

**Aspects of change in the syntax of Māori -
a corpus-based study**

BY

KARENA GWEN KELLY

A thesis submitted to the Victoria University of Wellington

in fulfillment of the requirements for the degree of

Doctor of Philosophy

Victoria University of Wellington

(2015)

ABSTRACT

In the current climate of Māori language revitalisation, there is ample anecdotal evidence to suggest that not only the vocabulary, but also the syntax of modern Māori is markedly different from its traditional roots, and that it shows significant influence from English syntax. However, syntactic change in Māori has not hitherto been rigorously studied. This thesis aims to provide material evidence of change in Māori syntax, through a corpus-based study of grammatical change in te reo Māori over the period of contact with English.

My methodology involved the compilation and comparison of two synchronic corpora representing the two ends of the contact period to provide a diachronic perspective on the language. Each corpus consists of approximately 102,000 running words of material written originally in Māori. The early corpus contains items published pre-1900. The modern material was written post-1990. The thesis is not only an exploration of the possibility of documenting syntactic change through the use of such corpora, but also tests whether it is possible to do this using corpora significantly smaller than the multi-million word corpora typical in corpus linguistics.

The scope of this methodology is tested by examining three distinct types of grammatical features: a grammatical particle (the preposition *mō*), a pair of semantically related lexemes that appear to be undergoing a process of grammaticalisation (the verbs *taea* and *āhei*), and a widespread grammatical construction (certain types of relative clauses). In each instance, the two corpora are compared for features such as the frequency of occurrence, the associated constructions, and the contexts of use.

In relation to the methodological questions, the thesis concludes that while these corpora are too small to provide adequate data on individual lexical items like *taea* and *āhei*, the methodology did make it possible to document change in the other, relatively high-frequency grammatical features.

The thesis also raises the questions of whether the changes identified result from the direct adoption of English usages and constructions, whether they result from insufficient exposure to traditional Māori as a result of the dominance of English, or whether they are perhaps instead the result of internally-motivated Māori language evolution. This leads to a discussion of the likely implications of the changes documented here for the future of the language and the language community. I argue that acceptance of all such change as natural and unavoidable is likely to be detrimental to the future of te reo Māori.

ACKNOWLEDGEMENTS

Kei ngā rangatira o te tuaka, ko te reka o ō koutou reo e kōhimuhimu tonu mai ana i ngā whārangi o ā koutou nā tuhinga. E kui mā, e koro mā, okioki mai rā. Kei ngā rangatira o te pene, o te papa pātuhi, mokori anō kia rere a mihi ki a koutou e hiki nei i te hoe kua mahue mai i a rātou mā, me te urungi atu i tō tātou waka reo ki te pae tawhiti e muramura mai nei. Kei ngā kaitiaki o te karahipi a Philippa rāua ko Morvyn Williams, tēnā rā koutou i tā koutou tautoko i puare mai ai te tatau ki te ao rangahau - ko te tuhinga nei tētahi anō o ngā tini hua o ō koutou whakaaro rangatira ki te hunga akoako pēnei me au nei.

Kei ngā tini manu kai mātauranga, kei taku kāhui hōmiromiro i tutuki ai tēnei tuhinga, me te manutaki, a Winifred, i ora ai tēnei kaupapa, tēnei au e mihi kau atu nei.

I would like to acknowledge a number of people for their assistance and support during the long journey that concludes with the submission of this thesis. Firstly to Dr Winifred Bauer, without whom this study never would have been undertaken, and on whose linguistic foundation this thesis was built, I am indebted to you for your support, and your generosity in sharing your expertise and passion for this subject with me. Both my reo and my career are profoundly changed for having met and worked with you. I also thank Associate Professor Meredith Marra and Emeritus Professor Janet Holmes for their guidance and support as secondary supervisors, and Associate Professor Paul Warren for his administrative support, especially throughout the latter stages of this PhD. I am grateful for the tautoko offered by my colleagues at Te Kawa a Māui, and for the support of te whānau o Te Herenga Waka in the many years since I first began to learn te reo under the shelter of our whare, Te Tumu Herenga Waka. To my teachers and my students who have so often guided the course of this research, and whose love and enthusiasm for te reo reinforces mine, I thank you. Dr Jennifer Martin, Dr Catherine Trundle, Fiona Leathart, Kelly Keane-Tuala, Nadia Te Huia, Naoko Yui, mei kore ake koutou. To my other friends and colleagues

who I have not named, but who have encouraged and supported me throughout the duration of this marathon of a task, I send my heartfelt appreciation. Lastly to my family, I send my love – e taku kati taramea, e taku toka tū moana - this is for you.

TABLE OF CONTENTS

ABSTRACT	i
ACKNOWLEDGEMENTS	iii
LIST OF TABLES AND FIGURES	vii
LIST OF ABBREVIATIONS	viii
1. INTRODUCTION	1
2. METHODOLOGY	6
LITERATURE REVIEW	6
CORPUS CONSTRUCTION	8
SELECTION CRITERIA	10
THE DATA COLLECTION PROCESS	16
THE ANALYSIS PROCESS	21
3. CASE STUDY ONE: <i>MŌ</i>	23
DATA SELECTION AND CATEGORISATION	23
Category 1: Future possessor	23
Category 2: Benefactive	24
Category 3: Topic of discourse	25
Category 4: Reason	25
Category 5: Position	27
Category 6: Transaction	28
Category 7: Duration	29
Category 8: Purpose	29
Category 9: Future location	30
Category 10: Performer of intransitive nominalisation	31
Category 11: Emphatic refusal	31
RESULTS	32
4. CASE STUDY TWO: <i>TAEA & ĀHEI</i>	44
<i>TAEA</i>	45

<i>ĀHEI</i>	54
5. CASE STUDY THREE: RELATIVE CLAUSES	57
LITERATURE REVIEW	58
PARAMETERS FOR RELATIVE CLAUSE ANALYSIS	65
The sentence function relativised on	65
The relativisation strategy used	65
The TAM used in the relative clause.....	66
The particle used in the relative clause.....	66
RELATIVISATION STRATEGIES	67
The deletion strategy	67
The particle strategy	68
The pronoun strategy	69
The possessive strategy	70
DATA SELECTION	72
RESULTS FROM THE DATA ANALYSIS	75
Sentence function relativised on	75
Relativisation strategy used	82
Subordinate TAM	83
SUMMARY OF RELATIVE CLAUSES RESULTS	89
6. IMPLICATIONS	90
Theme 1 – Shift toward English	91
Theme 2 – Specificity shift	94
7. CONCLUSION	103
METHODOLOGY	103
CHANGES OBSERVED	104
AREAS REQUIRING FURTHER STUDY	106
APPENDIX 1: Sample of <i>mō</i> data collection (traditional & modern)...	108
APPENDIX 2: <i>Taea</i> & <i>āhei</i> data collection.....	111
APPENDIX 3: Relative clauses - problematic examples	113
REFERENCES	116

LIST OF TABLES AND FIGURES

Table 1: Overall frequencies of different forms of <i>mō</i>	32
Table 2: Relative clauses grouped by sentence function relativised on	76
Table 3: Relative clauses on Subjects by relativisation strategy used.....	77
Table 4: Relativisation strategy used	83
Table 5: Occurrence of non-past TAMs and accompanying particles	85
Table 6: Frequency of TAMs used in relative clauses	87
Figure 1: Example of traditional corpus item photograph	17
Figure 2: Preposition used for complement to <i>tatari</i>	37
Figure 3: Frequency of senses of <i>taea</i> in the corpora	46
Figure 4: Percentage distribution of sentence functions relativised on.....	79
Figure 5: <i>E...ana</i> in relativisations on different sentence functions.....	84

LIST OF ABBREVIATIONS

1SG/PL	first person singular/plural
2SG/PL	second person singular/plural
3SG/PL	third person singular/plural
1DLINCL/EXCL	first person dual inclusive/exclusive
1PLINCL/EXCL	first person plural inclusive/exclusive
CAUS	causative prefix
CLASS	classifying particle
DEI	deictic particle
DET	determiner
DIR	directional particle
DL	dual
DO	Direct Object marker
INTENS	intensifier
LOC	locative particle
MAN	manner particle
NEG	negator
NMLZ	nominalising suffix
NUM	numeral particle
PART	particle
PASS	passivising suffix
PERS	personal article
PL	plural
POSS	possessive
PREP	preposition
REDUP	reduplicated form
SG	singular
TAM	tense/aspect/mood marker
VC	verb constituent

1. INTRODUCTION

Ko te reo te mauri o te mana Māori

Tā Hēmi Henare

The Māori language of the Māori people of Aotearoa/New Zealand is seen by many world-wide as a model of highly successful minority-language revitalisation (McCartey, 2010; Stiles, 1997). Indeed, the Māori-speaking community of today benefits from the concerted efforts of a dedicated group of Māori language proponents throughout the period of the renaissance of Māori language and culture. This renaissance first began to gain traction in New Zealand society in the 1970s (Statistics New Zealand, 2014; Waitangi Tribunal, 2013), after decades of cultural assimilation policies (Ngaha, 2014; Spolsky, 2003; Winitana, 2011) and the resulting “gradual loss of the language over a century of contact with colonizing English” (Spolsky, 2003, p. 553).

Māori language revitalisation efforts thus far have concentrated primarily on producing more speakers of te reo (Te Taura Whiri i te Reo Māori, 1992), and on providing new vocabulary to enable the language to be used for a wide range of contemporary topics (Harlow, 1993). This effort has caused a noticeable change in the structure of the Māori language community, as evidenced by the changing composition of the Māori-speaking population according to the statistics available – the most recent survey results available from *Te Kupenga* (2013) show that just 8% of Māori adults say te reo Māori is their first language, and only 32% of Māori adults with Māori as a first language report themselves as being able to speak it very well or well (Statistics New Zealand 2014). This means it is likely that fewer than 3% of Māori adults are competent L1 speakers of the language. However, although L2 speakers make up the overwhelming majority of the Māori-speaking community (Statistics New Zealand 2014), the effects on the language of the influx of L2 speakers of Māori in the modern period has not been formally investigated, though the level of proficiency attained by new L2 speakers of

Māori has been the source of comment, for example in relation to the proficiency of teachers of the language (cf. Education Review Office, 1995; 2002). My study is an attempt to begin to fill this gap, by providing empirical evidence of change in the language, and investigating the possible sources of the changes detected.

There is ample anecdotal evidence to suggest that te reo Māori has changed noticeably in the modern period. Examples of this type of anecdotal evidence include observations of language teachers in both immersion and mainstream education settings, the expressed frustrations of the L1 Māori elders who experience difficulty understanding the language used by their L1 or L2 Māori grandchildren or spoken on Māori television, and the coining of the terms ‘new Māori’, ‘book Māori’ or ‘school Māori’ to differentiate the variety of the language produced by heritage-language learners in formal Māori language-learning programmes from that spoken by their L1 elders in the kitchens and on the benches of their own marae. Some directly attribute this perceived change to the increasing proportion of L2 speakers in the Māori language community – thus Tīmoti Kāretu said in an interview in *The Listener* in 1990, “Unfortunately, you hear bad Māori everywhere. The trouble with the second-language learner is that their incompetence becomes the norm” (cited in Garlick, 1998, p. 44). The obvious changes in the vocabulary of Māori of the modern period are both expected and reasonably easily detected for the average speaker, but the often subtle nature of syntactic change means it is more difficult to identify and therefore less commonly discussed. An informal review of Māori-language publications suggests that traditional texts contain examples of word-types used in functions that seem to have disappeared in modern Māori (cf. Bauer, 2014) and modern texts contain examples of word-types in functions that are unprecedented in traditional Māori texts. A great many of these examples in modern Māori seem to mirror English constructions. The following quotation from Harlow (2005) elaborates:

It is striking that many people, particularly learners, who take great pains to avoid borrowing words from English, show no such aversion to using English syntax and idiom in their Māori. This can be observed in both spoken and written Māori even among advanced learners to a level which should be the despair of their teachers. Some instances of the influence of English have already almost become “correct” Māori, such as the perhaps relatively harmless use of *mō* ‘for’ to mark the complement of *tatari* ‘wait’ instead of the older *i* or *ki*. ... other examples which abound are so clearly direct encodings of English and so un-Māori that their absorption into Māori would represent a severe compromising of its genius. (pp. 137-8)

My whakapapa and my personal experience in Māori language teaching and learning underpin both the motivation for and direction of this research, and so my positioning is briefly outlined here. I am a heritage language learner of Māori, and the only Māori speaker in my immediate family. I began to learn te reo Māori when I enrolled in a teaching degree at Victoria University of Wellington in 2003. A Māori linguistics course was prescribed in the final year of my undergraduate degree, and as a learner of te reo I found this linguistic perspective invaluable, in that it could pinpoint and demystify some aspects of the language that had been otherwise difficult to identify, understand or have explained in my language-learning experience. This linguistic perspective also informs my practice as a Māori language teacher. My study is built on a desire to test the value of linguistic research as a means to shed light not only on the situation of te reo Māori in the modern period, but also to meaningfully contribute to quality Māori language teaching and learning, and thus Māori language revitalisation.

This thesis details the trial of the comparison of data from two small corpora, or principled collections, of Māori-language texts from different periods, in an attempt to provide robust, measurable evidence of language change. Where evidence of change is found, an investigation is conducted into the source

and/or cause of the change, to determine whether it may result from the influence of English, or whether it is brought about by some other language-internal processes. The consequences of the observed change for the Māori language and its speakers are then considered.

The literature drawn on to inform this thesis was so diverse overall and particular to certain sections that the review of literature required for this thesis is not contained within a single separate chapter. Instead, an individual review of the relevant literature has been included in each chapter. A review of the available literature pertaining to this methodology and the three chosen language forms upon which the three case studies are based can be found near the beginning of each of those respective chapters.

In Chapter 2 of this thesis the methodology being trialled is described in detail, and some of the critical decisions made and difficulties encountered in the project design and the corpus construction processes are discussed.

Chapter 3 is the first of the case studies for this project, which focusses on the use of a grammatical particle, namely the preposition *mō*, searching for evidence of change in the frequency of its use overall, and the proportional distribution across the different senses of *mō* between the two corpora.

Chapter 4 is a case study based on a pair of lexemes, *taea* and *āhei*, where both the frequency of the different senses of each and the syntactic environments in which the lexemes are found are analysed to determine whether there has been change in their use in the modern period. This case study also provides a test for the use of small corpora for lower frequency items, and shows the limitations of corpora of this size.

Chapter 5 tests the use of this methodology for a case study based on a grammatical construction. The occurrences of a certain subset of relative clauses are analysed, in order to try and detect change in the frequency of usage, the relativisation strategies used, and the distribution of the key grammatical particles within relative clauses in Māori.

In Chapter 6 the results from the case studies are summarised, and the overarching themes in the detected changes are described. The potential implications of the observed changes for the Māori language are then explored in Chapter 7, which leads to a discussion of the tenability of different positions represented in the Māori language community on the spectrum of attitudes towards change in te reo Māori.

The three case studies brought to light different issues with the methodology, and thus present a fair test. The findings from the case studies do not in all instances confirm the anecdotal evidence or my own intuitions, and thus demonstrate the usefulness of the methodology and the importance of this research. The implications chapter is intended to raise the level of the discussion of change in the Māori community, and while some the arguments may be contentious, it is hoped that this will serve to provoke thought and argument about a topic which is under-represented in language education, language policy and planning, and general debate about the language and its revitalisation.

2. METHODOLOGY

This research project is guided by the hypothesis that it is possible to use two synchronic corpora of Māori language material from different periods to give a diachronic perspective on the language, and identify aspects of change in Māori syntax over the period of time between the two corpora. This chapter describes the corpus construction process, detailing the fundamental literature, linguistic principles and critical decisions made. The process of analysis of the Māori-language features in the three case studies is then outlined.

LITERATURE REVIEW

Corpus linguistics is defined by Conrad (2000) as “the empirical study of language relying on computer-assisted techniques to analyze large, principled databases of naturally occurring language” (p. 548). The field of corpus linguistics began before the widespread use of computers: as Leech (2011) states, “it can be said that the corpus revolution in linguistics began with the completion and distribution of the Brown Corpus in 1964” (p. 10). Since then, the use of corpora for linguistic study has increased exponentially, alongside the increased availability of computers and the technological advancement of various text recognition tagging programmes and word-form analysis software (see for example Conrad (2000) and Gardner & Davies (2007)).

Boyce (2006) describes a corpus as a “large and principled collection of natural texts” (p. 37). A synchronic corpus includes a temporal restriction as one of those principles, restricting the collection to relevant natural texts from a particular period, as “an attempt to represent a language at a particular time” (Kennedy, 1998, p. 22). A diachronic corpus, by comparison, focuses upon tracking the development or evolution of language through time.

Constructing a diachronic corpus of written Māori in order to track change in Māori syntax is not possible; a marked decrease in Māori publications between 1900 and 1960 was noted by the Turnbull librarian, Sheila Williams (1990), and the dearth of Māori language publications in the modern period is also acknowledged by others, such as Garlick (1998). This trend continues despite a small increase in the material available since the publication of her book. The scant material available during this period would create a gap in the diachronic corpus, compromising its effectiveness in representing the entire period, and therefore the validity of the results. This research therefore aims to construct two synchronic corpora from different periods, and to compare the data on particular language forms from each corpus as a means of documenting change in these elements of the Māori language across the period between the two corpora. This methodology has a precedent in L. Bauer's (1994) corpora-based study of linguistic change in English. In that study, two corpora of material obtained from certain newspapers from different periods were compared with regards to many different types of language features, including some grammatical ones, such as the form of comparatives and relative clauses, in order to pinpoint loci of change in the language between the two periods.

The use of corpora for the study of te reo Māori is by no means unprecedented – corpora have been fundamental to a number of influential research projects regarding the Māori language. Biggs' PhD thesis entitled *The structure of New Zealand Maaori* (1957) was based upon his corpus of recordings of spoken Māori language, and it was upon this data that he based his grammar (1973). The MAONZE project, started in 2004, is “an acoustic sociophonetic investigation of sound change in Maori” (Watson, 2009), and has produced a large body of research into change in the pronunciation of the Māori language, using four synchronic corpora of audio recordings; three were of small cohorts of contemporary speakers of a certain age group, and their pronunciation of Māori was compared with that demonstrated in an archive of recordings of seven speakers born in the late 19th century (cf. Harlow, Keegan, King, Maclagan, Quinn & Watson, 2004). Mary Boyce's (2006) PhD study involved the construction of a 1-million word

corpus of modern spoken Māori, and the analysis of its contents with regards to word frequency. Her corpus has provided the basis for further vocabulary-based study, including Keane-Tuala's (2013) investigation into high frequency homonyms. The Legal Māori Project, based at the Law Faculty of Victoria University of Wellington, created an 8-million-word corpus of legal Māori language, and used that corpus as the foundation for the Legal Māori Lexicon and *He Papakupu Reo Ture: A Dictionary of Māori Legal Terms* (Stephens & Boyce, 2013).

Despite the number of corpora that have already been built for the analysis of the Māori language, none of them were appropriate for this study. While the MAONZE project corpora have provided ample data for capturing the pronunciation of the various speakers, they are unlikely to contain suitable syntactic data for this research project. Each of the MAONZE project's corpora is based on a very small number of speakers, and a significant proportion of the historical material involves the discussion of a limited range of topics, like the speakers' genealogies, thus relying heavily upon certain formulaic expressions and sentence types, for example '*Ka moe a X i a Y, ka puta ko Z...*' or '*Nā X ko Y, nāna ko Z...*', thereby producing a limited range of grammatical constructions. In fact, spoken data may not provide the most reliable basis for a study of syntax such as this, due to the increased potential for the use of contractions, and the increased likelihood of grammatical words being elided or undetectable in the flow of natural speech. For this same reason Boyce's Māori Broadcast Corpus was also inappropriate as the principal data source for this study. The Legal Māori Corpus contains a massive amount of written material in Māori, but a significant proportion is comprised of translated material written originally in English. Such material did not meet the criteria I established for my corpora.

CORPUS CONSTRUCTION

This study aims to provide measurable evidence to test the anecdotal observations of native speakers, who claim that the Māori language is changing to become more like English. For this reason, establishing the period of each synchronic corpus was critical. The first corpus is intended to

represent the Māori language at the point where there was no, or minimal influence from English. While anecdotal evidence from native speakers suggests that there has been considerable change in the Māori language since as recently as the 1960s (T. R. Higgins, personal communication, 18 May, 2008), the Benton survey (1991) provides the most comprehensive information available on Māori fluency statistics, and suggests that in all the districts with a high Māori population, and for the respondents that were measured, the fluency level in te reo Māori for those Māori born in 1900 was very near to 100%. Soon after this time in many of the areas surveyed, the proportion of fluent speakers of Māori decreased dramatically, at the same time as the number of Māori people speaking English rose. It would appear then that taking extracts from sources produced before 1900 would give the highest possible probability that the author of the extract was a native speaker of Māori. The first corpus, labelled the traditional corpus, was therefore restricted to Māori language material produced pre-1900.

The second corpus is intended to provide an accurate representation of the Māori language of today. The resurgence and revitalisation of te reo Māori which began in the 1970s (Ngaha, 2014; Statistics New Zealand, 2014) was formally recognised in New Zealand legislation with the establishment of the Māori Language Act in 1987. The adult Māori-speaking demographic began a period of change in the early 1990s, when the first students educated in Māori-medium schools began to reach adulthood. By this time various Māori-language teaching and learning programmes were established in both the education sector and the wider communities. For this reason the second corpus, labelled the modern corpus, was restricted to Māori language material produced post-1990. The types of speakers included in the modern corpus are discussed in detail later in this chapter.

A key concern in the development of corpora is the concepts of representativeness and balance; these are important ideas with regards to my own corpora. Kennedy (1998) mentions the importance of clearly specifying the part of the language that the corpus aims to represent, and defining that domain, genre, topic or subject field. Some of the parameters

and selection criteria mentioned are irrelevant given the intended use of my corpora for syntactic analysis, such as the ‘influentialness’ or academic merit scale suggested by Summers (1991) (cited in Kennedy, 1998). The principles suggested by Sinclair (1991) for a general written corpus include “the distinction between fiction and non-fiction; book, journal or newspaper; formal or informal; with control of age, gender and origin of the author” (cited in Kennedy, 1998, p. 63). These principles fall into two overarching groups, namely characteristics of the texts, and the characteristics of the text authors.

SELECTION CRITERIA

Considering first those principles pertaining to characteristics of the language material itself, the corpora constructed for this research project were restricted with regards to the formality of the language event: my corpora include only semi-formal to formal material, to minimise the likelihood that recorded language behaviours could be attributed to careless language use. The fiction/non-fiction distinction was irrelevant for this study, as it is desirable to include both genres to fully represent the syntax of the language. As wide a variety of types of material was included as possible in my corpora, though a few types of sources were expressly excluded from the collection. Materials written for young children and elementary language learners were excluded, to avoid the inclusion of simplified structures and other linguistic accommodations made for a juvenile or inexperienced audience. Informal correspondence like emails, Facebook entries and text messages were also excluded, given the increased likelihood of careless speech in most examples of those types. Some heavily illustrated material was likewise excluded: where the text was structured as a spoken conversation, the text was often contracted to fit within limited physical spaces in items like cartoon strips, and the text was dependent upon the accompanying pictures to convey meaning.

Translated material was excluded from these corpora; therefore all material included in both corpora was written originally in Māori. This is to avoid the possibility of skewing the data for either corpus by including language material perhaps inevitably influenced by English through the process of

translation. This excluded a huge amount of the material available, including some of the longest available texts from each period, for example the Bible, and the *Dictionary of New Zealand Biography*. As mentioned previously, both corpora are restricted to include only written material, given the paucity of spoken data available due to the technological limitations of the pre-1900 period of the traditional corpus, and the need for comparability between the two corpora.

The Alexander Turnbull Library was the principal source of material for the traditional corpus. This was largely a pragmatic decision, as the Turnbull's location ensured easy access to its resources, and there is a large body of pre-1900 published Māori language material within the Books in Māori collection. The collection contains a wide variety of types of texts, with various topics, from a large number of different contributors, and so provided a good cross-section of published Māori language available from that period. Moreover, Parkinson & Griffith's (2004) annotated bibliography for the collection proved to be an invaluable resource, and facilitated the selection process by providing the majority of the information required regarding selection criteria without my having to request and examine each of the items of the collection individually. The entries detailed all available information on the authorship, content, and often historical context of the thousands of books, pamphlets, single sheet items and serials available at the Alexander Turnbull Library for the traditional period. This greatly facilitated the elimination of the massive proportion of translated material, and helped to avoid over-sampling work from individual authors (cf. Biber, Conrad & Reppen, 1998).

Initially my intention was to include only published material, in an attempt to ensure that only language that had been carefully constructed and deliberately used by the author was included. An exhaustive search was conducted of the Books in Māori collection for the available material published during the traditional period. However, the resulting collection was too small to enable the construction of a 100,000 running-word corpus according to the principles I had established. It also became clear on review

of the published material for the traditional corpus that the material was more restricted in type and length than that of the modern corpus, due perhaps to the technological limitations, limited availability and greater expense of the publication process pre-1900. Both corpora were therefore expanded to include non-published material, but remained restricted to material that was semi-formal and formal in register, to give some assurance that the author would be less likely to use contractions and other linguistic accommodations like colloquialisms - hence the inclusion of material such as some of the material from the Alexander Turnbull Library's collection, Letters written to Donald Maclean (in Māori), in the traditional corpus, and some unpublished stories and university essays in the modern corpus.

Extracts included in the corpora were limited to a maximum of 1000 words in order to avoid the overrepresentation of the linguistic preferences of authors who produced longer pieces of writing. Where the sources were longer than 1000 words, a 1000-word sample was taken from each, starting from the third line of text. This starting point was arbitrarily chosen, but was intended to offset frequency imbalances caused by the overrepresentation of certain constructions which appear as part of a stylistic norm at the beginning or ends of some types of writings, for example salutations in letters. The *nā*+author phrase found frequently at the end of publications and letters in the traditional period was retained when it occurred within the 1000 word limit of a text in the data collection process, and although this may contribute to a perceived greater frequency of the particle *nā* in the traditional corpus, this will have no bearing on the results of this project, as *nā* within those syntactic contexts is not expressly studied in this thesis.

The characteristics of the authors of texts raised far more controversial and difficult issues than the characteristics of the texts. Not all characteristics caused problems: for example, gender was unimportant for this study. Furthermore, the only exclusion in terms of author age was child authors, which was only likely in the modern period, where technological advancement has greatly facilitated the publication process, making it feasible for people of any age to produce and disseminate their own material.

Accounting for author ethnicity posed some problems. The initial intention was to include only material produced by Māori authors, in the interest of ensuring language proficiency of the author, at least for the traditional period. However, determining author ethnicity proved difficult in many instances, a Māori author name being a dubious assurance; a number of entries within the Books in Māori annotated bibliography mention Pākehā referents with Māori aliases, for example John Sheehan, a.k.a. Hone Hihana or Hone Hiini (Parkinson & Griffith, 2004). There are also a considerable number of sources where it is not possible to determine the proficiency or ethnicity of the author at all. For example, in situations where a governmental office or station is credited with the publication, it is difficult to determine whether the publication was written in Māori by a non-Māori issuing officer, or instead on his behalf by a native speaking translator (Parkinson & Griffith, 2004). In fact, it is feasible that in the pre-1900 period, some Pākehā were born and raised in New Zealand immersed in and proficient in the Māori language. The ethnicity of the author was therefore shown to be a less significant selection criterion, particularly in the modern period, where the reported fluency levels of the Māori population mean that an author's Māori ethnicity, even where possible to ascertain, gives little independent assurance of their proficiency in the language.

With regards to the language proficiency of the authors, as previously mentioned, the traditional corpus was set to provide a benchmark of the state of the Māori language before the influence of English, and therefore material was taken from Māori authors before 1900, during which time a Māori individual was very likely to be an L1 speaker of Māori. For the modern corpus, materials were taken from L1 speakers of Māori, and also L2 speakers of Māori who were confident and frequent users of the language, and/or had Māori language qualifications of a level that would permit them to teach Māori in New Zealand secondary schools. This may appear on the surface to create two incongruent corpora, in that they each represent a different profile of speakers. However, as Mauranen states in her discussion of the usefulness of learner corpora “there is no reason to assume that the only speaker groups of linguistic interest should be native speakers” (2011,

p. 155). The inclusion of language material from L2 speakers of this type is warranted not just by their overwhelming majority in the population of Māori speakers in the modern period (see the Introduction), but also by the fact that the Māori spoken by this group is deemed suitable for those in roles as language models in the community.

It was in considering the author's language proficiency that a key issue in corpus construction was raised, namely balancing the principles of representativeness and comparability. As Vaughn & Clancy (2013) state, "Issues of balance and representativeness are salient no matter the size of the corpus" (p. 56). This methodology is centered on the comparison of the data from the two corpora; maintaining the comparability of the two corpora is thus vital to ensuring the validity of the results. This theoretically means that the materials included in each corpus should be as alike as possible with regards to the characteristics of both the author and the language item itself. For example if the traditional corpus contains only items from a native-speaking Māori author, then to be strictly comparable only items from L1 speakers of Māori should be included in the modern corpus. It is clear, however, that while doing this would result in a modern corpus that was optimally comparable with the traditional corpus, it would not be representative of the users of the Māori language in the modern period. As mentioned in the introduction to this thesis, less than 3% of Māori adults in 2013 are competent L1 speakers of Māori, and they comprise only about 5% of the population of Māori-language speakers, that is, those who can use Māori to at least talk about basic/ simple things (Statistics New Zealand, 2014). The primary need was for the two corpora to be highly representative of the Māori language for their respective periods in order for the findings to be useful. In fact, the corpora's representations of the Māori language of their respective periods would ultimately form the basis of their comparability.

Given the speakers included from the modern period it may be argued that the modern corpus is in fact more of a learner corpus than a corpus of the Māori language of the modern period, and the corollary of that is that this study provides not a diachronic perspective on the language but a

comparison of L1 and L2 speakers from different periods. It is important to note that this modern corpus does contain material from L1 speakers of Māori from the modern period in order to represent the entire demographic of adult Māori language speakers in the modern period, although I believe that an L2 corpus arguably could provide a reasonable representation of the Māori language of the modern period, given that L2 speakers are estimated to make up around 95% of the adult Māori speaking population (Statistics New Zealand, 2014).

The representativeness of each corpus was especially important given that this research project is a trial not just of the comparison of two synchronic corpora for a diachronic perspective, but also of the use of corpora a fraction of the size of the multi-million running-word corpora typically prescribed for use in corpus linguistics (Boyce, 2006). Building multi-million word corpora was not feasible for this study, especially given the limited material available and the laboriousness of the data-collection process for the traditional corpus. In general the dictum that 'bigger is better' applies in regards to corpora, so the usefulness and legitimacy of the use of small corpora for linguistic study has been the source of debate. Thus Vaughn & Clancy (2013) state,

What we are implying is that it has not always been a given that corpora considered 'small' had full legitimacy in the field of corpus linguistics. A major reason for this reluctance to fully admit small corpora to the fold was rooted in, as previously mentioned, the predominant research agenda in corpus linguistics in its 'early modern' period, lexicography, and the remediation of concerns in relation to 'representativeness' and 'balance' in commercial corpus building. Corpora used for lexicographical research need to be as large as possible in order to generate sufficient occurrences which reflect how lexical items are used. (p. 56)

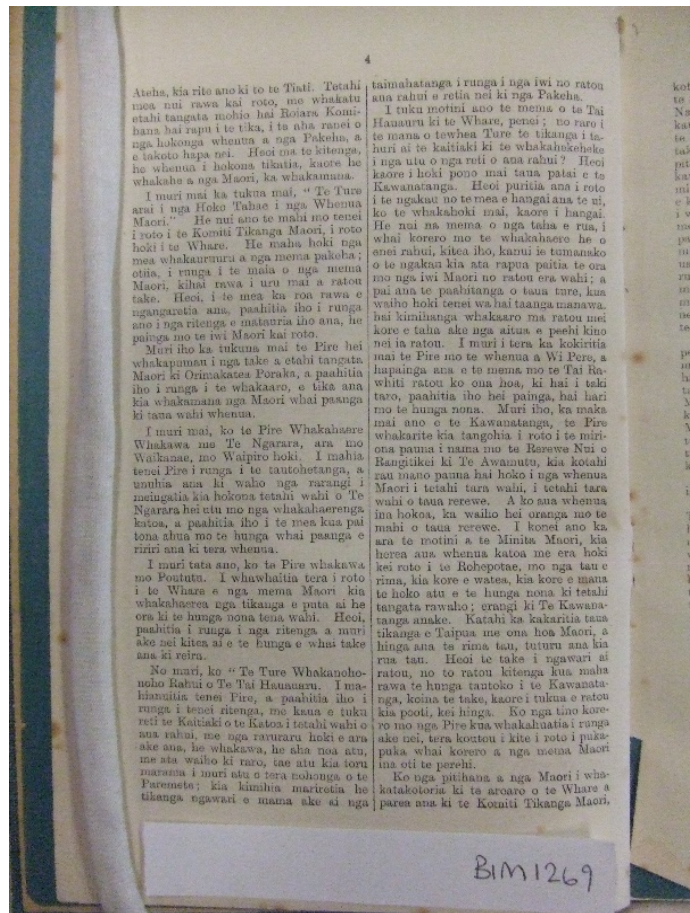
The methodology for this research project is guided by the hypothesis that while larger corpora are required for the study of vocabulary and lower-frequency language forms, corpora of a modest size are still able to provide valid observations on aspects of syntactic change in relation to high-frequency language forms (Hunston, 2002). There is evidence of the legitimacy of the use of smaller corpora: Flowerdew (2002) states, “The field [of corpus linguistics] has widened considerably to include the recognition of much smaller, specialised genre-based corpora” (p. 96). Neff van Aertselaer & Bunce (2011), for example, detail the use of the SPICLE corpus of around 200,000 running words, which was compared with two sub-corpora of around 30,000 running words in total to analyse the writing of Spanish university students. Vaughn & Clancy (2013) also describe the usefulness of small corpora with regards to the study of pragmatics.

Aston (1997) describes small corpora as between 20,000–200,000 running words of text. The two corpora for this study were therefore constructed with an aim of reaching 100,000 running words of Māori language text.

THE DATA COLLECTION PROCESS

As indicated above, the process of data collection for the traditional corpus was particularly laborious. After an exhaustive search for appropriate material from the Books in Māori collection at the Alexander Turnbull Library, the result was a list of 190 individual items. The fragile condition of the books meant they had to be handled under the strictest conditions, and so were requested one by one at the library. Each page was individually tagged with a reference number indicating the item’s reference details in the Books in Māori annotated bibliography (see BIM1269 in the image below), weighted down in position and photographed in low light, as shown in Fig.1 below.

Figure 1: Example of traditional corpus item photograph



Unfortunately, due to the condition of the original texts, the variety of texts and fonts used and the varying quality of the photographs, text recognition software was too inaccurate to be effectively used with the photographs, and a professional transcription service based overseas could not provide the required degree of accuracy for a reasonable price, given their transcribers' lack of Māori-language expertise. This meant that each item had to be transcribed manually from the photographs taken. A small proportion of the material that I photographed for the traditional corpus was also set to be digitised for the Legal Māori Project, and in return for my providing those images, the Legal Māori Project Leader, Māmari Stephens, kindly prioritised the digitisation of those items and provided them for inclusion in the traditional corpus before the release of the Legal Māori Corpus. No editing was done of any of the items during the transcription process, which meant that macrons were not added to the texts included in the traditional corpus. The collection of letters to Donald Maclean also housed at the Alexander

Turnbull Library provided the unpublished material for the traditional corpus. The digitisation process here was much more straightforward, as many of the items in that collection were already digitised by the Library and made available electronically.

There was a much wider variety of material available for the modern period, and a relatively broader range of different types of sources, from tertiary level academic submissions to collections of short stories, to descriptions online of competitions on Māori television. This therefore required a much broader search in order to accurately sample the variety of types of material available for the modern corpus. It is worth noting however that despite this broad range of sources and resource types, with regards to topic the majority of the material available in the modern period written in Māori was actually about the Māori language itself, from reports on Māori language proficiency to personal narratives about individual language-learning journeys, to opinion pieces on observations of Māori-language teaching, learning and use in the community: ironically, some of the very anecdotal evidence that this study looks to test. This was an interesting point of difference between the two corpora – the traditional corpus contained a limited variety of types of items, but on a comparatively wider range of topics, while the modern corpus contained a wider variety of types of items, on a much more limited range of topics.

It was difficult in some instances to ascertain the author and original language of a piece of material. In those instances if an original English text could not be found, the item was assumed to be written originally in Māori and a single sample was taken from the source in question. In many instances it was also difficult to determine whether editorial corrections had been made by a third party, such as in published collections of short stories, although material was not excluded from the corpus on those grounds alone. In most instances the material for the modern corpus was already digitised, although transcription was required of extracts from some published material, and some of the donated unpublished material, such as old essays completed for higher university level Māori language courses.

Some difficulties arose in the data-collection process for the modern corpus in relation to the sensitivities of living authors. A number of individuals who were made aware of this research project and met the selection criteria donated appropriate material for inclusion in the modern corpus, but most did so only on the assurance of their absolute anonymity, concerned that exposure of their language behaviours would cause them embarrassment if they were judged to be untraditional or incorrect. For this reason, every effort has been made to protect the anonymity of the author – for example, when modern corpus material is included in this thesis, trivial changes have sometimes been made to preserve author anonymity; as Hasund stated (1998) on the subject of informant anonymity in relation to the Bergen Corpus of London Teenage Language, “Informants’ rights to privacy should and must override other concerns” (p. 25). Some contributors for the modern corpus readily donated material for inclusion in the collection, and then subsequently requested the withdrawal of their material, even after being assured that their anonymity would be protected, because they were uneasy about any potential risk, however small, of exposure or embarrassment about a perceived lack of Māori language expertise. The relationship between concepts of cultural confidence and linguistic insecurity are further discussed in Chapter 6 in regards to the overall trends observed in the case studies and the implications for the language of the results obtained through this study.

In relation to the corpora, and to the examples from the corpora cited in this thesis, it is necessary for the reader to know that in the course of its comparatively short history as a written language, three different orthographic conventions have been used to show vowel length. The first system was developed in the early 1820s by Professor Samuel Lee of Cambridge University, and further developed by William Williams (see Williams, 1961) and other missionaries in the early 1840s. This orthography does not indicate the length of vowels, other than double vowels which are used to communicate distinct morphemes, as shown in example 1 below, and in instances of reduplication, as illustrated in example 2. This was the system used in many of the original traditional texts included in this study. In this system, words which are distinguished in pronunciation by a short-vowel,

long-vowel difference are written identically, e.g. *ki* (prep 'to') (pronounced /ki/) and *ki* 'say' (pronounced /ki:/).

- (1) whakaaro
whaka·aro
CAUS·focus
'to think'
- (2) haereere
haere·ere
go·REDUP
'to travel about'

The second orthography, where long vowels are indicated by double vowels, was strongly promoted by Bruce Biggs around the middle of the 20th century, and continues to be used by some, especially the Tainui people of the North Island in Aotearoa New Zealand, so what would have been written as *ki* 'to say' according to the first orthography is instead written *kii* according to this one. The third orthography has been promoted by the Māori Language Commission, Te Taura Whiri i te Reo Māori, since the 1990s, and uses macrons to indicate vowel length, except in word forms which contain distinct morphemes, in which case the double vowel is written, so *kii* 'to say' instead is written as *kī*, but *whakaaro* 'to think' is not written as *whakāro*. The vast majority of the modern corpus follows the orthographical conventions recommended by Te Taura Whiri i te Reo Māori (2012). The traditional corpus is largely written without the marking of vowel length, and the citations from my corpora retain the conventions of the originals. Similarly, where glossed examples are drawn from other descriptions in this thesis, the original glosses are retained, and they do not necessarily follow my own glossing conventions. Note that throughout this thesis, when corpus-sourced examples are used in text, they are labelled according to their corpus of origin, (TC) for the traditional corpus data, and (MC) for the modern corpus. Where neither a corpus tag nor a reference appears for an example, the examples are my own.

The data collection process therefore resulted in the two corpora of semi-formal/formal material written originally in Māori for adult audiences by competent, confident adult Māori language speakers – the traditional corpus

of 102,226 running words, and the modern corpus of 102,234 running words.

THE ANALYSIS PROCESS

The use of the two corpora is trialled through three separate case studies on three different types of language feature. The first case study analyses the use of a grammatical particle, namely the preposition *mō*. The second case study investigates the use of two semantically-linked lexemes which have a loosely 'grammatical' function, in that they can indicate a particular type of modality, *taea* and *āhei*. The final case study focusses on a grammatical construction, and examines a subset of relative clauses.

In each case study, all examples of the particular language feature were extracted from the two corpora. The available linguistic descriptions of the language feature were reviewed to provide the basis for the categorisation of the examples. All examples of each language feature were then extracted from each corpus, typically using WordSmith Tool lexical analysis software to produce a concordance run, although a manual search was required for the case study on relative clauses, which is explained further in Chapter 5.

The individual examples were then analysed according to the relevant categories. In order to answer the question of whether the language form is more or less frequently used in the modern period, the frequency of each category in the two corpora was compared. The comparative frequencies for each relevant category in the analysis were also compared, to provide evidence of an increase or decrease in the functions for which the language form is used. A chi-square test of independence with Yate's continuity correction used as appropriate (for 2x2 contingency tables), to determine whether there was a statistically significant difference between observed and expected values, that is, whether the corpus the examples were taken from was likely to have an impact on the number of occurrences of a particular language feature. P-values < 0.05 were considered significant (Clark & Randal, 2004).

This allowed me to document some ways in which Māori has changed during

the period between the two corpora, thus fulfilling the main objective of this research. In some instances, the data contained examples which did not fit within the analytical categories defined for the case study, and this provided further evidence of language change. Any changes detected were then examined to determine the likely cause/s for the change, and to assess whether the change in Māori is the result of language contact with English, or whether the change can be otherwise accounted for by language-internal processes or other factors. The first of the case studies is detailed in the following chapter.

3. CASE STUDY ONE: *MŌ*

The first case study of this thesis investigates the use of the grammatical particle *mō*. *Mō* is a preposition in Māori, and has a number of meanings, some of which include “for/on account of/ for the benefit of/ in preparation for” as mentioned in the Williams dictionary (1971, p. 203).

When *mō* combines with singular personal pronouns there is a special cliticised form of the pronoun, hence *mōku* (*mō* + first person singular) *mōu* (*mō* + 2nd person singular) and *mōna* (*mō* + third person singular). These cliticised forms and dialectal synonyms (for example *mōhou*, synonymous with *mōu*) have been included in the search for examples of the usage of *mō* within the two corpora of this study.

DATA SELECTION AND CATEGORISATION

While many of the various grammars of Māori include examples of *mō* and brief explanations of its usages, they usually only describe its use for future possessives, and for marking the topic of discourse. Bauer (1997) and Harlow (2001) however discuss *mō* and its various syntactic roles in the most detail. Their classifications of the usages of *mō*, and the explanations found in *He Pātaka Kupu* (Te Taura Whiri i te Reo Māori, 2008) and Williams (1971) dictionaries form the basis for the ten categories that were used to classify the examples extracted from the two corpora for this study. These categories will be explained in detail here.

Category 1: Future possessor

The first usage of *mō* identified as a classificational category for this study is its use as a future possessive preposition, where the intended owner of the specified possession is the complement. It may be translated as ‘for’, as in ‘for us’ in example 3 below.

- (3) Mō tātou tēnei taonga.
 PREP 1PLINCL this treasure
 ‘This treasure is for us.’ (MC)

This preposition *mō* is one part of a four-part system of possessive prepositions in Māori, the others being *mā*, *nō*, and *nā*. Within this system, the *m-* or *n-* of the preposition marks the irrealis/ realis distinction, and the *-ō* or *-ā* denotes the class of the possession, as per the a/o classification system in Māori (Bauer, 1997). This preposition *mō* therefore indicates a future possessive relationship between the ‘o’ class possession and the intended owner, specified in the complement.

This type of *mō* is found both predicatively and non-predicatively. Example 3 above shows this preposition used predicatively in a non-verbal possessive predicate for specific possession. In this example, *tātou* is the future possessor complement in the predicate, and *tēnei taonga* is the sentence Subject. Example 4 shows *mō* used in an adverbial, used to mark the intended recipient of the gift.

- (4) I hoko-na tētahi perehana mō-na.
 TAM buy-PASS a present PREP-3SG
 ‘A present was bought for him.’ (MC)

Category 2: Benefactive

The second usage of *mō* is to mark phrases with what Harlow calls “a benefactive meaning, that is, which name a person or group for whose benefit something occurs or is intended” (2001, p. 81). This usage of *mō* may often be translated as ‘for’ or ‘on behalf of’, for example *mō Mere*, ‘on Mere’s behalf’ or ‘for Mere’. This usage of *mō* is very similar to that described in Category 1, and in fact is not expressly distinguished by Harlow from the future possessive use. However, I have distinguished between the two in this study, as it is clear that in some instances this usage of *mō* does not denote futurity, nor the typical a/o class distinction specified in the description of the previous category. I have found no examples or discussion of this type of *mō* in predicative settings; all of the examples found are in adverbial phrases.

Example 5 illustrates the use of *mō* in an adverbial, to mark the beneficiary of an action in the predicate. It is an example of this use of *mō* where ‘on behalf of’ might be used as an appropriate translation – ‘She stood as spokesperson on behalf of the class.’

- (5) I tū ia hei māngai mō te karaihe.
 TAM stand 3SG PREP mouthpiece PREP theSG class
 ‘She stood as spokesperson for the class.’ (MC)

Category 3: Topic of discourse

The third usage of *mō* is to mark the topic of discourse. It can be translated as ‘about’ or ‘concerning’ (Harlow, 2001, p. 170). The complement is the topic of discourse, that is, the topic of the essay in example 6, and the topic of the conversation in example 7 below.

- (6) He tuingaroa tēnei mō Parihaka.
 DET essay this PREP Parihaka
 ‘This is an essay about Parihaka.’ (MC)

- (7) I kōrero ngā koroua mō ngā tikanga o te
 TAM talk the-PL elder PREP the-PL custom PREP theSG
 marae.
 marae
 ‘The elders talked about the customs of the marae.’ (MC)

This usage of *mō* is not possessive, neither does it mark a/o classification or futurity, as illustrated in both examples 6 and 7. Only modifier and adverbial examples of this usage were found in the examples taken from the corpora, as demonstrated in 6 and 7, but there is evidence of this usage of *mō* in predicative settings, such as the example given in Bauer’s (1997) discussion of subject-matter phrases:

- (8) Mō Puhihuia rāpea tō kupu?
 about Puhihuia indeed POSS-SG word
 ‘Are your words actually about Puhihuia?’ (p. 207)

Category 4: Reason

The fourth category of *mō* is its use in expressions of reason, as described by Harlow (2001). In these settings *mō* can be translated as ‘for/ due to’, but once

again, does not denote futurity, or a/o classification. This usage of *mō* was only found in adverbials. The complement is the reason or cause for the action in the predicate.

Example 9 below contains an expression commonly used in Māori to express remorse, or to say ‘sorry’. The preposition *mō* appears in sentence initial position, and might appear to be predicative. However I would argue that this expression is in fact an adverbial, where the predicate of the sentence is inferred and therefore omitted. An example of the projected full sentence is given in example 9a below.

(9) Mō taku hē.
 PREP POSS·1SG error
 ‘I’m sorry (lit. for my mistake).’

(9a) Ka whakapāha au ki a koe mō taku hē.
 TAM apologise 1SG to PERS 2SG PREP POSS·1SG error
 ‘I apologise to you for my mistake.’

A comparable contraction process in modern Māori is the contracted form of saying goodbye illustrated in examples 10 and 10a below, where example 10 shows the contraction now understood to mean ‘goodbye’ and 10a is the full inferred sentence. It is interesting to note that the omitted portion of the full sentence is syntactically different in each example; the contraction process which produces example 10 involves the omission of the Subject and Direct Object of the full sentence, whereas the opposite happens in the full sentence proposed in example 9a - the predicate and goal adverbial are omitted, leaving only the adverbial of reason in the contraction.

(10) Ka kite.
 TAM see
 ‘See you (lit. Will see!)’

(10a) Ka kite anō au i a koe.
 TAM see again 1SG PREP PERS 2SG
 ‘I will see you again.’

With this usage of *mō* there was a variety of different constructions represented in the complement. Some of the complements, such as those in example 9 above and example 11 below, were noun phrases containing stem

nominalisations. The performer of the nominalisation in these types of complements was the t-class possessor of the nominalisation, namely *rātou* in example 11.

- (11) I riri-a rātou mō tō rātou āwhina kore i
 TAM scold·PASS 3PL PREP POSS·SG3PL help NEG PREP
 a ia.
 PERS 3SG
 ‘They were scolded for not helping him.’ (TC)

- (12) Ka nui ngā mihi ki a koe mō·u i tae
 TAM many thePL thank to PERS 2SG PREP·2SG TAM arrive
 mai ki te tautoko i tēnei mahi.
 DIR to DET support PREP this work
 ‘Thank you very much indeed for coming to support this activity.’
 (Harlow, 2001, p. 248)

Example 12 above demonstrates *mō* with a noun clause complement. The complement to *mō* has the underlying structure *i tae mai koe ki te tautoko i tēnei mahi*, and the Subject of this clause *koe* is compulsorily raised to follow ‘*mō*’; when it follows ‘*mō*’, the independent form ‘*koe*’ is replaced by the required clitic, ‘-u’ (Bauer, 1997, pp.617-8)

Category 5: Position

The fifth category of *mō* is its use as complement to a location. The complement in this type of usage is the item which belongs in the location, such as the towels in example 13 below. It can be translated as ‘for’, but does not indicate futurity. It is not clear whether this usage of *mō* carries the a/o possessive category, as locations are typically classified under the o category. This particular usage of *mō* is not explicitly mentioned in any of the grammars of Māori; however it is identified as distinct from the other usages of *mō* in *He Pātaka Kupu* (2008), and therefore treated separately in this study.

- (13) Kei whea te wāhi mō ngā taora?
 PREP where theSG place PREP the·PL towel
 ‘Where is the place for the towels?’

Category 6: Transaction

Mō is also used in expressions of transaction, where the complement is either the price of the item mentioned in the preceding noun phrase, or the item that may be purchased for the price mentioned in the preceding noun phrase (Bauer, 1997). In these settings it could be translated as ‘for/of’. Example 14 shows the *mō* prepositional phrase as a modifier to the noun *utu* in the Subject slot of the sentence. Example 15 shows *mō* in an adverbial.

(14) Kōtahi mano tāra te utu mō tana rorohiko.
one thousand dollar theSG cost PREP POSS·3SG computer
‘The cost of his computer was one thousand dollars.’ (MC)

(15) I hoko-na tō māua whare tuatahi mō te
TAM buy·PASS POSS·SG 1DLINCL house first PREP theSG
rua rau mano tāra.
two hundred thousand dollar
‘Our first house was sold for two hundred thousand dollars.’ (MC)

Bauer identifies this usage of *mō* as a feature of modern Māori (1997) in the sense that it emerged with the arrival of Europeans and their forms of currency – note that this includes the period in which my traditional corpus is set. Indeed there are many examples in the traditional corpus involving these types of *mō* phrases, the vast majority of which were extracted from classified advertisements or sale notices in Māori newspapers and the like. In considering alternative expressions for cost or price in Māori, with some hesitation my consultants accepted example 14a below, although it was noted by some to be convoluted. Some agreed that the suggested use of *ki* as illustrated in 14b below might be sufficient, with a caveat that its use would be very uncommon, and somewhat different semantically from the original sentence in example 14. The pragmatic difference involved the emphasis upon the action that had taken place, namely that his computer was paid for, and the price was additional information, not integral to the sentence.

(14a) He kōtahi mano tāra te utu hei hoko i
DET one thousand dollar theSG cost PREP buy PREP
tana rorohiko.
POSS·3SG computer
‘The cost of purchasing his computer is one thousand dollars.’

- (14b) Ka utu-a tana rorohiko ki te kōtahi mano
 TAM pay-PASS POSS-3SG computer to theSG one thousand
 tāra.
 dollar
 'His computer was paid for with one thousand dollars.'

Category 7: Duration

Another usage of *mō* found in the extracted examples is for expressions of duration. This usage denotes neither futurity nor possessivity, and its complement is a specified period of time. This adverbial usage of *mō* could also be translated as 'for', as illustrated in the examples below. Example 16 illustrates the use of this type of *mō* in sentence initial position, and example 17 shows the *mō* phrase in sentence final position.

- (16) Mō te tau tuatahi i noho au ki te whare
 PREP theSG year first TAM stay 1SG to theSG house
 o Weir.
 POSS Weir
 'For the first year I lived at Weir House.' (MC)

- (17) I kura-ina ia ki reira mō ngā tau e toru.
 TAM school-PASS 3SG to there PREP the-PL year NUM three
 'He was educated there for three years.' (MC)

Other than instances of the idiom *mō ake tonu atu* 'forever', this usage of *mō* was only found in examples taken from the modern corpus. Given the size of my corpora, this may have been coincidental. However, in her discussion of temporal expressions in Māori, Bauer also queries whether this is a modern construction, and contrasts it with an alternative construction below:

- (18) E rua ngā wiki i noho ai ia ki Hāmoa.
 NUM two thePL week TAM stay PART 3SG to Samoa
 'He stayed in Samoa for two weeks.' (Bauer, 1997, p. 187)

Category 8: Purpose

Mō is also used in expressions of purpose. There are two distinct types of expressions of purpose, one where the expression of purpose is an adverbial phrase modifying the predicate, as in example 19. Bauer (1997) describes this usage of *mō* as modern, or at least more frequently used in modern Māori.

The complement is the purpose, in the form of a stem nominalisation with a t-class determiner. The performer of the action is expressed as the possessor of the stem nominalisation. The other type of purpose expression is a phrasal modifier to the Subject of the matrix clause, such as in examples 20 and 21.

- (19) Kua rite ahau mō taku haere ki tāwāhi.
 TAM ready 1SG PREP POSS·1SG go to overseas
 'I am prepared for (my) travelling overseas.' (MC)
- (20) He harakeke pai tērā mō te mahi whāriki.
 DET flax good that PREP DET make mat
 'That is good flax for making mats.' (MC)
- (21) Kārekau he toki mō te tapahi i te wahie...
 NEG DET axe PREP DET chop PREP theSG firewood
 'There was no axe to chop the firewood with.' (MC)

Category 9: Future location

Where *mō* is used as a marker of future location, the complement is a location, typically in time, as shown in examples 22 and 23. Locatives in Māori may be used to communicate location in time or space, but there was no mention of the use of *mō* to mark future spatial location in the various grammars, and only a single occurrence which might fit this description was found in the modern corpus, which will be discussed further later in this chapter (see new applications). *Mō* in these expressions of time shows futurity, but not possessivity, and there is no *a/o* class distinction. This type of *mō* adverbial occurred both in sentence initial position, as in example 22, and in sentence final position, as in example 23 below. Although this usage is defined in *He Pātaka Kupu* (2008), it is not widely discussed in the various grammars, except for Harlow's (2001) brief descriptions of the use of *mō* to mark future location in time.

- (22) Mō āpōpō rāua wehe atu ai.
 PREP tomorrow 3DL leave DIR PART
 'They will leave tomorrow.'
- (23) Waiho·tia tēnei mahi mō āpōpō.
 leave-PASS this work PREP tomorrow
 'Leave this task for tomorrow.'

Category 10: Performer of intransitive nominalisation

Another usage of *mō* which appeared in the classification of the examples extracted from my corpora was to mark the performer of an intransitive nominalisation, as illustrated in example 24 below. The complement is the performer of the intransitive action, and this follows a purpose clause with the preposition *hei* and the intransitive nominalisation, such as *wehenga* in example 24.

- (24) I whakatau·ria te wā hei wehe·nga mō te
TAM decide·PASS theSG day PREP depart·NMLZ PREP theSG
kapa haka ki te whakataetae.
haka-troupe to theSG competition
'The time for the haka troupe to depart for the competition was decided.' (MC)

Category 11: Emphatic refusal

The final category of *mō* was its use in expressions of refusal, an emphatic statement that the mentioned Subject will definitely not action the verb in the verbal clause (Harlow, 2001). This appears to be a formulaic expression, where the complement to *mō* is the action that is being refused, as there is no affirmative version of the expression. This usage of *mō* is found only after the negator *Kāore*, which is interesting as *mō* predicates are typically negated using the negator *ehara*. This may be part of the reason that Harlow classes this usage of *mō* as a verbal particle (2001), although I believe *mō* functions as a preposition in these settings, and is used to mark a noun phrase comprised of a stem nominalisation of either an action intransitive or canonical transitive verb, preceded by the determiner *te*.

Another native speaker consultant described this usage of *mō* as an expression of refusal, implying that it was a more emphatic version of the sentence given in example 25b below (T. R. Higgins, personal communication, 18 June, 2009).

- (25) Kāore au mō te haere ki te hui.
NEG 1SG PREP DET go to theSG meeting
'I will definitely not be/ have no interest in going to the meeting.'

(25b) Kāore au e haere ki te hui.
 NEG 1SG TAM go to theSG meeting
 'I will not go to the meeting.'

RESULTS

Overall frequency

The traditional corpus contained a total of 952 examples of the preposition *mō*, including the special cliticised pronouns. The modern corpus contained 731 occurrences in total. The respective frequencies of each form of *mō* are detailed in Table 1 below.

Table 1: Overall frequencies of different forms of *mō*

FORM	TRADITIONAL	MODERN
mo/mō	914	711
mōku	22	2
mōhou	1	0
mōu	7	3
mōna	9	15
TOTAL	952	731

Note that both *mo* and *mō* appear in a single category; because of the different orthographic systems of the two periods it was necessary to search for both forms. It was interesting to note that the examples in the modern corpus were not consistent in spelling; all modern texts in which examples of *mō* were found did appear to subscribe to the use of macrons, as recommended in the orthographic conventions of Te Taura Whiri i te Reo Māori (2012), but 10% of modern examples of *mō* were spelled without a macron. In a number of instances the spelling of *mō* was inconsistent within a single piece of text, or even within a single sentence, as shown in example 26 below.

- (26) mo ngā mahi i mahia e rātou, mo te
 for thePL work TAM work.PASS PREP 3PL for DET
 māia, mo te kaha, mo te aroha, mō
 bravery for DET strength for DET love for
 te titikaha me te rangatiratanga
 DET determination and DET nobility
 ‘for the work they did, for (their) bravery, for (their) strength, for (their)
 love, for (their) determination and (their) nobility.’ (MC)

Given the technological limitations and orthographic conventions of the traditional period it is not possible to conclusively attribute the 10% of the examples in the modern corpus that omitted the macron to a change in the language overall, although the observation does document issues some speakers are having in the modern period in determining vowel length. The decreasing length of the pronunciation of long vowels, and the lengthening of the pronunciation of short vowels is documented by Harlow et al. (2009).

IS THE USAGE OF *MŌ* DIFFERENT IN MODERN MĀORI?

In looking for evidence of change in the usage of *mō* between traditional and modern Māori, there are perhaps three main types of change which might be identified by this type of study: the emergence of new applications of *mō* in modern Māori, the disappearance of a traditional application of the language form in modern Māori, or the increase or decrease in the frequency of a particular usage of the language form between traditional and modern Māori.

To identify either of the first two types of change, one must first identify which usages of *mō* are traditional, and which are modern. When consulting the grammars, as mentioned in the explanation of the various usages of *mō* in Māori above, a few usages in particular are identified by Bauer (1997) in her grammar as features of modern Māori. As previously mentioned, Bauer (1997) identifies the use of *mō* in duration clauses and in expressions of transaction as modern. She also mentions an increase in the frequency of the usage of *mō* in adverbials of purpose in modern Māori. It is important to state that Bauer’s use of the term ‘modern’ may well refer to a different time period than the period identified in this study as ‘modern’, or in fact not to a specific time period as such, but instead to represent something other than the Māori

language as it was prior to European contact. Harlow (2001) does not discuss change in the usages of *mō* detailed above, but he does express his opinion that the commonly heard use of *mō* as the preposition to the complement of the verb *tatari* ‘to wait’ is “probably an Anglicism ... following the English usage ‘wait **for**’” (p. 164).

Another means of identifying which usages of *mō* are traditional is to look for those which are readily found in the traditional corpus. In the traditional corpus, examples abound for all of the usages described above, except for expressions of duration, purpose, and position, no examples of which were found in the traditional corpus. This suggests that perhaps the duration, purpose, and position uses of *mō* are modern applications of *mō* developed or accepted by the language community post 1900, or since the traditional corpus period. However, this absence from the traditional corpus may also potentially be explained if it happens that, in the texts collected for the traditional corpus, there were no semantic contexts where expressions like these would have been required. This is particularly likely for lower frequency usages of *mō*.

PERSONAL PRONOUNS AFTER *MŌ*

There were a number of occurrences of *mō* taken from the modern corpus which were classifiable according to the categories given above, but where the example indicated an emerging shift in the syntactic rules of traditional Māori. One such example involved singular personal pronouns appearing after *mō*, as in example 27 below. This example clearly shows *mō* used to mark the topic of discourse, but the personal pronoun *koe* is used, not the cliticised *mōu* which would usually be expected to indicate the second-person. Example 28 exemplifies a similar situation, but instead features the first person *ahau* instead of the cliticised form *mōku*.

- (27) I te māharahara ahau mō koe!
 TAM worried 1SG PREP 2SG
 ‘I was worried about you!’ (MC)

- (28) Nā·na ngā tīkiti i hoko mō ahau.
PREP·3SG thePL ticket TAM buy PREP 1SG
'He was the one who bought the tickets for me.' (MC)

Let me first qualify this by stating that the vast majority of co-occurrences of *mō* and a first-, second-, or third-person complement within the modern corpus featured the cliticised form; there is no evidence of a dramatic shift in the modern corpus away from the cliticised form of the pronoun. However, this was not a one-off anomaly; in the modern corpus there were five examples, from four different authors, which featured the use of *mō* with a singular personal pronoun. While these numbers may appear small, in consideration of the small corpus from which they are taken, the number is significant enough to merit discussion here. These occurrences may be too small in number to be statistically significant, but they are of considerable practical significance, especially given that these corpora are restricted to contain only examples of deliberate, written language from speakers who are qualified to be language teachers, and take on other language model roles in the Māori language community.

Why the non-cliticised forms were used by the authors in these instances is unclear, but there are a number of potential explanations. Perhaps the usage of the non-cliticised form was simply an oversight, a grammatical error that would have been corrected if noticed during the editing process, but which was somehow overlooked in these instances.

It may be that the use of the separate pronoun indicates an example of code-copying in Māori as described by Johanson (2002), which involves “the insertion of elements copied from one code within the context of another code” (p. 288). In these examples the speaker’s use of *mō* with an independent personal pronoun mirrors the English syntactic norm, where two free morphemes are required to represent ‘for you’ or ‘about you’, *mō* and *koe* respectively being perhaps the most common renderings of those free morphemes in Māori. As previously mentioned, the Māori syntactic norm is for three bound morphemes to form a single word to communicate the same information; *m-* for the irrealis distinction, *-ō* for the benefactive or ‘o’ class

possessive relationship between the possession and possessor, and *-ū* to indicate the second-person intended owner or beneficiary.

Example 29 is noteworthy in that the non-cliticised form is used with the second-person pronoun *koe* in the predicate, and yet the cliticised form involving *mō* and the third-person *-nā* is used within the same sentence. Perhaps it is significant that the non-cliticised form was used within the sentence constituent that was being emphasised, not only through the specific possessive predicate, but also with the postposed contrastive intensifier *kē*.

- (29) Mō koe kē te perehana nā, ehara mō·na.
PREP 2SG MAN theSG present DEI NEG PREP·3SG
'That present is in fact for you, not for him.' (MC)

It may be that the most significant factor is the frequency of usage of the cliticised form *mōu* compared with the other forms *mōku* and *mōna*. There is a list of the one thousand most frequent words in Māori, available as a teaching resource on the Ministry of Education website, *Te Kete Ipurangi*. This list was based on two Māori corpora: the Corpus of Māori Texts for Children (MTC), compiled by Huia Publishers, and the Māori Broadcast Corpus (MBC), a one million word representative corpus of broadcast material, recorded off-air in the mid-1990s, compiled by Mary Boyce. It is interesting that of the three cliticised *mō* forms, *mōna* is most frequent, then *mōku*, but *mōu* does not feature at all in the list (NZ Ministry of Education, 2010). This may explain why four of the five examples used the second-person pronoun *koe*, and only one contained the first-person pronoun *ahau*.

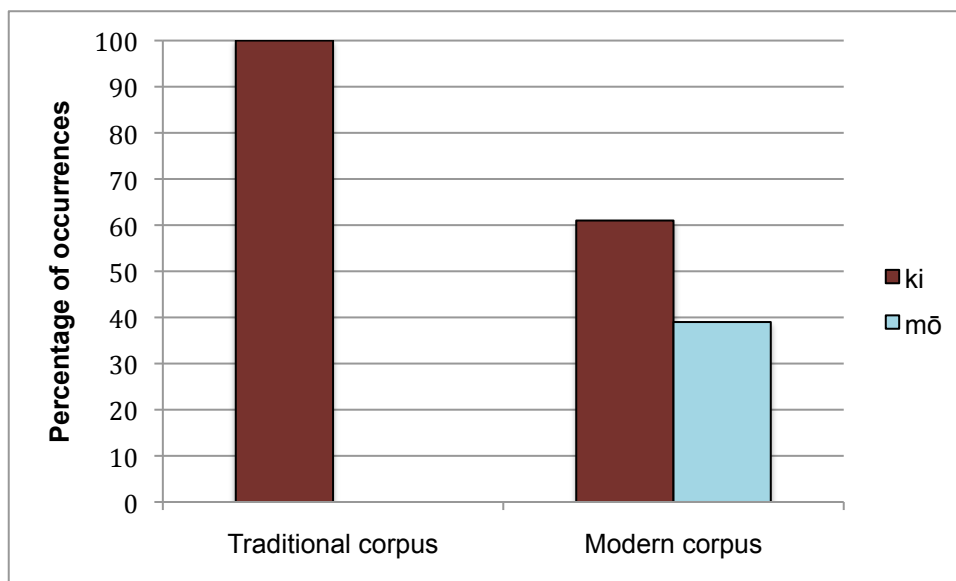
EMERGENCE OF NEW APPLICATIONS OF MŌ

- (30) I te tatari rātou mō·ku!
TAM wait 3PL PREP·1SG
'They were waiting for me!' (MC)

Example 30 above shows the use of *mō* in an adverbial marking the person that was being waited for by *rātou*, the Subject of the clause. There were twelve examples of this usage of *mō* in the modern corpus, but none in the traditional corpus. This may indicate that perhaps this is a new semantic setting for *mō* in modern Māori, marking the complement to *tatari* 'wait', which

would traditionally be marked by the preposition *ki*. As previously mentioned, Harlow puts the use of *mō* as the complement to *tatari* ‘wait’ instead of the traditional *ki* down to a copying of the English phrase ‘wait **for**’ (2001, p. 64). In a later publication Harlow speaks of the increase in frequency and acceptability of this new usage of *mō* with *tatari* as one of the “instances of the influence of English [that] have already almost become ‘correct’ Māori” (2005, p. 137-8). As a rudimentary test, I extracted from the two corpora all of the occurrences of *tatari*, and looked at which preposition was used – see Figure 2 below for the relative frequencies of each of the prepositions for each corpus. It would appear from this graph that the evidence in my study supports Harlow’s assertion that this usage of *mō* has become more frequently used by the language community overall in modern Māori.

Figure 2: Preposition used for complement to *tatari*



Example 31 below shows the use of *mō* as a future locative preposition for the locative noun phrase *waho i te whare* ‘outside the house’. Finding *mō* in front of a locative noun is not unusual; a number of examples from both the traditional and modern corpora feature *mō* with the locative nouns *mua* ‘in front/ before’ and *muri* ‘behind/ after’. What is different in this case is that *waho* can only refer to a spatial location, whereas all of the *mō* + *mua* or *muri* occurrences were temporal, not spatial references.

- (31) Mō waho i te whare tērā kēmu.
 PREP outside PREP theSG house that game
 'That game will be held outside the house.' (MC)

Semantically it seems that this example is more similar to the expressions of position in examples 13 – however in those instances the complement to *mō* was the item which belonged in the position stated in the predicate, whereas the complement to *mō* in example 31 is the designated place where it is permitted for the game to be played. Perhaps, as with the expressions of transaction where *mō* could be used to mark either the price or the item for purchase, *mō* in these semantic contexts may be used to mark either the location or the object or activity which belongs in the stated location.

There were also examples of *mō* taken from the modern corpus where the usage did not fit any of the classification categories, and which were not alluded to in any grammars of Māori, nor in the various dictionaries. Example 32 may well be the most unusual new-application of *mō*; it shows *mō taku hē*, the formulaic expression for 'I'm sorry' illustrated in example 9. However, it is in the syntactic setting of the phrase that this usage is unusual, as it appears that the phrase unit *mō taku hē* has been lexicalised to form a compound verb, meaning 'to say sorry' or 'to apologise'. My consultants rejected this usage as non-Māori, and suggested the verb *whakapāha* 'to apologise' as the most acceptable direct translation, as shown in example 32a.

- (32) I haere ahau ki te mō taku hē atu
 TAM go 1SG to DET PREP POSS·1SG error DIR
 'I went to apologise.' (MC)

- (32a) I haere ahau ki te whakapāha atu.
 TAM go 1SG to DET apologise DIR
 'I went to apologise to him.'

This case appeared to indicate a simple lexical gap, as a simple verb for verb substitution made an acceptable alternative for conveying the intended meaning; the syntax of other alternative expressions offered by my native speaker consultants has often been markedly different from the original sentence taken from the modern corpus. This may indicate a simple lexical gap in the vocabulary of the author of this particular sentence.

It is interesting that the author clearly believed that this usage of *mō taku hē* as a compound verb would be acceptable, or at least comprehensible to the intended audience. There are many examples of the verbalisation of nouns and the nominalisation of verbs in Māori, for example *waiata* meaning either ‘to sing’ or ‘song’, and the use of the loan word *kura*, ‘school’, as a verb, meaning ‘to be schooled’ or ‘to be educated’. There are also countless compound verbs in Māori created by the use of state intransitive verbs as adverbs, as in the compound verb *ua pūkōhukohu*, ‘to drizzle’, formed from *ua*, ‘to rain’ and *pūkōhukohu*, ‘to be spongy/ misty’. Other compound verbs are formed by the process of object incorporation, for example *hoko kai*, ‘grocery shopping’, formed by the incorporation of the Direct Object *kai* ‘food’ into the canonical transitive verb *hoko* ‘to buy’. I could find no evidence however of prepositional phrases or full idiomatic expressions from within Māori used as compound verbs.

The nearest comparison I could find for this usage of a phrasal unit as a verb is the Māori verb *hariru* ‘to greet (by shaking hands or hongi)’ which is the Māori adaptation of the English greeting ‘How do you do?’. What was a phrasal unit in English has been adapted phoneme by phoneme, lexicalised and naturalised to form a verb in Māori. This process is described by Picoche & Marchello-Nizia (1991) as the third and final stage of borrowing, where the target language (in this case Māori) adopts, then adapts a loan word, then naturalises the use of the loan word within the target language, independent of its original use in the donor language.

Another example of a new application of *mō* is exemplified below in example 33, where the preposition *mō* is found hanging in sentence final position, which is ungrammatical in Māori. Sentences 33a and 33b below were suggested as acceptable options to communicate the same sentiment, neither of which feature *mō*.

- (33) Kei wareware i a koe he aha koe i haere
 PREP forgotten PREP PERS 2SG DET what 2SG TAM go
 mai ki kōnei mō
 DIR to here PREP
 ‘(Take care) lest you forget what you came here for’ (MC)

- (33a) Kei wareware I a koe te take i haere mai
 PREP forgotten PREP PERS 2SG theSG reason TAM go DIR
 rā koe ki kōnei.
 DEI 2SG to here
 '(Take care) lest you forget the reason you came here.'
- (33b) Kei wareware i a koe nā te aha koe
 PREP forgotten PREP PERS 2SG AG DET what 2SG
 i haere mai ai ki kōnei.
 TAM go DIR PART to here
 '(Take care) lest you forget what caused you to come here.'

Mō was also found as a hanging preposition in sentence final position in the expression of the intended use/ purpose of the Subject of the matrix clause in example 34 below. This structure was rejected by my consultants, although there is anecdotal evidence of its increased usage by younger speakers. 34a below illustrates a possible correction for example 34, note that the correction offered once again does not include *mō*. It is also interesting to see the syntactic mirroring between the Māori sentence in example 34 and its English translation, for which a hanging preposition like this is grammatical.

- (34) Kīhai au i mōhio he aha taua rākau mō!
 NEG 1SG TAM know DET what that stick PREP
 'I didn't know what that stick was for!' (MC)
- (34a) Kīhai au i mōhio hei aha taua rākau!
 NEG 1SG TAM know PREP what that stick
 'I didn't know what that stick was for!'

This seems to be a common thread across many of the usages of *mō* for which a change of some kind is noticed in modern Māori – in a number of instances the alternative, more traditional Māori expression instead employs the preposition *hei*. This is true for not only example 34 above, but also for the future time expressions, and a number of purpose constructions as well.

Considering future time expressions first, both Harlow (2001) and Bauer (1997) mention that *mō* is not commonly used in temporal expressions. Given that the descriptions of this usage of *mō* consistently mention its 'occasional' use, it seems sensible to investigate what the other options are for marking future time adverbials. Bauer notes the use of *ā* to mark future temporal location, and comments that *hei* is normally used to mark a future spatial

location, and occasionally a future temporal location (1997). Harlow identifies *ā* as the most common preposition for this type of expression, and mentions that *hei* is also common (2001).

Thus *ā* and *hei* are alternative prepositions that might be used in adverbials of future temporal location. It appears that only *hei* may be used to mark future spatial locations; *mō* and *ā* would not be acceptable. It is clear that the three prepositions are not always interchangeable in future temporal location adverbials; compound temporal locative nouns which begin with *ā* such as *āpōpō* ‘tomorrow’, *ātahirā* ‘the day after tomorrow’ may co-occur with the prepositions *hei* and *mō*, or alternatively without a separate preposition, as shown in the examples below.

(35) *Āpōpō tāua e hui ai.*
 tomorrow 1DLINCL TAM meet PART
 ‘We will meet tomorrow.’

(35a) *Hei āpōpō tāua e hui ai.*

(35b) *Mō āpōpō tāua e hui ai.*

(35c) ***Ā āpōpō tāua e hui ai.*

It is not clear exactly if and how *mō*, *hei* and *ā* differ in their interpretation. Williams’ (1971) dictionary states the purpose of *mō* in this capacity as “fixing at a future date” (p. 203), which may suggest that the use of *mō* might imply more emphasis, a stronger belief, or commitment to ensuring that the stated activity would definitely occur at the stated time, although there are no grammatical descriptions available to confirm or refute this possibility.

Another usage of *mō* for which *hei* is an alternative is for expressions of purpose. Example 32 has already been discussed, and while that example was easily identified as unusual in Māori by the hanging preposition, perhaps there is more subtle change that may be identified through the analysis of purpose expressions in Māori, for which there is already anecdotal evidence.

The use of *mō* in expressions of purpose is rarely mentioned in the grammars of Māori, Bauer (1997) being perhaps the only exception. She mentions the use of *mō* in numerous purpose expressions, especially in modern Māori, with a caveat that while this type of expression may be common and have

precedent in some older texts, some older speakers disapproved of this usage, regarding it as English-influenced. This suggests that some alternative preposition could also be used to mark purpose expressions. Harlow (2001) discusses the use of *hei* in expressions of purpose of an object or person, as in example 36 below (note the glosses are my own).

- (36) Tīki-na he naihi hei tapahi i te parāoa.
fetch·PASS DET knife PREP chop PREP theSG bread
'Fetch a knife for cutting the bread.' (Harlow, 2001, p. 254)

In the example he provides, the *hei* clause is an adverbial of purpose. This sense of *mō* parallels that in examples 20 and 21 above, which instead feature *mō* with the stem nominalisations *tapahi* 'to chop' and *mahi* 'to make', preceded by the determiner *te*. From this it would appear that perhaps in modern Māori, *mō* is being accepted in purpose clauses where traditionally only *hei* was acceptable. *Hei* is readily used predicatively in Māori for expressions of purpose, but there were no examples of *mō* used predicatively in the same sense in my corpora. This may be a matter of there being inadequate data available due to the relative infrequency of *mō* in that context given the limited size of the two corpora used for this study.

Another possible was of translating the purpose expression in any of the given examples above is with an English 'for' phrase – 'for cutting the bread' in Harlow's example, and 'for chopping the firewood' in example 21. *Mō* is commonly translated in modern Māori dictionaries as 'for'; perhaps an assumption of one-to-one word matching between the two languages explains the modern Māori tendency to use *mō* instead of *hei* in expressions of purpose?

To conclude this chapter, my investigation into the usage of *mō* has provided observations of changes in frequency of the usage of *mō*, such as an increase in the use of *mō* in some purpose expressions, and in adverbials of duration, and a decrease in the use of *mō* for future time adverbials. I have also found some evidence of other changes, like the appearance of *mō* with the non-cliticised form of singular personal pronouns, and new applications as the preposition of the complement to *tatari*, and within a lexicalised phrase used

as a compound verb. I have suggested some explanations for the observed changes, some of which include lexical gaps in the speaker's knowledge of vocabulary, mirroring of English syntactic norms, and lexical adaptation. The implications for the language of the changes observed will be further explored in Chapter 6, after the two remaining case studies which follow this chapter.

4. CASE STUDY TWO: *TAEA* & *ĀHEI*

This chapter investigates the use of *taea* and *āhei* in Māori. Both *taea* and *āhei* have a number of meanings – three Māori dictionaries were referenced for the purpose of this study, namely Williams (1971), Te Taura Whiri i te Reo Māori's *He Pātaka Kupu* (2008) and Moorfield's online Māori dictionary (2003). The various definitions of *taea* can be summarised as follows. It is used as the passive form of *tae* 'to arrive/ to reach'; it is also commonly used, traditionally only in passive contexts, to indicate the physical ability to do something, equivalent with 'to be capable of' in English. Although its use as a loanword for 'tyre' is also described, this is a separate lexeme, and so those examples were excluded from this study, and will not be discussed further here.

Āhei is identified as a verb of ability, translated by 'to be able/ possible/ within one's power' but, unlike *taea*, *āhei* is attributed with an additional deontic function: it can also be used to mean 'to be allowed/ permitted'. Lyons (1977) describes the term deontic as referring to the logic of obligation and permission. He also differentiates deontic and epistemic modality by attributing futurity to deontic utterances. There is a homophonous lexeme *āhei* meaning 'collarbone', and the name of a particular type of snare used to catch birds (Williams, 1971, p. 2), but examples of this lexeme are irrelevant to this study, and so were excluded from the data collected.

Instruction and correction on the use of *taea* is commonplace in advanced-level Māori language-learning classes (cf Jacob, 2012), and its frequent and inappropriate use by young speakers, both L1 and L2 alike, is often the source of comment and frustration for older, or more traditional Māori speakers who are irritated by its use in semantic contexts copying those of deontic 'can' in English, and in syntactic environments that are unfamiliar to them. In many instances, *āhei* is taught as an alternative to *taea*, but in my observations of Māori language teaching and learning environments it is also often used with equal uncertainty with regards to both the range of meanings associated with

āhei, and the syntactic environments in which it may be used. Both have therefore been investigated here, and will be discussed separately.

TAEA

Of the various grammars of Māori, only Bauer (1997) and Harlow (2001) make any significant mention of *taea*. Harlow's (2001) grammar includes the most detailed description of the traditional use of *taea* as a modal verb of ability, where he describes the passive nature of its syntactic setting. He makes particular mention of the use of *e* to denote the agent, which is typical for passive constructions in Māori. He also notes that the Subject of *taea* in examples like this is the action that the agent is capable of performing – such as *pānui* 'to read', in example 37 below. Furthermore, Harlow mentions that when the Subject of *taea* is a transitive verb, it appears to have two Subjects (i.e. it is accompanied by two NPs with no preposition), illustrated in example 38 below by *te whakaoti* 'to complete' and what would have been the Direct Object of *whakaoti*, namely *ā rātou tuhingaroa* 'their essays'.

(37) Ka tae·a e ia te pānui.
 TAM attain·PASS PREP 3SG DET read
 'He can read.' (TC)

(38) I tae·a e rātou ā rātou tuhingaroa te whakaoti.
 TAM able·PASS PREP 3PL POSS-DET essay DET complete
 'They were able to complete their essays (Harlow, 2001, p. 191).'

In his 'Excursus on *taea*', Harlow (2001) also makes mention of the tendency in modern Māori for *taea* to be used in a similar syntactic setting to that typically used for *āhei*, as illustrated below in example 39. In examples like these, *taea* does not appear in a passive setting; it has an agentive Subject *ia*, and the subsequent action that can be performed is marked with the preposition *ki*, and its Direct Object marked with the preposition *i*. Harlow states that this use is widespread, but does not recommend it.

(39) Ka tae·a ia ki te āwhina i a au?
 TAM able·PASS 3SG PREP DET help PREP DET 1SG
 'Can he help me?' (Harlow, 2001, p. 191)

In her grammar of Māori, Bauer (1997) provides less detail for the description of *taea* – she notes its use to indicate physical ability, where the ability is stressed. Interestingly however, within her description Bauer classes these usages of *taea* for ability as polysemes of the passive form of ‘to arrive/ to reach’ – suggesting perhaps a sense of ‘to attain’ as the link between two senses of *taea* differentiated as separate by the dictionaries cited above.

Data collection and categories

Three categories were used to classify the examples of *taea*, each based primarily on semantic criteria:

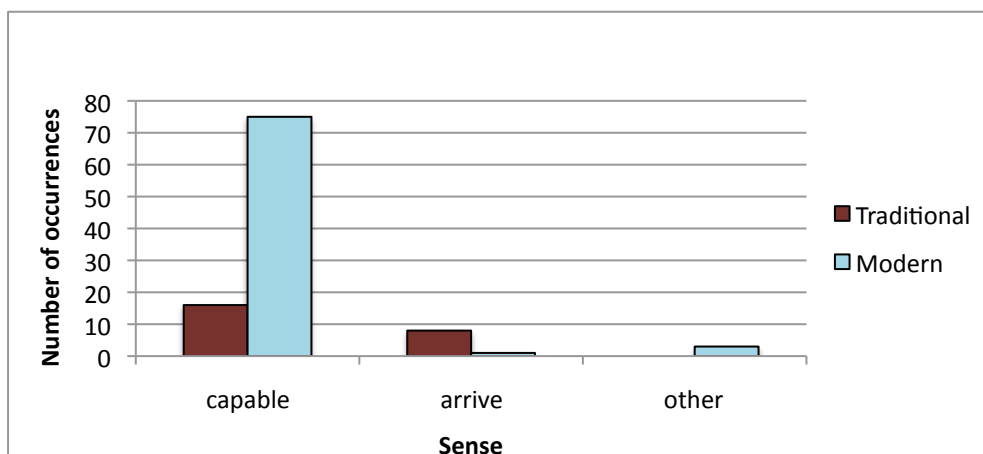
Category 1 – ‘to be capable’

Category 2 – passivised ‘to arrive/ to reach’

Category 3 – neither of the above (this is discussed further below)

Lexical analysis software Wordsmith Tool was used to produce a concordance run for each corpus, and produced 24 traditional examples, and 79 modern examples. The relative frequencies of each category are shown in Figure 3 below, which shows the general trends for the three categories across the two corpora. Note that the third category contains any examples which did not fit either ‘capable’ or ‘arrive’ categories – these examples will be discussed in detail later in this chapter.

Figure 3: Frequency of senses of *taea* in the corpora



A few things are obvious from this graph but it is important to first note that the numbers for all categories are very small, and the validity of these findings is therefore limited. There are, however two statistically significant observations which can be made: firstly, there is a dramatic increase in the number of occurrences of *taea* overall in the modern period, with a total of 79 examples in the modern corpus compared to 24 traditional examples. Secondly, the distribution of the occurrences of *taea* was noticeably different between the two corpora; only 1.3% of the modern examples used *taea* in its ‘arrive’ sense, compared with 33% of the traditional examples, and correspondingly 94.9% of the examples of *taea* in the modern corpus were used to mean ‘capable’, compared with 66.6% of the traditional examples.

A chi-square test of independence was performed to examine the relationship between traditional and modern Māori and the distribution of senses of *taea*. The relationship between these variables was considered significant, with the chi-square value = 24.24, and $p < 0.00001$.

There are three possible explanations for this observed increase in the frequency of *taea*. It may be that it is purely coincidental, and the texts included in the modern corpus just happened to include more expressions of ability. This may account for a significant part of the difference, especially given that a large proportion of the modern corpus material was written on the topic of Māori language, as mentioned in the discussion of the data collection process detailed in Chapter 2. It is possible that this would provide a context where a statement of ability is more likely to be made - resulting in examples like 40 below. A review of the modern examples of the ‘capable’ sense of *taea* showed 15 of the 75 examples were like 40, in that they involved stating a person’s ability to speak either Māori or English.

- (40) te maha o te tangata ki Aotearoa e taea
 theSG many of theSG person at New Zealand TAM able.PASS
 te kōrero i te reo Māori
 DET talk PREP the language Māori
 ‘the population in New Zealand able to speak Māori’ (MC)

Alternatively it may be that this increase in the frequency of *taea* indicates it has become the default form in modern Māori for expressing a particular ability. As a comparison, commonly-used possible alternatives for the communication of an ability include those listed in examples 41, 42 and 43 below.

- (41) He kōrero Māori ia.
 DET talk Māori 3SG
 'He is a Māori speaker/ He can speak Māori'
- (42) E kore ngā kiwi e rere.
 TAM NEG thePL kiwi TAM fly
 'Kiwis can't fly/ Kiwis don't fly' (Bauer, 1997, p. 139)
- (43) E mōhio ana ahau ki te kaukau.
 TAM know TAM 1SG to the swim
 'I can swim' (Bauer 1997:139)
- (44) E āhei ana koe ki te kōrero Māori?
 TAM- able -TAM 2SG PREP the talk Māori
 'Are you allowed to speak Māori?'

Example 44 illustrates the use of *āhei*, which is discussed further later in this chapter, where a search is conducted to determine whether there is a decrease in the use of *āhei* to correspond with the observed increase in the use of *taea*. Bauer (1997) also details a number of alternative constructions. Notice that the examples in 41 - 43 all express ability without the use of either *taea* or *āhei*. This suggests another hypothesis for the change in frequency: perhaps these alternative constructions are being used less frequently in modern Māori, with speakers choosing instead to default to *taea* or *āhei* to convey ability. In order to check this hypothesis, a pragmatic search of the two corpora for all statements which communicate ability would need to be done. Given the fact that there are no common key words to use as headwords for an automated search, this search would have to be a manual canvas of all of the texts in both corpora. These modest corpora might well be used for this purpose, as the use of small corpora for pragmatic research is attested by Vaughn and Clancy (2013), who said "the primary benefit of small corpora to the study of pragmatics is a fundamental one: they can enable the researcher to access authentic, naturally occurring language and to maintain a close connection between language and context" (p. 57). This further study would no

doubt provide some interesting results but such a search is outside of the scope of this study of syntax, and therefore will not be addressed further here.

In the examples of *taea* ‘capable’ identified in the modern corpus there was also an inconsistency in the syntactic environments in which *taea* was found. The grammars of Māori suggest that a stem nominalisation of the specified ability should form the Subject of the passivised verb *taea* (see for example Harlow, 2001, p. 191). This means that one would traditionally expect the stem nominalisation of ability to be a basic noun phrase in the Subject position of the verbal clause containing *taea*, as shown in example 45. This was true for all of the examples in the traditional corpus that expressly stated the ability within the sentence. However, five of the 75 modern corpus examples featured an agentive Subject, and the stem nominalisation of ability followed in a prepositional phrase marked by the preposition *ki*, as in example 46 below. This too is consistent with the example of modern usage given in Harlow’s (2001) ‘Excursus on *taea*’. Note this syntactic setting does not reflect a change in sense – those who consider this sentence grammatical would give an identical translation for both examples 45 and 46 – ‘He can swim’.

- | | | | | | | |
|------|-----|-----------|-----|-----|-----|-----------|
| (45) | Ka | taea | e | ia | te | kaukau |
| | TAM | able.PASS | by | 3SG | DET | swim |
| (46) | Ka | taea | ia | ki | te | kaukau |
| | TAM | able.PASS | 3SG | to | DET | swim (MC) |

A 2x2 contingency table was run to determine whether there was a statistically significant association between the corpus period and the appearance of an agentive Subject. The results were as follows: Chi-squared = 0.582. The two-tailed P value = 0.4457, meaning the association between the corpus period and the appearance of an agentive Subject is not considered to be statistically significant. Despite the fact that the number of occurrences was too small to be statistically significant, the fact that they occur only in the modern corpus at least confirms that this is unlikely to be a traditional usage.

The marked decrease in instances of *taea* ‘arrive’ in the modern corpus, in combination with the marked increase in frequency of examples of *taea* ‘capable’ might suggest a narrowing in the semantic contexts for *taea*, so that

in the modern period only its sense as ‘capable’ is represented in any significant numbers. At the same time, however, a small number of new applications of *taea* were found in the modern corpus, hinting at an expansion of the modern senses of *taea* to cover functions that were neither represented in the traditional corpus, nor mentioned in the grammars of Māori.

Two of the examples, although somewhat different in their respective syntactic environments, both featured a deontic sense in contexts involving permission, similar to deontic *can* in English, as seen in example 47, and the request for permission in example 48 below.

- (47) te mānatunatu kāore e taea e te tangata me te
the concern NEG TAM able by the person and the
whānau te haere ā noa nei ki runga i
family DET go with freely DEI to upon LOC
ngā oneone o te moana
thePL sand of the sea
‘Concern grew that members of the public and families would not be able to freely access the seaside.’ (MC)
- (48) Ka taea e koe te homai te pene?
TAM able.PASS by 2SG DET give the pen
‘Can you pass me the pen?’ (MC)

These examples did not refer to anyone’s capability of performing a particular action, but instead formed the basis of a request for the stated agent to perform an action that was clearly within their capabilities. Although the number of examples like this was too infrequent in the corpus to be of any real statistical significance, I have often observed *taea* used in this way in my Māori language classroom, by both L2 learners and L1 Māori-medium school graduates.

When asked about the meaning of *taea* in contexts identical to that in example 48, and when asked to describe its semantic difference from a request using other imperative constructions, as shown in examples 48a and 48b below, my students generally account for the difference by indicating that the use of *taea* made the request more polite, and the other imperatives, by comparison, could seem rude or demanding. It is well-known that L2 speakers often find differences in politeness norms in the L2 difficult to deal with, and will attempt

to compensate if they feel something is missing. Winifred Bauer (personal communication) tells me that Danish speakers, for instance, will say “Please” as they give you something, because they are accustomed to saying the equivalent of ‘Be so good’ in these situations in Danish. It is beyond the scope of this thesis to explore the extent to which the different norms of politeness in Māori and English play a part in the increased use of *taea* and *āhei*, but it would be a worthwhile topic to explore.

(48a) Homai te pene?
 Give theSG pen
 ‘Pass me the pen?’

(48b) Tēnā, homai te pene?
 VOC give theSG pen
 ‘Please, pass me the pen?’

Another unexpected use of *taea* found in the modern corpus is shown in example 49 below. In this one example, *taea* is used as a noun, indicating the ‘achievability’ of the specified task. The meaning of the sentence itself can be guessed, although this usage is not recorded in any of the available dictionaries of the Māori language consulted for this study, and my consultants rejected this as ungrammatical and peculiar.

(49) ka noho mai ko koutou hei whakatauirā i te taea
 TAM stay DIR PREP 2PL as demonstrate DO the attain.PASS
 o tēnei kaupapa
 of this venture
 ‘you will be the ones who remain to demonstrate the achievability of
 this venture’ (MC)

With regards to the other traditional sense of *taea*, it is also noteworthy that the use of *taea* ‘to arrive/ to reach’, while infrequent in both corpora, was only found once in the modern corpus, where it appeared in a relative clause, illustrated in example 50. Example 51 was taken from the traditional corpus. Although *taea* is translated differently in examples 50 and 51 below, they are in fact the same sense - example 50 refers to a physical location being reached, whereas example 51 involves a temporal location as its destination to be reached, namely ‘the end’.

- (50) kia tino kitea ai hoki i konei te wāhi ka
 PREP very see.PASS PART INTENS LOC here the place TAM
 taea atu.
 arrive.PASS DIR
 ‘In order to see clearly from here the place we will reach.’ (MC)
- (51) me tiaki tonu taea noatia te mutunga.
 TAM care still arrive.PASS freely.PASS the end
 ‘It should continue to be looked after until the end.’ (TC)

As previously mentioned and shown in Figure 4 there were a total of 16 traditional examples and 75 modern examples of the use of *taea* to mark ability. The traditional examples were uniform in the syntactic environments in which *taea* was found, and aligned with the descriptions in both Harlow (2001) and Bauer (1997). The modern examples were not as uniform; of the 75 examples, 70 appeared in passive environments with the Subject NP containing a stem nominalisation of ability. The other 5 examples had agentive Subjects, as shown in example 52, suggesting the inherent passivity of *taea* had been ignored by the speaker.

One of the examples of *taea* with an overt agentive Subject involved *taea* in a negative construction with *kāore*, as shown in example 52. In this example, the Subject *ia* of the embedded positive sentence *e taea ki te kōrero Māori* is raised to the Subject position of the negative *kāore* in the matrix clause. This Subject raising is typical in negative verbal sentences in Māori. Contrast this with example 53 which shows the negation for *taea* used as would be expected in traditional Māori given the inherent passivity of the verb. Interestingly, while some of my consultants dismissed the usage of *taea* in example 52 as ungrammatical, a few accepted its usage this way, and when presented with 53 deemed that also correct. This may suggest that there is at least familiarity with this new syntactic environment for *taea* in modern Māori, even amongst mature adult L1 speakers of Māori.

- (52) Kāore ia e tae·a ki te kōrero Māori.
 NEG 3SG TAM able·PASS PREP DET talk Māori
 ‘She is unable to speak Māori.’ (MC)
- (53) Kāore e tae·a e ia te kōrero Māori.
 NEG TAM able·PASS PREP 3SG DET talk Māori
 ‘She is unable to speak Māori.’

As has been discussed in the previous chapter on *mō*, due to the small size of the corpora and the relatively low frequency of the use of these language forms, the statistical significance of these findings as a representation of the usage of *taea* amongst the population of modern Māori speakers is very limited. However, it is at least significant that *taea* is used more frequently in the modern corpus. There was also evidence of a syntactic setting for *taea* which is not represented at all in any record or description of traditional Māori, nor does it align with the behaviour of the relevant English translation, as Māori has no auxiliary verbs with which to copy the English parallel. This suggests that this change is not the result of the influence of English, but has come about through some other means. It would appear that for these speakers, *tae.a* has become lexicalised (see Bauer, 2003) and is no longer perceived as the passive of *tae*, but as an independent, non-passive lexical item. Although the scant data available showed the numbers of this type of occurrence to be statistically insignificant, this usage is much more frequently observed in my language-learning classes, and commented on by other teachers and Māori language traditionalists.

To investigate this further, I conducted an additional search for the use of *taea* within the corpus constructed for the Legal Māori project. This corpus is a multi-million word collection of legal documents in Māori published in New Zealand since the 19th century. A Wordsmith general search for *taea* resulted in over 9200 examples. It is noteworthy that all 15 of the examples of *taea* with the preposition *ki* marking the expressed ability as the right collocate came from publications from the modern period, namely between 1988 and 2009. This suggests that marking the ability complement of *taea* with *ki* is not a traditional form in Māori, and has possibly emerged quite late in the period between the two corpora. This would appear to support some of the anecdotal claims of language change since the 1960s, mentioned earlier in the explanation of the data selection criteria in Chapter 2.

ĀHEI

There is very little mention of *āhei* within any of the grammars of Māori: Harlow (2001) mentions it only in his description of the modern use of *taea*, as previously stated. Bauer (1997) mentions the use of *āhei* in expressions of physical ability, and also notes its use in expressions of permission, such as example 54 below.

- (54) Kāore i āhei te tangata... kia haere i te
neg TAM able the man TAM move at the
tahatika o te one...
bank of the beach...
'... people were not allowed to go to the shore...' (Bauer, 1997, p. 140)

The various definitions of *āhei* in Māori were described previously, but there were very few examples of the use of *āhei* as a verb in the corpora. In fact, there were very few examples of *āhei* found in either corpus: only five occurrences in total in the traditional corpus, and fifteen in the modern corpus.

There were two different categories used to classify the examples of *āhei*, once again largely based on semantic criteria:

Category 1 – 'to be able/ capable'

Category 2 – 'to be permitted/ allowed'

In many instances it was not possible to distinguish between the two senses of *āhei*, even when considering the wider contexts of the examples given in their original texts: many of the situations could have been appropriately described with either translation, and most examples could not be conclusively categorised. The overall numbers were also so small that the difference between the corpora was not statistically significant.

- (55) Kāore i āhei te tūroro te haere ki te hui
NEG TAM ? the invalid DET go to the meeting
'The invalid was not able/allowed to go to the meeting' (TC)

Despite the small number of examples, I found there was still substantial variation between the corpora in the syntactic environments in which *āhei* was

found. In *He Pātaka Kupu*, Te Taura Whiri i te Reo Māori (2008) provides just one single sense of *āhei*, namely one's ability to complete a specified task or activity, due either to one's own skills or the permission granted by someone else. The verb appears in two distinct syntactic settings in the three examples given of its use. These are shown in examples 56 and 57 below.

- (56) I tēnei taumata ka āhei ngā tamariki ki te
 PREP DET level TAM able thePL child PREP DET
 tatau whakamuri mai i te rua tekau ki te kore.
 count backward DIR PREP DET two ten PREP DET zero
 'At this level children will be able to count backwards from twenty to zero.' (Te Taura Whiri i te reo Māori, 2008, p. 3)
- (57) Ka āhei i a ia te mahi rā, engari e kore
 TAM able PREP DET 3SG the work DIST but TAM NEG
 e pahure i a au!
 TAM accomplished PREP DET 1SG
 'He will be able to do that task, but it would never be accomplished by me!' (MC)

It appears then that *āhei* in example 56 functions as an action intransitive verb, with an agentive Subject *ngā tamariki* and the expressed ability in an adverbial marked by the preposition *ki*. In example 57, *āhei* appears to function as a state intransitive verb, where the task which is able to be accomplished *te mahi rā* is functioning as the Subject noun phrase, and the cause in an adverbial marked with *i*, the causative preposition for state intransitive verbs. It appears then that this may be an example of the use of *āhei* to convey the sense of something being 'possible', as mentioned in Moorfield's (2003) online dictionary. Despite the fundamental syntactic difference in the two examples of usage in *He Pātaka Kupu* (2008), especially regarding the nature of the Subject of the verb in each example, both are assigned to the same sense in the dictionary.

In contrast, the Williams' (1971) dictionary was consulted, where two senses of the verb *āhei* are differentiated. One sense is defined as 'able'; the other is 'possible, within one's power'. There was no mention of permission. Once again, there were two different types of syntactic settings for *āhei* in the examples. The entry for 'able' showed *āhei* used as in example 56 above, that is, as an action intransitive verb with an agentive Subject and the clause of

ability marked by *ki*. Interestingly the Williams' dictionary example for 'possible/ within one's power' featured *āhei* used as a state intransitive verb as in example 57. These grammatical distinctions did not seem to conclusively align with the examples for either corpus, so there appeared to be no syntactic clues to help determine which sense of *āhei* applied to a given sentence.

The data in this case study may have been too limited to make detailed or statistically significant observations of the use of *āhei* and *taea*. The data that was obtained from the modern corpus does show examples of usage that may be new, and thus suggests that the study of these lexemes would merit further investigation, although possibly done by elicitation, rather than corpus work, to compensate for the reasonably low frequency of both lexemes.

5. CASE STUDY THREE: RELATIVE CLAUSES

In the previous two case studies the investigation has concerned first a grammatical particle (the preposition *mō*) and then a lexical item (*taea*). In order to test the possibilities and limitations of the methodology, the final case study of this thesis focusses not on a lexical form, but on a grammatical construction. This chapter examines the data on certain types of relative clauses.

Some older native speakers of Māori have suggested that there has been an increase in the frequency of relative clauses in Māori in the modern period, and they also comment on the changes they perceive in the complexity of the constructions that are used by modern speakers. As both a second language learner and a teacher of te reo Māori, I am also all too familiar with the issues concerning relative clauses in the language-learning classroom. Commonly asked questions (although they are not usually worded in this manner) include - How are they constructed? Which TAMs are possible? When should *ai* be used? How do the deictic particles *nei*, *nā* and *rā* differ from each other and from *ai* in their use and semantic effect? Other questions abound, and little is offered by way of answer in language-learning textbooks to aid the learner in mastering this complex construction. The linguistic literature provides some insight, but the variation in both terminology and description poses further difficulty for the learner.

This case study aims to test whether this corpus-based methodology, using two relatively small corpora, can actually produce adequate examples of a particular grammatical construction to enable meaningful comment on and/or quantifiable observation of change in the language. In choosing relative clauses as the grammatical construction to be studied, my intention is to check the validity of the anecdotal observations of change, and also to see how the difficulties commonly experienced by learners of the language in mastering relative clauses may be manifesting in the language output of the modern

period; that is, whether the paucity of data and lack of understanding in teaching material have led to change in the language.

LITERATURE REVIEW

What follows is a review of the literature concerning relative clauses in Māori, in order of publication. This does not provide exhaustive coverage of all grammars of Māori; instead I have reviewed only those which include comment or instruction on relativisation - hence the exclusion of some well-known Māori language grammars and texts, such as those of Hohepa (1967), Head (1989), and Kārena-Holmes (2006). I have also excluded from this review those grammars which are simply derived from the principles of one of the texts I have reviewed, hence the exclusion of the various early school grammars of Māori, based entirely on Williams (1862).

Before turning to the literature, a word of caution is needed in relation to the terminology necessary for this discussion. The grammars and texts included in this review span a period of more than 170 years, and one of the challenges here has been to compile and reconcile the various accounts of relative clauses and the principles that govern their construction, despite differing terminology between authors, and even occasionally differing views of what exactly can be considered a relative clause.

The wide variety of constructions and language features that are involved in this complex construction means that there are a number of terms mentioned here which may be particular to te reo Māori, and therefore unfamiliar to those readers who have little or no knowledge of the Māori language. I have included either a brief summary description of those terms where they appear in this thesis, or a reference to where a further, more detailed description of the term or language feature in question may be found. Note that Bauer's (1997) terminology has been followed unless otherwise stated, and, in particular, the Subject of any construction is taken to be the zero-marked NP in that construction. Where unglossed examples have been quoted from the linguistic literature, I have adhered to the translation given in the original text, but have provided interlinear glosses to assist readers unfamiliar with Māori.

In order to facilitate the discussion of the literature, I will begin by discussing the structure of two representative relative clauses, the first relativising on the Subject, and the second on an adverbial of goal. In both cases, the matrix clause is a classifying sentence, with the antecedent for the relative clause as its Subject. The relative clauses are in bold.

- (58) He nui te tama i **tae mai ki te hui.**
 DET big theSG boy TAM arrive DIR to theSG meeting
 'The boy who attended the meeting was big.'
- (59) He nui te tama i **kōrero ai au.**
 DET big theSG boy TAM talk PART 1SG
 'The boy I talked to was big.'

The phrase order in both matrix and relative clauses is VS(O) where the clauses contain verbs, and in verbless clauses, such as the matrix clause in examples 58 and 59, the nominal predicate is in initial position in the clause. The relative clauses follow their antecedents (as do almost all other modifiers in Māori). In the second example, the relative clause construction includes the particle *ai*, which will be discussed further later. Although it is not illustrated in the above examples, one of the deictic particles *nei*, *nā* or *rā* ('near speaker', 'near hearer', 'distant from speaker and hearer') may in some circumstances be used instead of *ai*. Note also that in Māori nouns are not inflected for number; instead number is marked in the determiner. I will mark this in the glosses where relevant. Verbs are not inflected for tense, aspect, etc.; instead the verb is preceded by a particle, glossed TAM, which conveys this information.

In the discussion that follows, I have drawn my examples wherever possible from my corpora. In order to preserve the anonymity of my informants, I have also made trivial changes to place and personal names where necessary. Where the corpora did not contain any usable examples of a particular construction I have invented my own, so that any unattributed examples are mine. All relative clauses are in bold. I have also retained the original orthography in all examples (see Chapter 2).

Maunsell's (1842) *A Grammar of the New Zealand Language* was the first grammar of Māori to mention relativisation, albeit indirectly, where he

attempts to describe the use of *ai* as ‘chiefly employed as a substitute for the relatives *who*, *which*, *what*’ (p. 90). Maunsell exemplifies this use of *ai* with the following examples:

te whare i moe ai ia
 theSG house TAM sleep PART 3SG
 ‘the house in which he slept’

te pēhea-tanga i mea-tia ai
 theSG how-NMLZ TAM do-PASS PART
 ‘the way in which it was done’

te take i **patu-a** **ai**
 theSG reason TAM hit-PASS PART
 ‘the cause for which he was beaten’ (p. 90)

Maunsell also briefly notes that ‘the place of *ai* may be often supplied by *nei*, *na*, or *ra*’ (p. 91), but he does not identify how or whether the particles are distinguished from each other, or the process by which the relative clause itself is formed.

Williams’ (1862) *First Lessons in the Māori Language of New Zealand* is the first grammar to discuss relative clauses in detail. An entire chapter is devoted to relative clauses, beginning with the statement that ‘there are no Relative Pronouns in Māori’ (p. 58), followed by his description of how the relative pronouns of English may be expressed in Māori. He includes the deictic particles and *ai* as possible adverbs in relative clauses on Subjects, and as necessary in situations where ‘the relative in English is governed by a verb or by one of these prepositions: *by*, *on*, *at*, *in*, *with*, *by-means-of*, *on-account-of*, *by-reason-of*’ (p. 59-60).

Williams mentions the possible use of the deictic particles within relative clauses, but states that in those contexts they function as markers of physical location (p. 59). *Ai* is the only particle to which he attributes a temporal restriction, identifying it as useful only for past and future tenses (p. 60).

The 1973 revised edition of Bruce Biggs’ grammar *Let’s Learn Maori – A Guide to the Study of the Maori Language* includes discussion of what he terms ‘active, stative and passive relative constituents’, and other ‘subordinate constituents’ which include a number of other types of constructions

recognised as relative clauses in this thesis, for example those featuring the Agent Emphatic, T-class possessives, which are here called headless relative clauses, and M- and N-class possessive nominal relative clauses (p. 122-3). It is unclear why Biggs uses the term ‘relative constituents’ only for those with a verbal predicate, and not those with non-verbal predicates.

Biggs (1973) does not overtly describe the relative clause contexts in which the use of *ai*, *nei*, *na*, or *rā* is obligatory, nor does he differentiate between them, beyond echoing Williams’ statement that *ai* may only be used for past and future time. He does, however, note in his discussion of active, passive and stative relative constituents that their use indicates the subordination of what he calls the ‘verbal phrase’, and all nine of his examples feature one of those particles, both in relativising on Subjects and Direct Objects, as illustrated in two of his examples below:

Koia nei te poaka, i puu·hia ai e
 this DEI theSG pig TAM shoot·PASS PART by

taku matua.
 POSS·1SG father
 ‘This is the pig which was shot by my father.’

He taariana te poaka, i pupuhi ai taku
 DET boar theSG pig TAM shoot PART POSS·1SG

matua
 father
 ‘The pig my father shot was a boar.’ (p. 122)

The examples Biggs uses to illustrate the Agent Emphatic subordinate constituents have *nei* in all examples involving past tense, but no particle at all in the two examples in the future tense, and all of the examples of subordinate constituents with T-class possessives contain either *na*, or *ai*. This suggests patterns of usage for these particles which do not reflect native speaker norms as represented in my corpora.

In his thesis on complex sentence formation in Māori, Tamati Reedy (1979) rejects Williams’ claim that there are no relative pronouns in Māori, identifying *naana*, *maana*, *nei*, *naa*, and *raa* as relative pronouns. He furthermore argues that all verbal relative clauses are restrictive, and non-verbal relative clauses are non-restrictive. He also evaluates two distinct theoretical models for the

analysis of the formation of relative clauses, the NP-S or NOM-S analysis and the Deep Structure Conjunction analysis, and the advantages and distinct usages for each.

Because the particle *ai* is common in relative clauses, it is also worth mentioning in this review Chapin's (1974) paper on Proto-Polynesian **ai*, where he states:

Anaphoric **ai* was not lexical but grammatical; generally speaking, it was a substitute for a noun phrase which was in the oblique case ... and which was identical to and coreferential with some other noun phrase in the same sentence or a preceding sentence. (pp. 59-60).

This conclusion explains the function of *ai* in clauses in Māori which relativise on adverbials, and makes it clear that the use of *ai* in relative clauses on other syntactic functions (especially on Subjects) breaks away from the proto-Polynesian use of *ai*.

John Foster's (2012) *He Whakamārama - A full self-help course in Māori* also provides extensive comment on relative clauses, identifying them as 'groups of words, containing a verb, to distinguish or describe some person or thing' (p. 44). The majority of the relative clauses he discusses are incidentally relativisations on Subjects, and involve the use of *e...ana* as the subordinate TAM. Foster first mentions the use of deictic particles where he states, in relation to *e...ana*, 'it is possible to introduce additional shades of meaning by replacing *ana* with one of three particles of location: *nei*, *nā* or *rā*' (p. 46).

Foster goes on to mention the temporal sense conveyed by the use of the deictic particles by saying 'sometimes, in certain contexts, it is appropriate to translate 'nei' as 'now' and 'rā' as 'then (in times past)'' (p. 46).

Although the vast majority of his examples feature *e...ana*, Foster does mention that other 'verb signs' (TAMs) are possible in relative clauses in Māori. However, he mentions only two of them, stating that *i* is most frequent, and *kua* is also common.

He whakamārama includes a brief mention of headless relative clauses, mainly through presenting a number of examples featuring T-class possessives and either *e...ana* or *e + nei*. Foster also introduces the use of *ai* in this way: ‘used when past or future time is involved, in which the particle “ai” is used in place of the particles “ana”, “nei”, “nā”, or “rā”’ (p. 83).

In his chapter devoted to *ai* and its various senses and grammatical roles in Māori, Foster differentiates the use of *ai* for specifying location, time or occasion as distinct from the use of *ai* in past and future relative clauses, although most of the examples given for location and time or occasion involve relativising on adverbials of place and/or time. This is illustrated in two of the examples he gives as shown below; the first is a relativisation on an adverbial of place, the second a fronted adverbial of place – note the glosses are my own, but the translations appear as in Foster’s text:

Ko te whare tēnei i **noho ai** **ngā** **rangatira**
 PREP theSG house this TAM stay PART thePL chief
 ‘This is the house did stay therein the chiefs (This is the house in
 which the chiefs stayed).’ (p. 157)

Ko tērā te whare kai; ko reira huihui ai ngā
 PREP that theSG house food PREP there gather PART thePL
 tāngata katoa.
 people all
 ‘That is the dining room; it will be there that all the people will
 assemble.’ (p. 157)

Jacob’s *Mai i te Kākano* (2012) is the only text I reviewed which was written originally in te reo Māori. Although she does not expressly discuss relative clauses in detail, Jacob does include a breakdown of a number of the contexts in which *ai* may be used, and one of the contexts clearly describes some verbal relative clauses containing *ai*, although she assigns her examples to categories based on the function of the underlying sentences relativised on, including adverbials of time and place, adverbials of cause or reason, and groups together both Subjects and Direct Objects in a single category, referring to both functions as people somehow related to the verb in the relative clause [my translation], and recommending *ai* in all of those instances.

Of all linguists to comment on relative clauses in Māori, Winifred Bauer has covered relative clauses most extensively in various publications. Her 1982 paper on relativisation in Māori was a response to Keenan and Comrie's (1979) cross-linguistic study of relative clauses, pointing out inaccuracies in some of the information presented about Māori, and discussing some of the ways that the Māori data did not conform with the tendencies that Keenan and Comrie proposed as potential language universals. Of her other publications, the *Reed Reference Grammar of Māori* (1997) involves the most comprehensive linguistic account of relative clauses in Māori to date. In her chapter on relative clauses, Bauer describes in detail the processes involved in relativisation. She emphasises the significance of the sentence function relativised on in determining which relativisation strategy may be used. Her discussion of restrictive versus non-restrictive relative clauses contradicts Reedy's claim that all verbal relative clauses are restrictive, as evidenced by her non-restrictive verbal relative clause example given below:

ka tae ki a Puhihuia **e** **tū** **mai** **rā**
TAM arrive to PERS Puhihuia TAM stand DIR DEI
'... and reached Puhihuia, who was standing there' (p. 586).

Harlow's (2001) *A Māori Reference Grammar* also contains a comprehensive section on relative clauses and the strategies used in relative clause formation. It is closely based on Bauer's work, although the exemplification is his own, and there is some variation in the terminology used, eg. 'zero strategy' for Bauer's 'deletion strategy' and his use of the term 'comment' derived from Biggs (1973).

Summary of literature review

In summary, the literature review reveals that while a large number of authors have mentioned relative clauses in Māori, the information available in the vast majority of texts is by no means clear or comprehensive. Very few language-learning texts contain detail on the processes and rationale involved in relative clause construction: most contain little more than a list of examples from which the learner must infer the rules governing relative clauses for themselves. Detailed description of the rationale and processes of construction of relative

clauses is only to be found in a few key linguistic publications, ironically those least likely to be used in the vast majority of Māori language-learning programmes. Bauer (1997) provides the best model available for the analysis of relative clauses, and it will therefore form the basis of my preliminary analysis of the relative clauses extracted from my two corpora for this case study. The question that forms the basis of this chapter is whether this lack of clarity in teaching material and reliance on speakers' deductions based on their own language exposure and experience is reflected in a change in the form and frequency of relative clauses between the two corpora.

PARAMETERS FOR RELATIVE CLAUSE ANALYSIS

From my survey of the literature on relative clauses in Māori, I conclude that the parameters discussed in this section impact upon the final structure of the relative clause, and are necessary for their detailed description. These parameters form the basis of my analysis of the relative clauses I extracted from my two corpora.

The sentence function relativised on

Bauer (1997) and Harlow (2001) discuss this in the most detail, but Williams (1862) also alludes to the significance of what is relativised on in the final structure of the relative clause, by grouping his descriptions of relative clauses according to the relative pronouns which would be used in the equivalent English expressions. When coding for this feature, I differentiated between Subjects, DOs and Adverbials. To facilitate further analysis of the types of Adverbials being relativised on, I also differentiated where possible between Adverbials of reason or cause, those of place and time, and those expressing the agent of a passive verb.

The relativisation strategy used

It is at least implied in all of the descriptions discussed above that there is more than one process involved in the formation of relative clauses. Bauer (1997) and Harlow (2001) both overtly discuss this in terms of relativisation strategies. They agree that there are four different strategies, described in

detail below. They also agree that in many instances more than one relativisation strategy is possible to relativise on a particular sentence function. Because there is not a one-to-one correlation between the sentence function relativised on and the strategy used, it is necessary to code for both in my data analysis.

The TAM used in the relative clause

Bauer (1997) outlines the differing distribution of TAMs in matrix and relative clauses in traditional Māori. From my own observation these rules are no longer strictly followed by many modern Māori speakers. Jacob (2012) does not detail all of the TAMs permissible in relative clauses, but her examples do include a number of relative clauses with the TAMs *i*, *e*, and *ka*, the last of which is not typically used as a subordinate class TAM (see Bauer (1997) for example). As previously mentioned, Foster (2012) comments on the comparative frequency of certain TAMs in relative clauses. I have therefore coded for the TAM used in the relative clause, with the aim of verifying these observations with empirical evidence.

The particle used in the relative clause

Many of the grammars mention the variation between *ai* and the deictic particles, and their absence in the presence of the verbal particle *ana*. Bauer (1997) and Harlow (2001), and to some extent Foster (2012) attribute a temporal sense to the use of the deictics in relation to relative clauses, although Foster (1987) and Biggs (1973) mainly refer to the deictics in their primary sense as markers of physical location. I have coded for the particle which appears in the relative clause, in the hopes of clarifying the variation between *ai* and the deictic particles, and checking for change in this variation between the traditional and modern corpora.

It remains to be seen whether this lack of clarity in teaching material and reliance on speakers' deductions based on their own language exposure and experience will lead to a change in the form and frequency of relative clauses between the two corpora. One of the key differences between the two corpora is the level of exposure to te reo Māori of their respective contributors – the

traditional corpus contributors are assumed to have developed their Māori language skills through intergenerational transmission from extensive exposure to L1 speakers without the influence of English. Given the statistics available on the state of the Māori language in the modern period (Statistics New Zealand, 2014), the average Māori speaker from the modern period is exposed to much less Māori, and is also much less likely to have the opportunity to observe and engage with L1 speakers of Māori. The average modern Māori speaker is therefore more likely to engage or have engaged in more formal Māori language instruction, in order to compensate for the lack of exposure. If the formal instruction is unclear, without ample opportunities to engage with reliable L1 language models, linguistic insecurity among the speakers of the modern period would seem a predictable result.

RELATIVISATION STRATEGIES

The relativisation strategies outlined by Bauer are summarised below.

The deletion strategy

In this strategy the NP in the sentence underlying the relative clause matching the antecedent is deleted from the relative clause. Consider, for example, the following:

- (60) Ko tēnei te tangata i tae tōmuri mai.
 PREP this theSG person TAM arrive late DIR
 'This is the man who arrived late.'

Below is the matrix clause of example 60. 60b is the sentence from which the relative clause is derived. In example 60 the antecedent of the relative clause is *te tangata*. In the deletion strategy, the constituent of 60b which matches the antecedent is omitted, leaving 60c, which is the relative clause in 60.

- (60a) Ko tēnei te tangata
 PREP this theSG person
 'This is the man.'

- (60b) I tae tōmuri mai te tangata.
 TAM arrive late DIR theSG person
 'The man arrived late.'

(60c) i tae tōmuri mai
 TAM arrive late DIR
 'arrived late'

Bauer (1997) and Harlow (2001) agree that this strategy is typically used to relativise on subordinate clause Subjects, as shown in 60, NP possessors of Subjects, as shown in 61, and DOs of experience verbs, as shown in example 62 below.

(61) Tokomaha ngā tāngata **he roa ngā makawe.**
 many thePL people DET long thePL hair
 'There are many people whose hair is long.'

(62) Kua wehe atu te wahine i **pīrangi a Hone.**
 TAM leave away the woman TAM desire PERS Hone
 'The woman Hone desired has left.'

The particle strategy

In this strategy, as in the deletion strategy, the subordinate clause constituent matching the antecedent is omitted. In this strategy, however, one of the particles *ai*, *nei*, *nā* or *rā* is added to the verb constituent of the relative clause, appearing after the verb. Example 63 illustrates the particle strategy.

(63) Ko tēnei te whare i **noho ai au.**
 PREP this theSG house TAM stay PART 1SG
 'This is house where I stayed.'

63a is the matrix clause of example 63. 63b is the sentence from which the relative clause is derived. In example 63 the antecedent of the relative clause is *te whare*. In the particle strategy, the constituent of 63b which matches the antecedent is omitted, including its preposition (here *ki te whare*), and the particle *ai* is inserted in the post-verb slot, leaving 63c which is the relative clause in example 63.

(63a) Ko tēnei te whare.
 PREP this theSG house
 'This is house.'

(63b) I noho au ki te whare
 TAM stay 1SG PREP theSG house
 'I stayed at the house.'

- (63c) i noho ai au
 TAM stay PART 1SG
 'where I stayed'

Bauer (1997) states that this strategy is typically used to relativise on adverbial phrases of various types (like the adverbial of location in example 63), DOs as in example 64, and occasionally for Subjects, especially when non-agentive, as in the passive in example 65 below.

- (64) Ki tā tātou kōrero i **kōrero ai** **tātou** i
 to POSS-SG 1PLINCL talk TAM talk PART 1PLINCL at

Warepu

Warepu

'towards our speech that we delivered in Warepu' (TC)

- (65) i muri anō ngā iwi i **pei-a** **atu nei**
 LOC behind also thePL tribe TAM expel-PASS DIR DEI
 'the tribes who were expelled were also behind' (TC)

The pronoun strategy

In this strategy the relative clause constituent matching the antecedent is not deleted, but instead replaced with an appropriate pronoun. *Ai* or one of the deictics may appear in the relative clause. Example 66 illustrates the pronoun strategy. The verb *mahue* 'to be abandoned' is a neuter verb (Hooper (1984), and its Subject is the semantic patient. 66a is the matrix clause of example 66. 66b is the full sentence from which the relative clause is derived. In 66c the constituent matching the antecedent, in this case the NP in the cause phrase *i te kaiako* is replaced with the appropriate personal pronoun. Note that the personal article *a* is required as the determiner in the resulting noun phrase.

- (66) Ko tēnei te kaiako i **mahue** i **a** **ia**
 PREP this theSG person TAM abandoned PREP PERS 3SG

te tama.

theSG boy

'This is the teacher who left the boy behind.'

- (66a) Ko tēnei te kaiako.
 PREP this theSG teacher
 'This is the teacher.'

- (66b) i mahue te tama i te kaiako
 TAM abandoned theSG boy PREP theSG teacher
 'the boy was left behind by the teacher'

(66c) i mahue te tama i a ia
 TAM abandoned theSG boy PREP PERS 3SG
 'the boy was left behind by him'

Bauer states that this strategy is used for nominal predicate NPs, as illustrated in example 67, possessive phrases as in example 68, and human NPs in adverbials excepting passive agent phrases, as in the cause phrase in example 66 above.

(67) Ko tēnei te kaiako i a ia taku waea
 PREP this theSG teacher PREP PERS 3SG my phone
 pūkoro
 pocket
 'This is the teacher who has my cell phone.'

(68) Tokomaha ngā tāngata he roa ō rātou makawe.
 many thePL people DET long POSS3PL hair
 'There are many people whose hair is long.'

The possessive strategy

The fourth and final strategy involves the same processes as the particle strategy, with an additional process of removing the Subject of the relative clause and expressing it as the possessor of the antecedent in the matrix clause, typically with an A-class possessive relationship.

Bauer (1997) states that the contexts in which this strategy is used include DOs of canonical transitive verbs, as shown in example 69, and occasionally DOs of experience verbs, and NPs in adverbial phrases.

In example 69 below, 69a is the matrix clause of 69. 69b is the full sentence from which the relative clause is derived - note the use of *ka*, typical in a verbal matrix clause to indicate future tense. In 69c the constituent matching the antecedent is omitted (in this case the DO prepositional phrase), and the particle *ai* is inserted into the post-verb slot of the relative clause VC. *Ka* is replaced with the TAM *e* which typically encodes non-past in subordinate clauses. In 69d the relative clause Subject 'Mere' is expressed as the possessor of the antecedent 'te whare'.

- (69) He nui **tā** **Mere** **whare** **e** **hanga** **ai**.
 DET large POSS·SG Mere house TAM build PART
 'The house that Mere will build is large.'
- (69a) He nui te whare.
 DET large the-SG house
 'The house is large.'
- (69b) ka hanga a Mere i te whare
 TAM build PERS Mere DO the-SG house
 'Mere will build the house'
- (69c) e hanga ai a Mere.
 TAM build PART PERS Mere
 'that Mere will build'
- (69d) tā Mere whare e hanga ai
 POSS·SG Mere house TAM build PART
 'the house that Mere will build'

This possessive strategy is also used for headless relative clauses on DOs, where the T-class possessive phrase then functions pronominally as the antecedent of the relative clause, as in example 70 below. 70a is the underlying matrix clause. The Subject of 70a is unspecified, but is known to be singular and definite, given the *t-* of *tā Mere*, and so is represented as *te [...]* since *te* cannot be used as a pronoun in Māori. 70b is the sentence underlying the relative clause, where the constituent matching the antecedent appears in the DO slot. 70c is the result of the deletion of the DO prepositional phrase, marked by the insertion of the particle *ai* after the verb. 70d shows the relative clause with its unspecified antecedent. 70e shows the removal of the relative clause Subject *a Mere* and its expression as the A-class possessor of the unspecified antecedent.

- (70) He pai **tā** **Mere** **i** **mahi** **ai**.
 DET good POSS·SG Mere TAM do PART
 'What Mere did is good.'
- (70a) He pai te [...]
 DET good the-SG
 'The [...] is good.'
- (70b) i mahi a Mere i te [...]
 TAM do PERS Mere DO the [...]
 'Mere did the [...]

- (70c) i mahi ai a Mere
TAM do PART PERS Mere
‘that Mere did’
- (70d) te [...] i mahi ai a Mere
thesG TAM do PART PERS Mere
‘the [...] that Mere did’
- (70e) tā Mere i mahi ai
POSS-SG Mere TAM do PART
‘what Mere did [lit. Mere’s that was done]’

DATA SELECTION

In this study of relative clauses, I was searching for a grammatical construction, not a grammatical or lexical item as for the first two case studies. For this reason it was impractical to do a WordSmith search to extract the data I needed. As has been discussed in the literature review, there is no set of pronouns particular to relative clauses, and any of the particles which do commonly feature in relative clauses (*ai*, *nei*, *nā* and *rā*) are so commonly found in other grammatical settings that this search would be impractical. It would also inevitably exclude arguably the most common type of relative clause, which does not contain any of those particles. Instead, the only practical option was to search both corpora manually, extracting all examples of relative clauses, and collating all of those examples in the Excel files I used for data analysis.

This manual search was laborious, and not without its own problems. An initial manual search of the corpora resulted in more than 1200 examples from the traditional corpus containing what appeared at first sight to be relative clauses. Upon proper analysis of the examples however it became clear that a large proportion of the examples were not actually relative clauses, but were instead other constructions with a somewhat similar surface structure, as in the fronted Subject with a numerical pre-modifier in example 71, which is the result of topicalisation, and the fronted Subject phrase with the determiner *he* in example 72, and the fronted adverbial of reason in example 73.

- (71) e 200 tangata e noho herehere ana i
 NUM 200 person TAM- remain detain -TAM CAUS
 te ture o te whawhai
 theSG rule POSS theSG fight
 '200 people remain imprisoned due to the rules of engagement' (TC)
- (72) ka puta tona reo ano he reo tangata e
 TAM appear his voice like a voice person TAM-
 talk -TAM
 korero ana
 'Its voice came out as though it was a human voice speaking' (TC)
- (73) he ngenge o taku hoiho i hoki ai ahau
 DET tired POSS my horse TAM return PART 1SG
 'it is because my horse was tired that I returned' (TC)

These examples and others like them were therefore excluded from the collection before the statistical analysis was completed – resulting in a final count of 826 traditional relative clause examples.

This case study does not include all types of relative clauses. As mentioned previously there are both nominal and verbal relative clauses in Māori, but I have excluded from my study all relative clauses with non-verbal predicates, including those involving the Agent Emphatic construction (note the analysis of the Agent Emphatic construction in Māori is controversial - see Bauer (1997, pp. 513-514) for further information).

There are a number of reasons for the exclusion of nominal relative clauses. There is no anecdotal evidence of change in the structure or use of nominal relative clauses, probably due to their highly frequent, predictable and consistent nature: they usually involve relativising on the Subject using the deletion strategy, as is shown in example 74 below, which features a locative relative clause. 74a is the underlying matrix clause. 74b is the underlying locative sentence from which the relative clause is derived. In 74c the constituent matching the antecedent is omitted, leaving just the locative preposition *i* and the locative pronoun *konei*. This construction is identical to that of an Adverbial of place.

- (74) He ātaahua ngā wāhine i konei.
 DET beautiful the-PL women LOC here
 'The women here are beautiful.'

- (74a) He ātaahua ngā wāhine.
 DET beautiful the·PL women
 'The women are beautiful.'
- (74b) i konei ngā wāhine
 LOC here thePL women
 'the women are here'
- (74c) i konei
 LOC here
 'here'

Including such a high-frequency construction in this study would greatly increase the size of the data extraction and analysis task, with a comparatively small chance of contributing anything new to the body of knowledge on relative clauses in Māori. There is a drawback to this decision; in excluding all nominal relative clauses I have excluded those cases where nominal relative clauses involve relativisation on their Predicates. This is one of the primary environments in which the pronoun strategy occurs, and thus I have very little data on this strategy. This is also one of the options for relativising on Direct Objects, so their exclusion may mean relativisations on Direct Objects are under-represented in the results based on the sentence functions relativised on.

I also excluded from this study relative clauses with numeral predicates. The issues around the linguistic categorisation of numerals in Māori are discussed by Bauer (1997). The problems and inconsistencies in the classification of numeral predicates, and therefore numeral relative clauses, would mean that their inclusion would add little to the discussion of verbal relative clauses in this chapter.

Reduced relative clauses were also excluded from this study, as I am most interested here in the structure and use of function words within the relative clause, see for example 75 below, where the TAM is elided in the reduced relative clause.

- (75) Ko Mereana te wahine pupuri i ngā moni
 PREP Mereana theSG woman hold DO thePL money
 'Mereana is the woman who holds the money.'

There were only a small number of reduced relative clauses in my corpora, and they are predictably low on function words. They also almost exclusively involve relativising on Subjects using the deletion strategy, and therefore were not included in this study.

What follows then is a case study of verbal, non-reduced relative clauses in Māori, focussing on the relativisation strategy used, the TAMs used within the relative clause, and the absence or inclusion of *ai*, *nei*, *nā* or *rā* and their associated meanings.

RESULTS FROM THE DATA ANALYSIS

Overall frequency

My manual search of the two corpora resulted in 826 relative clauses from the traditional corpus, and 894 relative clauses from the modern corpus. With regards to the overall frequency the modern corpus contained just 68 more verbal non-reduced relative clauses than the traditional corpus. The general anecdotal claim that relative clauses are much more commonly used in modern Māori is therefore unsupported by the data in this study.

Sentence function relativised on

Relativisations on Direct Objects showed the most significant difference in frequency between the two corpora – the traditional corpus contained almost twice the number found in the modern corpus. Figure 5 below shows the number of relative clauses for each corpus, grouped according to the sentence function relativised on. A chi-square test of independence was performed to examine the relationship between traditional and modern Māori and sentence functions relativised on. The chi-square statistic was 62.7853. The p-value was < 0.00001. The relationship between these variables was therefore statistically significant, meaning that the differences in frequency are unlikely to be the accidental result of the corpus material, and are thus likely to identify real changes in usage between the corpora.

Table 2: Relative clauses grouped by sentence function relativised on

	TOTAL	Subject	DO	Adverbial	Unsure
Traditional	826	524	117	185	0
Modern	894	637	59	164	34

With regards to the distribution of different relative clause types, as shown in Table 2, a slightly larger proportion of the relative clauses in the modern corpus were relativisations on Subjects, although the difference is unlikely to account for the observations of those speakers who believe there is an increase in the frequency of relative clauses in modern Māori. The relative simplicity of the relativisation process for Subjects also means that in many instances the untrained language user does not recognise relativisations on Subjects using the deletion strategy as relative clauses, and instead identifies relative clauses by the appearance of *ai* or the deictic particles in the subordinate verb constituent. Subjects in both traditional and modern corpora were most commonly relativised using the deletion strategy, as shown in Table 3 below. The particle strategy was used for a significantly greater proportion of the traditional corpus examples than the modern examples - the chi-square test with the Yates correction showed the association between corpus and the strategy used to relativise on the Subject of the underlying sentence was statistically significant, with chi-square test statistic of 31.348 with 1 degree of freedom, and the two-tailed P value < 0.0001.

Table 3: Relative clauses on Subjects, grouped by corpus and relativisation strategy used

Relativisation strategy used	TRAD	% of examples	MOD	% of examples
TOTAL	524		637	
Deletion strategy	367	70%	537	84.3%
Particle strategy	154	29.4%	100	15.7%
Possessive strategy	1	0.2%	0	0
Pronoun strategy	2	0.4%	0	0

The vast majority of the traditional relativisations on Subjects using the particle strategy involved non-agentive Subjects, in line with Bauer (1997); only 3 of the 154 traditional examples involved an agentive Subject, as shown in example 76 below. All of those examples featured the particle *nei* in combination with the non-past TAM *e*. It is important to note here that there may be an alternative explanation for these relative clauses: it is possible that these examples may in fact be instances of the deletion strategy, and that the particle *nei* may be used in its deictic temporal sense ‘here/now’, as opposed to its use as an indicator of relativisation.

- (76) ki tō tātou kuini atawhai e tiaki nei
to POSS:SG 1PLINCL queen kind TAM care DEI
i a tātou
DO PERS 1PLINCL
‘to our kind queen who cares for us’ (TC)

The particle strategy examples from the modern corpus were not as consistently non-agentive as those in the traditional corpus; in fact 28 of the 100 examples featured agentive Subjects, as exemplified in 77 below.

- (77) He nui ngā taura i whakatūtū ai i
DET many thePL student TAM raise PART DO
te puehu
thesG dust
‘There were many students who raised the dust’ (MC)

18 of the 28 examples involved the combination of the TAM *e* with the particle *nei*, which once again could possibly be accounted for by its temporal sense. The other ten examples however could not be feasibly accounted for with this analysis – two examples used *ai* which has no applicable temporal alternate sense, and eight examples involved the non-past TAM *e* in combination with the particle *rā*, as shown in 78.

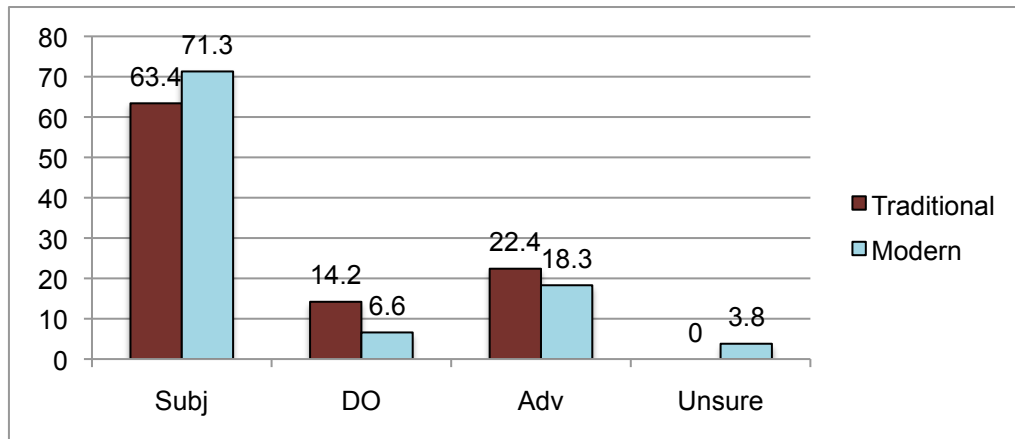
- (78) ki ō mātou hoihō **e** **tūtū** **mai**
to POSS·PL 1PLEXCL horse TAM stand DIR

rā **i** **te** **taepa**
DEI LOC theSG gate
'to our horses that are standing at the gate' (MC)

It may once again be possible that *rā* is being used in its locative deictic sense 'over there', but this could be neither conclusively determined nor excluded for any of the relevant examples. The numbers here are too small to be statistically significant, but they do suggest a difference in the usage of the particles in relative clauses in the modern period.

Perhaps the most noteworthy difference between the two corpora, however, was an issue that arose in the analysis of a number of the examples from the modern corpus. In the traditional corpus data, the sentence function relativised on could easily be identified in all examples, and for the vast majority of the examples from the modern corpus. The distribution of relative clauses amongst the three specified sentence functions is illustrated in Figure 4 below, and showed no significant difference between the two corpora with regards to the types of sentence functions of the underlying sentences that were relativised on. However, as shown in Figure 4 below, an issue arose in the categorisation of a number of the relative clause examples from the modern corpus.

Figure 4: Percentage distribution of sentence functions relativised on



As shown in Figure 4 almost 4% of the modern corpus relative clauses (34 examples in total) could not be coded in this way, as it was not possible to determine the function of the antecedent in the sentence underlying the relative clause from the syntax of the relative clause. In each situation, although the role of the noun phrase relativised on could usually be clearly identified semantically, the syntax of the relative clause was inconsistent with the semantic analysis. Example 79 below shows one such case, where the antecedent *mahi* is clearly the patient in the relative clause. Note that in the discussion of these problematic examples, I have included an intended translation, inferred from the context of the original text, and the closest literal translation of the actual phrase.

- (79) e pā ana ki ngā mahi e whakarite nei
TAM- affect -TAM to thePL work TAM arrange DEI
e rātou
by 3PL
INTENDED: 'about the work they are arranging' (MC)
ACTUAL: 'about the work arranging by them'

The relative clause contains the preposition *e*, the typical agent marker in passive contexts in Māori. Conflicting with this passive analysis however is the active verb in the relative clause VC. The particle *nei* which appears after the verb is consistent with the particle strategy, which suggests that *ngā mahi* is unlikely to be the Subject of the underlying sentence. These conflicting

features mean that the relative clause cannot be clearly categorised as active or passive, as it contains conflicting features of both constructions. A clearly active parallel relative clause is shown below in 79a, and 79b shows the parallel passive construction, which would be expected from a speaker of traditional Māori.

(79a) e whakarite nei rātou
 TAM arrange DEI 3PL
 'that they are arranging'

(79b) e whakarite·a e rātou
 TAM arrange·PASS by 3PL
 'that is being arranged by them'

Example 80 below shows a different situation. In the relative clause the agent is marked by *i*, which is not traditional for either active or passive sentences in Māori. Given the use of the canonical transitive verb in the past tense setting, traditionally a passive relative clause would be expected, as shown below in 80a.

(80) pērā i te waiata i tito mai i a Rewi
 like PREP theSG song TAM compose DIR DO PERS Rewi
 INTENDED: 'like the song that Rewi composed' (MC)
 ACTUAL: 'like the song that composed from Rewi'

(80a) i tito·a mai e Rewi
 TAM compose·PASS DIR by Rewi
 'that was composed by Rewi'

These 34 examples were therefore categorised according to the semantic roles relativised on, as opposed to the sentence function - see Appendix 3 for the complete list of these examples and their respective analyses. The examples were coded for the TAM that appeared in the subordinate clause, the type of verb in the subordinate VC, the use of a post-posed particle in the subordinate VC, the phrase type in which the agent was found, and the phrase type in which any specified patient was found. The most significant feature identified through this analysis was the contradictory use of active and passive grammatical features - 22 of the 34 examples involved relativising on a patient,

but used the deletion strategy, typically used for sentence Subjects, and featured an active verb in the relative clause VC.

This second framework greatly facilitated analysis in the majority of examples, but there were a few examples which still could not be conclusively categorised, even within the new categories, as the sentence was ambiguous, and remained so even on examination of the full original text. Example 81 below shows one such example, where two elements appear in Subject-like NPs, the VC features a passivised verb, but neither of the NPs is specified or obviously identifiable as agent.

- (81) Tiro·hia ngā āhuratanga hou e whāngai·a ana a
 look·PASS thePL feature new TAM- adopt·PASS -TAM PERS
 Microsoft Windows7.
 Microsoft Windows7
 ‘*Look at the new features that Microsoft Windows7 is being
 adopted’ (MC)

Upon closer examination of example 81 the verb *whāngai* which appears in the subordinate VC has multiple senses, including ‘to feed’ and ‘to care for’. Given the context of the utterance ‘adopt’ seems to be the most appropriate translation, but it is worth noting that although *whāngai* can be used to describe the action of ‘adoption’ in familial contexts, its use for ‘adopt’ in the context of computer software requires a dramatic semantic extension from the traditional Māori sense, given that it is typically used to indicate ‘adopt’ not in the sense of employing a different system, but instead to refer to the action of one entity caring for, or taking responsibility for another. Even accepting *whāngai* as ‘adopt’ in this type of context, the passive suffix means the most accurate translation of the utterance is still nonsensical - ‘Look at the new features that Microsoft Windows7 is being adopted.’ In order to construe any comprehensible meaning from this sentence, the reader is therefore required to ignore both the grammatical marking (or lack thereof) of sentence function and the passive suffix, and instead decipher the sentence solely through identifying the content words and applying their logic and understanding of the broader context to infer what the speaker was trying to communicate. This issue will be discussed further in the following chapter.

Relativisation strategy used

Since the relative clauses in the two corpora might differ in the choice of strategy for any particular syntactic function, the relationship between strategy and function was also examined. An issue arose when coding the examples containing the discontinuous TAM *e...ana* for the relativisation strategy used; in examples with *e...ana* it was not possible to conclusively determine which strategy had been used, as when *e...ana* was used no other particle was possible in the post-verb slot, which prevented the appearance of the primary indicators of the particle strategy. The examples featuring the discontinuous TAM *e...ana* were all coded under the deletion strategy by default. This decision has potentially reduced the numbers of particle-strategy clauses, and boosted the numbers of deletion-strategy clauses. The only alternative was to call all *e...ana* clauses relativising on an adverbial instances of the particle strategy. The problem with that solution is that it is not possible to provide evidence for the correctness of that solution. Note, however, that the choice I have made implies that the deletion strategy is, in the modern period, at least, a possibility for adverbial relative clauses, which is a change from the rules described by Bauer (1997) and Harlow (2001).

As shown in Table 4 below, the modern corpus contained significantly more relative clauses constructed using the deletion strategy. The traditional corpus featured a significantly larger number of particle strategy and possessive strategy relative clauses. There were only three examples of the pronoun strategy in the traditional corpus, and just one example in the modern corpus. This lack of data for the pronoun strategy may well be explained by one of two factors, firstly the corpora may have been too small to provide adequate data for these lower frequency constructions. Secondly, as previously mentioned, the data set was restricted to exclude Agent Emphatic relative clauses and other types of nominal predicates, which are some of the syntactic environments in which the pronoun strategy is most frequently used. An exploration of the full range of syntactic environments in which the pronoun strategy is used is outside of the scope of this thesis, but does merit further research. The same 34 examples from the modern corpus previously

discussed were problematic here, for the same reasons as above. These examples have therefore once again been grouped separately.

Table 4: Relativisation strategy used

Relativisation strategy used	Traditional	Modern
Deletion	393	602
Particle	352	215
Possessive	78	42
Pronoun	3	1
Unsure	0	34
TOTAL	826	894

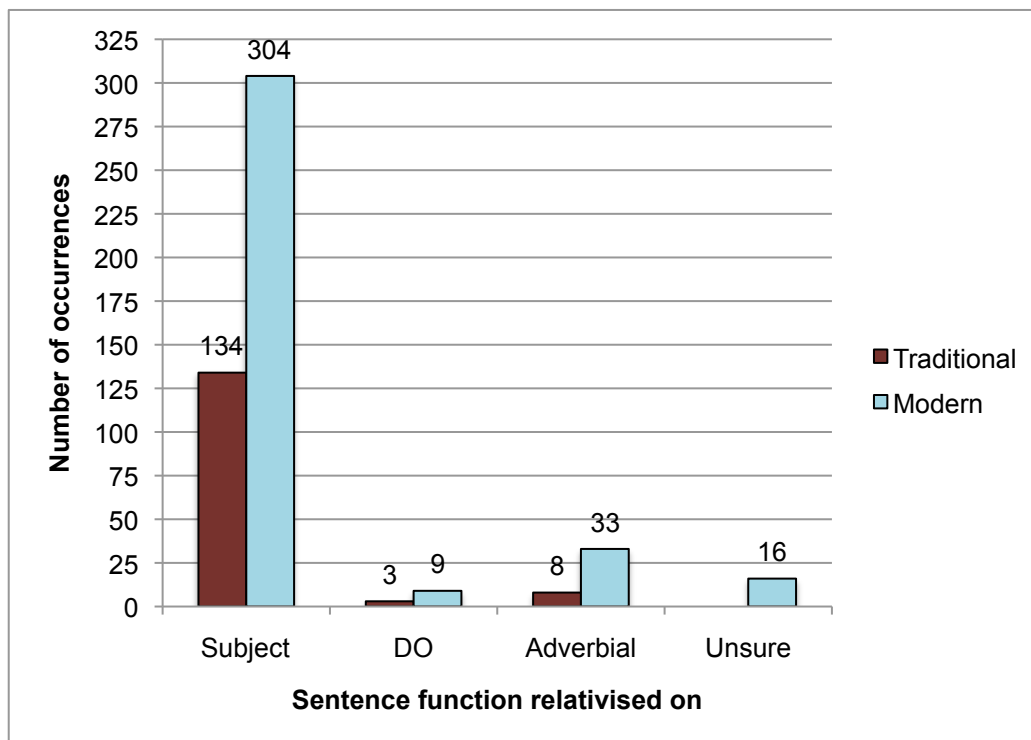
Subordinate TAM

Perhaps the most significant difference between the two corpora was the marked increase in the frequency of the TAM *e...ana* in the relative clauses from the modern corpus, as shown in Figure 5 below. The most significant difference between the corpora was in the relativisations on Subjects, where the modern examples containing *e...ana* outnumbered the traditional examples by more than 2:1. This increased frequency of use of *e...ana* was consistent across all categories of sentence function relativised on; although the actual numbers were considerably smaller than for the Subjects, the occurrences of *e...ana* in the Direct Object and Adverbial categories increased three-fold and four-fold respectively.

This suggests that *e...ana* is being more readily used in the modern period to relativise on all sentence functions. It is worth noting that the increasing use of *e...ana* inevitably corresponds with a decrease in the use of *ai* and the deictic particles in relative clauses on Direct Objects and other adverbials. It can be argued that this change in behaviour may contribute to a substantial decrease in exposure for the modern Māori speaker to the use of *ai* and the deictic particles, and therefore may lead to uncertainty for speakers in the modern

period as to the correct and specific use of the particles. This uncertainty is evidenced by the inconsistency of the semantic environments in which the particles appear in the examples from the modern corpus, as mentioned earlier in the chapter in the discussion of the particle-strategy examples. The implications of this linguistic uncertainty with regards to grammatical particles are further discussed in the following chapter.

Figure 5: Frequency of *e...ana* in relativisations on different sentence functions



Alternatively, it may be that the difference in the frequency of the use of *e...ana* is merely the coincidental result of a larger number of contexts in the modern corpus for which the continuous TAM was required. The size of the increase in the modern corpus makes this seem less likely, so in an effort to find an alternative explanation to account for the increase, I reviewed the alternative TAMs to see whether there was a corresponding decrease in the use of other non-past TAMs and TAM+particle combinations in the modern corpus. I therefore counted the examples of the non-past TAM *e*, both on its own and in combination with *ai* and the deictic particles. Table 5 below shows the relative frequencies of the comparable non-past TAMs.

Table 5: Occurrence of non-past TAMs and accompanying particles

TAM (+particle)	Traditional	Modern	Difference	Percentage difference
e...ana	145	364	+219	+150%
e	30	20	-10	-33%
e & ai	44	50	+6	+14%
e & nei	72	64	-8	-11%
e & nā	13	4	-9	-69%
e & rā	2	15	13	+650%
ka	30	18	-12	-40%

As is shown in Table 5, there was no substantial difference between the corpora in the frequency of usage of *e* in combination with *ai* or the deictic particle *nei*. Although there was a considerable percentage change shown in the usage of *e* on its own, and in combination with *ai* and *nā*, each of those categories involved an actual difference of around 10 examples, limiting the statistical significance. The usage of *e* with *rā* may be shown to have increased six-fold, but the numbers are small, so this may only represent a coincidental increase of applicable contexts. A chi-square test of independence was performed to examine the relationship between traditional and modern Māori and the non-past TAMs (except *e...ana*) listed in Table 5. The relationship between these variables was significant, with chi-square = 19.514 and the P-value of 0.00154. The change raises the possibility that speakers in the modern period are using *e...ana* in lieu of other non-past TAMs, either *ka* or *e* on its own, or with *ai* or another deictic particle. There were also some examples in the modern corpus which used *e...ana* in contexts other than the continuous sense that was traditionally its only use, such as 82 below.

- (82) i ngā tau e rua e whai mai ana
 LOC the-PL year NUM two TAM- follow DIR -TAM
 ‘in the two years to follow...’

One factor which has impacted on the increase of *e...ana* in the modern corpus relative clauses is the 42 occurrences of *e pā ana* ‘about/ relating to’, where there were no examples at all found in the traditional corpus. The prolific use of this expression in the modern period may correspond with the decrease in use of the preposition *mō* to introduce topic of discourse, discussed in Chapter 4, although there are other potential explanations for that perceived difference. It is difficult to determine whether there is a corresponding decrease in the modern corpus of the use of other markers for the topic of discourse, like the sense of the preposition *ki*, exemplified in 83 below, and the preposition *i* in example 84; the high frequency of *ki* and its frequent use in a wide range of other senses mean that a corpus wordsearch for examples relating to topic of discourse would be unfeasible.

- (83) Kua kōrero ia ki tana rautaki
 TAM talk 3SG DO POSS-3SG thought
 ‘She has spoken about her strategy.’

- (84) Kei te ako ia i te hītori Kiriki
 TAM learn 3SG DO theSG custom Greek
 ‘He is learning about Greek history.’

In regard to the other TAMs, a significant decrease of 31% was noted in the use of *i* in the modern corpus, as shown in Table 6 below. I searched the modern corpus examples for evidence of the use of *e...ana* in lieu of *i* in clearly past tense environments but could find no examples in my data that could be conclusively labelled as referring to the past tense. It may be that the traditional corpus simply contained more material set in the past – certainly a significant part of the traditional corpus material comprised tribal historical narratives where the use of the past tense would be likely. Regardless, the historical narratives did not make up such a large proportion of the corpus as to be likely to account for a difference as significant as this.

Table 6: Frequency of TAMs used in relative clauses

TAM	Traditional	Modern
e	164	165
e...ana	144	362
i	395	270
ka	45	24
kua	69	65
i te	3	1
kei te	3	7

It was interesting to find in both corpora a number of examples of relative clauses containing the TAMs *kua* and *ka*, as these are usually reserved for matrix clauses in Māori. The two corpora contained almost equal numbers of occurrences of *kua*, but the traditional corpus contained almost twice the number of relative clauses with *ka*. It appears that this variation may have just been happenstance; a chi-square test of independence was performed to examine the relationship between traditional and modern Māori and the use of *ka* and *kua*, but this produced a chi-square statistic of 3.4847, and a P-Value of 0.061937, showing that the relationship was not statistically significant. *Ka* occurred in a wide variety of environments in relative clauses; examples from the traditional corpus include the relativisation on a passive Subject shown in 85, relative clauses modifying the Subject of a presentative matrix clause, as shown in example 86 below, and modifying adverbials of place/time, in either past or future tense, as illustrated in examples 87 and 88 below.

(85) e nui rawa ērā ka whakahoki-a e au
TAM many intens those TAM return·PASS by 1SG
‘there were too many of those that I returned’ (TC)

(86) Tēnei anō te karere ka haere atu ki a koe
this INTENS theSG messenger TAM go DIR to PERS 2SG
‘This is the messenger who will go to you’ (TC)

- (87) I taturi atu mātou i ngā wiki ka pahemo ake nei
 TAM wait DIR 1PLEXCL LOC thePL week TAM pass DIR DEI
 'we waited in the weeks that passed' (TC)
- (88) i te 15 o ngā rā o te Maehe
 LOC theSG 15 of thePL day of theSG March
 ka whai ake nei
 TAM follow DIR DEI
 'on the 15th of next March' (TC)

The 24 examples from the modern corpus however showed *ka* used not only in relativisations on passive Subjects, but also used in other contexts where the subordinate TAM *e* would typically be expected, such as the habitual aspect conveyed in example 89 below. There were no examples of *ka* in relative clauses following presentative constructions, nor were there any examples of the use of *ka* in adverbials of place/time.

- (89) ngā āhuatanga o te ākonga ka puta i
 thePL feature of theSG student TAM emerge LOC
 ngā kura reo Māori
 thePL school language Māori
 'the features of students who come from the Māori language schools' (MC)

Although once again the numbers are small, there were also more examples in the modern corpus of the use of *kei te* in relative clauses, where traditionally *e* or *e...ana* would be expected, as in examples 90 and 91.

- (90) mō tētahi hui kei te tū ki reira
 for a meeting TAM stand LOC there
 'for a meeting that is being held there' (MC)
- (91) He maha rawa te tangata Māori kei te kai hikareti
 DET many INTENS theSG person Māori TAM eat cigarette
 'There are many Māori people who smoke cigarettes'

These corpora have therefore documented the ready use of matrix clause TAMs in subordinate clauses in both the traditional and modern periods. While their use is undoubtedly less frequent than the other TAMs *e* and *i* which are typically described in grammars for subordinate clauses (see Bauer (1997) for example), the evidence of their use in these contexts in the traditional corpus examples indicates that this distinction between matrix clause and subordinate clause with regards to TAMs may have been

dissolving in the traditional period, which suggests this may be a change in the language that was developing internal to the Māori language, as opposed to being as a result of language contact with English. Differentiating between types of language change observed in this study is further discussed in the following chapter.

SUMMARY OF RELATIVE CLAUSES RESULTS

This study of relative clauses has shown this methodology to be effective in providing a number of observations of statistically significant change in the language. The most significant change noted is the dramatic increase in the use of the TAM *e...ana* in relative clauses in the modern period. There was also a statistically significant change observed in the sentence functions relativised on, with an increase in relativisations on underlying sentence Subjects, and a decrease in relativisations on underlying sentence Direct Objects in the modern corpus. With regards to the anecdotal claims of an increased frequency of relative clauses in the modern period, these observations were found to be incorrect, as the difference in frequency of relative clauses between the two corpora was not statistically significant. Another statistically insignificant but nonetheless important observation is the evidence which points to issues emerging for speakers around the accurate and consistent use of grammatical particles to distinguish between active and passive voice and clearly differentiate roles of agent and patient within subordinate clauses, with the emergence of a cluster of relative clauses showing contradictory grammatical features. The implications of these observations for the Māori language and Māori language speakers will be explored in the following chapter.

6. IMPLICATIONS

This chapter examines trends in the results of my case studies, and explores the relevance of my findings to the revitalisation of te reo Māori in Aotearoa New Zealand. The chapter contains a review of the results obtained in the three case studies and discusses the nature of the changes that have been observed. The potential impact of these types of changes on the Māori language are analysed, supposing the acceptance of these changes by the Māori language community. What follows is the examination of some of the common attitudes towards language change that are currently represented in the Māori language community, based on my own observations and categorisations from my experiences. The tenability of each point of view is discussed, and a possible compromise is suggested that allows for a meaningful discussion and negotiation of potential change in the language. A number of linguists have engaged with the topic of language change in Māori (see for example Boyce, 1995, Bauer, W. 2007, Harlow, 2005, Harlow et al., 2004 & 2010), but there is little formal discussion about language change within the Māori community, and what there is centres on the relatively trivial matter of vocabulary. Because this thesis focusses on syntactic change, change in vocabulary is only discussed here where it is immediately relevant to the results of my study.

The case studies on *mō*, *taea* and *āhei*, and relative clauses have produced a number of pieces of empirical evidence of change in the Māori language, although the statistical significance of the evidence in some cases is limited by the size of my corpora. In reviewing the types of changes observed in my study, two key themes emerged: some changes showed Māori shifting to be more like English, and other changes showed a shift from specificity to generality. These two themes will be discussed in detail below.

Theme 1 – Shift toward English

Given the history of contact between te reo Māori and English, the importation of some language features is to be expected; Heine & Kuteva (2010) state ‘all the information that is available on language contact suggests that contact-induced grammatical replication in general and grammaticalization in particular are far more common than has previously been assumed’ (p. 101). The inevitability of change in a language is widely accepted by many, both linguist and lay-person alike. However, discussion of the acceptability of different types of change is less common. This is especially true in the context of a minority, heritage language in the process of revitalisation, where the primary focus is often increasing the number of speakers in the population and promoting the use of the language in new domains, and there is less emphasis placed on critically assessing the quality of the language that is produced.

The first type of language change identified is one where the language form in question increasingly conforms in some way to its semantic equivalent in English. This was observed on both a lexical and syntactic level; as Thomason (2001) states, ‘all aspects of language structure are subject to transfer from one language to another, given the right mix of social and linguistic circumstances’ (p. 11). In some instances the observed change involved the expansion or contraction of the meaning of a Māori word until it paralleled the English equivalent. This was observed in the case study on *taea*. *Taea* was traditionally used for talking about capability or possibility, meaning ‘the ability of something to be accomplished’ (Williams, 1971, p. 356). However, a new sense emerged in the modern corpus, where *taea* was also used to seek permission or indicate politeness in a stated request, thus mirroring the English equivalent *can* in having both senses (cf. Quirk, Greenbaum, Leech & Svartvik, 1972). The use of *taea* in requests is both often observed and often overtly remonstrated against by teachers of Māori (see for example Jacob, 2012), yet the phenomenon is clearly common enough to have been captured, albeit infrequently, in this modestly-sized corpus of the semi-formal language of competent speakers of Māori.

The case study on the preposition *mō* showed evidence of Māori syntactic features being replaced with features which copy equivalent English expressions. This is exemplified in the complement to the verbs for 'to search' and 'to wait', where the traditional object markers *i* and *ki* are replaced by *mō*, mirroring the English use of *for* as the complement marker with these verbs. There was also evidence of the infiltration of English structure into Māori syntax in several modern examples which were unprecedented in traditional Māori, such as the stranded preposition in '*He aha tēnā mō?*' – 'What is that for?' – in place of the traditional '*Hei aha tēnā?*': 'What is the purpose of this?' Martin (1995) describes a comparable situation with the Brunei Murut (Lun Bawang) community in Temburong, "A limited amount of data ... shows that borrowings from Malay are a common feature in everyday speech. Such borrowings by the minority language groups are not a new phenomenon. However, in the past, this strategy was used mainly when there was a need to describe a new concept or item ... The tendency today is for Malay (particularly Brunei Malay) lexis to be used even in cases where an equivalent word exists in Murut" (p. 37). Just as Martin describes Malay (Bahasa Malayu) infiltrating the minority Brunei Murut language, the instance of *i* and *ki* being frequently replaced with *mō* as a complement marker, and the shift to stranded preposition use, proves that English has influenced and provided the pattern for alterations to the grammatical structure of at least some Māori forms in the modern period.

The nature of the change in these instances is the most significant issue, particularly as some of these observations are not examples of grammatical importation from English to fill a language 'gap' in Māori. Instead, traditional Māori words and structures are being replaced by a loan translation from English. My own experience suggests that this is more likely to involve a lack of speaker awareness about the traditional Māori precedent, rather than the deliberate abandonment of traditional Māori language forms in favour of English-like equivalents. However, regardless of the awareness or otherwise of the speaker, if language forms of English origin are substituted for traditional Māori forms, there is an inevitable corresponding decrease in both the use of and therefore also the exposure to the traditional Māori form, which

sets up an optimal environment for the loss or endangerment of the traditional language feature.

Most of the examples of importation from English identified in this study are isolated to a limited set of specific semantic contexts, and may be dismissed as learner errors by advanced speakers of the language. Heine & Kuteva (2010) note:

There typically is *spontaneous replication* in bilingual interaction, where an individual speaker – consciously or unconsciously – propagates novel features in the replica language that have been influenced by some other language ... the vast majority of instances of it will have no effect on the language concerned, being judged as what are commonly referred to as ‘speech errors’. But some instances may catch on: being taken up by other speakers and used regularly, they may become part of the speech habits of a group of speakers (early adopters), and they may spread to other groups of speakers – in exceptional cases even to the entire speech community. (p. 88)

Two important factors of the situation in Māori make it likely that these types of errors will be replicated. Firstly, errors like this, termed ‘covert errors’ by Corder (1973, p. 272-3), typically result in a sentence which is still comprehensible to the listener, or at least decipherable, in the case of some older native speakers of Māori, who say the only way they are able to understand young speakers is by translating what is said into English and understanding the English. The covert nature of the errors decreases the likelihood that the speaker will have issues with communication, or otherwise have the error highlighted as unacceptable in the flow of normal communication (cf. change from below, Labov, 2001). Secondly, these errors were found in deliberate language produced by a group of individuals who would be expected to take roles as language models in the Māori language community; all the contributors for the modern corpus were speakers who either were raised speaking Māori, or who are considered to have the requisite

Māori language expertise to be qualified to teach the Māori language at any secondary school in New Zealand. The language behaviours demonstrated by this cohort then are arguably less likely to be corrected, and more likely to be replicated by other speakers and learners.

Theme 2 – Specificity shift

The second type of change documented indicated the dissolution of a grammatical category or distinction, and a shift towards a non-specific or general default option for the language form in question. This was illustrated in a number of observations from different case studies, and often involved different syntactic aspects of the language.

The first example is the use of the possessive preposition *mō* rather than the traditional *mā* in the context of an A-class possessive relationship, which suggests a loss in the specificity of the possessive system which is likely also to manifest itself in the possessive determiner system.

The relative clause case study provides further instances. Analysing the use of TAMs provided evidence of a shift away from the use of TAMs as a marker of the distinction between matrix and subordinate clauses, with the increased use of traditionally matrix-clause-only TAMs like *kei te* in relative clauses. The increased frequency with which the TAM *e...ana* was used in relative clauses suggests that it has become the default TAM in those contexts. The effect here is two-fold: there is a loss of specificity in the indiscriminate use of *e...ana* without regard to its usual aspectual sense (i.e as a marker of imperfective or progressive aspect (see Bauer, 1997, p. 89-90, for example)), and the use of *e...ana* also precludes the use of *ai* or a deictic particle. This inevitably affects the level of exposure of a speaker/ learner to the particle strategy for relative clause formation, and hence decreases their opportunities to deduce for themselves the appropriate contexts and the rules governing the particles, and the use of the particle strategy. This is perhaps supported by the emergence of relativisations on agentive Subjects using the particle strategy, showing the erosion of the distinct environments for the different relativisation strategies, and thereby also removing syntactic clues for the listener as to the

sentence function of the antecedent to the relative clause in the underlying sentence.

This shift away from specificity is also manifested in those examples that did not differentiate the roles of agent and patient within the syntax, illustrated by the relative clauses from the modern corpus that could not be categorised by the sentence function relativised on (see chapter 5). This indicates a shift from the requirement for consistent grammatical marking of sentence function in the sentence. It is worth noting that the vast majority of examples which showed irregular use of grammatical particles involved a passivised verb in the subordinate VC. Speakers who produce language like this and who can readily construe meaning from sentences with contradictory marking, or with marking missing, must be doing so by ignoring grammatical marking altogether. This implies that some speakers rely solely upon their knowledge of the content words, pragmatic inference from the context and knowledge of the world to construe meaning from a sentence, as opposed to marking the semantic roles with syntactic features. This situation is not unique to Māori: Dorian, (2002), NeSmith (2007) and Palosaari & Campbell (2011) for example discuss observations of change in syntax and other language features for endangered languages in contact situations. Grinevald (1997) talks about several languages which have lost grammatical/morphological complexity through the levelling of paradigms after contact situations. She refers to grammatical features of Gaelic, Dyirbal and Oklahoma Cayuga which were formerly used to differentiate tense, case markers and verb structures respectively, but which are now used far less confidently by speakers of those languages, if at all.

Some of the changes described previously are clearly the result of Māori language contact with English, from semantic expansion of vocabulary to match English equivalents (as with *taea* in Chapter 4 and the *whāngai* example in Chapter 5), to the importation of facets of English syntax. It also may be due to the influence of English that a number of the grammatical distinctions are at risk of disappearing, such as the A/O distinction, which does not exist in English. However, some of the changes observed in this study

cannot be attributed to the direct influence of English; they have no model in English to which to conform. Instead these changes indicate an issue internal to the Māori language and its speakers. The inconsistencies in the use of passivised verbs and other syntactic features is perhaps the most significant example of this. Those examples show such inconsistencies with traditional Māori and confusion in the grammatical marking of different sentence functions that the reader/listener/addressee is forced to rely entirely on their knowledge of the meaning of the content words and the use of context to retrieve meaning. Where adequate contextual information is unavailable, speakers apply common-sense to guess meaning, as shown by my language-learning classroom observations discussed in Kelly (2014). I believe many of the changes like this pose the greatest threat to the survival of the Māori language, in that they ultimately undermine the language by increasing the levels of ambiguity and rendering the syntax unreliable, thereby inhibiting its effective, efficient and accurate use for the full range of purposes required of a vital language.

Attitudes to language learning and language change

As both a language learner and a language teacher myself, it is easy to see the appeal of a fail-safe default option in the grammar of the target language. Often in my Māori language classes, when learning a new construction or grammatical feature, my students will show great enthusiasm for any default option that means they 'can't be wrong'. It is understandable that the desire to not make a mistake is especially keen for heritage language learners, particularly those who see their mastery of the Māori language as a measure of their 'Māori-ness', so that a perceived lack of Māori language ability is also perceived as compromising their cultural integrity. Te Huia (2013) for example, in her PhD on establishing psychological foundations for higher levels of Māori language proficiency, references authorities such as Kāretu, Dewes, Brewer and Mead who describe proficiency in te reo Māori as 'central to many Māori feeling authenticated in their identity as Māori' (p. 71). This was observed in this study in the apprehension of some of the modern corpus contributors and the refusal of others to contribute material for the modern corpus, for fear their

self-assessed lack of Māori language skill might be exposed for scrutiny, causing them embarrassment. The link between Māori language ability and perceived 'Māori-ness' attracts many L2 speakers to learn the language - as Skerret (2011) explains, "Regarding the impetus to learn te reo Māori some of the key motivators referred to were: cultural identity; iwi epistemologies; values and forms of Māori language; desire for one's children to be Māori language speakers; whakamā or being embarrassed when unable to speak te reo Māori ..." (p. 77). It is perhaps ironic that the same link which motivates a speaker to learn the language may also contribute to the speaker (L2 or otherwise) hedging their bets with default, non-specific language forms, thereby both avoiding and also missing out on having to know how to use grammatical forms and other language features that would enable them to use the language with discernment to its fullest potential.

Given the fact that L2 learners make up the vast majority of the population of Māori language speakers (Statistics New Zealand, 2014), it is reasonable to expect that language behaviours common to them will have a significant effect on the language overall. When their behaviours lead to a decrease in the frequency of the use of context-specific language features, this in turn results in less exposure to these features for the language community, thus impinging on the ability of young L1 speakers and L2 learners to determine their appropriate use without overt teaching. The power of the L2 speaker to affect change in a language through the sheer force of numbers has been observed in other language revitalisation scenarios. As NeSmith (2007) states with regards to the situation in Hawaiian, "This is the power that L2 speakers wield in their control of formal education: the power to define the terms, and to change the language. Instead of native speech being the default, it is the Hawaiian language of L2 speakers that is now the default" (p. 6).

The corollary of acknowledging the power of the L2 speaker in affecting change in the Māori language is acknowledging the power of quality language teaching and learning as a catalyst for both language revival and language change. As noted in the introduction to this thesis, older native speakers of Māori have noticed change in the language, and frequently comment on the

changes they perceive, but the critical mass of L2 speakers which now dominates the Māori language population has fewer and fewer opportunities to engage with these older native speakers. Formal language instruction has been pivotal to the revitalisation of the Māori language thus far, by enabling Māori language-learning for those with limited natural exposure to the Māori language in the modern period. Likewise, the acceptance of or resistance to language changes in formal language instruction will inevitably be reflected in the language behaviours of the L2 speakers. Māori language teachers therefore bear an enormous responsibility to not only model quality Māori language for their students, but also to equip them with the necessary skills to be accurate and discerning users of the language. This responsibility is compounded when the classroom is the only environment in which those students can be made critically aware of how their knowledge of English can lead to inappropriate assumptions about how the Māori language should be used. The teacher and L2 learner attitudes towards change are therefore critical factors in the development of the Māori language in the modern period.

In order to describe the complexity of attitudes towards change in the Māori language, I find it helpful to consider two separate factors, namely the speaker's awareness of change in the language, and their acceptance or otherwise of that change. In relation to the first factor, a speaker's ability to notice non-traditional forms in te reo Māori and in their own language use may be influenced by factors like the following: the extent of their exposure to the Māori language, the nature of their own language acquisition, and their formal knowledge of the grammar of Māori. From my experience in the Māori language community, it is my assessment that the clear majority of Māori language speakers of the modern period do not perceive these non-traditional forms in their own language use, although some are aware of these issues in observing the language of others. This is consistent with Harlow's (2005) observation of learners who strenuously avoid borrowing vocabulary, but freely borrow elements of English syntax (p. 137-8).

With regards to the acceptance of change, there is a spectrum of attitudes in the community towards change in Māori language, on which I think it is

reasonable to distinguish two clusters. The first I will label 'traditionalist': although this cluster of people recognises the inevitable need for new vocabulary, they are highly resistant to any other perceived change in the language. They see the language revitalisation process as a means to restore the Māori language to its pre-European-contact state. The second I will label 'permissivist'; this cluster is situated at the opposite end of the spectrum of attitudes. They are accepting of change in the language. As O'Shannessy (2011) states in her discussion of how language change is evaluated, 'A new way of speaking may be seen as a threat to traditional languages, but can also be seen as a form of language maintenance' (p. 94). This permissivist position with respect to the Māori language community may be characterised by statements like the following:

"Adapt or die - language change is necessary for language survival."

"It's better to hear new Māori than no Māori."

"We need these new constructions because you just can't say that in Māori."

An advantage of the traditionalist position is that it is focused on the conservation of the linguistic traditions and features that distinguish Māori from English, and it is committed to preserving its linguistic diversity. The traditionalist position is therefore more likely to resist change resulting from the influence of English, thereby helping to conserve the linguistic distinctiveness of the Māori language. However, the issue with this position it is that while it does not accept change in the language as a result of language contact, nor does it allow for any natural internally-motivated change.

There is evidence that the Māori language was in fact in the process of changing in many areas of its syntax at or before the point of first contact with English (see for example Hohepa, 1969, Chung, 1978, and Harlow, 2007). This is not surprising, given that all living languages do and must change (Boyce, 1995). Examples of this include the use of matrix clause TAMs in

subordinate clauses, as discussed in the case study of relative clauses (see Chapter 5).

The extreme traditionalist position and its resistance to all language change then looks to fix the Māori language at a relatively arbitrary stage in its development, purely for the reason that contact with English occurred at that point. This does not allow for the natural development and continued evolution of Māori that would have been expected if language contact with English had never occurred. In my view this position is untenable: it is impractical to expect the Māori language to return to the state it was in prior to European contact, especially after a sustained period of dramatic language loss such as that observed between the periods represented by the two corpora.

In an educational context, the traditionalist's staunch opposition to all language change also provides little assistance to teachers to aid them in determining which elements of change should be prioritised for overt teaching in language classrooms. In fact, when teachers adopt the traditionalist standpoint, the result is often that they concentrate on the correction of those aspects of change which are most obvious to the speaker, but which are relatively trivial in that they have little bearing on the fundamental syntactic fabric of the language - thus Leech (2011) reinforces the potential for corpus-based study to inform language teaching practise when he points out

“[a]uthoritative’ figures in language teaching, whether teachers, materials writers or just native speakers, are very poor at guessing relative frequencies of grammatical classes and structures. If the time wasted teaching rather uncommon structures and weak rules is to be avoided, the ‘more frequent = more important to learn’ principle should be applied to grammar. This is where corpus evidence again becomes crucial” (p. 18).

It is ironic that in adopting a traditionalist's staunch opposition to all change in the interest of language maintenance, a teacher could spend too much time focussing on correcting changes in low-frequency items, and not have an opportunity to focus on the less obvious language issues and more insidious

aspects of syntactic language change which actually threaten the vitality of the Māori language by compromising its 'genius' (Harlow, 2005) and therefore its usefulness for the language community.

The permissivist view is a laissez-faire approach, and is therefore perhaps no more helpful in an educational setting. The broad acceptance of all change in the language does not foster critical awareness of one's language use, and inevitably results in a compromise of the integrity of the traditional syntax of Māori, resulting in ambiguity, contradictory syntactic rules, and a general breakdown of the system of language that is meant to be clarified through language instruction (cf. replacement of inherited grammar, Thomason & Kaufman, 1988).

While it is useful to separate out these two perspectives on language change, it is an individual's movement along the traditionalist – permissivist spectrum which produces the complex range of attitudes and behaviours evidenced in the community. This may be illustrated by the language speaker who is vocal in their opposition to change in the language, and who will actively avoid borrowed vocabulary, but whose own language use often includes non-traditional constructions [eg. *tatari mō*] (cf. Harlow's (2001) comment on seeing borrowed vocabulary vs borrowed grammar). A second illustration is the speaker who inadvertently uses a non-traditional form, and when made aware will actively try to avoid it in the future. In this situation although the speaker's attitude towards change remains the same, their awareness has changed, and therefore so does their output. One last illustration is that of the speaker who also unwittingly uses a non-traditional form, but who persists in its use, even when made aware of the issue.

What is clear from my observations is that a speaker's attitude towards language change may have little bearing on the nature of their language output – the fossilised or habituated language use of both L1 and L2 speakers of Māori for example is a common experience for many language teachers and learners. The commonly accepted view is that the influence of other languages upon Māori should be avoided wherever possible, but this seems inconsistent with the large number of changes observed in the language and

exemplified in the previous chapters. It is clear that change is in fact happening, regardless of the community's overt attitudes which advocate resistance to it.

There is potentially a middle ground to be negotiated between these two positions of traditionalist and permissivist, that would enable the preservation of the integrity of the Māori language while maintaining its ability to continue to develop and evolve to meet the linguistic needs of the language community. This requires the issue of language change to be addressed directly, whereby changes in the language are identified and a means devised for determining the acceptability or otherwise of the changes observed. Key factors for consideration have been proposed by Bauer (2007) and may include:

- the source or reason for the change – is the change the result of the influence of English, or is the change motivated by some factor internal to the Māori language?
- the extent of the impact of the change in the language – does this change have wider implications for the language system of Māori, or is it confined to limited contexts only?
- does the change make the language more or less regular?
- does the change involve replacing a traditional Māori form/ structure?

Other potential factors for discussion abound, but the intention here is to raise the issue of language change in Māori, and to raise the need for the Māori language community to address the phenomenon and begin to be deliberate and discerning in identifying and accepting, rejecting, or at least critically assessing different aspects of language change observed in the language community.

7. CONCLUSION

This thesis has successfully identified a number of aspects of syntactic change in te reo Māori over the period between 1900 and 1990, and tested a corpus-based methodology for its effectiveness in providing valid and reliable evidence to compare with anecdotal observations of language change.

METHODOLOGY

With regards to the methodology, the use of two small synchronic corpora has been shown to provide a useful diachronic perspective on the language, capturing statistically significant empirical evidence of aspects of change in some high frequency language forms and constructions. The issues in corpus construction have been detailed, including the difficulties encountered in the digitisation of archival texts, in determining data selection criteria and in maintaining the balance of the principles of representativeness and comparability. If the corpora were not representative and comparable, then the findings would not be valid; indeed some results were tempered by discrepancies in the corpora which could not be avoided, like the increase in the frequency of *taea* in the modern corpus, partially caused by the many texts based on the topic of Māori language-learning, or the increase of *mō* in the traditional corpus, due in part to the large number of sale notices in the older periodical texts. While these factors were taken into account in reporting the findings, there were also enough of my findings which were clearly statistically significant, enough to be confident of their validity even with some margin for variability. Other findings agree with anecdotal evidence, which is another assurance, albeit tentative, that the two corpora used were appropriately balanced and representative.

CHANGES OBSERVED

This comparative methodology has been comprehensively tested on three distinct language structures: a grammatical particle in the first case study on the preposition *mō*, a pair of semantically-linked lexemes in the second case study on *taea* and *āhei*, and a grammatical construction in the third case study on verbal non-reduced relative clauses. The *taea* and *āhei* case study highlighted the limitations on the effectiveness of the use of corpora of modest size, in that they failed to provide adequate data for these lower frequency lexemes to enable statistically valid evidence of systematic change, although differing patterns of use were suggested in the available data. The *mō* and relative clause case studies produced a much larger data set, and provided a number of statistically significant observations of change in the language. A significant increase was identified in some cases, such as in the use of the TAM *e...ana* in relative clauses. A significant decrease was observed in some cases, such as the use of *taea* as the passive form of ‘to arrive/ to reach’, and the frequency of use of the particle strategy for relativisations, although this may have been an artefact of the increasing use of *e...ana*. Evidence was also found of no significant change, such as in the use of matrix clause TAMs *ka* and *kua* in subordinate clauses, and perhaps more importantly, no significant increase in the use of relative clauses, refuting the anecdotal evidence claiming a marked increase in relative clauses in modern Māori.

Some observations were less statistically significant, but provided an important documentation of the emergence of potential aspects of language change. Evidence such as the increased use of *e...ana* and the decrease in the use of *ai* and other particles in relative clauses in the modern period suggest that the paradigm for relative clauses may be undergoing simplification. This study also provided evidence of the development of new senses of traditional Māori words. The first of these was the use of *taea* to indicate the politeness of a request, mimicking the semantic range of ‘can’ in English; the second example is the use of *mō* to mark the complements of the verbs *tatari* ‘to wait’ and *kimi* ‘to search’, replacing the traditional object markers *ki* and *i*, and most likely a direct transfer from the preposition ‘for’ used for both verbs in English.

From the results of the three case studies, two different trends in language change emerged. The first of these trends involved Māori becoming more like English, such as the observed shift in the use of *taea* from the traditional sense of 'ability to do something' or the accomplishment of an action, to express a polite request, mimicking the pragmatic effect of 'can' in English. This influence of English on te reo Māori of the modern period may result not only from historical language contact with English, but also from Māori language loss and/or interrupted transmission, evidenced by the different demographic of the Māori language community of the modern period, with L2 learners forming the overwhelming majority of the population. In his summary on contact-induced change, Trudgill (2011) discusses the relevance of the language acquisition profile of the speaker community, and the effects of L2 status on the loss of language complexity.

This leads to the second trend observed, toward changes which involved a loss of specificity and the dissolution of certain syntactic distinctions in Māori, like the marking of sentence function in relative clauses, and the use of the passive suffix in distinguishing active versus passive voice. The implications of these examples in te reo Māori were explored, and it was concluded that in fact it may not be the direct influence of English, but instead the potential for dissolution of important syntactic distinctions that poses the greatest threat to the revival and ultimately the survival of the Māori language. Trudgill argues that the loss of language complexity does not inevitably reduce the communicative power of a language (2011); it will be recalled that in Chapter 6 of this thesis I conclude with some aspects for consideration in order to potentially differentiate between those changes which pose no threat to the communicative power of te reo Māori, and those which potentially undermine its integrity, such as the issues mentioned previously regarding marking of sentence function and active versus passive voice. This thesis maintains that if left unchecked, this type of language erosion in te reo Māori has the potential to eventually render the language vague and ambiguous, and therefore inefficient and ineffective for communication with the clarity, subtlety and linguistic variety that a vital language community requires, and that inevitably influences their choice of which language to use.

AREAS REQUIRING FURTHER STUDY

There are a number of potential extensions to this study that were outside the scope of this project but would merit further investigation. The two corpora for example could be readily expanded to include some of the additional material for the traditional period that has been digitised since these corpora were constructed, and there is also much more material available for inclusion in the modern corpus in order to maintain equal corpora size to aid comparison. This may enable the effective study of some of the lower frequency language features, like *āhei*, that was not possible here. A second corpus of translated traditional material would also provide an interesting point of comparison with the modern corpus, especially with regard to some of the English-influenced language features being commented on anecdotally within the Māori language community.

This study has documented the laborious nature of the manual search, even of corpora as small as these. However, as mentioned previously, an expansion of the manual search for relative clauses to include Agent Emphatic constructions would provide a clearer insight into the use of the pronoun strategy, which merits further investigation. This study has also highlighted the value of the insights provided by such research about grammatical constructions which are often difficult for the typical language user to detect, and are overlooked by many teaching programmes and in the course of day-to-day language use. This speaks to the need for tagging software to be developed, to enable the effective automated tagging of large amounts of Māori language material, to facilitate the study of Māori grammar through the use of corpora. In this way the large body of language material available that exemplifies competent, specific and discerning use of the language can potentially continue to guide language learners of the modern period, who do not have the opportunity for exposure to the requisite mass of Māori-language material from adept native speakers.

This study has also challenged the reliability of speaker intuition and anecdotal evidence with respect to language trends - some of the aspects of the language that were anecdotally alleged by native speakers to have changed

have been shown to be no different from the traditional period, and my own intuitions about what was happening are not all supported either. Conversely some observations from the classroom have been verified in the language of the modern corpus. Many of the examples from the three case studies highlighted the issues that speakers experience in mastering the passive, especially in subordinate clauses. They indicate that passives are a major locus for change in the language, and warrant further study. There is other evidence of speakers struggling with the grammar of Māori, not due to the importation of English per se, but instead due to a lack of skills, or instruction, or exposure to Māori resulting in a grammar gap, which leads to a lack of specificity in the language, and that in turn forces the addressee to rely instead upon common-sense to interpret the utterance. This is an important area for future research, and another means by which linguistic research can highlight a need and make a meaningful contribution to quality Māori-language teaching and learning.

This study has highlighted the influence of the L2 speaker on the Māori language as a whole, and the potential for the L2 speaker to meaningfully advance the successful revitalisation of the language through quality teaching and learning and their overwhelming statistical dominance in the Māori language population. This has in turn highlighted the need to identify aspects of language change (hence the usefulness of this methodology), and determine which language changes are acceptable and which should be rejected, to help inform language teaching and enable learners to use the language with clarity, specificity, and linguistic confidence – thereby contributing to language revitalisation by reinforcing some of the very features of Māori that would serve as markers of its continued and renewed vitality.

APPENDIX 1: Sample of *mō* data collection

Classification	Source (corpus)	Examples
Position	Traditional	kei au nga patiki MO nga hoiho. Homai te ki o te pataka MO nga hamanu
	Modern	ka whakairohia he tūnga MŌ te Paipera Tapu i Rangiātea he paraki ki runga hei tūnga MŌ ngā kēna. te wāhi i whakaritea MŌ te whare karakia
Nominalisation	Traditional	rawa te wa hei haerenga atu MO matou rangatira tera he whawhaitanga MO Ngati Apa mo te pupuri hei nohoanga MO nga Maori
	Modern	ko Ōtaki hei taunga atu MŌ ngā iwi kua tae ki te wā hei haerenga MŌ Kui ki te kura i Tāmaki-makaurau whakatū hāpori hei waihotanga MŌ te hunga e whai ake.
Future possession	Traditional	kia haere mei hei hoa MO ratou. hakatakoto e koe tetahi whare MO matou, Homai tetahi koti moku, kia wha nga koti,
	Modern	te reo Pākehā hei reo matua MŌNA rangi o te Paipera, he kāmata MŌ te minita. Hangā he whare karakia MŌ tātou...
On behalf of/for	Traditional	kai mate a hau MO tenei tangata e takoto atu nga Pakeha kia tukua kautia MO nga Maori? ko wai hei Kingi MO ratou
	Modern	roto i taku kawenga hei Toihau MŌ Te Taura Whiri i te Reo Māori hei minita MO Ngāti Raukawa me tōna pakari hei kaihautu MŌ te iwi.

Reason	Traditional	<p>aromia ana nga tangata whenua MO a ratou kino.”</p> <p>Ka tohe au, a Pita hoki MO to teka ki a au.</p> <p>ka nui te pouri o taku ngakau MO te korero i korerotia</p>
	Modern	<p>wā i rīria rā a Naida Glavish MŌ tana kī “Kiaora”</p> <p>me mihi ka tika ki a koe MO ngā mahi kua mahia e koe!</p> <p>uruhau ki a koe, MO ngā korero, MO ngā whakaaro i whārikihia</p> <p>whakamihi a te whānau ki a ia MŌ tana hua pākeke mo te māia, MO te kaha, MO te aroha, MŌ te titikaha</p>
About	Traditional	<p>tou i a koe, kia korero tatou MO Patea, MO Otara.</p> <p>i taku korero MO nga tangata, MO te tikanga o tau korero,</p> <p>Kahore aku kupu MO te whenua i tai...</p> <p>Ka korero atu au ki a koe MO te whenua</p>
	Modern	<p>kōrero a Tama MŌ tōna hōiho, MŌ Tamati.</p> <p>maha tonu ngā waiata i titoa MO te mauheretanga nei.</p> <p>pūrongo whānui e kōrero ana MŌ Microsoft Vista me Office 2</p>
Transaction	Traditional	<p>kotahi pauna, L1. 11. MO nga taura, paraire, kopare,</p> <p>ia e pono tonu: tekau ma rua MO te peke parawa,</p> <p>He utu tenei MO nga whakaaro o Kawana.</p> <p>e £9 pauna MO te takiwa i tu ai nga whare</p>
	Modern	<p>au hēkona, e rima tekau taara MŌ ia kupu ka mau i a ia.</p> <p>I hokona mai e Kāwana Kerei MŌ te £15/15/-.</p>
Purpose	Traditional	<p>i nga Maori kia keria te awa MO te wira o te wai.</p> <p>ki te Pakeha nana hei hanga MO te mira.</p> <p>i te oranga ki a maua, MO te whakanuinga hoki i a mau nga Hoia me a ratou pu MO te whawhai.</p>
	Modern	<p>ngā āhuatanga katoa e tika ana MŌ te taka kai.</p> <p>i ake i ngā wāhi e tukuna ana MŌ te ako i te reo.</p> <p>haere tonu te whakatika a Hori MŌ te hoki me tana kite atu</p> <p>mata tika anō to te karawhaea MŌ te tāpuke i ngā parareka.</p> <p>Ko te wā i whakaritea MŌ te haere ehara i te wā tika te mānukura kua whakatakotoria MŌ te tautoko i ngā mahi</p> <p>Kua reri anō te wharekuau MŌ te miraka o te ahiahi</p>

Duration	Traditional	<p>hoki kia mutu ta raua whawhai MO nga wiki e toru. wahine e tiakina maretia ana MO tetehi takiwa A kei te niunga o taua moni MO te tau kotahi, a whakatako[to] ai te tuatahi MO te tau 1850,</p>
	Modern	<p>He haerenga MŌ te roanga o te rā tēnā, e toru, i whakariterite haere MŌ te tau 2011. ko ētehi whakangungu MŌ te tekau wiki ki te whakapa I whakarerea te Whare Wānanga MO te wā poto. i taua ngārara, nā, ia wiki MŌ ngā pō e toru</p>
Future location	Traditional	<p>tonu ana hei whakatupu whawhai MO a mua atu? E hoa, MO ahea koe hoki mai ai? ka tapahia ka waiho marie, MO a mua ka rui ano ki te whenua aro noa te tini o nga mea pai MO roto o te whare Ka whakaae mai a Pare, MO te Parairi ka haere atu au.</p>
	Modern	<p>me ngā kākaho te kohikohi MŌ roto i ngā tukutuku. ngā tukutuku me ngā kowhaiwhai MŌ roto i te whare.</p>

APPENDIX 2: 'Taea' & 'Āhei' data collection

Example	Sense	C Subj = nominalisation of ability?	C Agent = PP?	C Agent prep?	Ability DO = PP?
te matakitaki haere nei ki to lwi ka ngaro nei e TAEA hoki koa te aha e taua	capable	yes	y	e	
Otiia me pehea e TAEA ai i te nui o tona mana ki Ingarani me Aerana Nuitireni me era atu wahi katoa o te ao e noho nei i raro i tona Kingitanga.	capable	not specified	/		
tena ano etahi take e mate ai ratou, erangi ka TAEA ano hoki aua take te whakakore atu	capable	yes	n		
Na, ka whakahaerea mai e ahau taku mohiotanga ki enei, ki te tahuri ke i ahau, ki te riri, ki te aha, ki te aha, a, TAEA noatia taku kororonga tuatahitanga atu ki a koe	arrive				
E mohio ana koutou e kore a koutou ngakinga e TAEA te tango e te Pakeha, me nga wahi whenua kihai i hokona	capable	??			
E TAEA hoki e ahau te pehea?	capable	yes	y	e	
Otira kei a koe ano te whakaaro, e TAEA koa te aha i a koutou tikanga?	capable	yes	/		i
Kia wawe au te riro ke atu i te he, kia TAEA atu au ki tetahi whenua kore he.	arrive				
Hei te 10 Noema te TAEA ai tana pukapuka.	tāia?				
Ki taku mahara, e kore e rite he whakaaro utu ki a koutou, no te mea kua roa ke te takotoranga, e kore e TAEA nga utu o a koutou poaka i nga Pakeha, i a ia	capable	no	y	i	
Me aha i te aroha ka kaipuke, i TAEA te huna iho?'	capable	yes	/		
Me ki atu koe ki a Kawana, ka tupu te raruraru i konei, e kore e TAEA te whakaaro matou ko nga Pakeha.	capable	yes	n		

E hoa, tenei ano nga korero nunui e kore e TAEA e au te korero ki tenei pukapuka.	arrive				
Taku mohio, TAEA noatia tenei he, kei te takoto kau noa iho te he o Rawiri	arrive				
Ko tenei pukapuka me pupuri tonu e koutou, kei maka, kei ngaro ranei, me tiaki tonu TAEA noatia te mutunga.	arrive				
E TAEA hoki koa te aha i aitua, i te ringa kaha o te Atua ki te tango atu?	capable	yes	/		i
Kaore he tara i au, erangi nga moni kaore e TAEA te tiki e au i roto i te peke	capable	yes	y	e	
Ka mea atu au, 'E kore e TAEA te wiu puta noa i nga whenua katoa.'	capable	yes	/		i
E TAEA hoki koa, e te iwi, te aha, na te kino tona ka wenga mai ki whenua o tangata	capable	yes	y	e	
E kore au e pai ki te whawhai ki te Maori, aha koa TAEA noatia te 22 o nga tau e kore ano au e kino, ara e whawhai ki te Maori	arrive				
I te whitinga atu o nga hoia i te awa o Mangatawhiri a TAEA noatia tenei mahi, kaore kau he kohuru a te Maori ki te Pakeha	arrive				
ka whakaaro hoki au, ki te marino te moana ka TAEA atu he oranga moku	arrive				
na te taimaha o nga hanga totohu tonu iho, no te timunga o te Tai ka TAEA ano te tari mai ki uta.	capable	yes	/		

APPENDIX 3: Relative clauses - problematic examples

Example	Antecedent	subTAM?	Verb?	Particle?	Agent in rel cl? - Marking?	Patient in rel cl? / Marking?
Kia tere te hunga i tō tauira ka tuku ai ki a <i>Urban Pasifica</i> hei tuku atu ki te rōpū pūoro	patient	ka	active	ai	-	-
He pūtoi paepae hou kua whakapai ake a <i>Windows7</i>	unsure					
He wāhanga Aromatawai hou, he tāpiri whakamāhere i hangaia houtia e whakaritea ana i a koe te whakaemi rauemi	adverbial	eana	passivised	*	-	i
Tirohia ngā āhuatanga hou e whāngaia ana a <i>Microsoft Windows7</i>	unsure					
ehara i te mea kua takahi ngā kupu nei i te ara tē taea te whakahoki mai	a?	tē	passivised	-		
ko te whakahua tika i te reo Māori, me ērā momo āhuatanga katoa e taea te ako te reo Māori ki te hunga e hiahia ana	agent	eana	passivised	*	-	-
E tangi ana Te Taura Whiri i te Reo Māori i te korowai o aituā e kākāhutia ana e te manawa me te roimata	adverbial	eana	passivised	*	-	e
I ēnei rā ko te Haka he huarahi hoki e whakaputaina e te rōpū tāngata ngā tuāhua kaupapa katoa e pā ana ki tōna ao	adverbial	e	passivised	-	e	X
Kei te rongo hākiri tonu ki ngā tapuwae o rātou i tīmatahia ai te whakanui i te rā mō te reo rangatira	agent	i	passivised	ai	-	X
kei te tangi tonu ki a koutou mā i takina ai tēnei kaupapa	agent	i	passivised	ai	-	-
Ki te kite koe i ētahi rōpū i roto i tō rohe e whakanuihia pēnei ana te reo	adverbial	eana	passivised	*	-	X
He tino huarahi e whakaatuhia atu te mana o tōna iwi	adverbial	e	passivised	-	-	X
Hei tā Te Whakaruruhau, ngā kaiwhiwhi tohu mō Te Tohu Hapori, ko te take e whakanuia ai e Te Taura Whiri i te Reo Māori i ngā mahi i oti i a rātou katoa i te Wiki o Te Reo Māori te mea nui kē	adverbial	e	passivised	ai	e	i

he nui ngā mea i ako au	patient	i	active	-	X	-
Ko te rōpū tuarua i tūtaki au ko Wharerongo Pacifica me ā rātou tonu i rite nei ki ā te hunga o ngā kōhanga reo	patient	i	active	-	X	-
Ka whakairihia ēnei kōrero i runga i tā te Minita i whakaputa i taua pō rā	patient	i	active	-	-	-
kua tutuki te wāhi i hōmai rātou māhau e mahi	patient	i	active	-	X	-
i ngā kau ka whiu ki te wharekau	patient	ka	active	-	-	-
Ka tae ki te rā i whakamaua ngā tīni ki taku kara,	adverbial	i	passivised	-	X	-
ngā kaihiaki e rongonuitia ana, ko ngā manu, ngā kararehe...	patient	eana	passivised	*	-	-
he tangata ia ka taea te taraiwa kāpō mai i Wharekāhika ki Te Araroa	adverbial	ka	passivised	-	-	-
pērā i te wā i kitea tuatahitia e Saint Columbia	adverbial	i	passivised	-	e	-
I te wā i uru ahau ki te whakapono Karaitiana..	adverbial	i	active	-	X	-
puta noa i te rā nei ngā kōrerorero, ngā kauhau e whakamāramatia ana i ngā whakaaro	agent	eana	passivised	*	-	i
me te whakawhētai a te whānau mōna ka horahia ngā kai mō te hākari whānui	agent	ka	passivised	-	X	-
te hangaia he māhere e kapohia ana i ngā painga o tēra ao	adverbial	eana	passivised	*	-	i
e hia kē nei ngā tau āu e tōtā ana te rae kia mau i a ia	agent	eana	state intransitive	*	X	-
koirā tērā te whakaaro e ngāwari ai ki te tauawhi i ngā mahi o te kāwana i ēnei rā	adverbial	e	state intransitive	ai	ki	i
i kōrero ia mō te whiwhinga o tētahi pukapuka i tuku atu ki tōna koro	patient	i	active	-	-	-
ki te whakawhiti kōrero e pā ana ki ngā mahi e whakarite nei e rātou	patient	e	active	nei	e	-
ko ngā tikanga ture e whai ana tēnei motu	patient	eana	active	*	X	-
te utu i kawē ngā mōrehu nei	patient	i	active	-	X	-
kia eke ai mātou ki ngā taumata e hiahia ana	patient	eana	active	*	-	-
Paihere te aroha, paihere te mana o ngā tauria i whakakotahi ai i runga i te karanga o ngā tuākana ki ngā teina	patient	i	active	ai	-	-

e kore rawa tēnei āhuatanga e whakahē ana i taua wā	patient	eana	active	*	-	-
pērā i te waiata i tito mai i Te Panekiretanga o te Reo	patient	i	active	-	i	-
AE, KAO rānei ko tāu e rapu ko te Waka Ama!	patient	e	active	-	-	-
Nā te aukati o te wheketere, tokomaha te hunga i whakarerea e te mahi	patient	i	passivised	-	e	-
Tokomaha tātou te iwi Māori e pāngia nei tēnei mate	patient	e	passivised	nei	X	-
ki tōna hoa rangatira ki a Mere Rangi i mate ai ia i te mate manawa pērā ki a Taini	agent	i	state intransitive	ai	X	
ko ngā mōteatea e wetekina ana i ngā kōrero whaiāpō, kōrero ekeeke hoki	agent	eana	passivised	*	-	i
i te reo Māori e kaha whakarere ana i te arorangi	patient	eana	active	*	-	-
te ahorangi a Teina Smith me te whānau whakamarumarū i hāpai, i tautoko hoki te kaupapa nei	agent	i	active	-	X	-
ki ngā pou o te reo Māori, te hunga kua roa nei e 'tutungia te hatete' o te reo	agent					
i ngā whānau Māori e whakarerea ana i te hunga kua whai i tēnei ara	agent	eana	passivised	*	-	i
e taea ai te kōrero, te rongō, te kite hoki i te reo e whakamahi ana	patient	eana	active	*	-	-
I runga anō i te maha o ngā kaupapa whakaora reo e karapotia nei e te motu	agent	e	passivised	nei	-	e
ngā mahi kua tutukia e mātou	patient	kua	p-si	-	e	-
te autaiā hāngi kua whakatīnana haere nei	patient	kua	active	nei ?	-	-

REFERENCES

- Aston, G. (1997). Large and small corpora in language learning. In B. Lewandowska-Tomaszczyk & P. J. Melia (Eds.), *PALC97: Practical applications in language corpora* (pp. 51-62). Łódź, Poland: Łódź University Press.
- Bauer, L. (1993). *Manual of information to accompany The Wellington corpus of written New Zealand English*. Wellington, New Zealand: Dept. of Linguistics, Victoria University of Wellington.
- Bauer, L. (1994). *Watching English change: An introduction to the study of linguistic change in the twentieth century*. London, New York: Longman.
- Bauer, L. (2003). *Introducing Linguistic Morphology* (2nd ed.). Edinburgh: Edinburgh University Press.
- Bauer, W. (1982). Relativisation in Māori. *Studies in Language* 6, 305-342.
- Bauer, W. (1997). *The Reed Reference Grammar of Māori*. Auckland, New Zealand: Reed Publishing.
- Bauer, W. (2007). *Typology to the rescue: Halting the infiltration of English into Māori syntax*. Plenary paper presented at the AFLA XIV Conference, Montreal, Canada, May 2007.
- Bauer, W. (2014). In A. Onysko, M. Degani & J. King (Eds.), *He Hiringa, He Pūmanawa - Studies on the Māori Language* Wellington, New Zealand: Huia Publishers.
- Benton, R. (1991). *The Māori language – Dying or reviving?* Honolulu: East-West Center Association.
- Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing*, 8(4), 243-57.
- Biber, D., Conrad, S. & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Biggs, B. G. (1958). *The structure of Maaori* (PhD thesis). University of Indiana, USA.

- Biggs, B. G. (1968). The Maori Language Past and Present. In E. Schwimmer (Ed.), *The Maori People in the Nineteen-Sixties* (pp. 65-84). Auckland, New Zealand: Blackwood and Janet Paul.
- Biggs, B. (1973). *Let's learn Maori: A guide to the study of the Maori language*. Auckland: Auckland University Press.
- Bowern, C. (2008). *Linguistic fieldwork: A practical guide*. Basingstoke, Hampshire; New York: Palgrave Macmillan.
- Boyce, M. (1995). Languages in Contact I: Creating new words for Maori. *New Zealand Studies* 5(2): 13-18.
- Boyce, M. (2006). *A corpus of modern spoken Māori* (PhD thesis). Victoria University of Wellington, New Zealand.
- Chapin, P. G. (1974). Proto-Polynesian *ai*. *The Journal of the Polynesian Society* 83(3), 259-307.
- Chung, S. (1978) *Case Marking and Grammatical Relations in Polynesian*. Austin and London: University of Texas Press
- Clark, M. J. & Randal, J. A. (2004). *A first course in applied statistics: With applications in biology, business and the social sciences* (2nd ed.). North Shore, New Zealand: Pearson.
- Conrad, S. (2000). Will Corpus Linguistics Revolutionize Grammar Teaching in the 21st Century?*. *TESOL Quarterly* 34(3), 548-560.
- Corder, S. P. (1973). *Introducing applied linguistics*. Harmondsworth: Penguin Education.
- Dorian, Nancy C. (2002) Commentary: Broadening the Rhetorical and Descriptive Horizons in Endangered-Language Linguistics. *Journal of Linguistic Anthropology* 12 (2): 134-140
- Education Review Office (1995). *Kura Kaupapa Māori*. Retrieved from <http://www.ero.govt.nz/ERO/Publishing.nsf/Print/Kura%20Kaupapa%20Maori>

- Education Review Office (2002). *The performance of kura kaupapa Māori*. Retrieved from <http://www.ero.govt.nz/ERO/Publishing.nsf/Print/The%20Performance%20of%20Kura%20Kaupapa%20Maori>
- Flowerdew, L. (2002). Corpus-based analyses in EAP. In J. Flowerdew (Ed.), *Academic discourse* (pp. 95-114). London: Longman.
- Foster, J. (1987). *He whakamārama: A new course in Māori* (2nd ed.). Auckland, New Zealand: Reed Publishing (NZ) Ltd.
- Foster, J. (2012). *He whakamārama - A full self-help course in Māori* (4th ed.). Auckland, New Zealand: Pearson Education New Zealand.
- Gardner, D., & Davies, M. (2007). Pointing Out Frequent Phrasal Verbs: A Corpus-Based Analysis. *TESOL Quarterly* 41(2), 339-359.
- Garlick, J. (1998). *Māori language publishing: Some issues*. Wellington, New Zealand: Huia Publishers.
- Grinevald, C. (1997). Language contact and language degeneration. In F. Coulmas (Ed.), *The Handbook of Sociolinguistics* (pp. 257-270). Oxford: Blackwell.
- Harlow, R. (1993). Lexical Expansion in Maori. *Journal of the Polynesian Society* 102(1), 99-107.
- Harlow, R. (2001). *A Māori reference grammar*. Auckland, New Zealand: Pearson.
- Harlow, R. (2005). Covert attitudes to Māori. *International Journal of the Sociology of Language* 172, 133-147.
- Harlow, R. (2007). *Māori: A linguistic introduction*. Cambridge: Cambridge University Press.
- Harlow, R., Keegan, P., King, J., Maclagan, M., Quinn, E. & Watson, C. (2004). *NZE influence on Māori pronunciation over time*. Paper presented at the Language and Society Conference, Palmerston North, New Zealand, September 2004.
- Harlow, R., Keegan, P., King, J., Maclagan, M., & Watson, C. (2010). The changing sound of the Māori language. In J. N. Stanford & D. R.

- Preston (Eds.), *Quantitative sociolinguistic studies of indigenous minority languages* (pp. 129-152). Amsterdam, The Netherlands: John Benjamins Publishing Company.
- Hasund, K. (1998). Protecting the innocent: The issue of informants' anonymity in the COLT corpus. In A. Renouf (Ed.), *Explorations in Corpus Linguistics* (pp. 13-28). Amsterdam, The Netherlands: Rodopi.
- Head, L. (1989). *Making Māori sentences*. Auckland, New Zealand: Longman Paul Ltd.
- Heine, B. & Kuteva, T. (2010). Contact and grammaticalization. In R. Hickey (Ed.), *The handbook of language contact* (pp. 86-105). Chichester, West Sussex; Malden, MA: Wiley-Blackwell.
- Hohepa, P. W. (1967). *A profile generative grammar of Maori*. Supplement to *International Journal of American Linguistics* 33(2). Indiana University Publications in Anthropology and Linguistics Memoir 20. Baltimore: Indiana University Press.
- Hohepa, P.W. (1969) The accusative to ergative process in Polynesian languages. *Journal of Polynesian Society* 78, 295-329.
- Hooper, R. (1974). Neuter verbs, stative aspect, and the expression of agency in Polynesian. *The Journal of the Polynesian Society* 93(1), 39-70.
- Huia Publishers (2008). *Te reo whakaputa o te tamaiti: Corpus case studies of children's productive Māori*. (Unpublished research report; principal researcher). Wellington: Huia Publishers.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Jacob, H. (2012). *Mai i te kākano*. Ōtaki, New Zealand: Te Takupu, Te Wānanga o Raukawa.
- Johanson, L. (2002). Contact-induced change in a code-copying framework. In M. C. Jones & E. Esch (Eds.), *Language change: The interplay of*

- internal, external, and extra-linguistic factors* (pp. 285–313). Berlin, New York: Mouton de Gruyter.
- Kārena-Holmes, D. (2006). *Māori language: Understanding the grammar*. Auckland, New Zealand: Reed.
- Keane-Tuala, K. (2013). *Māori vocabulary: A study of some high-frequency homonyms* (MA thesis). Victoria University of Wellington, New Zealand.
- Keenan, E., & Comrie, B. (1979). Data on the noun phrase accessibility hierarchy. *Language* 55, 333-351.
- Kennedy, G. (1998). *An introduction to corpus linguistics*. Harlow, Essex: Addison Wesley Longman.
- Kelly, Karena (2014). Iti te Kupu, Nui te Kōrero: The Study of the Little Details that Make the Māori Language Māori. In R. Higgins, P. Rewi, & V. Olsen-Reeder (Eds.), *The Value of the Māori Language: Te Hua o te reo Māori*, (pp. 255-268). Wellington, New Zealand: Huia.
- Labov, William (2001). *Principles of linguistic change, Volume 2: Social Factors*. Blackwell: Oxford
- Leech, G. (2011). Frequency, corpora and language learning. In F. Meunier, S. De Cock, G. Gilquin & M. Paquot (Eds.), *A taste for corpora: In honour of Sylviane Granger* (pp. 7-32). Amsterdam, The Netherlands; Philadelphia: John Benjamins Publishing Company.
- Lyons, J. (1977). *Semantics* (Vol. 2). Cambridge: Cambridge University Press.
- Maclagan, M., Harlow, R., King, J., Keegan, P., & Watson, C. (2004). New Zealand English influence on Māori pronunciation over time. *Te Reo* 47, 7-27.
- Martin, P. W. (1995). Whither the Indigenous Languages of Brunei Darussalam? *Oceanic Linguistics*, 34(1), 27-43.
- Maunsell, R. (1842). *A grammar of the New Zealand language*. Auckland, New Zealand: J. Moore.

- Mauranen, A. (2011). Learners and users – Who do we want corpus data from? In F. Meunier, S. De Cock, G. Gilquin & M. Paquot (Eds.), *A taste for corpora: In honour of Sylviane Granger* (pp. 155-172). Amsterdam, The Netherlands; Philadelphia: John Benjamins Publishing Company.
- McCartey, T. L. (2010). Revitalising indigenous languages in homogenising times. *Comparative Education* 39(2), 147-163. DOI: 10.1080/03050060302556
- Ministry of Education. (2010). *High frequency word lists*. Retrieved from <http://tereomaori.tki.org.nz/Teacher-tools/Te-Whakaipurangi-Rauemi/High-frequency-word-lists>
- Moorfield, J. C. (2003). *Te Aka Māori dictionary*. Retrieved from <http://www.maoridictionary.co.nz/>
- Neff van Aertselaer, J. & Bunce, C. (2011). The use of small corpora for tracing the development of academic literacies. In F. Meunier, S. De Cock, G. Gilquin & M. Paquot (Eds.), *A taste for corpora: In honour of Sylviane Granger* (pp. 63-83). Amsterdam, The Netherlands; Philadelphia: John Benjamins Publishing Company.
- NeSmith, R. K. (2007). *Tūtū's Hawaiian and the Emergence of a Neo Hawaiian Language*. University of Hawai'i.
- NeSmith, R. K. (2007) *Tūtū's Hawaiian and the Emergence of a Neo Hawaiian Language* (2009 ed) Retrieved from: http://www.academia.edu/7564417/Tūtūs_Hawaiian_and_the_Emergence_of_a_Neo_Hawaiian_Language
- Ngaha, A. (2014). Te reo Māori and Māori Identity: What's in a Maunga? In M. Kawharu (Ed.), *Maranga Mai! Te reo Māori and Marae in Crisis?* (pp. 71-96). Auckland, New Zealand: Auckland University Press.
- O'Shannessy, C. (2011). Language contact and change in endangered languages. In P. Austin & J. Sallabank (Eds.), *The Cambridge Handbook of Endangered Languages* (pp. 78-99). New York: Cambridge University Press.

- Palosaari, N. & Campbell, L. (2011). Structural aspects of language endangerment. In P. Austin & J. Sallabank (Eds.), *The Cambridge Handbook of Endangered Languages* (pp. 78-99). New York: Cambridge University Press.
- Parkinson, P. G., & Griffith, P. (2004). *Books in Māori, 1815-1900: An annotated bibliography*. Auckland, New Zealand: Reed.
- Picoche, J., & Marchello-Nizia, C. (1991). *Histoire de la langue française*. Paris, France: Nathan.
- Quirk, R., Greenbaum, S., Leech, G. & Svartvik, J. (1972). *A Grammar of Contemporary English*. London: Longman.
- Reedy, T. (1979). *Complex sentence formation in Maori* (PhD thesis). University of Hawai'i.
- Saville-Troike, M. (1982). *The ethnography of communication: An introduction*. Oxford: Blackwell.
- Skerrett, M. (2011). Whakamanahia Te Reo Māori: He Tirohanga Rangahau – A review of literature on the instructional and contextual factors likely to influence Te Reo Māori proficiency of graduates from Māori Medium ITE programmes. Retrieved from:
<http://www.teacherscouncil.govt.nz/sites/default/files/Whakamanahia%20Te%20Reo%20Maori%20-%20He%20Tirohanga%20Rangahau%202011.pdf>
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Spolsky, B. (2003). Reassessing Māori regeneration. *Language in Society* 32(4), 553-578.
- Statistics New Zealand (2014). *Ka māro te aho tapu, ka tau te korowai: Te reo Māori findings from Te Kupenga 2013*. Wellington: Author.
- Stephens, M., & Boyce, M. (2013). *He papakupu reo ture: A dictionary of Māori legal terms*. Wellington, New Zealand: LexusNexus.

- Stiles, D. B. (1997). Four successful indigenous language programmes. In J. Reyhner (Ed.), *Teaching indigenous languages* (pp. 428-462). Flagstaff, Arizona: Northern Arizona University.
- Summers, D. (1991). *Longman/Lancaster English language corpus: Criteria and design*. Harlow, England: Longman.
- Te Huia, A. (2013). *Whāia te iti kahurangi, ki te tuohu koe me he maunga teitei: Establishing psychological foundations for higher levels of Māori language proficiency* (PhD thesis). Victoria University of Wellington, New Zealand.
- Te Taura Whiri i te Reo Māori (1992). *Te matatiki: Ngā kupu hou a te taura whiri i te reo Māori*. Wellington, New Zealand: Author.
- Te Taura Whiri i te Reo Māori (2008). *He pātaka kupu: Te kai a te rangatira*. North Shore, New Zealand: Raupo.
- Te Taura Whiri i te Reo Māori (2012). *Guidelines for Māori language orthography*. Wellington, New Zealand: Te Taura Whiri i te Reo Māori.
- Thomason, S. (2001). *Language contact: An introduction*. Edinburgh: Edinburgh University Press.
- Thomason, S. & Kaufman, K. (1988). *Language Contact, Creolization, and Genetic Linguistics*. California: University of California Press.
- Trudgill, P. (2011). *Sociolinguistic Typology: Social Determinants of linguistic Complexity*. Oxford, United Kingdom: Oxford University Press.
- Vaughn, E. & Clancy, B. (2013). Small corpora and pragmatics. In J. Romero-Trillo (Ed.), *Yearbook of corpus linguistics and pragmatics 2013* (pp. 53-73). New York; London: Springer.
- Waitangi Tribunal (2013). *Matua rautia: The report on the kōhanga reo claim*. Lower Hutt, New Zealand: Legislation Direct.
- Watson, C. (2009). *MAONZE*. Retrieved from <http://homepages.engineering.auckland.ac.nz/~cwat057/MAONZE/purpose.html>

- Williams, H. W. (1961). *The early journals of Henry Williams*. L. M. Rogers (Ed.). Christchurch, New Zealand: Pegasus Press.
- Williams, H. W. (1971). *A dictionary of the Māori language*. Wellington, New Zealand: Government Printer.
- Williams, W. L. (1862). *First lessons in the Maori language: With a Short Vocabulary*. London: Trubner & Co.
- Williams, S. (1990). The Maori Language Printed Collections. *The Turnbull Library Record*, 23, 12.
- Winitana, C. (2011). *My language, my inspiration*. Wellington, New Zealand: Huia Publishers.