Trends in New Zealand English: Some Observations on the Presence of Maori Words in the Lexicon

John Macalister

1. Introduction

The English language is recognised as the nearest the late twentieth century has to a global *lingua franca*. However, this should not mask the fact that there exist numerous regional varieties of English, and that a regional variety, by definition, contains unique elements. For New Zealand English, the main feature that distinguishes it from other regional varieties is the lexical borrowing from the indigenous Maori language. Indeed, Deverson (1991:18) identified this distinguishing feature as being 'in reality the most unmistakably New Zealand part of New Zealand English'. He speaks further of a 'current renaissance of Maori language and culture' with the result that 'Maori is making its presence in English more strongly felt than ever before' (1991:19) and identifies a consequent 'strengthening of the Maori component' in New Zealand English (1991:21). Orsman also perceives 'ongoing bilingual interchange with Maori' as one of the factors ensuring 'the future of a distinctive written and spoken New Zealand English' (1997:vii).

In this paper, then, two chronologically distinct corpora are compared and corpus-based analysis is used to examine two issues to establish whether the presence of Maori in New Zealand English is indeed an ongoing and strengthening process:

- are there changes in the pattern of use, or their distribution in particular semantic categories, of Maori words?
- are there changes in the nature of lexical borrowing from Maori?

2. The corpus[1]

Since its first publication in 1907, the *School Journal* has become a New Zealand icon. In various ways it encapsulates the evolution of New Zealand since 1907. As a written medium which has been a potential influence in shaping and reinforcing notions of language use, it is also a resource for measuring trends in English. Furthermore, as the *Journal* is written, overwhelmingly, by New Zealand writers for New Zealand children the assumption is that the language used is familiar to and understood by the readers, that it is, in fact, New Zealand English. From corpus analysis of the *Journal*, therefore, it is possible to identify changes in the patterns of use, and the treatment of Maori words in New Zealand English.

[1] The creation of this corpus was made possible with the co-operation of the *School Journals* editorial staff at Learning Media. I would like to acknowledge, in particular, the support, encouragement and involvement of Patricia Glensor.

3. Methodology

A computer corpus of *Journals* was created for two years of publication of each part from 1997–1999.[2] In creating the corpus the following were excluded:

- covers
- contents
- acknowledgements inside back cover
- any item written wholly in Maori rather than English[3]

In addition, although present in the corpus, extended stretches of Maori, such as waiata, where Maori words were clearly not being used as part of New Zealand English[4], were not included in the count of word tokens and types.

Furthermore, as there is some overlap between Maori and other Polynesian languages, decisions had sometimes to be made as to whether a word was Maori or not. While personal names that could be Maori but for which there was no identifying context were included, collocation work did result in the exclusion of a number of types and tokens. Tokens of *Tama*, for example, were not counted when collocation work revealed them to be a shortened form of *tamagouchi*. Similarly, types such as *ika* and *kai* were excluded when they were found to be present as Tongan rather than Maori.

The same exclusions and guidelines applied to the *Journals* for 1967–1968, for which a computer corpus was not, however, created. A manual search for and count of Maori words was performed.

4. The *Journals* of the 1990s: initial findings

The initial exploration of the corpus treated Parts One and Two and Parts Three and Four as separate corpora. This is in line with the editorial division at Learning Media, the publishers of the *School Journals*, and the popular idea of there being 'junior' and 'senior' journals. The real interest, however, lay in treating the *Journals* as a single corpus of 142,647 word tokens.

When the words of Maori origin were categorised, the categories used by Kennedy and Yamazaki (1999) were retained, with the addition of one extra category, 'Other proper nouns'. This added distinction allowed a more precise description of actual use, so that, for example, *kiwi* when used for the bird

---

[2] The corpus for Parts One and Two ran from the first issue of 1997 to the final issue of 1998. The corpus for Parts Three and Four began with the second issue for 1997 and concluded with the first issue for 1999. Thus, each corpus encompassed two complete years. The contents of each corpus were determined by the material available in an electronically readable form.

[3] This provision resulted in the exclusion of only one item, 'He aha au?' in Part One, Number 2, 1997; it had no effect on the 1960s corpus

[4] In the 1990s corpus, this provision resulted in the exclusion of the waiata in 'Can It Be A Gannet?', Part Two, Number 3, 1998, and the proverb in 'A New Friend', Part Two, Number 1, 1998; it also resulted in the exclusion of several short passages in the 1960s corpus

was recorded separately from *Kiwi* when employed to designate New Zealand origin. However, in calculating the total number of word types, such examples were only counted once.

It should also be noted that there were two exceptions in the classifying of proper nouns. Following Kennedy and Yamazaki's practice, both *Maori* and *Pakeha* were included in the General category.

Table 1 shows the distribution of words of Maori origin in the corpus as a whole.

Table 1: Distribution of words of Maori origin in the *School Journals* 1990s corpus

Word categories	**Tokens**	**(%)**	**Types**	**(%)**
Flora & fauna	**110**	**(12.2%)**	**27**	**(13.0%)**
Trees & plants	31		7	
Fish & shellfish	11		4	
Birds	23		11	
Other creatures	45		5	
Place names	**107**	**(11.9%)**	**50**	**(24.0%)**
Names of persons	**394**	**(43.9%)**	**67**	**(31.7%)**
Other proper nouns	**42**	**(4.7%)**	**13**	**(6.3%)**
Structural words	**65**	**(7.3%)**	**10**	**(4.8%)**
Tikanga Maori	**91**	**(10.1%)**	**23**	**(11.1%)**
General	**89**	**(9.9%)**	**19**	**(9.1%)**
TOTAL	**898**	**(100%)**	**208**	**(100%)**

It can be seen that words of Maori origin are used in three main ways:

- as proper nouns; this is the largest single use
- in the Tikanga Maori / General category; this is the second largest use
- to name the flora and fauna of the country; this is the third largest use

This pattern of use will be discussed in more detail later in the paper.

5. The *Journals* of the 1990s: representative of New Zealand English?

In seeking to determine if the pattern of use of Maori words in New Zealand English has changed over time, it is relevant to inquire whether the *School Journals* are representative of the written language.

5.1 Frequency

Kennedy and Yamazaki (1999) investigated the presence of Maori in both spoken and written New Zealand English. Their finding for the frequency of words of Maori origin in the *Wellington Corpus of Written New Zealand English* (Bauer, 1993) is strikingly similar to that for the *Journals* 1990s corpus. In Kennedy and Yamazaki's analysis almost 6 words per 1000 were found to be

of Maori origin while in the *Journals* the frequency of occurrence was a little over 6 words per 1000.

5.2 Patterns of Use

There are also broad similarities in the patterns of use, as shown in Table 2, which compares the findings for the *Journals* with those of Kennedy and Yamazaki for the *Wellington Corpus of Written New Zealand English.*

Table 2: Comparison of % distribution of words of Maori origin in the *School Journals* 1990s corpus and the *Wellington Corpus of Written NZ English*

Word categories	School Journals		Written NZ English	
	Tokens	Types	Tokens	Types
Flora & fauna	**12.2**	**13.0**	**6.3**	**5.7**
Trees & plants	3.4	3.4		2.8
Fish & shellfish	1.2	1.9		1.1
Birds	2.6	5.3		1.2
Other creatures	5.0	2.4		0.6
Place names	**11.9**	**24.0**	**33.5**	**40.4**
Names of persons	**43.9**	**31.7**	**21.0**	**23.2**
Other proper nouns	**4.7**	**6.3**	**---**	**---**
Structural words	**7.3**	**4.8**	**11.2**	**2.9**
Tikanga Maori	**10.1**	**11.1**	**12.8**	**14.2**
General	**9.9**	**9.1**	**15.3**	**13.6**
TOTAL	**100.0**	**100.0**	**100.0**	**100.0**

It should be noted that the figures for the *Wellington Corpus of Written New Zealand English* given in Table 2 exclude tokens of the word *Maori.* In their analysis of that corpus, Kennedy and Yamazaki commented on the frequency of the word in the General category, accounting for 998 of 1754 tokens (almost 57%). This finding was not duplicated in the *School Journals,* where *Maori* accounted for only 11 of the 87 tokens in that category (around 10%), and approximately 1% of the entire occurrence of words of Maori origin in the corpus. Therefore, as the effect would be insignificant and would not affect the patterns of use, the figures for the *School Journals* 1990s corpus were not recalculated. [2]

This difference in the frequency of the word token *Maori* can perhaps be accounted for by the fact that, while the *Wellington Corpus* includes topicality in its embrace so that news stories about, for instance, Maori land claims and statements by the Maori Affairs Minister on Maori housing issues will have an impact on the corpus, the *School Journals* do not make decisions about suitability for publication on the ephemeral grounds of topicality.

[2] It should also be mentioned that the inclusion of *Maori* in the figures for the *Wellington Corpus* does not alter the broad picture of patterns of use, with Tikanga Maori/General as the second largest category.

As Table 2 shows, there are similarities between the two corpora in patterns of use. Most significantly, the same distributional ranking is found in both corpora. Thus:

- proper nouns (place names, names of persons, and other) represent the largest single use. The percentage of proper nouns was roughly between 55 and 60% in both corpora.
- in both corpora the Tikanga Maori / General categories, between which the division is sometimes arbitrary, together accounted for between 20 and 30% of tokens and represented the second largest use of words of Maori origin.
- flora and fauna ranked third in importance in both corpora. While these terms are more common in the *School Journals*, the difference parallels that between the junior and senior corpora, possibly for the same reasons. If the transition from junior to senior *Journals* can be regarded as a transition from childhood to early teenage, then the transition from senior *Journals* to the *Wellington Corpus* can be seen as a move from early teenage to adult language and interest.

In general terms, therefore, both the frequency and the distribution of words of Maori origin in the two corpora are sufficiently similar to confirm the assumption that the *School Journals* would be a representative sampling of current New Zealand English.

6. The 1960s: a generation and a world apart

The 1960s were chosen as the period for historical comparison for a number of reasons.

- 30 years is generally accepted as being the measure of a generation. Thus, by comparing the late 1990s with the late 1960s, we are comparing the changes from one generation to the next.
- the 1960s and the 1990s are marked by their differences rather than their similarities. The 1960s were the Holyoake Years, a time popularly remembered as 'the golden 1960s' (Chapman, 1981: 365) when 'the number of registered unemployed was so small ... it was jokingly said that they were all known by name to the Prime Minister' (Sinclair, 1991: 298). Comparable data for the 1990s presents a grimmer picture.
- the late 1960s and the late 1990s are roughly equidistant from the consciousness-raising of the 1981 anti-Springbok Tour movement and the publication of works which challenged conventional thinking and aroused debate, such as Donna Awatere's *Maori Sovereignty* (1984) and Claudia Orange's *The Treaty of Waitangi* (1987).

7. The *Journals* of the 1960s: initial findings

The *School Journals* of the 1960s were somewhat different from their counterparts of the 1990s. The four Parts structure was already bedded-in by the 1960s, but Part One had four rather than the current five issues a year. Illustrations and colour were a much less dominant feature, with a corresponding increase in the quantity of text in an issue. There was less uniformity between issues, with variations in length from 32 to 80 pages, and also changes in format, from portrait to landscape.

The differences between the decades do not pertain to physical presentation alone. The content was much more obviously driven by curriculum considerations than is the case today. However, they were still vehicles for New Zealand writers writing for New Zealand school children, and as such can be assumed to be reasonably representative of written New Zealand English at that time.

The corpus for the 1960s consisted of all the issues of each Part for 1967 and 1968, a total of 28 issues. A manual search of the corpus was undertaken to identify and count the words of Maori origin, the distribution of which is shown in Table 3.

Table 3: Distribution of words of Maori origin in the *School Journals* 1960s corpus

Word categories	**Tokens**	**(%)**	**Types**	**(%)**
Flora & fauna	**298**	**(16.2%)**	**56**	**(17.0%)**
Trees & plants	181	(9.8)	38	
Fish & shellfish	16	(0.9)	9	
Birds	96	(5.2)	6	
Other creatures	5	(0.3)	3	
Place names	**364**	**(19.8%)**	**125**	**(38.0%)**
Names of persons	**506**	**(27.5%)**	**71**	**(21.6%)**
Other proper nouns	**53**	**(2.9%)**	**16**	**(4.9%)**
Structural words	**78**	**(4.2%)**	**8**	**(2.4%)**
Tikanga Maori	**110**	**(6.0%)**	**23**	**(7.0%)**
General	**432**	**(23.4%)**	**30**	**(9.1%)**
TOTAL	**1841**	**(100.00)**	**329**	**(100.00)**

At first glance, it would appear that there is no difference in the distributional ranking from the 1960s to the 1990s. However, 294 of the tokens in the General category, or 68%, are one word, the word *Maori*. When the figures are recalculated to exclude this word, a different picture emerges, as shown in Table 4 and discussed in 8.2 below.

8. The 1960s and the 1990s compared

8.1 Frequency

An accurate comparison of the frequency of words of Maori origin in New Zealand English of the 1960s and 1990s cannot be gained, primarily because the manual search of the 1960s corpus did not allow an accurate total word count to take place.

8.2 Patterns of Use

To gauge changes in the patterns of use of Maori words over a generation, it is necessary to compare their semantic distribution in the corpora.

Table 4: Comparison of distribution of words of Maori origin in the two *School Journals* corpora

Word categories	**1960s including *Maori***	**1960s excluding *Maori***	**1990s**
Flora & fauna	**16.2**	**19.2**	**12.2**
Trees & plants	9.8	11.7	3.4
Fish & shellfish	0.9	1.0	1.2
Birds	5.2	6.2	2.6
Other creatures	0.3	0.3	5.0
Place names	**19.8**	**23.5**	**11.9**
Names of persons	**27.5**	**32.7**	**43.9**
Other proper nouns	**2.9**	**3.4**	**4.7**
Structural words	**4.2**	**5.1**	**7.3**
Tikanga Maori	**6.0**	**7.1**	**10.1**
General	**23.4**	**8.9**	**9.9**
	100.00	**100.00**	**100.00**

When tokens of the word *Maori* are excluded, the following features emerge:

- proper nouns, in their various manifestations, remain the single largest category
- in the 1960s corpus, the flora and fauna category replaces Tikanga Maori / General as the second largest use of Maori words
- inevitably, therefore, in the 1960s words in the Tikanga Maori / General category assume least importance among the content word groupings

Table 5 summarises Table 4 and presents in a simplified form the information about changing patterns of use.

Table 5: Patterns of use (%) of content words in the two *School Journals* corpora

	1960s excluding *Maori*	**1990s**
Proper Nouns	59.6	60.5
Tikanga Maori / General	16.0	20.0
Flora & fauna	19.2	12.2

9. Discussion of the findings

9.1 Proper Nouns

Proper nouns, whether they be place names, personal names, or some other type of name, can be treated as one grouping because they do not carry a

significant semantic load. The initial capital letter signals to the reader that *Wairarapa* is a less problematic unknown word than *waiata.*

While the semantic distribution categories used in this paper have followed Kennedy and Yamazaki, to allow comparability, it could, of course, be argued with some validity that *Maori* ought to be considered and treated as a proper noun, rather than counted in the General category. To do so presents no difficulty, and does not alter the argument regarding changing patterns of use, as Table 6 makes clear. In Table 6 the figures have been recalculated to include *Maori* in the Proper Noun category. In both the 1990s corpus and the *Wellington Corpus* the flora and fauna category has still been relegated to third place by Tikanga Maori / General.

Table 6: Patterns of use (%) of content words, including *Maori,* in the two *School Journals* corpora and the *Wellington Corpus of Written New Zealand English*

	1960s	**1990s**	**Wellington Corpus**
Proper Nouns	66.1	61.5	62.2
Tikanga Maori / General	13.5	18.9	23.3
Flora & fauna	16.2	12.2	5.2

It is noteworthy that in all corpora the proportion of proper nouns, whether tokens of *Maori* are included or excluded, among words of Maori origin was around 60% or more. Thus, up to about 40% of the Maori word tokens, and an identical 33% of Maori word types in the two *School Journals* corpora, require greater understanding than the recognition of, usually, a place or personal name.

However, the similarity in the occurrence of proper nouns should not be allowed to disguise an important difference. Although personal names were more frequent than place names in both corpora, the ratio of one to the other altered markedly. Place names were significantly more frequent in the 1960s than in the 1990s and accounted for 38% of all Maori word types encountered by the reader, while in the 1990s personal names occurred much more often than in the 1960s. Thus, the Maori presence in the 1960s is strongly geographical, whereas the pages of the *Journals* of the 1990s are much more obviously inhabited by Maori characters.

An examination of the actual types of personal names present in the two corpora refines this observation further. The following is a selection of the names of historical figures found in the 1960s corpus but not in the 1990s. Even more tellingly, there were no names of obvious historical significance in the 1990s corpus.

Hongi Hika
Te Rauparaha
Te Kooti
Tawhiao
Rewi Maniapoto
Apirana Ngata
Maui Pomare
Timi Kara (James Carroll)
Hemi Maki (James Mackay)

In addition, the 1960s corpus also contains names from myth and legend, such as Maui, Kupe and Tane, that have no representation in the 1990s corpus.

This suggests differences in the treatment of Maori in the two corpora. As already noted, Maori more evidently people the 1990s *Journals*. Furthermore, Maori in the 1960s corpus are more likely to be found in informative prose passages as subjects of historical study than in the 1990s, whereas Maori in the 1990s corpus are more likely to feature as the actors in imaginative prose passages.

9.2 Flora and Fauna

As Deverson (1991: 19) and others have pointed out, lexical borrowing from one language to another 'to describe a previously unknown natural environment' occurs at an early stage in the contact between two cultures. The practice of identifying New Zealand's flora and fauna by Maori names was established during the first half of the nineteenth century, and at the end of the twentieth century every New Zealander's vocabulary includes some words of Maori origin that name the country's flora and fauna. Sometimes the Maori words co-exist with an equally acceptable English language alternative, as in the examples *hapuku (hapuka) / groper* and *koromiko / hebe* but sometimes these Maori words are the exclusive term for particular birds, trees, fish, or other creatures. For example, English language alternatives for *kiwi, kahawai, kauri* and *tuatara* do not readily spring to mind.

Both the 1960s and the 1990s corpora contain a considerably higher proportion of flora and fauna words than was evident in the *Wellington Corpus of Written New Zealand English*, which may in part be a result of the age group factor, as well as the difference in text categories represented.

It is interesting, however, that while the broad pattern of use identified in both the *Wellington Corpus* and the 1990s corpus is consistent, with flora and fauna ranking third, in the 1960s corpus flora and fauna occupies second place, ahead of Tikanga Maori / General.

9.3 Tikanga Maori/General

The decision to classify a word as either Tikanga Maori or General was sometimes arbitrary. A word such as *whare* was regarded as General, although its connotations are specifically Maori, while a word such as *pa* was

classified as Tikanga Maori, because *pa* incorporates concepts of social relationship. Thus, a general distinction was made between words that described a concept and words that described an aspect of the material culture. However, the arbitrariness and fuzziness of this distinction justifies the treating of the two categories as one when discussing patterns of use.

As already noted, words in the Tikanga Maori/General category, whether measured as tokens or types, assume a greater role in the 1990s corpus than in the 1960s, both when *Maori* is excluded and when *Maori* is regarded as a proper noun rather than a General word. The exclusion of *Maori* in the 1960s corpus reduced the number of tokens in the General category from 432 to 138. Of those remaining 138, a further 68 are the word *pakeha* (whereas only 3 tokens in the 1990s corpus are the word *pakeha*). This is a significant presence which will be discussed in a later article.

Not only is the frequency of Tikanga Maori/General words greater in the 1990s corpus, but an examination of the words classified as Tikanga Maori in both corpora suggests changing patterns of use, as shown in Table 7.

Table 7: Tikanga Maori word types common to, and exclusive to, the two corpora

Exclusive to the 1960s corpus	**Common to both 1960s and 1990s corpora**	**Exclusive to the 1990s corpus**
kainga	pa	karakia
haka	marae	kaumatua
tapu	mokopuna[a]	koro
Kotahitanga	moko[b]	tangata whenua
mana	taniwha	tangihanga
hui		taonga
hapu		tekoteko
rangatira		tipuna
tohunga		waiata
tangi		whaikorero
powhiri		whakapapa
tiki		whanau
hipitoi		aroha
e ropi		hongi
hei tama tu tama		korero
		rangatiratanga

The implication of Table 7 is that there has been a shift from words referring to

(a) historical and hierarchical aspects of Maori society (*rangatira, tohunga, Kotahitanga*), and

[a] *mokopuna* includes the contraction *moko*, referring to a grandchild/grandchildren
[b] *moko* here refers exclusively to tattoo

(b) artefacts of the material culture (*hipitoi, e ropi,* and *hei tama tu tama* are hand games, for example),

to words reflecting

(c) a sense of belonging to a community and of continuity with a human rather than an historical past (*whanau, kaumatua, koro, tipuna, whakapapa, tangata whenua*), and

(d) aspects of the social culture and social interaction (*hongi, waiata, karakia, korero, whaikorero*).

Furthermore, when the words in Table 7 are compared to those found by Geering (1993) in nineteenth century newspapers, it is tempting to suggest, and not impossible to imagine, that changes in the presence of words of Maori origin in New Zealand English over the generation since the late 1960s have been more rapid and more radical than in the preceding one hundred years. Eight of the 15 words exclusive to the 1960s corpus were found in Geering, but only four of the 16 words exclusive to the 1990s corpus.

It is, moreover, difficult to avoid the conclusion that the Tikanga Maori / General category is the most sensitive barometer of change. In terms of word types, both the proper noun grouping and the flora and fauna category are relatively closed classes of words. Growth in types of words of Maori origin within the New Zealand English lexicon, therefore, is most likely to be within the Tikanga Maori / General category. Growth in the number of these types within New Zealand English will also have an impact upon the tokens present in a corpus, and so will contribute to changes in both patterns of use and frequency

10. Conclusions

This study has sought to describe trends in the Maori word lexicon in New Zealand English, rather than the lexicon itself. Study of the *School Journals* corpora has provided evidence of an evolving presence of Maori in New Zealand English, but it must be emphasised that this is a process of gradual change in New Zealand English.

The evidence allows the formulation of a number of conclusions.

- proper nouns of Maori origin typically account for around 60% of the presence of Maori words in written New Zealand English. While they add a certain distinctive 'colour' to the language, they do not carry a significant semantic load and thus do not require any special semantic knowledge.
- the Tikanga Maori/General category of words is growing and strengthening. Over one generation this category has replaced flora and fauna in importance in the New Zealand lexicon. Growth in the presence of Maori words in New Zealand English will most probably occur in this category.
- the nature of the Tikanga Maori words unique to the two corpora suggests a change in attitude to the Maori and the Maori world, with a shift from an

historical and anthropological interest to a recognition of Maori as belonging to a living culture.

- a living language never remains still. The frequency, the patterns of use, and the nature of Maori words in the New Zealand English lexicon will continue to change. The current trend suggests that Maori words in the lexicon will become increasingly significant in the future.

References

Awatere, Donna 1984. Maori Sovereignty. Auckland: Broadsheet

Bauer, Laurie 1993. The Wellington Corpus of Written New Zealand English. Wellington: Department of Linguistics, Victoria University of Wellington

Chapman, R. 1981. From Labour to National. In Oliver, W. H. with Williams B. R. (eds) The Oxford History of New Zealand. Wellington: Oxford University Press: 333-368

Deverson, Tony 1991. New Zealand lexis: the Maori dimension. English Today 26: 18-25

Geering, Elaine 1993. The use of Maori in the late nineteenth-century Auckland press. In Bauer, Laurie & Franzen, Christine (eds) Of Pavlova, Poetry and Paradigms. Wellington: Victoria University Press: 250-260

Kennedy, Graeme & Yamazaki, Shunji 1999. The Influence of Maori on the New Zealand English Lexicon. In Kirk, John M. (ed) Corpora Galore: Analyses and Techniques in Describing English. Amsterdam: Rodopi: 33-44

Orange, Claudia 1987. The Treaty of Waitangi. Wellington: Allen & Unwin/Port Nicholson Press

Orsman, Harry (ed.) 1997. The Dictionary of New Zealand English. Auckland: Oxford University Press

Sinclair, Keith 1991. A History of New Zealand. Auckland: Penguin