1602) and Biber et al. (1999: 396–7) both consider *light* to be a regular verb in AmE and an irregular verb in BrE. Quirk et al. (1985: 113) comment on the adjectival usage of the word, stating that *lighted* is the only acceptable variant there.

2.2.2 *New Zealand English Grammar: Fact or Fiction?*

In her book *New Zealand English Grammar: Fact or Fiction?*, Marianne Hundt presents a series of comparative corpus studies focusing on irregular and regular verbs in NZE newspaper texts (1998: 29–38). A major goal for this thesis is to expand upon Hundt’s study by adding more verbs, a larger corpus, and taking diachronic change into account. According to Hundt, NZE is not a mere reflection of British or American norms: for certain verbs (*smell, spell*), the NZE usage is British, while for others (*burn*), NZE is said to take after AmE (ibid.: 29).

Hundt’s first study, regarding the verbs *burn, learn, and dream*, compares their past tense usage in the New Zealand based newspapers *Dominion* and *Evening Post* (both of which are coincidentally featured in the corpus that is analyzed in this thesis) to those in *The Guardian* (BrE) and *Miami Herald* (AmE) (Hundt 1998: 29–30). In the use of these verbs, NZE was found by Hundt to be very close to BrE in usage, with both varieties using regular and irregular forms extensively for each verb (ibid.: 30). American English seems to be the outlier, exhibiting universal or near-universal preference for regular forms (ibid.). The irregular form *burnt* was found by Hundt to be used more often as a participle than in simple past tense, while no such difference exists for *learnt* (ibid.). In adjectival

usage, *burn* shows a greater preference for irregular form, except in the highly regularized AmE (ibid.: 31). *Learn*, on the other hand, assumes the regular form *learned* almost exclusively when used as an adjective (ibid.).

In her second study of irregular verbs, Hundt (1998: 31–3) expands the scope into nine verbs, adding *lean*, *leap*, *smell*, *spell*, *spill*, and *spoil*, and focusing on the differences between British, Australian, and New Zealand English. The main finding of the study is that AusE and NZE bear striking resemblance to each other regarding the use of irregular past tense forms, while regularization seems to be the most advanced in BrE (ibid.: 32). This greater conservatism in the postcolonial varieties is interpreted by Hundt to be a manifestation of colonial lag (ibid.: 33). It is, however, worth noting that the dataset used by Hundt is very small: for each corpus, there are only a little over 200 tokens in total, spread between nine verbs with two possible variants each (Hundt 1998: 32). This means that the results would be easily swayed by only a few statistical outliers.

In the third study, Hundt (1998: 33–6) focuses on the participial forms of *prove* in NZE, BrE, and AmE. Hundt notes that *proven* is the more common form in American English, and that the form is actually gaining ground despite being irregular (ibid.: 33). The results show that NZE takes the intermediate position in the adoption of *proven* – the proportion of *proven* is not as large as in AmE, but larger than in BrE (ibid.: 34). According to Hundt (ibid.), adjectival use of *proven* is more readily accepted than participial use in NZE. Overall, the rise in the form *proven* is seen by Hundt to be an exception to a general trend of regularization, and Hundt even goes as far as theorizing that *proven* might completely replace the regular form, thus creating more irregularity (ibid.: 35).

The final study of irregular verbs in Hundt's book (1998: 36–8) examines the participle form *gotten*. In this case, American English has retained an archaic form which persists despite prescriptivist forces advocating the use of *got* (ibid.: 36). Hundt found very little evidence that this particular Americanism has made any significant headway in NZE: of the eight tokens of *gotten* in the NZE newspaper corpus, five are quoted direct speech of Americans (ibid.: 37). *Gotten* was not found to be any more common in spoken than written language, though teenagers rated it as acceptable more than adults (ibid.). Thus, Hundt raises the question whether or not the youth of 1990s will retain the form as adults (ibid.: 38), a question for which I am seeking to provide an answer in this thesis.

Quinn (1999: 179–80) postulates that the choice of participle form for verbs with variant *-t* or *-ed* ending is governed by phonology and duration of the event in question. According to Quinn (ibid.), Hundt's results indicate that NZE shows a marked preference for *-t* over *-ed* with verbs ending in /l/. For the study at hand, this would predict a strong preference for irregular forms with *spoil*, *spell*, and *smell*. Verbs ending in nasal sounds /m/ or /n/, on the contrary, are predicted by Quinn (ibid.) to prefer *-ed* to *-t*, with the degree of regularity lessening if the stem vowel of the verb is /ɜ/. This indicates a preference for the regular ending for *lean* and *dream*, with *learn* and *burn* showing less of a regular preference. However, Quinn (ibid.) does note that previous studies such as Hundt (1998) have focused on written rather than spoken language. (This is true of corpus linguistics in general.) This complicates matters because some English speakers spell and pronounce participial endings differently: /t/ is a common ending in speech even for verbs that are commonly spelled with *-ed* (Quinn 1999: 179). Regarding duration, Quinn (ibid.: 179–80) notes that events which take place over a short period of time tend towards the irregular, as in the following example from Bauer (1987):

(5) “When the flame caught, the curtains **burnt** immediately.”

On the other hand, events lasting a longer time tend to result in higher proportions of irregular forms (example from Bauer (ibid.)):

(6) “The fire **burned** for hours.”

The influence of syntactic function (adjectival or participial) on the choice of verb form has been studied by both Hundt and Bauer with inconclusive results: some verbs are more regular as adjectives than as participles and vice versa (Quinn 1999: 180).

2.2.3 Regularization

In regularization, irregular grammatical forms fall out of use and are replaced with regular forms, creating further regularity in the linguistic system. In the context of this study, regularization of verbs is of particular interest. Simply put, it refers to a process in which irregular verbs become regular. That is, the original irregular past tense of a verb is replaced with the regular past tense, formed using the suffix *-ed* (Gray et al. 2018: 1). Verbal regularization has been a dominant historical trend: according to Cheshire (1993: 117), the number of irregular verbs in English has decreased greatly over the years, presently being around one-sixth of what it was in Old English. The pace of regularization has never been constant in different varieties of English: AmE tradition advocates regularization to a greater degree than BrE (Peters 2009: 14). For example, Webster’s *American Dictionary of the English Language* (1828) disendorsed the now-obsolete participial forms *bounden*, *bursten*, and *sitten* for *bind*, *burst*, and *sit*, respectively, while British grammarians of the same era held on to them (ibid.). British tradition has thus been more in favor of retaining older irregular forms, with irregular participles being labeled as “more proper and more elegant” by Johnson’s *Grammar* in 1755 (ibid.).

What about the current state of regularization, then? Cheshire (1993: 117) claims that regularization all but seized in standardized varieties of English (particularly BrE) around 500 years ago due to codification. New regular forms typically spread from the lower strata of society: people with no desire to prove their superiority by adhering to conservative, high-prestige, ‘cultivated’ forms generally would not hesitate to use an (initially) non-standard form such as *seed* for the past tense of *see* (ibid.: 119). Another explanation for the relative lack of regularization in modern English given by Cheshire is the desire of some higher-class people to sound as little like ‘uneducated’ children as possible by using irregular forms – after all, children have been known to overgeneralize the regular verb ending *-ed* to verbs like *give* to form **gived* (ibid.: 118). These factors might have led to advanced regularization (or lack thereof) becoming a marker for social differentiation (ibid.: 119). Cheshire (ibid.) and Peters (2009: 14) argue that the few irregular verbs which persist despite regularization are ones with extremely high frequencies. These would, after all, occur in everyday speech so often that they become automatic for practically all speakers (Cheshire 1993: 119).

One further thing worth noting is that regularization does not affect all forms equally: participial forms of irregular verbs have been observed to be more resistant to regularization than simple past forms (Cheshire 1993: 127). This is particularly conspicuous in verbs with alternate *-t/-ed* forms such as *burn*, *dream*, and *spoil* (ibid.). A tendency to mark aspect with different past tense forms has been noted in these verbs by Quirk (1970: 308), whose research concluded that *-ed* is used to mark durative (continuous, habitual, or permanent) aspect, while *-t* is used for other meanings. In other words, when the meaning of past tense is closer to perfective, the form is more likely to resemble the participial *-t* form (Cheshire 1993: 128). Unfortunately, researching the

marking of aspect in New Zealand English is beyond the scope for this thesis, as the checking of aspect would have to be done manually and the search results in the corpus number in the thousands.

In her analysis of the English verb system, Anderwald (2009) applies the framework of natural morphology. According to Anderwald, natural morphology assumes that language is governed by the Principle of Higher Naturalness. That is, it moves towards a direction that increases dominant patterns (Anderwald 2009: 186). According to natural morphology, the dominant verb class by sheer volume is the class of regular verbs, the dominant past tense marker is *-ed*, realized as /t/, /d/, or /ɪd/, and the dominant pattern is distinguishing between past and present forms. This predicts four types of change: either irregular verbs change into regular ones, adopt the dominant past tense marker *-ed* in addition to their irregular endings, or change into a more dominant irregular verb type. The fourth option is an abstract change to the currently dominant pattern where simple past and past participle are identical and set apart from present tense. According to Anderwald (2009: 51), all of these changes except the second one (double-marking of past tense) are attested in non-standard dialects. The first type of change is of particular interest, as it is the most common of the four – for example, the past tense of *know* is regularized into **knowed* in certain dialects (ibid.: 62). Anderwald argues that this regularization witnessed in (non-standard) dialects represents a natural process unhindered by codification and prescriptivism (ibid).

It would certainly seem tempting to view regularization as a linear process, one with a clear-cut goal. However, the history of grammatical forms is far more than that, being marked with new innovations that create more complexity within the language system. For example, verbs like *bend*, *lend*, and *send* were originally regular before the

omission and devoicing of the suffix (*-ed* to *-t*) (Anderwald 2009: 57). The reclassification of *-end* type verbs, in turn, may have inspired previously exclusively regular verbs like *burn* to become variable regular/irregular verbs by analogy (ibid.). A similar, more recent development is the creation of the form *snuck* for *sneaked* in American English (ibid.: 62–3). Thus, new irregular verbs can be expected to form, and gain prominence from time to time.

The re-emergence of the form *gotten*, which is of particular interest in this study, does not seem to fit the pattern of regularization: with the word *get*, something resembling the fourth type of regularization listed by Anderwald (2009: 51) is happening in reverse, from *get/got/got* to *get/got/gotten*. Since Anderwald (ibid.) and Peters (2009: 16) view the paradigm with two different irregular forms as less regular than the one with only one irregular form for both past tense and participle, *get/got/gotten* will be treated as more irregular than *get/got/got*, and any further use of the form *gotten* in the corpus data will be seen as a step away from regularization.

3 METHODOLOGY AND MATERIALS

In this chapter I will outline the materials used to conduct the research, in addition to providing a concise description of the research methodology employed. The first section of this chapter contains an overview of corpus linguistics which serves as the justification for using this particular methodological approach to the research at hand. The following section introduces the corpus used in this study, Corpus of New Zealand Newspaper English, including an overview of the contents of the corpus and the rationale for using this particular corpus to provide answers to the research questions. Then, the actual process of obtaining analyzable data from the corpus is described. Limitations of the corpus, and its part-of-speech tagging in particular, are discussed, along with the ways in which these challenges can be mitigated. Finally, some key methodological choices are explained with the concepts of precision, recall, statistical significance, and the chi-squared test.

3.1 *Corpus linguistics*

Corpus linguistics is a methodological approach to linguistics that employs corpora, large collections of natural texts (Biber 2010: 159). Biber (*ibid.*: 159–61) outlines some basic principles corpus linguistics operates under:

- Corpus linguistics is empirical; it analyzes actual usage of patterns as found in natural texts.
- It makes heavy use of computers for analysis which employs both automatic and interactive methods.
- It is both qualitative and quantitative in its analytical techniques; corpus linguists both document previously unrecognized constructs and provide numerical data on the distribution of certain patterns.

- Corpus linguistics operates under the assumption that linguistic variation is systematic. That is, the distribution of grammatical features is never truly random, but is instead governed by factors and circumstances unique to each linguistic variety.

These factors make a corpus-based approach fruitful for the present study. With large amounts of quantitative data, it is possible to obtain statistically significant results on the distribution of regular and irregular verb forms and the change of that distribution over time.

Two key concepts in corpus linguistics are precision and recall. Precision means the proportion of retrieved tokens that are relevant, while recall means the proportion of relevant information that was retrieved (Ball 1994: 295). A corpus linguist will have to strike a balance between these two, as no query will result in a perfect precision with perfect recall. This is referred to by Ball (*ibid.*) as the recall problem: a search with perfect recall will contain a large amount of irrelevant data that will have to be sorted manually. If the search criteria are narrowed, precision will obviously improve, but there is a risk of missing relevant data – after all, it is impossible to know what data is omitted from the search results if more restrictive search terms are used (*ibid.*). The recall problem has no easy solution, but as perfect or near-perfect recall is essential to achieve any sort of representability, precision has to be compromised, and irrelevant or mistagged search results will have to be discarded manually. In the context of this study, this means performing a large number of simple searches, counting the number of irrelevant tokens, and subtracting them from the total number of search results. Because of the very large number of tokens (17,780 without the forms *got* and *gotten*), perfect precision cannot be guaranteed. The numbers presented in the thesis should be fairly accurate, especially

given the fact that there is always some room for interpretation in some borderline cases. Some of these key methodological issues are touched upon in section 3.4.

3.2 *Corpus of New Zealand Newspaper English*

The data used in this study comes from the Corpus of New Zealand Newspaper English, hereafter referred to as CNZNE, compiled by Paul Rickman of Tampere University. As indicated by its name, CNZNE is a corpus of New Zealand English which consists exclusively of newspaper texts. In its present state, the corpus consists of 796,572,762 words, meaning that it is a rather extensive corpus – most available corpus research on NZE has been carried out with smaller corpora such as the Wellington Corpus of Written New Zealand English and the ICE corpus (Rickman 2017: 171). A larger corpus allows for research to be done on relatively small segments of the corpus which nonetheless contain millions of words and makes it possible to obtain statistically significant results on even some lower-frequency words and structures. As stated by Rickman, “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” (ibid.: 175).

CNZNE consists of texts from 13 New Zealand newspapers: *Daily News / Taranaki Daily News, Dominion, Dominion Post, Evening Post, Evening Standard / Manawatu Standard, Nelson Mail, Press, Southland Times, Sunday News, Sunday Star Times, Timaru Herald, Truth, and Waikato Times*⁶ (Rickman 2017: 173). The featured papers have been chosen to adequately represent regional variation within NZE – there are major newspapers from the three biggest cities in New Zealand, Auckland, Wellington, and Christchurch, as well as smaller, provincial papers (ibid.: 171–2). The

⁶ The example sentences from the CNZNE are tagged by paper and date of publication. The abbreviations for the papers are found in the appendix.

contents were chosen by Rickman to mirror those of the British National Corpus (BNC), with an emphasis on news reports and a mixture of sections from sports to science and topics from real estate to rugby union (ibid.: 172). Access to CNZNE currently restricted so the corpus is not available to the general public.

A thing to note is the ownership of the papers. All of the papers featured in the corpus belong to the same Australia-based media conglomerate, Fairfax Media (Rickman 2017: 171), which owns a significant share of newspapers in Australia and New Zealand (Rickman and Kaunisto 2018: 76)⁷. The dominant position of Fairfax has led to a lack of competition among newspapers in New Zealand – no city in the country has more than one daily newspaper published in it (Gibbons 2014: 184). However, there are some major papers, such as the Dunedin-based *Otago Daily Times*, which are not owned by Fairfax (ibid.: 185) and are thus not included in the CNZNE. Featuring papers from other publishers would be ideal to account for the possible variation, but the advantages of using exclusively Fairfax papers are clear: the material from the Fairfax archives comes as full texts and with tags denoting section, topic, and sub-topic (Rickman 2017: 174). The material coming from a single source also facilitates comparisons between different part of the corpus – the differences found in different parts of the corpus are unlikely to be caused by differences in data collection, tagging or other technical factors. Various kinds of lists were omitted from the corpus: television listings, stock prices, weather forecasts, and other such material were deemed by Rickman (ibid.) to contribute almost nothing to the study of morphology and serve as a very poor representative of NZE in general.

⁷ Fairfax Media has been recently acquired by the Australian media conglomerate Nine and renamed as *Nine Publishing*. Fairfax New Zealand is currently known as *Stuff Limited*.

3.3 Obtaining corpus data

In order to study diachronic variation, two subcorpora representing texts from different time periods were created from the CNZNE. The time periods chosen were 1996–97 and 2011–12, the two earliest and latest years of release for the texts in the corpus. The former subcorpus consists of 34,697,350 words, 8,553,662 from the year 1996 and 26,143,688 from the year 1997, while the latter subcorpus contains 66,030,779 words, of which 46,379,258 are from 2011 and 19,651,521 from 2012. Since the study of distribution is chiefly based on proportions instead of overall frequencies, the results obtained from analyzing the two subcorpora are comparable without normalizing frequencies⁸. The subcorpora were created to achieve the highest possible amount of variation, as well as to continue Hundt’s work from the 1990s by considering the possible changes that have happened since then. There is a 16-year gap between the earliest and latest texts in the corpus so the changes occurring between the two subcorpora are unlikely to be anything dramatic in any direction, but the results of the study nonetheless allow us to observe some overall trends in the distribution of variable past tense and participle forms.

In order to obtain the necessary data from the corpus, searches were performed using simple search strings featuring the desired form and a part-of-speech tag in the CLAWS4 tagset. The relevant tags for the present study were *_VVD* for past tense verbs, *_VVN* for participles, and *_JJ* for adjectives. For example, the following searches were performed for the verb *burn*: *burned_VVD*, *burned_VVN*, *burned_JJ*, *burnt_VVD*, *burnt_VVN*, and *burnt_JJ*. All of the aforementioned searches were performed for both subcorpora, meaning that there were twelve searches for each of the 18 verbs, for a total

⁸ The form *gotten* is the sole exception due to the extremely high number of tokens for *got*. See section 4.18 for more details.

of 216 searches. All search results were examined manually, and mistagged or duplicate tokens were subtracted from the total number of search results as described in subsection. From these numbers, two operations were performed. First, the share of regular and irregular verbs as a percentage will be calculated for each verb, part-of-speech tag, and year. This is done to visualize the changing proportions of regular and irregular forms. Then, the possible changes between the corpora will be checked for statistical significance with the chi-squared test as outlined in section 3.5.

3.4 *Methodological issues*

The corpus is tagged for part of speech with the CLAWS4 tagger, facilitating effective searches with simple search strings. The verb forms studied in this thesis are fortunately very simple and easily identifiable by artificial intelligence: simple past forms are predicates following a noun which is the subject of the sentence, participle forms are preceded by a form of the verb *have* or *be*, and adjectives have an entirely different (attributive or predicative) function in the sentence. However, as noted by Biber et al. (1998: 262), no automatic tagger is 100% accurate. It is thus of utmost importance to check the automatically tagged results for relevance (*ibid.*) and discard the irrelevant tokens.

The tagging errors which are present in CNZNE arise from ambiguities of various types. For example, the same lexical item may be tagged with more than one part-of-speech tag due to the automatic tagger not being able to definitively place the word into one category. Such double-tagging increases recall at the expense of precision, so these cases had to be manually identified and correct judgments made of their actual part of speech.

Technical reasons may also contribute to false tagging, as is the case in the following example:

- (7) Picture: BRUCE MERCER **Knit** one, purl one GRAPHIC: electronic version unavailable. Please see hard copy. (WT_25_06_2011_77)

In this example, the automatic tagging has falsely recognized *Bruce Mercer*, the person credited with providing a picture for the article, as the subject of the headline sentence *Knit one, purl one*. This interpretation of the corpus data, extremely unlikely for human readers, has the verb *knit* used in the simple past tense instead of the present, which would have the third person singular *-s* ending. Such tokens were promptly discarded from this study.

Simple typos contribute to incorrect tagging, too. In the following example, one misplaced letter has completely changed the meaning of the verb:

- (8) All Black Alama Ieremia and Wellington Hurricane Filo Tiatia **leant** a hand in training this week to their Western Suburbs club side which meets Hutt Old Boys-Marist today in a clash of unbeaten teams in Wellington club rugby. (DO_19_04_1997_8)

In (8), the verb *leant* is most likely a misspelling of *lent*, as it is used as part of the common phrase *to lend a hand* ‘to assist’. Misspellings and typos, like the one featured in the above example, were discarded from the study. Another tagging issue comes in the form of words from other languages being mistaken for English-language words by the tagger, as in the following sentence:

- (9) Judge John Bisphan fined **Hung** Tan Nguyen, 29, a fish shop owner, who appeared on a charge of leaving a child without supervision, \$250 plus \$95 costs. (DO_04_11_1997_68)

In (9), the word *Hung*, in this case a name for a man of Vietnamese origin, is automatically tagged as a form of the verb *hang*.

The CNZNE contains some duplicate tokens which distort the search results. As the papers included in the corpus are all owned by the same conglomerate, a news item might get included in the corpus more than one time due to the same news item being reprinted in several newspapers. During the process of compiling the corpus, Rickman (2017: 174–5) did run an automatic check for duplicates which he promptly deleted from the corpus, but some duplicates did nonetheless remain CNZNE. Problems like this are found in other, more high-profile corpora as well (ibid.). These duplicate entries, whether from same or different newspapers, were ignored, and only one instance of each sentence appearing multiple times was counted. Other reasons for duplicate entries in the corpus are headlines and lead paragraphs whose contents are sometimes repeated verbatim in the body of the article. In cases with very similar but not completely identical tokens, such as (10) and (11), both instances were included:

- (10) Bus travellers will now be dropped off into a **lighted** city centre at night, rather than the more isolated museum setting. (ST_27_12_2011_27)
- (11) It would also allow bus travellers to be dropped off in the **lighted** city centre at night instead of being dropped beside the poorly **lit** and isolated museum on the city fringe. (ST_22_12_2011_50)

In cases where the same form appears multiple times in a single text, such as the following example in which repetition is used as a rhetorical device, all instances were counted:

- (12) “I **learned** so much there. I **learned** about setting a goal and reaching it which I never, I don't think, could have learned but there. I set the goal to get out. I **learned** to test myself and to contain myself and to discipline myself and to set goals, visualise the goals and attain them.” (EP_25_08_1997_1)

3.5 *Statistical significance and the chi-squared test*

Statistical significance is obtained from numerical data by performing calculations whose results are compared to criterion values dependent on the test of significance used (Chow

1996: 43). If the values are more extreme than the criterion values, the results are statistically significant (ibid.). In other words, statistically significant results have a low probability of occurring by random chance (ibid.). Being statistically significant does not, in itself, make any distribution linguistically significant, but an absence of statistical significance means that the differences in distribution are not of linguistic interest. As Stefanowitsch (2004: 1) puts it, “statistical significance is a precondition for linguistic significance, but not a guarantee.”

The chi-squared, or χ^2 , test is a frequently used test of statistical significance for comparing numerical data from different sources (Oakes 2009: 163), and is the test chosen for the present study. The test works by comparing a sample observation against a predicted value which is assumed to be binomially distributed (Wallis 2013: 351). The chi-squared test must be performed without normalizing frequencies, as it takes different sample sizes into account by design (Oakes 2009: 165). In the chi-squared test, the values examined are set out in a contingency table and the totals are counted for each row and column (Oakes 2009: 163–4), as in the following example table with the distribution of participial *burnt* and *burned* in the CNZNE:

	<i>burnt</i>	<i>burned</i>	row totals
1996–97	219	133	352
2011–12	443	239	682
column totals	662	372	1034 (grand total)

Table 1.

Contingency table for participial forms of burn in CNZNE.

The first things to calculate are the expected value if the frequencies of the studied items were the exact same in all datasets. This is done with the following formula (Oakes 2009: 163):

Expected frequency = (Row total * Column total) / Grand total

From this calculation, we get the following expected values:

	<i>burnt</i>	<i>burned</i>
1996–97	225.36	126.64
2011–12	436.64	245.36

Table 2.

Expected values for participial forms of burn in CNZNE.

Next, the chi-squared values of each cell are calculated with this formula (Oakes 2009: 164):

$$\text{Chi-squared value} = (\text{Observed value} - \text{Expected value})^2 / \text{Expected value}$$

The values in the example table come out as follows:

	<i>burnt</i>	<i>burned</i>
1996–97	0.18	0.32
2011–12	0.09	0.16

Table 3.

Chi-squared values of participial forms of burn in CNZNE.

Finally, the values of the cells are added together to give an overall chi-squared value, which in the case of our example is 0.75.

In order to tell something about the statistical significance of distribution, another value, namely degrees of freedom, has to be obtained. The number of degrees of freedom is calculated with the following formula (Oakes 2009: 164):

$$\text{Degrees of freedom} = (\text{number of rows} - 1) * (\text{number of columns} - 1)$$

In this study, the all the matrices contain two rows and two columns (for the two subcorporpora and the two possible forms), meaning that there is exactly 1 degree of freedom in all cases. The chi-squared value is then compared to values on a chi-squared table in which the rows represent degrees of freedom and the columns some common thresholds of statistical significance. In our example with one degree of freedom, the chi-

squared value 0.75 falls short of the minimum threshold of statistical significance, 3.84, so the observed change is not statistically significant.

The value achieved via the chi-squared test, p , shows the probability that the distribution is the same as or more extreme than observed if the null hypothesis were true (Levshina 2015: 103). In the case of this study, null hypothesis can be formulated as “There has been no change between the 1990s and the 2010s in the usage of irregular and regular verbs”. In other words, the lower the p-value, the less likely the observed changes are to be a result of random variation. It is worth noting that the null hypothesis can never be outright *disproven* (no matter how small a chance it has of being true, there is still a non-zero possibility), but a p-value below the significance level gives a valid reason to *reject* it (Levshina 2015: 104), and to assume that actual change has taken place.

For our (non-significant) example, the p-value is 0.384332, meaning that a random sampling from an unchanged dataset would result in changes that are at least as big as those between the two CNZNE subcorpora around 38 times out of 100. The most common significance levels are 0.05, 0.01, and 0.001 (Levshina 2015: 13), which mark statistically significant, *very* significant, and *highly* significant results, respectively (Stefanowitsch 2004: 3). For the purpose of this study, 0.05 is considered to be the maximum p-value with which the results can be considered as statistically significant in categorizing the verbs by whether or not the distribution of their past tense forms has changed.

One must keep in mind that statistical significance is not the be-all and end-all of quantitative corpus linguistics. Kopleinig (2017: 18) claims that since corpora are intentionally compiled “samples of convenience”, they can never represent a truly random sample of language in use. This is a problem since the null hypothesis is based on the

assumption of randomness (ibid.: 17). If the sampling is not random, the results of any research cannot be generalized into the language as a whole (ibid.: 18). It is therefore unwise to jump to conclusions on the entirety of the linguistic system based any and all statistically significant results (Koplenig 2017: 22). The results which are statistically significant in the context of the corpus analyzed may not meet the criteria for significance in another sample obtained from the same variety of language (NZE) or even the same register (New Zealand newspaper English). As follows, it is not appropriate to draw any far-reaching conclusions (outside of the CNZNE) based on the results obtained in this thesis. That is to say, the results of this study are suggestions of possible linguistic change instead of definite proof.

4 RESULTS

In this chapter, I will present the main findings of the study. Each of the 18 verbs analyzed is displayed in its own section containing a table with the raw frequencies and the percentage shares of regular and irregular forms in past tense, participle, and adjectival usage. The findings are then analyzed and compared with those in previous literature. All significant changes in distribution are commented on. Qualitative findings about the use of these verbs are presented in the sections as well, particularly for forms and senses rarely attested in the corpus or not described in background literature. The final section recapitulates some of the main trends evident in the corpus data.

4.1 *Sneak*

	1996–97	2011–12
past tense		
<i>sneaked</i>	49 (79.0%)	97 (46.9%)
<i>snuck</i>	13 (21.0%)	110 (53.1%)
participle		
<i>sneaked</i>	19 (86.4%)	16 (30.8%)
<i>snuck</i>	3 (13.6%)	36 (69.2%)
adjective		
<i>sneaked</i>	0	0
<i>snuck</i>	0	0
total		
<i>sneaked</i>	68 (81.0%)	113 (43.6%)
<i>snuck</i>	16 (19.0%)	146 (56.4%)

Table 4.

Distribution of regular and irregular forms of sneak in the CNZNE.

The verb *sneak* has seen a dramatic shift in preferred past tense form between the two subcorpora: while *sneaked* is clearly the preferred form in the subcorpus with texts from the 1990s with around 80 per cent share in both past tense and past participle forms, *snuck* is the more common form in the latter subcorpus. The increase of *snuck* has been especially notable in participles, but the results are extremely statistically significant for simple past forms, too ($p < 0.00001$ for both). This suggests that the large-scale shift in form preference is extremely unlikely to have been caused by random chance. In other words, the preferred verbal paradigm in New Zealand newspaper English has shifted from *sneak/sneaked/sneaked* to *sneak/snuck/snuck* in a matter of 15 years. Forms of *sneak* are not used adjectivally in the corpus.

The irregular form *snuck*, considered by Huddleston and Pullum (2002: 1604) to be jocular in tone, is used in the corpus in contexts that do not seem lighthearted in the slightest, such as (13):

- (13) One female Kiwi journalist had **snuck** into Myanmar to cover the aftermath of the Cyclone Nargis. (WT_01_06_2011_48)

As the regular form *sneaked* can be used in “jocular” contexts as well, it seems that context has very little bearing on the preferred form of *sneak* in the corpus:

- (14) Mrs King also remembered fondly the time an elderly woman in a rest home enjoyed a tune so much she **sneaked** the song sheet from the director. (WT_25_08_2011_46)

Snuck, a form noted by Anderwald (2009: 62–3) as a relatively recent American invention, appears to be making headway in New Zealand English, as indicated in the corpus data. Overall, the spread of the previously rare form has been extremely rapid, going from a share of less than one fifth to more than 50 per cent in just 15 years. As is clearly evident in the data, the use of *snuck* in NZE newspaper texts is far too major to be considered a novelty of any sort; it has emerged from relative obscurity in the 1990s

subcorpus and actually overtaken the regular *sneaked* in usage in both simple past and participial uses. The trajectory of change is obvious, now it remains to be seen whether or not the change will continue in the forthcoming years and decades, and whether the irregular *snuck* will replace the regular *sneaked* altogether.

The proximity of *snuck* to the form *struck* for the (invariably) irregular *strike* and the influence of American English through the spread of American culture might be possible factors explaining the adoption of this new irregular form. However, American influence is not as straightforward an explanation as one might think: in their 1990s data, Biber et al. (1999: 397) found *sneak* to exhibit a very strong predilection towards regular forms in both British and American English. The earlier subcorpus in the CNZNE is in the same timeframe and the preference in it is equally regular. The innovative irregular form must therefore have started spreading rapidly after that point. The advancing of the form *snuck* would be interesting to observe in American English within a similar timeframe: does NZE lead the way in irregularization or has the change been even faster in the United States, leaving NZE to lag behind AmE in this regard?

4.2 Dive

	1996–97	2011–12
past tense		
<i>dived</i>	79 (98.7%)	142 (97.9%)
<i>dove</i>	1 (1.3%)	3 (2.1%)
participle		
<i>dived</i>	16 (100%)	32 (100%)
<i>dove</i>	0 (0.0%)	0 (0.0%)
adjective		
<i>dived</i>	0	0
<i>dove</i>	0	0
total		
<i>dived</i>	95 (99.0%)	174 (98.3%)
<i>dove</i>	1 (1.0%)	3 (1.7%)

Table 5.

Distribution of regular and irregular forms of dive in the CNZNE.

Based on corpus data, the form *dived* seems to be the universally preferred one in NZE.

The very few uses of *dove* such as (15) and (16) do, however, affirm that the form is still technically possible and in infrequent use in NZE:

- (15) Mark Thompson said he wanted to see whether the airspeed of his son's plane was increasing or decreasing immediately before it **dove** to earth, as this could show whether it had been affected by ice. (DO_19_07_1997_87)
- (16) Dagg's defensive qualities were also to the fore in the second spell, never more so than when he **dove** in to secure a loose ball as the Force poured through on attack. (PR_02_05_2011_19)

The minor rise observed in the frequency of past tense *dove* between the two subcorpora is not statistically significant.

As Biber et al. (1999: 398) outline, only the past tense of *dive* seems to exhibit variation in the corpus, while the participle is exclusively regular. In fact, the entire

CNZNE only has one relevant token for participial *dove*, which appears in an interview of a rugby player suffering from repeated injuries:

- (17) I felt like I had **dove** straight into a time machine but because I have been through this whole experience before I had a strange feeling of resignation and said “oh well”. (SS_08_05_2005_74)

As this participial use of *dove* comes from a transcribed spoken-language passage, its context of usage can be described as informal. No conclusions about participial *dove* making any headway in NZE can thus be drawn from this instance.

During the past few decades, *dove* has not gained any significant traction as a form of *dive* in New Zealand English, at least not in newspaper texts. Based on these results, the paradigm of *dive* in NZE seems to *dive/dived/dived*, with *dive/dove/dived* being an uncommon yet possible variant. This matches the results in Biber et al. (1999: 398), who found the past tense *dived* to register usage shares of over 90 per cent in news texts.

4.3 *Knit*

	1996–97	2011–12
past tense		
<i>knitted</i>	1 (100%)	13 (100%)
<i>knit</i>	0 (0%)	0 (0%)
participle		
<i>knitted</i>	5 (100%)	40 (100%)
<i>knit</i>	0 (0%)	0 (0%)
adjective		
<i>knitted</i>	13 (86.7%)	95 (100%)
<i>knit</i>	2 (13.3%)	0 (0%)
total		
<i>knitted</i>	19 (90.5%)	148 (100%)
<i>knit</i>	2 (9.5%)	0 (0%)

Table 6.

Distribution of regular and irregular forms of knit in the CNZNE.

Knit is a rather infrequent verb in CNZNE, particularly in the first subcorpus. The word seems to have greatly increased in frequency during the 15 years between the two subcorpora, from a total of 0.61 past tense, past participle and adjectival uses per million to 2.26 uses per million. This indicates an increase in knitting between the 1990s and 2010s in New Zealand society, or at least an increase in people talking about knitting. At any rate, the upsurge does not seem to have altered the distribution of the different forms – the verb remains remarkably regular in both verbal and adjectival usage. The following is an example of a typical use:

- (18) Mr Shipley said the last time he **knitted** was in 1987, when his wife was giving her maiden speech in Parliament. (EM_12_07_2011_71)

The scarce counterexamples to the otherwise regular paradigm of *knit* can be found in adjectives. The adjectival *knit* occurs in CNZNE as part of a highly specific set phrase *fisherman's knit jersey*. These adjectival uses were mistagged in the corpus as past

participles probably due to the possessive suffix 's in the previous word (*fisherman's*) being mistaken for a contraction of *has* by the artificial intelligence, as in example (19):

- (19) The second person in the bank wore a black beanie hat and full-face latex mask, a white fisherman's **knit** jersey, light coloured track pants and black shoes. (EP_19_06_1997_67)

4.4 *Lean*

	1996–97	2011–12
past tense		
<i>leaned</i>	21 (100%)	67 (75.3%)
<i>leant</i>	0 (0%)	22 (24.7%)
participle		
<i>leaned</i>	9 (81.8%)	15 (88.2%)
<i>leant</i>	2 (18.2%)	2 (11.8%)
adjective		
<i>leaned</i>	0	0
<i>leant</i>	0	0
total		
<i>leaned</i>	30 (93.7%)	82 (77.4%)
<i>leant</i>	2 (6.3%)	24 (22.6%)

Table 7.

Distribution of regular and irregular forms of lean in the CNZNE.

Among the quite small number of tokens with participial forms of *lean*, *leaned* and *leant*, both subcorpora feature instances of both possible forms. There appears to be a strong preference for the regular form *leaned*, which seems to be increasing slightly between the two subcorpora. The change is not statistically significant, though, and is likely to be caused by random chance and a low sample size than linguistic change. *Leaned* and *leant* are not used as adjectives in the CNZNE.

By far the most striking thing about the verb *lean* in this study is the emergence of the form *leant* in past tense: whereas *leaned* is the only past tense form of *lean* in the

1996–97 subcorpus, the 2011–12 subcorpus contains 22 instances of simple past *leant*.

Consider the following examples:

- (20) Police yesterday took a man from Parliament after he **leant** over the railing of the public gallery and had to be held up to stop him from falling metres on to MPs seated below. (TD_06_10_2011_17)
- (21) Many a time she **leaned** over the concrete wall at Island Bay beach hoping to spot the San Antonino II making its way back to the safety of its sheltered mooring in the bay. (DP_30_06_2012_68)

At a glance, there does not seem to be anything about sentences (20) and (21), such as the use of direct quotes or major semantic differences, to explain the newfound use of *leant* in the second subcorpus. Consequently, it is reasonable to assume that the form *leant* has legitimately gained ground in New Zealand newspaper English between 1996 and 2012. The change in the proportions of *leaned* and *leant* is statistically significant ($p < 0.05$) and linguistically interesting, perhaps warranting further inquiry.

If Quirk et al. (1985: 107) are to be believed, *leant* is an emblematically British form and *leaned* an Americanism. Hundt's (1998: 32) results from the 1990s indicate a high degree of regularity in NZE regarding the use of *lean*: in her data, New Zealand English used a higher proportion of *leaned* than Australian or British English. The results from the CNZNE match Hundt's data quite well for the earlier subcorpus but not for the more recent one. This would mean that according to the subcorpora, NZE has moved from an almost exclusively AmE-style use of *lean* towards a more even mixture of American and British-style forms, thus reversing a near-complete process of regularization.

4.5 *Dream*

	1996–97	2011–12
past tense		
<i>dreamed</i>	39 (73.6%)	106 (72.6%)
<i>dreamt</i>	14 (26.4%)	40 (27.4%)
participle		
<i>dreamed</i>	64 (79.0%)	145 (80.1%)
<i>dreamt</i>	17 (21.0%)	36 (19.9%)
adjective		
<i>dreamed</i>	0	0
<i>dreamt</i>	0	0
total		
<i>dreamed</i>	103 (76.9%)	251 (76.8%)
<i>dreamt</i>	31 (23.1%)	76 (23.2%)

Table 8.

Distribution of regular and irregular forms of dream in the CNZNE.

The results for *dream* are rather unceremonious: the regular form *dreamed* is preferred over the irregular *dreamt* in both subcorpora and both of the tenses under scrutiny. *Dreamt* is far from being a rare form, either, registering around a fourth of past tense uses and a fifth of participial uses. Thus, uses like (22) and (23) are by no means a rarity in the CNZNE:

- (22) Ms Dobbie, 22, said since she was 10 she had **dreamt** of becoming a police officer, and after completing her BSc in maths at Otago University she had decided to realise that dream. (DO_20_12_1996_68)
- (23) I never **dreamt** I'd ever get to see Public Enemy play an intimate show to just 300 punters in the genteel surrounds of the Isaac Theatre Royal. (PR_10_01_2011_40)

There have been no significant changes in the distribution of the forms in either past tense or participial usage; the proportions have remained very similar over time. Neither *dreamt* nor *dreamed* is used adjectivally in CNZNE.

Hundt's finding that irregular forms are used for *dream* 20 per cent of the time (1998: 30) seems hold true remarkably well in the CNZNE as well. According to Hundt's research (ibid.), this distribution places NZE in an intermediate position between British and American English in the use of *dream* – the irregular form *dreamt* is not used as much as in BrE but more than in AmE.

4.6 *Spoil*

	1996–97	2011–12
past tense		
<i>spoiled</i>	23 (74.2%)	27 (54.0%)
<i>spoilt</i>	8 (25.8%)	23 (46.0%)
participle		
<i>spoiled</i>	50 (70.0%)	69 (34.2%)
<i>spoilt</i>	21 (30.0%)	133 (65.8%)
adjective		
<i>spoiled</i>	9 (32.1%)	19 (29.2%)
<i>spoilt</i>	19 (67.9%)	46 (70.8%)
total		
<i>spoiled</i>	82 (63.1%)	115 (36.3%)
<i>spoilt</i>	48 (36.9%)	202 (63.7%)

Table 9.

Distribution of regular and irregular forms of spoil in the CNZNE.

For both past tense and participial uses, the distribution of the forms of *spoil* has changed greatly between the two subcorpora, with the share of irregular form *spoilt* more than doubling in participles and almost doubling in past tense. The increase in the use of *spoilt* in past tense fall just short of statistical significance, while the change in the distribution of the participles is extremely statistically significant ($p < 0.00001$).

The use of *spoiled* and *spoilt* are almost always metaphorical, commonly appearing in the phrase *spoiled/spoilt for choice*:

- (24) **Spoilt** for choice, I have 1400 square metres of bathing facilities at my disposal. (PR_14_03_2011_64)

Particularly with *spoiled*, there are some instances of more literal uses as well:

- (25) Just one drunk 23-year-old New Plymouth man **spoiled** the perfect night when he was ejected from Peggy Gordon's Irish Bar and taken to police cells after throwing a punch at a bouncer. (TD_13_09_2011_40).

In comparison with the verbal uses, the ratio in the adjectival use of *spoiled* and *spoilt* has remained relatively static: the irregular form is used roughly twice as much as the regular one. The typical use is as a part of the phrase *spoilt child/brat*, as in the following example:

- (26) After being tossed out of Parliament for behaving like a **spoilt** brat, Mr Peters stormed back to his office and summoned members of the press gallery for an emergency media conference, where he launched into an incoherent diatribe. (DP_25_08_2012_95)

Biber et al. (1999: 397) found *spoil* to display a marked preference for regular forms in American English. British English usage varied according to tense in the study (ibid.): the dominant past tense form was found to be *spoiled* and the participle form displayed a predilection for *spoilt*. All of the 9 instances of past tense *spoil* in Hundt's (1998: 32) study were irregular, creating a contrast with the chiefly regular British English data in the same study. The results from this study are thus not in line with those from Hundt (ibid.) regarding *spoil* – the distribution of past tense forms in the latter subcorpus is more comparable to that found in BrE by Biber et al. (1998: 397), suggesting that the use of *spoil* is not necessarily any more irregular in NZE than in BrE, at least in newspaper language.

4.7 Learn

	1996–97	2011–12
past tense		
<i>learned</i>	419 (76.3%)	990 (64.1%)
<i>learnt</i>	130 (23.7%)	554 (35.9%)
participle		
<i>learned</i>	518 (68.6%)	939 (48.9%)
<i>learnt</i>	237 (31.4%)	983 (51.1%)
adjective		
<i>learned</i>	24 (88.9%)	46 (97.9%)
<i>learnt</i>	3 (11.1%)	1 (2.1%)
total		
<i>learned</i>	961 (72.2%)	1,975 (56.2%)
<i>learnt</i>	370 (27.8%)	1,538 (43.8%)

Table 10.

Distribution of regular and irregular forms of learn in the CNZNE.

As a common verb, *learn* appears over a thousand times in both subcorpora. It is used most in past participle, where the proportions of the irregular *learnt* have climbed up almost 20 percentage points between the subcorpora. The change is extremely statistically significant ($p < 0.00001$) and indicates that the use of *learned* and *learnt* has seen real change between 1996 and 2012 in New Zealand newspaper English. The following is a typical example of the form *learnt*, the majority form by a slight margin in the 2010s subcorpus:

- (27) There are lessons to be **learnt** by everyone involved in rugby in this country.
(WT_16_07_2011_130)

As with many other verbs in this study, past tense forms of *learn* seem to be more resistant to diachronic change; the reverse regularization process seems to be stronger with participles. This is not to say there has been no change in the distribution:

the 12 percent point increase in the share of past tense *learnt* – and the corresponding decline in *learned* – are remarkably statistically significant ($p < 0.00001$) as well.

Adjectival uses of *learned* carry the meaning ‘scholarly’ which is exclusive to this particular form (Quirk et al. 1985: 105):

- (28) As it turned out, the **learned** ones were all wrong. There was a far more simple reason. (EP_03_03_1997_66)

Because this meaning is not shared by *learnt*, the cases in which adjectival *learnt* is used are fairly low in number. In these instances, the meaning of *learnt* is ‘acquired through social interaction (29), though these uses are found with *learned* as well (30):

- (29) In anthropological terms, culture represents the range of **learnt** human behaviour patterns and includes things such as art, knowledge, belief systems and world views, morals, customs, and other habits acquired by people as members of society. (TD_24_09_2012_50)
- (30) Pivotal to the programme is the belief that violence is never acceptable and anger is a **learned** behaviour which can be changed. (TH_19_11_1997_4)

Biber et al. (1999: 397) perceived that *learn* is primarily a regular verb in both BrE and AmE, with the preference being less distinct in BrE. Furthermore, Hundt (1998: 32) found *learn* to be very similar in use between BrE and NZE; in Hundt’s data, both variants favor the regular form *learned*. The results from the earlier subcorpus match the results of Hundt and Biber et al. quite well, but the latter subcorpus can no longer be said to be in line with previous research with its high proportion of irregular *learnt*. With *learn*, there has once again been a reverse regularization process as indicated in the corpus data. In this case, New Zealand English has not only moved towards British English in its degree of irregularity, but actually surpassed it.

4.8 *Burn*

	1996–97	2011–12
past tense		
<i>burned</i>	55 (79.7%)	120 (69.0%)
<i>burnt</i>	14 (20.3%)	54 (31.0%)
participle		
<i>burned</i>	133 (37.8%)	239 (35.0%)
<i>burnt</i>	219 (62.2%)	443 (65.0%)
adjective		
<i>burned</i>	4 (8.2%)	18 (16.4%)
<i>burnt</i>	45 (91.8%)	92 (83.6%)
total		
<i>burned</i>	192 (42.8%)	377 (39.0%)
<i>burnt</i>	257 (57.2%)	589 (61.0%)

Table 11.

Distribution of regular and irregular forms of burn in the CNZNE.

Based on the results, the most common paradigm for *burn* in CNZNE appears to be *burn/burned/burnt*, though there is plenty of variation and both regular and irregular forms are within the realm of possibility in all tenses. The use of the irregular form *burnt* has seen a slight increase in use for verbal uses and a decrease in adjectival use (and vice versa for *burned*), but none of the changes are statistically significant.

The strongest form preference for *burn* is found in adjectives, where *burnt* is the most common form by a large margin. As is evident in the following examples, the regular form *burned* (32) is still definitely acceptable, occupying much of the same contexts as the more common *burnt* (31):

- (31) Other odd props which sold for a handsome price included a replica of a dead horse that fetched \$600 and a **burnt** severed head which went for \$525.
(TD_12_12_2012_59)

- (32) Yesterday, teams were combing the Pyne Gould Corporation site, the **burned** remnants of the CTV building and Christchurch Cathedral.
(EM_04_03_2011_105)

The direction of change implied in the CNZNE does not suggest a diminishing share of adjectival usage for the form *burned*, so based on the corpus data, it can be assumed that the form will remain a possible, if somewhat marginal, variant.

The results for verbal uses are, curiously enough, essentially the opposite of Biber et al. (1999: 396–7), who found that *burn* displays a preference for irregular form in past tense and regular form in past participle in expository registers (news texts and academic writing). In contrast, the results of the present study fall in line with those of Hundt (1998: 30), who found that in NZE, the irregular form *burnt* is around three times more common in participles than past tense, and that *burnt* is generally the form of choice in adjectival use.

What might cause the discrepancy between the findings of Biber et al. on one hand and Hundt and the current study on the other? Two possible explanations are differences between different Englishes and differences between registers. Using more *burned* in past tense and *burnt* in past participle might simply be a New Zealand feature, which would have gone undetected by linguists like Biber et al. who focused on BrE and AmE in their studies. Furthermore, it might be a feature typical of newspaper language in general, or New Zealand newspapers in particular, as both Hundt's research and this thesis use New Zealand-based papers as source material. A topic of further study would be to compare other registers within NZE, for example, informal conversations or academic prose, in order to see if the proportions are characteristic of New Zealand newspaper English or NZE in general. Of course, the difference may simply be caused by some skewed results in the corpus data – the corpus used by Biber et al. (1999: 25–6) contains

only around 5 million words of academic writing and 11 million words of news texts, considerably less than the CNZNE. This means that one longer text with several instances of either *burned* or *burnt* (for example, a news report of an arson) would distort the search results in favor of a particular form.

4.9 *Smell*

	1996–97	2011–12
past tense		
<i>smelled</i>	32 (71.1%)	53 (58.9%)
<i>smelt</i>	13 (28.9%)	37 (41.1%)
participle		
<i>smelled</i>	6 (60.0%)	9 (36.0%)
<i>smelt</i>	4 (40.0%)	14 (64.0%)
adjective		
<i>smelled</i>	0	0
<i>smelt</i>	0	0
total		
<i>smelled</i>	38 (69.1%)	62 (54.9%)
<i>smelt</i>	17 (30.9%)	51 (45.1%)

Table 12.

Distribution of regular and irregular forms of smell in the CNZNE.

Smell is used in CNZNE mostly in simple past tense rather than past participle, with around 80 per cent of the uses from *smelled* and *smelt* coming from past tense. There is a fair amount of variation in usage, with both regular and irregular forms being used in both tenses. The 2010s subcorpus contains a higher proportion of the irregular *smelt* than the 1990s one, suggesting diachronic change. Perhaps owing to a particularly small sample size, neither observed change is substantial enough to reach the threshold of statistical significance. The results nonetheless conform to the overall trend of irregularization in

the corpus, or at least do not go against it. Both forms are used intransitively (the sense ‘to emit an olfactory signal’) and transitively (the sense ‘to receive an olfactory sensation’), as in the following examples:

- (33) Danielle’s room **smelled** of mould, but mine has got to the point where it is just this extreme toxic smell. (DP_04_08_2012_155)
- (34) “It **smelt** like someone had died in there,” he said. (SS_04_09_2011_81)
- (35) Police entered the house to search it when they **smelled** cannabis. (PR_11_11_2011_48)
- (36) But Mr Arthur **smelt** a rat, backed out of the purchase and alerted the Times to the scam. (WT_02_04_2011_93)

Smell was found by Biber et al. (1999: 397) to exhibit noticeable trans-Atlantic variation between the mainly irregular BrE and the more regularized AmE. In her study with a small sample size, Hundt (1999: 32) found a slight preference in NZE for *smelled*, with 7 of the 12 past tense uses of *smell* being regular, an observation at least somewhat reinforced by the findings of the present study. Based on these results, the distribution of the forms of *smell* in NZE seems to fall between the two prestige variants, perhaps a tad closer to BrE than AmE.

4.10 *Spell*

	1996–97	2011–12
past tense		
<i>spelled</i>	26 (55.3%)	34 (49.3%)
<i>spelt</i>	21 (44.7%)	35 (50.7%)
participle		
<i>spelled</i>	107 (71.8%)	55 (43.0%)
<i>spelt</i>	42 (28.2%)	73 (57.0%)
adjective		
<i>spelled</i>	0	0
<i>spelt</i>	0	0
total		
<i>spelled</i>	133 (67.9%)	89 (45.2%)
<i>spelt</i>	63 (32.1%)	108 (54.8%)

Table 13.

Distribution of regular and irregular forms of spell in the CNZNE.

The past tense forms of *spell*, *spelled* and *spelt*, are used in roughly equal proportions in both subcorpora in simple past tense. Irregular forms like the one in the following example have slightly gained prominence in the newer subcorpus:

- (37) An urgent message from him to me **spelt** only one thing - my worst nightmare.
(TD_17_06_2011_29)

While the change in the proportions of the past tense forms is nothing out of the ordinary, not even breaking the threshold of statistical significance, the participial forms of *spell* are a whole another story. The share of *spelt* in participles has more than doubled in the 15 years between the texts of the subcorpora, becoming the preferred form in the meantime. This change is, of course, extremely statistically significant ($p < 0.00001$). The following examples demonstrate typical uses of *spelled* (38) and *spelt* (39), respectively:

- (38) The Maori words will also be **spelled** on the screen, allowing the users to get the right pronunciation and spelling. (WT_08_12_1997_13)

- (39) “It was supposed to be *cactus* but I had **spelt** something else wrong,” she said.
(TD_11_01_2011_32)

One should keep in mind, though, that both forms are still in frequent use in both corpora, so counterexamples to the above should be relatively easy to find. *Spelled* and *spelt* are not used as adjectives in the corpus.

In their corpus-based study, Biber et al. (1999: 397) found that *spell*, along with *smell* are two of the verbs which display the greatest variation between the irregular-preferring BrE and the more regularized AmE. The present study indicates that at least in newspaper texts, New Zealand English takes the intermediate position; neither variant can be said to be the dominant one. Though *spelled* is preferred in the 1990s texts of the corpus, the preference is far from the over 75 per cent share the form was discovered by Biber et al. (ibid.) to hold in American English. Based on the results, New Zealand newspaper texts are moving from an intermediate position between British and American English towards British-style use in the use of participial forms of *spell*. Whether or not this direction will continue in the future remains to be seen.

4.11 *Leap*

	1996–97	2011–12
past tense		
<i>leaped</i>	36 (23.8%)	52 (19.3%)
<i>leapt</i>	115 (76.2%)	217 (80.7%)
participle		
<i>leaped</i>	7 (14.3%)	2 (2.6%)
<i>leapt</i>	42 (85.7%)	75 (97.4%)
adjective		
<i>leaped</i>	0	0
<i>leapt</i>	0	0
total		
<i>leaped</i>	43 (21.5%)	54 (15.6%)
<i>leapt</i>	157 (78.5%)	292 (84.4%)

Table 14.

Distribution of regular and irregular forms of sneak in the CNZNE.

Most of the uses of *leaped* and *leapt* in CNZNE are found in past tense. The irregular form *leapt* is used around 80 per cent of the time in the corpus, in contexts ranging from mundane to rather unorthodox (40):

- (40) However, this detail was lost on the throng of Asian businessmen queuing at the counter as Slav elegantly **leapt** over a dining table and let fly with his Super Soaker. (WT_29_12_1997_14)

There has been some change in the distribution of the past tense forms between the subcorpora (namely, *leapt* gaining prominence at the expense of *leaped*), but the change is not statistically significant.

The share of participial *leaped*, already a rather marginal form in the 1990s, seems to have diminished even further. Based on corpus data, *leaped* seems to be a form in decline, at least in NZE newspaper texts. Uses such as (41) do, however, prove that the form is still within the realm of possibility:

- (41) Hamilton judo fighter Sean Choi has **leaped** into the world's top 20 after defending his title at the Oceania championships last weekend.
(WT_23_04_2011_10)

Unlike with past tense use, the decline of the participle *leaped* is statistically significant ($p < 0.05$). *Leaped* and *leapt* are not used as adjectives in the corpus.

Biber et al. (1999:397) and Quirk et al. (1985: 197) discovered *leaped* to be the more common form in American English, and the irregular *leapt* to be the preferred form in British English. Hundt (1998:32) found *leapt* to be even more common in NZE than in BrE, though her study had a very small sample size (a total of 6 instances of *leap* in NZE and 10 in BrE) so the results should be taken with a grain of salt. Based on CNZNE data, New Zealand English takes after British English in the use of *leap*.

4.12 Hang

	1996–97	2011–12
past tense		
<i>hanged</i>	4 (3.7%)	6 (2.0%)
<i>hung</i>	104 (96.3%)	290 (98.0%)
participle		
<i>hanged</i>	44 (26.2%)	121 (30.9%)
<i>hung</i>	124 (73.8%)	271 (69.1%)
adjective		
<i>hanged</i>	4 (14.3%)	0 (0%)
<i>hung</i>	24 (85.7%)	26 (100%)
total		
<i>hanged</i>	52 (17.1%)	127 (17.8%)
<i>hung</i>	252 (82.9%)	587 (82.2%)

Table 15.

Distribution of regular and irregular forms of hang in the CNZNE.

The verb *hang* displays quite a small amount of variety in its past tense and adjective forms where *hung* is the form of choice, while past participles show a smaller yet substantial preference for *hung*. There is a clear factor explaining the distribution: the two past tense forms have two different senses. Quirk et al. (1985: 112) and Biber et al. (1999: 396) both state that *hanged* is used in reference to capital punishment or suicide by hanging, at least in British English. This seems to be the case in NZE as well, since all uses of *hanged* in both subcorpora contain (rather gruesome) references to hanging as suicide, as in (42), or punishment, as in (43):

- (42) A spurned man murdered his older lover then made it appear that she **hanged** herself with a skipping rope, a High Court jury trial has heard in New Plymouth. (TD_24_10_2012_81)
- (43) An Iranian man who attracted several disciples after claiming to be God has been **hanged** for apostasy in the southwest of the country, the semi-official Fars news agency reported today. (EM_01_02_2011_21)

This sense is not restricted to just the form *hanged* – there is also a case in the corpus where *hung* is used in a similar sense:

- (44) Harawira reminded Brash that the British had a liking for inflicting punishment by “**hung**, strung and quartering” people. (PR_11_07_2011_98)

This tendency to occasionally use *hung* in this sense is observed by Huddleston and Pullum (2002: 1604), so despite its rarity, this use is not unheard of in previous literature.

In adjectival use, *hang* assumes the form *hung* most of the time. The only examples where *hanged* is used as an adjective come from different versions of the same news item in which the sense is the expected ‘executed by hanging’:

- (45) Ledum Mitee, former cellmate of **hanged** Nigerian activist Ken Saro-Wiwa, says Nigeria’s human rights record is continuing to worsen. (EP_12_08_1997_67)

Every instance of the adjective *hung* in the 1996–97 subcorpus is part of one of two common phrases, *hung parliament* or *hung jury*:

- (46) Niue's elections on February 16 followed an extraordinary political stalemate, a **hung** parliament, with the Assembly divided 10-10 between the government and the opposition since December 1994. (DO_23_02_1996_81)

The newer subcorpus features more uses for *hung* besides the two mentioned above, such as *hung verdict* and the notably literal use *hung beef*:

- (47) Barrington is now stocking Harris Meats' tender **hung** beef for your dining pleasure. (PR_12_01_2011_24)

None of the changes between the subcorpora are statistically significant, so the use of the verb *hang* seems to have remained very similar in NZE newspaper texts. The division of the past tense forms into different senses might be a factor which govern the distribution of forms – an increase in the use of the form *hanged* would indicate an increase of hangings, or at least an increase in people talking and writing about them.

4.13 *Quit*

	1996–97	2011–12
past tense		
<i>quitted</i>	0 (0%)	1 (0.5%)
<i>quit</i>	149 (100%)	192 (99.5%)
participle		
<i>quitted</i>	0 (0%)	0 (0%)
<i>quit</i>	92 (100%)	134 (100%)
adjective		
<i>quitted</i>	0	0
<i>quit</i>	0	0
total		
<i>quitted</i>	0 (0%)	1 (0.3%)
<i>quit</i>	241 (100%)	326 (99.7%)

Table 16.

Distribution of regular and irregular forms of quit in the CNZNE.

As Biber et al. (1999: 397) found out in their study, *quit* is a verb which prefers irregular forms. The monolithic preference for the form *quit* is clearly visible in the NZE corpus data as well. *Quit* is most often used in the meanings ‘to resign’ (48) or ‘to stop smoking cigarettes’ (49):

(48) When she **quit** as a women’s affairs and local government minister in September, Mrs Fletcher denied she was eyeing the mayoralty.
(DO_17_12_1997_83)

(49) US First Lady Michelle Obama says the President has **quit** smoking.
(PR_10_02_2011_29)

Quit is not used adjectivally in either subcorpus. Several instances in which the word is used a part of a set phrase appear tagged as adjectives, as in the following example sentence:

(50) Keep going with your **quit** attempt. (EP_30_05_1997_25)

In this example, the phrase *quit attempt* means ‘an attempt to quit (smoking)’ instead of ‘an attempt which has been quit’. In cases like this, the word *quit* would best be interpreted as an infinitive which has the same outward appearance as its past tense and past participle forms, and is therefore not of interest for the present study.

There are only two instances of the form *quitted*, in any part of speech, in the entire CNZNE (one of which can be found in the 2011–12 subcorpus). Both uses are in past tense and demonstrate that its rareness notwithstanding, the form is still technically possible in NZE:

(51) If or when a smoker **quitted**, they could choose to “enforce” their new clean lifestyle by surrendering their ID card. (DP_25_04_2012_89)

Based on these results, *quit* can be said to be an irregular verb with the paradigm *quit/quit/quit* in New Zealand newspaper English, with a very infrequent variant form *quitted* at least in past tense.

4.14 *Speed*

	1996–97	2011–12
past tense		
<i>speeded</i>	0 (0%)	0 (0%)
<i>sped</i>	100 (100%)	181 (100%)
participle		
<i>speeded</i>	12 (42.9%)	9 (19.6%)
<i>sped</i>	16 (67.1%)	37 (80.4%)
adjective		
<i>speeded</i>	0	0
<i>sped</i>	0	0
total		
<i>speeded</i>	12 (9.4%)	9 (4.0%)
<i>sped</i>	116 (90.6%)	218 (96.0%)

Table 17.

Distribution of regular and irregular forms of speed in the CNZNE.

Simple past tense makes up the vast majority of instances of *speed* in the corpus, and with it, the results are clear: there are no instances of *speeded* used in the past tense in either subcorpus, while *sped* gathers at least a hundred tokens in both subcorpora. The entire CNZNE contains only 11 instances of past-tense *speeded*, 8 of which form a part of the phrasal verb *speeded up*. This is in line with Biber et al. (1999: 397), who list *sped* as the more common past-tense form.

Because they display actual variation in their distribution, the participle forms *speeded* and *sped* are of most linguistic interest with the verb *speed*. As is the case with several other verbs, the distribution appears to have changed over the years when judging by corpus data, and the direction of the change has been towards a greater use of irregular forms. The change is statistically significant, but only slightly ($p < 0.05$). Quirk et al. (1985: 112) claim that the regular form *speeded* is always used for forms of the phrasal

verb *speed up*. This seems to not be the case in the CNZNE, as there is no clear preference for either form with *speed up*:

- (52) We don't have the kind of timelines that have been suggested and this whole process needs to be **sped** up. (PR_04_06_2011_143)

In addition to *speed up*, other, less common phrasal verbs formed with *speed* can be found in the corpus as well, such as *speed off* 'leave quickly':

- (53) The then 20-year-old had **sped** off from a police checkpoint after failing a general result before she crashed into Ms Phillips' car at the junction of Nelson and Hutcheson streets in Blenheim. (EM_12_07_2011_53)

While *speeded* or *sped* are not used as adjectives by themselves, they do find some use as parts of the (relatively rare) adjectives *speeded-up* and *sped-up*:

- (54) Eerie in atmosphere, haunting in mood and touching in the affection between Owen and Abby, with the only jarring note being **sped-up**, jumpy special effects, *Let Me In* might give foreign-film fans a sense of déjà vu. (EM_07_07_2011_80)

The aforementioned adjectives are, of course, analogous with the forms *speeded up* and *sped up* of the phrasal verb *speed up*, which makes up the bulk of uses of the verb *speed* in the corpus.

4.15 *Wed*

	1996–97	2011–12
past tense		
<i>wedded</i>	0 (0%)	0 (0%)
<i>wed</i>	3 (100%)	5 (100%)
participle		
<i>wedded</i>	3 (37.5%)	7 (26.9%)
<i>wed</i>	5 (62.5%)	19 (73.1%)
adjective		
<i>wedded</i>	7 (100%)	21 (100%)
<i>wed</i>	0 (0%)	0 (0%)
total		
<i>wedded</i>	10 (55.6%)	28 (53.8%)
<i>wed</i>	8 (44.4%)	24 (46.2%)

Table 18.

Distribution of regular and irregular forms of wed in the CNZNE.

As a fairly rare verb, *wed* has a low frequency in the corpus. The numbers can consequently be swayed by dozens of percentage points by a single token, particularly in the older and less extensive subcorpus. This means that all results for this verb should be interpreted as chiefly qualitative, demonstrating possible forms found in the corpus instead of their numerical shares.

Once again, the preferred forms for past tense verbs and adjectival participles are clear: the irregular form *wed* is chosen for past tense use and the regular *wedded* for adjectival use. The past tense use being regular is also commented on by Biber et al. (1999: 397). With participles, there is a slight preference for the irregular *wed*, which seems to be slightly – but not statistically significantly – increasing in use. Based on data from CNZNE, the possible paradigms for this verb are *wed/wed/wedded* and *wed/wed/wed*.

As previous literature (Biber et al. 1999: 397) predicts, the verb *wed* has a marked preference for irregular forms in CNZNE. The corpus data, however, does contain multiple uses of the regular form *wedded* in participial use. Most uses of the form *wedded* in the corpus seem to be metaphorical, in stark contrast to the more literal sense of the form *wed* ‘to join in matrimony’:

- (55) When your uncles or mine fought “for king and country”, they were not **wedded** to one visual symbol of country. (PR_04_08_2011_110)

The exceptions to this trend are chiefly stock phrases such as “*lawful wedded wife/husband*” (57) and “*wedded bliss*” (56), which constitute a typical adjectival use of *wedded*:

- (56) Unhappy friends are drawn to their idyllic state of **wedded** bliss and harmony as if maybe it will rub off on them. (EM_28_04_2011_79)
- (57) Each party has to say: “I, Bla Bla Bla, take thee, Bla Bla Bla, to be my lawful **wedded** wife/hubby.” And that’s it. Hitched. (WT_10_08_2011_34)

Based on the (admittedly low) number of tokens, there appears to be variation in the participial forms of the verb *wed* that is larger than predicted in the previous literature. What is more, there seems to be a semantic difference between the forms that warrants further study with larger sample sizes and a more in-depth look on the semantic properties of *wed* and *wedded*.

4.16 *Light*

	1996–97	2011–12
past tense		
<i>lighted</i>	3 (5.0%)	1 (0.7%)
<i>lit</i>	57 (95.0%)	133 (99.3%)
participle		
<i>lighted</i>	3 (2.1%)	0 (0%)
<i>lit</i>	138 (97.9%)	232 (100%)
adjective		
<i>lighted</i>	12 (100%)	16 (100%)
<i>lit</i>	0 (0%)	0 (0%)
total		
<i>lighted</i>	18 (8.5%)	17 (4.5%)
<i>lit</i>	195 (91.5%)	365 (95.5%)

Table 19.

Distribution of regular and irregular forms of light in the CNZNE.

The results for *light* are particularly obvious: in CNZNE, *light* is irregular in almost all verbal uses and regular in adjectival uses. The results can be said to correspond to the findings in previous literature. As Quirk et al. (1985: 113) state, *lighted* is the only accepted form in adjectival usage – there is not a single instance of *lit* as an adjective in the whole CNZNE, not even in the slang sense ‘inebriated’ or the more recent sense ‘very good’. Because *lighted* is the American English form and *lit* the British English one (Biber et al. 1999: 396–7), NZE can be said to take after BrE regarding the use of the verb *light*.

In the 1996–97 subcorpus, a methodological challenge is encountered: one text contains a majority of uses for non-adjectival *lighted* in the entire subcorpus, greatly altering the visible proportions of regular and irregular forms of *light*. The instances appear in an editor’s response to a letter to the editor, a very typical medium for complaint tradition. The letter from a prescriptivist complaining about the tendency of Americans

to use regular forms for *light* in past tense received a response containing the following sentences:

- (58) English and New Zealand writers and speakers normally choose **lit**, but for no apparent reason other than that it sounds more familiar.
 I would no more say or write "He **lighted** the fire" than I would fly, and if a child learning to talk said it had **lighted** a candle I would correct it to say **lit** even before I rushed to put the flame out.
 Cliche enthusiasts could say I do not agree with the American preference for using **lighted**, but I will defend to the death their right to do so.
 (DO_16_08_1997_41)

This text alone accounts for a sizeable proportion of the non-adjectival uses of *lighted* in the earlier subcorpus, distorting the figures to an extent. This sort of metalanguage is a rather poor representative of New Zealand English, especially since the author of the text clearly marks the regular form as atypical of NZE. In light of this, the form *lighted* cannot be said to have been any more common in the 1990s in the 2010s. At any rate, the minor changes in the distribution of the past tense and participial forms of *light* are not statistically significant even when counting the metalinguistic text.

4.17 Prove

	1996–97	2011–12
past tense		
<i>proved</i>	718 (100%)	1,478 (100%)
<i>proven</i>	0 (0%)	0 (0%)
participle		
<i>proved</i>	833 (72.4%)	1,191 (62.2%)
<i>proven</i>	317 (27.6%)	725 (37.8%)
adjective		
<i>proved</i>	1 (0.4%)	1 (0.2%)
<i>proven</i>	242 (99.6%)	413 (99.8%)
total		
<i>proved</i>	1,552 (73.5%)	2,670 (70.1%)
<i>proven</i>	559 (26.5%)	1,138 (29.9%)

Table 20.

Distribution of regular and irregular forms of prove in the CNZNE.

The results for *prove* show clear preferences for *proved* in past tense and *proved* in adjectival usage, in addition to variable use of regular and irregular forms in participial verbs. This would make the paradigm of the verb *prove/proved/proved* or *prove/proved/proven*, the former paradigm being the preferred one in the corpus data. The preference shows a marked shift between the corpora, as the share of *proven* has increased by 10 percentage points, from 28 in 1996–97 to 38 per cent in 2011–12. The changes in the proportions of participial *proven* and *proved* are extremely statistically significant ($p < 0.00001$) and thus likely to be caused by a real change in usage between 1990s and 2010s. The results are very close to those by Hundt (1998: 33–6), who found that the regular form *proved* was used 64 per cent of the time in NZE. This share is higher than the 35% in American English but lower than the 80% share in British English (ibid.: 34).

The form *proven* has an association with legal terminology (Hundt 1998: 33), especially as part of the popular saying *innocent (un)til(l) proven guilty*. The corpus data features numerous instances of the phrase, as in the following example:

- (59) As in other areas of the law, an officer is considered innocent till **proven** guilty. (DO_21_03_1997_29)

However, a few counterexamples like (60) can be found in the corpus as well, suggesting that *proved* can be used in similar, law-related contexts as *proven*:

- (60) Sir, isn't the phrase "innocent till **proved** guilty" the basis of our justice system? (DO_01_07_1996_52)

Unlike the archaism *gotten*, *proven* is an innovative form (Hundt 2009: 22). It was originally introduced to BrE from Scots in the 16th century, while its use has developed in AmE as well (ibid.). For this reason, it is not feasible to attribute the perceived rise in the use of participial *proven* to influence from either Scottish or American English (Bauer 1997: 271) – it can be caused by either or both of them, or by extralinguistic factors beyond the scope of the present study.

In the CNZNE data, *proven* is used exclusively as a participle; not even a single instance of non-participial *proven* can be found in the entire corpus. When it comes to adjectival use of the participle, the positions switch: there is only one instance of *proved* used as an adjective in each subcorpus among hundreds of instances of adjectival *proven*:

- (61) It was a lack of exploration rather than poor prospects that accounted for the limited level of **proved** and probable reserves. (DO_24_05_1996_35)
- (62) There is a certain amount of **proved** science about some companion planting and the rest is unproven. (ST_06_10_2011_69)

It is to be noted that even the latter instance of *proved* is followed by the form *unproven*.

While derivative forms such as *unproven* are of no particular interest for this thesis, its use over *unproved* seems to be a part of a general pattern in which *proven* is used near-

universally for adjectival form of *prove* in the data. From this, it can be gathered that *proved* is a very rare (yet definitely possible) adjective in New Zealand newspaper English. As the form *proven* has been found to be retained more in adjectival than participial use in previous literature (Hundt 1998: 34), the result comes as no surprise.

4.18 *Get*

	1996–97	2011–12
past tense		
<i>got</i>	5,436 (100%)	14,640 (100%)
<i>gotten</i>	0 (0%)	0 (0%)
participle		
<i>got</i>	7,448 (99.8%)	21,580 (99.4%)
<i>gotten</i>	18 (0.2%)	124 (0.6%)
adjective		
<i>got</i>	0	0
<i>gotten</i>	0	0
total		
<i>got</i>	12,884 (99.9%)	36,220 (99.7%)
<i>gotten</i>	18 (0.1%)	124 (0.3%)

Table 21.

Distribution of regular and irregular forms of get in the CNZNE.

For this verb, the use of the AmE-influenced participle *gotten* is so rare that its uses make up less than 1 per cent of the total number of tokens. In order to obtain any significant number of results for *gotten*, no smaller samples can be taken from the subcorpora. Because the number of tokens featuring the usual form *got* obtained this way numbers in the tens of thousands, it becomes unfeasible to manually check for tagging mistakes and duplicates. The numbers for *got* should therefore be taken as approximate. Instead, normalized frequencies per million words are used to chart the presence of the participle

gotten in New Zealand English. It is necessary to normalize frequencies between subcorpora of different sizes to make direct comparisons possible between subcorpora of different sizes (Biber et al. 1998: 263). Frequencies are normalized by dividing the raw frequency by the total number of words in the corpus and multiplying the quotient by a number chosen as a basis for norming (ibid.: 263–4). For *gotten*, a basis of one million was chosen due to scale – the overall frequency of the form is very low, and the subcorpora are rather large.

In the 1990s data, the form *gotten* is used 18 times for a rather low frequency of 0.52 per million words. Over the decades, *gotten* has gotten more prominent – it is used 124 times in the newer subcorpus, meaning that its normalized frequency is 1.88 per million words. This means that the frequency of *gotten* has more than tripled in fifteen years, though when compared to *got*, *gotten* is still an undeniably marginal form in NZE. Despite the rapid increase in frequency, it would take several decades of similar upsurge for the form to be nearly as common in New Zealand English as it is in American English. Consistent with previous research (Biber et al. 1999: 398–9), *gotten* is used only as a participle and neither *got* nor *gotten* is used as an adjective in the corpus data.

With *gotten*, there is another example of the New Zealand English complaint tradition in the corpus. As was the case with *lighted*, this complaint concerns a perceived American influence in NZE:

- (63) We seem to be drifting towards speaking a new form of English which includes nasty Americanisms such as “**gotten**”. Hopefully this won't appear in schools and corrupt our young. (SS_06_11_2011_165)

If nothing more, the quote exemplifies the resistance held by the prescriptive tradition and its adherents towards changes to a direction perceived as “American”. Those speakers of NZE that are aware of these implications of using *gotten* might avoid using it at least

outside of spontaneous conversations, thereby presenting a potentially limiting factor for any widespread usage of the form. As the following response to the previous example shows, not all NZE speakers share the animosity towards *gotten*:

- (64) Geoffrey Horne calls the word “**gotten**” a nasty Americanism. However, the internet tells us that its origin comes from 1150-1200AD, and the word can also be found in Shakespeare’s plays. (SS_13_11_2011_63)

On the basis of the increase in the use of *gotten* in CNZNE, the number of people in New Zealand who see forms such as *gotten* as part of normal linguistic variation (as in (64)) instead of an encroachment by American linguistic imperialism (as in (63)) seems to be on the rise.

What might be the reasons for the growing popularity of *gotten*? It seems unlikely that the increase would be attributed solely to an increase in direct quotations by AmE speakers, though their role in the increase of the form cannot be ruled out, either. Hundt’s (1998: 37) findings that Māori use the form *gotten* more than Pākehā (white New Zealanders) might be relevant here: the substantial rise in the use of *gotten* between the two subcorpora might be caused (at least in some measure) by an increase in Māori representation in New Zealand newspapers. Unfortunately, this is virtually impossible to verify based on the corpus data alone, as CNZNE is not tagged according to the ethnicity of the speaker. The explanation for the change could simply be that *gotten* forms a part of an overall tendency in CNZNE to move towards further irregularity as time passes, regardless of whether the irregular forms are associated with American English (*snuck*, *proven*, *gotten*) or British English (the rest of the verbs in this study).

4.19 Discussion

The overall trend in the corpus data is that of irregularization: all significant changes in the previous sections have been towards further use of irregular forms. The changes are evident in the following table. Do note that the numbers for *got* and *gotten* are omitted from the overall numbers due to the extremely high frequency of *got* in the corpus and the fact that both participial forms of *get* are irregular.

	1996–97	2011–12
past tense		
<i>regular</i>	1,505 (67.0%)	3,187 (62.7%)
<i>irregular</i>	742 (33.0%)	1,897 (37.3%)
participle		
<i>regular</i>	1,829 (58.1%)	2,889 (47.3%)
<i>irregular</i>	1,317 (41.9%)	3,214 (52.7%)
adjective		
<i>regular</i>	74 (18.2%)	216 (27.2%)
<i>irregular</i>	333 (81.8%)	578 (72.8%)
total		
<i>regular</i>	3,408 (58.8%)	6,292 (52.5%)
<i>irregular</i>	2,392 (41.2%)	5,688 (47.5%)

Table 22.

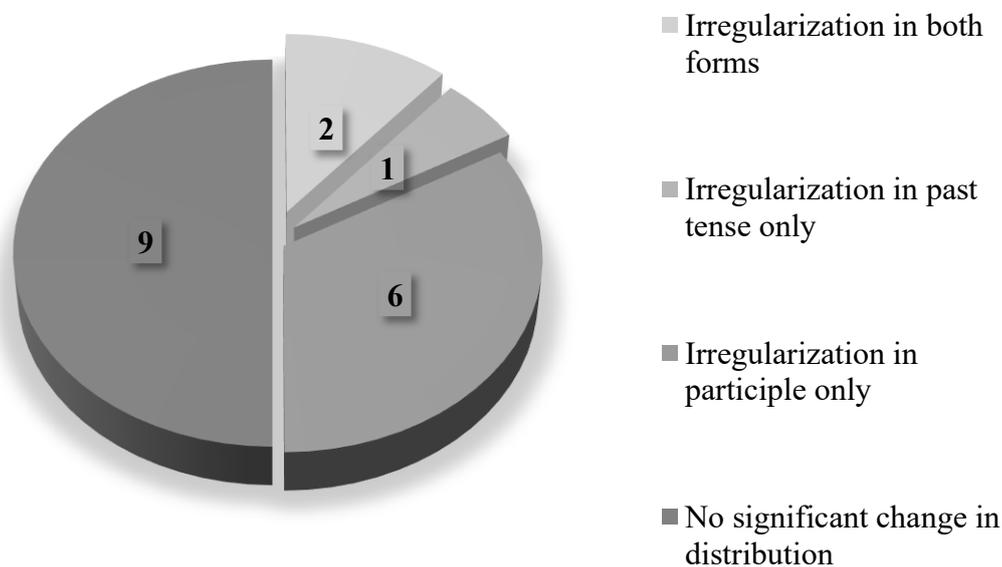
Combined distribution of regular and irregular forms of all the studied verbs except get in the CNZNE.

The biggest change in the above table is the switch in the preferred participle forms between corpora: the share of irregular participles has increased by more than 10 percent points to overtake regular participles as the more common form in the 2010s subcorpus. The change is, of course, extremely statistically significant ($p < 0.000001$).

Even the past tense forms exhibit an overall trend of irregularization when added together: the minor changes in distribution break the threshold of statistical significance ($p < 0.001$) with a large enough sample size. While the overall figures for adjectives does imply a

highly statistically significant ($p < 0.001$) regularization, this is mostly caused by the increase in prominence by the exclusively regular adjectival use of *knitted* – when the verb *knitted* is removed, the shares of regular adjectives in the corpus appear to have remained very similar (15.5% in the 1990s and 17.3% in the 2010s) with statistically insignificant change. The overall results therefore indicate that the use of verbs in New Zealand newspaper English has moved slightly towards irregularity in past tense, more so in participial verbs, and has not seen much change at all in adjectival participles.

None of the 18 verbs researched showed signs of regularization – the very slightly increasing shares of regular forms in some verbs (participial *hanged*, adjectival *burned*) are not statistically significant. On the other hand, half of the featured verbs (9/18) exhibited statistically significant irregularization in either past tense or participial forms. The types of change are visualized in the following graph:



Graph 1.
The verbs of the study by type of change.

Past tense forms appeared to be more resistant to change than participial ones. Only 3 of the 18 verbs displayed a statistically significant increase in the proportion of irregular forms in past tense: *sneak*, *learn*, and *lean*, of which *lean* was the only one not to display a similar change in the use of participial forms in addition to past tense. This contrasts with participles which showed statistically significant irregularization in 8 of the 18 verbs: *sneak*, *spoil*, *learn*, *spell*, *leap*, *speed*, *prove*, and *get*. Based on corpus data, the shares of adjectival usage for all verbs in this study have remained surprisingly static in New Zealand newspaper texts. Of the eight verbs whose participles can be used adjectivally, none reported a statistically significant change in the proportions of regular and irregular forms.

From this follows that the three possible grammatical functions in which the verb can take both regular and irregular forms can be placed in a clearly-defined order according to their tendency for diachronic change in the corpus: verbal participles irregularized more than past tense verbs, which in turn irregularized more than adjectival participles, which saw no significant change in either direction. As Cheshire (1993: 127) observed participial forms to be more resistant to regularization than past tense forms, the exact opposite seems to hold true for irregularization. When combined, the results from Cheshire (ibid.) and the present thesis suggest that at least for some irregular verbs, the paradigm is moving towards regular past tense forms and irregular participles. Out of the verbs studied in this thesis, *burn* exhibits definite preferences for this type of paradigm (*burn/burned/burnt*) in the corpus data, and the irregularizing *learn* (*learn/learned/learnt*) and *prove* (*prove/proved/proven*) seem to be moving in this direction as well.

The theory of colonial lag appears to have very little capacity to explain the shifts towards further irregularity in the corpus data. As previous literature (for example, Hundt

1998 and Cheshire 1993) has identified regularization as an ongoing trend around the world, a variety affected by colonial lag would regularize as well, at roughly the same speed but behind the original variety in its state of progress. In the CNZNE, no process of this sort happens – corpus results for the verbs *lean* and *leap*, for example, show a move from a more regularized variety towards a BrE-style irregular usage, instead of the other way around as predicted by the colonial lag hypothesis. If anything, the type of change witnessed in some verbs in the present study could be termed as *postcolonial backlash* – there is certainly no lag of any kind to be seen in the results. Furthermore, for the term to be applicable, there would need to be a common trajectory for both the prestige variety (BrE in this case) and the postcolonial variety (NZE), which does not seem to be the case with regularizing BrE and irregularizing NZE. To summarize, New Zealand English seems to be going against the grain among world Englishes in its tendency to use irregular verb forms in increasing numbers.

5 CONCLUSION

This chapter begins with explicit answers given to the research questions presented in Chapter 1. This is followed by a discussion of the possible reasons for the observed results. Finally, there is some reflection on the shortcomings of this thesis and suggestions on further research on the topic.

In conducting the research for this thesis, I set out to answer three research questions. The first question was “Which of the studied verbs prefer regular forms and which irregular forms in New Zealand English corpus data?” In the corpus data, *dive*, *knit*, *lean*, and *dream* were found to prefer regular forms in all their uses, while *leap*, *hang*, *quit*, and *speed* preferred irregular forms. The most common participial form of irregular *get* was *got*, as expected. The rest of the verbs exhibited at least some variety, bringing us to the second research question, “Do regular and irregular forms have different distributions in past tense, past participle, and adjectival usage?” The nine verbs with distributions that differed by grammatical function in the CNZNE data are *sneak*, *spoil*, *learn*, *burn*, *smell*, *spell*, *wed*, *light*, and *prove*. *Burned* was found to be the most used form for *burn* in past tense and *burnt* in participial and adjectival uses. For *sneak*, the preferred forms were regular in the earlier subcorpus and irregular in the latter, whereas for *spoil*, *learn*, *smell*, and *spell* the preferred participle changed from regular to irregular in the 15 years between the subcorpora. The most used form in adjectival use differed greatly from verbal uses with three verbs: *wed*, *light*, and *prove*. The irregular verbs *wed* and *light* took only regular forms in adjectival use, while adjectival uses of the otherwise mainly regular *prove* were almost exclusively irregular. The preferred form in adjectival usage was very definite with all verbs except *spoil* (even with *spoil*, *spoilt* was more than twice as common as *spoiled* in both subcorpora).

For the question “How have the proportions of the different forms changed over the years for each verb?”, the answers can be found in more detail in section 4.19. In brief, the expected regularization was nowhere to be found in the results, as the polar opposite happened: half of the 18 verbs in the study showed no significant changes in any direction and half registered statistically significant irregularization. The shares of irregular forms increased in participial use for *sneak*, *spoil*, *learn*, *spell*, *leap*, *speed*, *prove*, and *get*. In past tense use, similar changes happened with only *sneak*, *learn*, and *lean*. As there was found to be one clearly favored form for almost all adjectivally used verbs in the present study, no significant changes were found in the distribution of the adjectival forms.

As elaborated in section 4.19, the general trend of the results is irregularization. Even though irregular forms are more closely associated with British English (Biber et al. 1999: 397), calling the irregularization of the corpus results a British influence would be inaccurate. While it is true that diverging from BrE standards is not as important to the identity construction of a Phase 5 variety such as NZE (Schneider 2007: 53), one must keep in mind that of the nine irregularizing verbs of this study, three (*snuck*, *proven*, and *gotten*) are associated with American English. This gives a possible reason for the observed changes in the corpus data: new New Zealand-specific linguistic norms. The implication of the research results is that the NZE norms are mainly based on those from British English, but without the British-style resistance towards irregular forms that are either new innovations (*snuck*) or associated with American English (*gotten*). It is to be acknowledged that the results show a similar development than those of Rickman (2018) regarding the agreement of collective nouns – contrary to expectations, NZE seems to be moving back towards British-style use of irregular verbs and plural agreement.

During the writing of the thesis, it came to my attention that Fairfax Media, the company owning all of the newspapers in CNZNE, used to have an extensive style guide which the journalists writing for the papers were expected to follow. Due to being long out of print, unavailable in electronic form, and associated with a company that does not exist anymore, the style guide could not be obtained. This is a major limitation for the wider applicability of the results presented in this thesis, as it is unknown to which extent the guide covers regular and irregular verb forms. In other words, there is a possibility that any and all changes in the frequencies of the forms can be explained by reasons changes in prescription or closer adherence by the journalists to the style guide (or a more relaxed attitude towards the style guide, if the guide prescribes regular forms).

As the study of regional variation within NZE was beyond the scope of this thesis, an obvious improvement would be to take it into account as well. Could the southern tip of South Island, known for its high concentration of people of Scottish descent (Hay et al. 2008: 98–9), have retained some Scottish English influence in irregular verb usage? At least the form *proven* is said to have spread in BrE from Scotland (Hundt 1998: 36).

The results of the thesis could be further reinforced and expanded with the use of sociolinguistic methodology; does the gender, age or social class of the speaker have an influence on the choice of verb form? Would Pākehā, Māori, Asian or Pacific Islander speakers prefer different forms? Assuming Labovian sociolinguistics to hold true, younger people and women would use irregular forms more than older people and men (Labov 1990: 206), respectively, if irregularization is indeed taking place in NZE.

A different choice of register may yield different results as well – after all, CNZNE only contains newspaper texts. Adding other genres and registers would ensure

that the form preferences (and the direction of diachronic change) are not merely distinctive traits of newspaper language, thus being essential for any wider applicability of the results of this research. It might also be a worthwhile idea to compile a comparable corpus for British English or American English and compare the research results of this thesis to results obtained from those corpora. This would ensure that the comparisons to British and American Englishes made in this thesis are accurate and up to date, and would allow one to find out whether irregularization is a trend even outside of NZE.

REFERENCES

Primary source

Rickman, Paul. Corpus of New Zealand Newspaper English (CNZNE). Compiled by Paul Rickman at the University of Tampere.

Secondary sources

Anderwald, Lieselotte. 2009. *The morphology of English dialects: Verb formation in non-standard English*. New York: Cambridge University Press.

Ball, Catherine N. 1994. Automated text analysis: Cautionary tales. *Literary and Linguistic Computing* 9(4), 295–302.

Bauer, Laurie. 1987. New Zealand English morphology: some experimental evidence. *Te Reo* 30(1), 37–53.

Bauer, Laurie. 1997. Attempting to trace Scottish influence on New Zealand English. In Edgar W. Schneider (ed.), *Englishes around the World: Studies in honour of Manfred Görlach*. Volume 2: Caribbean, Africa, Asia, Australasia, 257–72. Amsterdam: John Benjamins Publishing Company.

Bayard, Donn. 1989. “Me say that? No way!”: The social correlates of American lexical diffusion in New Zealand English. *Te Reo* 32(1), 17–60.

Biber, Douglas. 2010. Corpus-based and corpus-driven analyses of language variation and use. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of linguistic analysis*, 159–92. Oxford: Oxford University Press.

Biber, Douglas, Susan Conrad & Randi Reppen. 1998. *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. London: Pearson Education Ltd.

Burridge, Kate & Simon Musgrave. 2014. It’s speaking Australian English we are: Irish features in nineteenth century Australia. *Australian Journal of Linguistics* 24(1), 24–49.

Cheshire, Jenny. 1993. Standardization and the English irregular verbs. In Dieter Stein & Ingrid Tieken-Boon van Ostade (eds.), *Towards a Standard English: 1600–1800*, 115–33. Berlin: De Gruyter, Inc.

- Chow, Siu L. 1996. *Statistical significance: Rationale, validity and utility*. London: Sage.
- Gibbons, Matthew. 2014. Newspaper quality, content and competition in New Zealand. *Pacific Journalism Review* 20(1), 181–99.
- Gray, Tyler J., Andrew J. Reagan, Peter Sheridan Dodds & Christopher M. Danforth. 2018. English verb regularization in books and tweets. *PLOS One* 13, e0209651.
- Hay, Jennifer, Margaret Maclagan & Elizabeth Gordon. 2008. *New Zealand English*. Edinburgh: Edinburgh University Press.
- Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Hundt, Marianne. 1998. *New Zealand English grammar, fact or fiction? A corpus-based study in morphosyntactic variation*. Amsterdam: John Benjamins Publishing Company.
- Koplenig, Alexander. 2017. Against statistical significance testing in corpus linguistics. *Corpus Linguistics and Linguistic Theory* (forthcoming).
- Labov, William. 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2(1), 205–54.
- Levshina, Natalia. 2015. *How to do linguistics with R. Data exploration and statistical analysis*. Amsterdam: John Benjamins Publishing Company.
- Oakes, Michael P. 2009. Corpus linguistics and language variation. In Paul Baker (ed.), *Contemporary corpus linguistics*, 159–83. London: Continuum.
- Peters, Pam. 2009. Irregular verbs: Regularization and ongoing variability. In Pam Peters, Peter Collins & Adam Smith (eds.), *Comparative studies in Australian and New Zealand English: Grammar and beyond*, 13–29. Amsterdam: John Benjamins Publishing Company.
- Quinn, Heidi. 1999. Variation in New Zealand English syntax and morphology. In Koenraad Kuiper & Allan Bell (eds.), *New Zealand English. Varieties of English around the world*, 173–97. Amsterdam: John Benjamins Publishing Company.
- Quirk, Randolph. 1970. Aspect and variant inflection in English verbs. *Language* 46(2), 300–11.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Pearson Education Ltd.

- Rickman, Paul. 2017. Aspects of verb complementation in New Zealand newspaper English. In Markku Filppula, Juhani Klemola, Anna Mauranen & Svetlana Vetchinnikova (eds.), *Changing English: Global and local perspectives*, 169–90. Berlin: De Gruyter, Inc.
- Rickman, Paul. 2018. Collective nouns and number agreement in New Zealand English. *New Zealand Studies in Applied Linguistics* 24(1), 40–55.
- Rickman, Paul & Mark Kaunisto. 2018. Aspects of the use of the transitive into -ing pattern in New Zealand English. In Mark Kaunisto, Mikko Höglund & Paul Rickman (eds.), *Changing structures: Studies in constructions and complementation*, 71–87. Amsterdam: John Benjamins Publishing Company.
- Schneider, Edgar W. 2007. *Postcolonial English: Varieties around the world*. Cambridge University Press.
- Stats NZ. 2013. *2013 Census totals by topic*.
<http://archive.stats.govt.nz/Census/2013-census/data-tables/total-by-topic.aspx> (8 May, 2019)
- Stefanowitsch, Anatol. 2004. Quantitative thinking for corpus linguists, 5 pp.
http://www-user.uni-bremen.de/~anatol/qnt/qnt_dist.html (22 March, 2019).
- 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), 227–39.
- Trudgill, Peter. 2000. *Sociolinguistics: An introduction to language and society*, 4th edn. London: Penguin Books.
- Wallis, Sean. 2013. Z-squared: The origin and application of χ^2 . *Journal of Quantitative Linguistics*, 20(4), 350–78.

Appendix: Abbreviations of the newspapers featured in the corpus.

DO	<i>Dominion</i>
DP	<i>Dominion Post</i>
EM	<i>Nelson Mail</i>
EP	<i>Evening Post</i>
ES	<i>Evening Standard</i>
MS	<i>Manawatu Standard</i>
PR	<i>Press</i>
SS	<i>Sunday Star Times</i>
ST	<i>Southland Times</i>
TD	<i>Taranaki Daily News</i>
TH	<i>Timaru Herald</i>
TR	<i>Truth</i>
WT	<i>Waikato Times</i>