The discourse particle *eh* in New Zealand English

MARTIN SCHWEINBERGER

University of Hamburg

Department of English and American Studies

Von-Melle-Park 6

D-20146 Hamburg

Germany

E-mail: martin.schweinberger@uni-hamburg.de

ABSTRACT

This paper investigates the use of the semantically bleached and syntactically optional discourse particle *eh* in New Zealand English (NZE). The analysis takes a quantitative, variationist approach and is based on the New Zealand component of the *International Corpus of English* (ICE New Zealand) (Bauer et al. 1999; Vine 1999). The analysis of the use of *eh* in NZE includes both sociolinguistic (ethnicity, age, gender, occupation type) and psycholinguistic variables (priming). A mixed-effects binomial logistic regression model shows that younger speakers use *eh* more often than older speakers, males use it more often than females, and Maoris (indigenous Polynesian people of New Zealand) use it more often than Pakehas (decedents of European settlers). The analysis neither confirmed a significant impact of priming on the use of *eh* nor significant interactions between predictors which contrasts with previous studies.

1 Introduction

The discourse-pragmatic particle *eh* (1a–d) is typically used in clause-final position in Canadian and New Zealand English. Similar to Canadian *eh* (cf. Columbus 2009; Gold & Tremblay 2006), the New Zealand counterpart has been extensively studied (e.g. Holmes 2005; Meyerhoff 1994; Stubbe & Holmes 1995).

(1) Speech Unit Final *eh* in NZ[1]

(a) oh we won't go that far *eh* (ICE-NZ:S1A-001#1:M)

(b) it's a really bad buzz *eh* (ICE-NZ:S1A-004#1:M)

(c) yes so pete moves on and and clinton comes in and the girls will still stay here *eh* (ICE-NZ:S1A-037#1:M)

(d) yeah yeah it is *eh* (ICE-NZ:S1A-047#1:A)

Although *eh*[2] is a salient feature of New Zealand English (Bauer 2007: 20), previous studies have exclusively used mono- or bivariate statistics (e.g. Kruskal-Wallis rank order tests in the case of Stubbe & Holmes 1995) while multivariate analyses are still lacking. The present analysis applies advanced statistical modeling, i.e. mixed-effects binomial logistic regression models, to address this shortcoming. The intention of the present study is to investigate which factors impact the usage of *eh* in New Zealand English, to determine how strong the relative influence of significant variables is, and to evaluate the findings of previous research which found that *eh* usage is affected by an interaction between gender and ethnicity: *eh* is used more often by females among Pakehas while it is more often used by males among Maoris (Meyerhoff 1994: 382).

Furthermore, the study elaborates on previous studies in so far as it is not restricted to sociolinguistic predictors such as the age, gender, social status, or ethnicity of speakers but also incorporates priming as a psycholinguistic determinant of variation (for a discussion cf. Szmrecsanyi 2006 or Gries 2013). Unfortunately, there have only been a few studies which have simultaneously considered socio- and psycholinguistic predictors of language change (notable exceptions are Gries 2005 and Szmrecsanyi 2006). Thus, beyond investigating the social profile of *eh*, the study intends to determine to which degree psycholinguistic factors affect language use.

In summary, to remedy these shortcomings the current study intends to attain three main objectives:

- Which sociolinguistic factors impact the usage of *eh* and how strong is their relative impact?
- Are there significant interactions between sociolinguistic predictors as suggested by previous studies (Meyerhoff 1994)?
- Is the usage of this semantically bleached, syntactically optional element, i.e. the discourse particle *eh* in NZE, affected by priming?

2 PREVIOUS RESEARCH ON EH

The following section provides an overview of previous research and introduces key concepts related to the discourse particle *eh*.

Over the past few decades, discourse particles have attracted a growing amount of attention among linguists (cf. e.g. Aijmer 2002) and this is also true for *eh* (e.g. Bauer 1997; Meyerhoff 1994; Stubbe & Holmes 1995; Holmes 2005; Holmes, Stubbe & Marra 2003). Similar to other syntacically optional discourse particles, *eh* is a feature of NZE vernacular (Stubbe & Holmes 1995: 74) and overtly stigmatized (Meyerhoff 1994: 367). Turner (1966) states that *eh* is particularly wide-spread on the North Island and hypothesizes that it entered NZE due to Maori influence. Functionally, *eh* resembles the French question tag *n'est-ce pas?* (Turner 1966: 170). Various authors state that *eh* is an identity marker among Maoris (Bell & Kuiper 1999; Meyerhoff 1994; Stubbe & Holmes 1995; Holmes 1997, 2005; Holmes, Stubbe & Marra 2003), which entered NZE due to its similarity to Maori *ne* (Bauer 1997: 426) and was only later adopted by Pakehas (cf. Holmes 1997). In terms of its social distribution, previous research shows that *eh* usage is particularly associated with younger, working-class, suburban, Pakeha women, and working-class Maori men (cf. Meyerhoff 1994: 374; Stubbe & Holmes 1995: 63). Although Stubbe & Holmes (1995) agree that *eh* is predominantly used by younger speakers, their data did not show a significant correlation between age and the usage of *eh* (Stubbe & Holmes 1995: 72). However, Stubbe & Holmes (1995) found that middle-class speakers tend not to use *eh* regardless of age, while *eh* was age-graded among working-class speakers (Stubbe & Homes 1995: 72). Stubbe & Homes (1995: 72) thus argue that the absence of a correlation between age and *eh* usage is caused by the low number of young working class subjects in their study. In addition, Stubbe & Homes (1995: 72) could not confirm an effect of gender on *eh* usage based on their data and conclude that *eh* is primarily a marker of working class speech (Stubbe & Homes 1995: 73).

The Wellington social dialect data and popular perception (described in Meyerhoff 1994) argue for the existence of an interaction between gender and ethnicity (cf. Meyerhoff 1994). However, she admits that three of the five Pakeha females in her data lived with male Maoris (Meyerhoff 1994: 368—69) and that this is likely to have affected the frequency of *eh* use by these females; particularly since the other two Pakeha females used *eh* substantially less often. Despite her reasonable explanation for this finding, Meyerhoff (1994) concludes that the interaction is caused by differences in the functional employment of *eh* by Pakeha females compared to Maori men. Similar to Meyerhoff (1994), Stubbe & Holmes (1995: 63) also argue for the existence of an interaction but between social class and age instead of between age and ethnicity as Meyerhoff (1994) does.

On a functional level, *eh* has been described as an addressee-oriented tag that is semantically empty and lacks epistemic meaning (Stubbe & Holmes 1995: 68). Among young Pakeha females, *eh* serves as a positive politeness marker (Meyerhoff 1994: 367) while it serves as an identity marker among Maori men for whom it signals in-group solidarity (Meyerhoff 1994: 384).

3 Data and Methodology

The analysis of *eh* in the current study is based on data drawn from the New Zealand component of the *International Corpus of English* (ICE) (cf. Vine 1999). The corpus data was tabulated so that each row contained exactly one speech unit (SU). It was then determined whether the last element of the speech unit was *eh*. In a next step, speech units within the scope of 10 words following an instance of *eh* were coded as being primed ('prime') – all speech units that lay outside that scope were coded with 'noprime'.

The biographical information, such as the age, gender, ethnicity, and occupation of the speakers was extracted from Vine (1999). Each speech unit was then coded as occurring in the speech of a Maori or a Pakeha and whether the speaker was female or male. Subsequently, the age of the speaker uttering a given speech unit was determined.

A preliminary display of the mean frequencies of *eh* per speech unit across text types confirms that *eh*, like other discourse particles, is used most in informal conversation (cf. Figure 1). The distribution matches the finding by Stubbe & Holmes (1995: 74) who arrived at the conclusion that *eh* is a vernacular rather than a standard feature in New Zealand English.

The analysis therefore only uses data from files representing private dialogue, i.e. face-to-face and telephone dialogues, because *eh* is extremely rare in other text types and registers.[3]

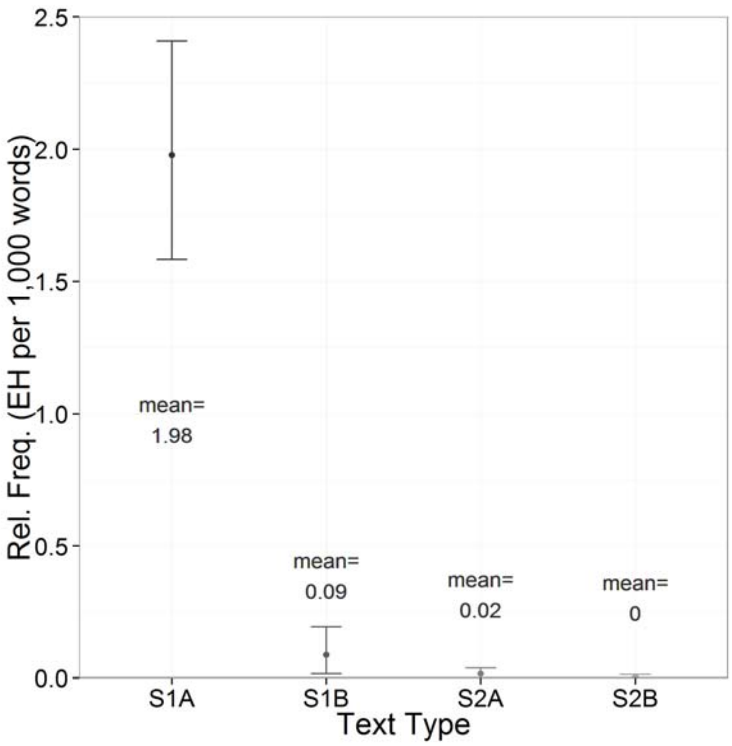

Figure 1: eh *by text type*

After removing all the data from files not representing private dialogue, all the data representing speakers for whom information regarding their age, gender, or ethnicity was not available were removed from the analysis. Table 1 provides a summary of the data during the three stages of data processing and cleaning.

Table 1: Overview of the data set before, during, and after data editing and cleaning.

Data	Speakers (N)	Words (N)	Speech Units (N)	*eh* (N)
All spoken files	1,085	653,186	68,189	421
Only private dialogue	250	213,555	31,544	410
Only private dialogue with complete cases (Age, Gender, Ethnicity, Occupation type)	140	130,960	17,770	217

3.1 Classification and coding of variables

The following section describes the coding of the variables in this study. The basis for the classification of the sociolinguistic variables is Vine (1999) which accompanies the corpus material of the New Zealand component of the ICE.

3.1.1 Speech Unit Final eh (dependent variable)

The dependent variable in this study is the occurrence of *eh* in the speech unit final position. Each speech unit in the data was coded as 0 (no *eh* in the speech unit final position) or as 1 (*eh* in the speech unit final position). The resulting factor has accordingly two levels (0, 1) and represents a nominal variable.

3.1.2 Age (independent variable)

The age classification of the New Zealand component of the ICE is quite fine-grained and consists of 12 age groups (16-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74) (Vine 1999). For the current study these age groups were collapsed based on their respective mean frequency of *eh* per speech unit in order to create cohorts with a sufficient number of speakers to increase the reliability of the results of the statistical analysis – which would not be the case if the cohorts consist of too few speakers. A cluster analysis was applied to the data and yielded two clusters (cf. Figure 9 in the Appendix): speakers between 16 and 39 on the one hand and speakers between 40 and 74 on the other. The age groups were accordingly re-categorized: young (39 and younger) and old (40 and older). The resulting factor has two levels (‘young’, ’old’).

3.1.3 Gender (independent variable)

The classification of gender is based on the self-reported sex in the guide accompanying the ICE New Zealand corpus. In the present study, gender is a nominal variable representing a factor with two levels (‘female’, ‘male’).

3.1.4 Ethnicity (independent variable)

Ethnicity is a factor consisting of two levels: Maori and Pakeha. The coding of a speaker’s ethnicity depended on the information provided by Vine (1999). The resulting factor has, accordingly, two levels (‘pakeha’, ‘maori’) and is a nominal variable.

3.1.5 Occupation type (independent variable)

The occupation type of speakers serves as a proxy for social class and represents a re-coding of the self-reported occupation of speakers provided in Vine (1999). The factor has two levels: On the one hand ‘sml’ which stands for skilled manual labor and refers to e.g. craftsmanship, bar tending, and shop assistants. The factor level ‘acmp’, on the other hand, encompasses academic, clerical, and managerial occupations as well as the professions.

3.1.6 Priming (independent variable)

Priming refers to the re-use of material that was used in previous utterances (cf. Tulving 1990: 301) -- a phenomenon also referred to as production priming (cf. Szmrecsanyi 2005: 113). The fact that speakers re-use material is widely acknowledged, but while there is a growing body of research dedicated to e.g. structural priming which builds on both psycholinguistic experimentation and corpus-linguistic analyses, various issues remain unsolved. For instance, there is much debate concerning the duration of priming effects as the decay time may vary between milliseconds and months or even years (Althaus 2006: 962). One of the factors determining the durability of priming is the type of priming: semantic or conceptual priming can for last extended periods of time, while both syntactic and form priming will decay very swiftly (within seconds). The scope of 18 words is based on the fact that form priming is short-lived and disappears soon after exposure to the stimulus prime (Althaus 2006: 962) in cases of phonological priming –

the effect of semantic or conceptual priming is arguably much longer and may even last months or even years (Althaus 2006: 962). The assumed scope of priming in this study is an 18-word span. The 18-word span was determined by counting the mean frequency of words uttered within 5 second intervals in the *Santa Barbara corpus of spoken American English* (DuBois et al. 2000-2005). The time span of 5 seconds follows Szmrecsanyi (2006) who found that priming effects (or the effect of α-persistence in Szmrecsanyi's terminology) declined after three to ten words -- which amounts to approximately 10 seconds – after the prime occurred (Szmrecsanyi 2006: 190).

The current study thus follows Szmrecsanyi (2006) in acknowledging that priming effects may occur within a 5-second window after the prime occurred but deviates from Szmrecsanyi (2006) with respect to the word-span: Szmrecsanyi (2006) used a span of 10 words while this study uses an 18-word span based on the mean frequency of words uttered in 5 seconds in the *Santa Barbara Corpus of Spoken American English*.

3.2 The final data set

The following section summarizes and displays the properties of the final data set, i.e. the data that the analysis of *eh* is based on. Table 2 displays the data without including priming but the number of speakers within each sub-cohort.

Table 2: Overview of the speakers and their parameters in the final data set.

Age	Sex	Ethnicity	Occ. Type	Speakers (N)	SU with *eh*	SU without *eh*	Mean *eh* per SU
Young	Female	Maori	ACMP	9	28	1,383	0.020
Young	Female	Maori	SML	0	NA	NA	NA
Young	Female	Pakeha	ACMP	33	31	3,671	0.008
Young	Female	Pakeha	SML	16	15	2,348	0.006
Young	Male	Maori	ACMP	2	15	375	0.038
Young	Male	Maori	SML	1	3	43	0.065
Young	Male	Pakeha	ACMP	24	42	2,559	0.016
Young	Male	Pakeha	SML	17	45	2,108	0.021
Old	Female	Maori	ACMP	6	13	1,047	0.012
Old	Female	Maori	SML	1	0	165	0.000
Old	Female	Pakeha	ACMP	15	1	1,635	0.001
Old	Female	Pakeha	SML	1	1	191	0.005
Old	Male	Maori	ACMP	7	21	1,172	0.018
Old	Male	Maori	SML	0	NA	NA	NA
Old	Male	Pakeha	ACMP	5	0	607	0.000
Old	Male	Pakeha	SML	3	2	249	0.008
Total				140	217	17,553	0.012

Table 2 shows that the final data set consists of 140 speakers, 217 instances of *eh*, and 17,770 (217 plus 17,553) speech units. The largest sub-cohort with 33 speakers consists of young female Pakehas who work in academic, managerial, clerical occupations, or the professions. The highest number of instances of *eh* within the sub-cohorts is 45 among young male Pakeha who work in academic, managerial, clerical occupations or the professions. Overall, the mean of *eh* by speech unit in the data amounts to 0.012, denoting that there are 0.012 instances of *eh* per speech unit. The highest mean frequencies of *eh* occur among young speakers (0.065, 0.038, 0.021) suggesting a correlation between the age of speakers and their usage of *eh*.

Table 3 provides an alternative overview of the final data set. In this tabulation, priming is included which forces a display that does not tabulate the number of speakers for each sub-cohort: since speakers uttered primed and non-primed utterances, Table 3 cannot meaningfully display the number of speakers.

Table 3: Overview of the speaker parameters and priming in the final data set.

Age	Sex	Ethnicity	Occ. Type	Priming	SU with *eh*	SU without *eh*	Mean *eh* (per SU)
Young	Female	Maori	ACMP	noprime	27	1,305	0.020
Young	Female	Maori	ACMP	prime	1	78	0.013
Young	Female	Maori	SML	noprime	NA	NA	NA
Young	Female	Maori	SML	prime	NA	NA	NA
Young	Female	Pakeha	ACMP	noprime	29	3,596	0.008
Young	Female	Pakeha	ACMP	prime	2	75	0.026
Young	Female	Pakeha	SML	noprime	15	2,303	0.006
Young	Female	Pakeha	SML	prime	0	45	0.000
Young	Male	Maori	ACMP	noprime	15	333	0.043
Young	Male	Maori	ACMP	prime	0	42	0.000
Young	Male	Maori	SML	noprime	3	43	0.065
Young	Male	Maori	SML	prime	NA	NA	NA
Young	Male	Pakeha	ACMP	noprime	37	2,444	0.015
Young	Male	Pakeha	ACMP	prime	5	115	0.042
Young	Male	Pakeha	SML	noprime	43	2,008	0.021
Young	Male	Pakeha	SML	prime	2	100	0.020
Old	Female	Maori	ACMP	noprime	12	1,027	0.012
Old	Female	Maori	ACMP	prime	1	20	0.048
Old	Female	Maori	SML	noprime	0	165	0.000
Old	Female	Maori	SML	prime	NA	NA	NA
Old	Female	Pakeha	ACMP	noprime	1	1,622	0.001
Old	Female	Pakeha	ACMP	prime	0	13	0.000
Old	Female	Pakeha	SML	noprime	1	188	0.005
Old	Female	Pakeha	SML	prime	0	3	0.000
Old	Male	Maori	ACMP	noprime	17	1,102	0.015
Old	Male	Maori	ACMP	prime	4	70	0.054

Old	Male	Maori	SML	noprime	NA	NA	NA
Old	Male	Maori	SML	prime	NA	NA	NA
Old	Male	Pakeha	ACMP	noprime	0	605	0.000
Old	Male	Pakeha	ACMP	prime	0	2	0.000
Old	Male	Pakeha	SML	noprime	2	244	0.008
Old	Male	Pakeha	SML	prime	0	5	0.000
Total					217	17,553	0.012

In addition to the information already summarized in Table 2, Table 3 suggests that the vast majority of instances of *eh* occur in non-primed speech units (primed 15 versus 202 not primed). However, the vast majority of speech units are also non-primed (primed 568 versus 16,985 not primed). The low number of primed instances of *eh* in certain sub-cohorts (e.g. young male Pakeha who work in academic, managerial, clerical occupations or the professions exhibit not a single instance of *eh* in primed speech units) could have been problematic for the model fitting process as these sub-cohorts exhibit complete separation, i.e. cases in which the outcome variable correlates perfectly with a certain predictor (Field, Miles & Field 2012: 323). During the model fitting process only higher level interactions which did not complete separation were included (cf. Tables 5 and 6 in the Appendix).

3.3 Regression modeling

The study uses a type of multivariate analysis to statistically test if any of the independent variables or interactions between them correlate with the usage of *eh* in the NZE data. More specifically, the primary tool to investigate which factors impact the occurrence of *eh* is a mixed-effects binomial logistic regression model. Binomial logistic regression models calculate the likelihood of a binary outcome (*eh* occurring versus *eh* not occurring) given the independent variables. For instance, a binomial logistic regression model calculates the likelihood of *eh* occurring given the speaker is a young Maori female.

To obtain accurate estimates for the size of the effect of a given variable, the model was fit using both step-wise step-up (independent variables and their interactions are added consecutively, i.e. the model is build up) and step-wise step-down procedures (independent variables and their interactions are removed consecutively, i.e. the model is shrunk) to arrive at the final minimal adequate model, i.e. the best model in the sense that a minimum of predictors explains a maximum of variation. Although criticism of model fitting has been raised (Johnson 2010), it remains crucial for at least three reasons:

– not removing insignificant predictors leads to unreliable results as variance is erroneously attributed to potentially insignificant predictors.
– not removing insignificant predictors leads to inaccurate p-values which results in interpreting potentially insignificant predictors as significant and vice versa.
– not removing insignificant predictors leads to inaccurate effect sizes which results in interpreting predictors with potentially weak effects as having a more substantial impact than they really have or strong predictors as weaker than they really are.

The essential difference between traditional fixed-effects regression models, i.e. the type of model used by Varbrul (cf. Cedergren & Sankoff. 1974) and GoldVarb analyses, and mixed-effects models is that mixed-effect models allow one to integrate within-speaker variation while fixed-effects models do not. In other words, mixed-effects models enable to integrate the widely acknowledged and generally uncontested fact that the linguistic performance of a speaker shows variation, i.e. variation does not only exist on the group level but also on an individual level. Technically speaking, mixed-models can handle nested or hierarchical data and variable structures while fixed-effects models can only handle non-nested data and variable structures, e.g. data where each data point is independent from any other data point in the data.

Both step-wise step-up and step-wise step-down model fitting arrived at the same final minimal adequate model (cf. Tables 5 and 6 in the Appendix). The analysis only included higher-level interactions which did not cause complete separation.

4 Visualization of the Results

This section displays and summarizes the results of the analysis. Figures 2, 3, 4, and 5 show the distribution of occurrences of *eh* across age groups (young = 39 or younger, old = 40 or older), genders, ethnicities, and with respect to the occupation type of speakers (sml = skilled maunal labor, acmp = academic, clerical, managerial or the professions).

Figures 2 and 3 display the distribution of *eh* across age groups and genders.

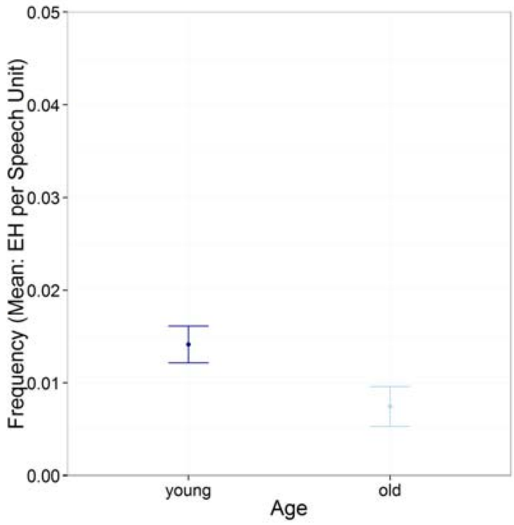

Figure 2: eh *by age.*

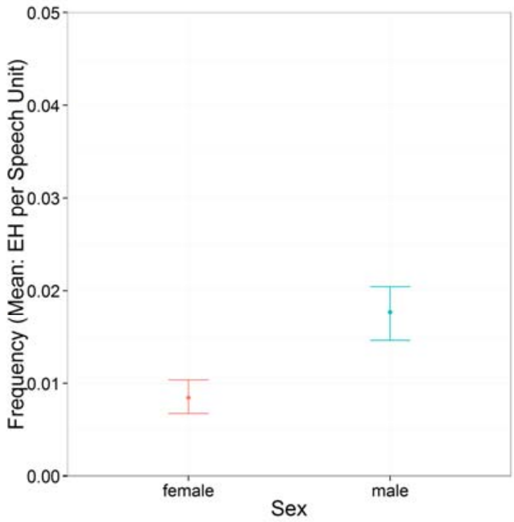

Figure 3: eh *by gender.*

Figure 2 suggests an effect of age on *eh* usage as the mean frequency of *eh* per speech unit is substantially higher than among older speakers. Figure 3 indicates that male speakers use *eh* substantially more than female speakers.

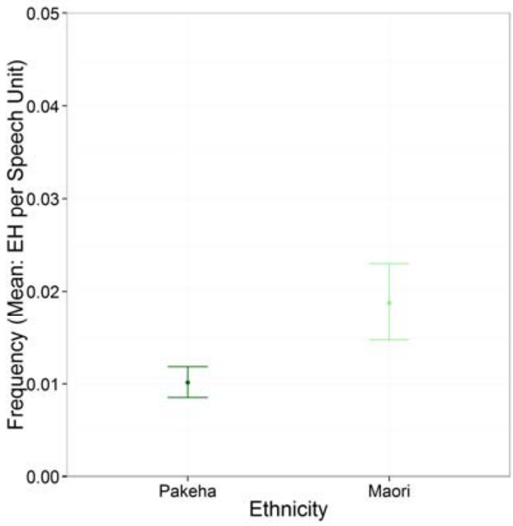

Figure 4: eh *by ethnicity.*

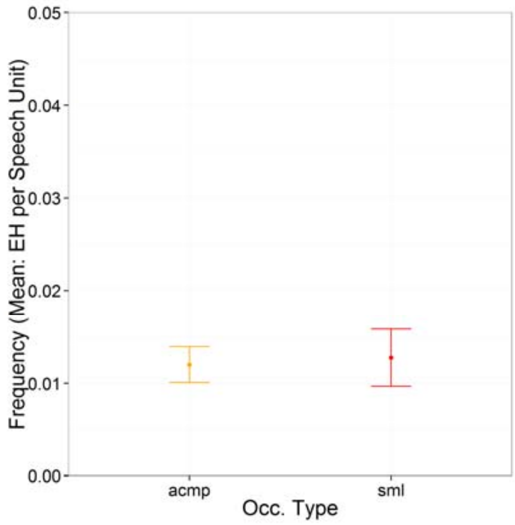

Figure 5: eh *by occupation type.*

Figure 4 shows that Maori speakers exhibit a higher mean frequency of *eh* by speech unit than Pakeha speakers. Figure 5 suggests that the occupation type of speakers does not significantly affect the occurrence of *eh* as the error bars clearly overlap.

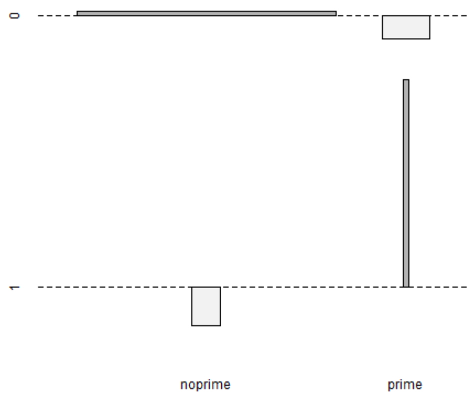

Figure 6: association plot of eh *by priming.*

Figure 6 shows the distribution of *eh* with respect to priming, i.e. whether *eh* occurred within an 18-word span following an instance of *eh* or not. It indicates that the likelihood of *eh* occurring is substantially higher if *eh* was used shortly before.

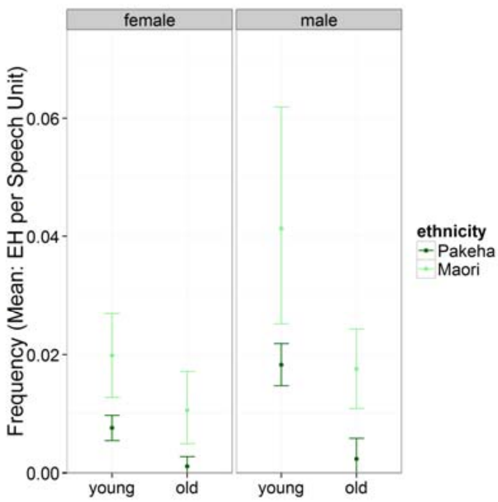

Figure 7: eh *by gender, age, and ethnicity.*

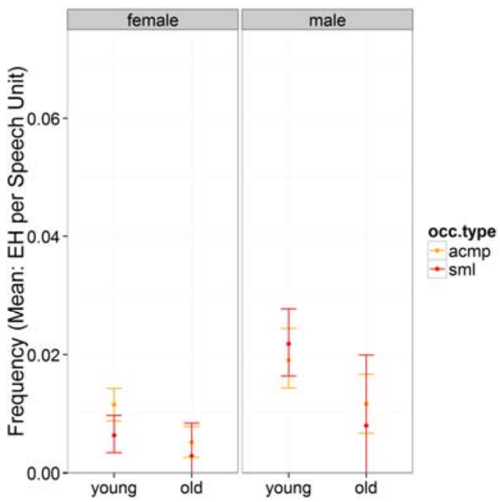

Figure 8: eh *by gender, age, and occupation type.*

The distribution of mean frequencies of *eh* per speech unit displayed in Figure 7 indicates that the social group which uses *eh* most often is young male Maoris while young female Maoris exhibit the second highest, young male Pakehas the third, and old male Maoris the fourth highest mean. The lowest mean frequencies occur amongst old female and male Pakehas. This distribution suggests that ethnicity has the strongest effect on the usage of *eh* while the gender and age of speakers also correlates with *eh* usage but their impact is less clear. Figure 8 suggests that age, gender, and occupation type do not interact significantly as the error bars mostly overlap.

5 Results of the Mixed-Effects Binomial Logistic Regression Model

The final minimal adequate mixed-effects binomial logistic regression model[4] performs significantly better than a base-line model[5] and reports the age[6], gender[7], and ethnicity[8] of speakers as significant predictors. The model neither confirmed that the occupation type of speakers – or social class if you will -- affects the occurrence of *eh* nor did it confirm a significant impact of priming on the occurrence of this particle. In addition, the model did not report any significant interactions[9]. Including the random effect (random intercepts) was justified as it significantly improved model fit.[10]

The ethnicity of speakers has the strongest impact on the usage of *eh*: Maoris use *eh* 4.51 times for every instance of *eh* uttered by a Pakeha. Age also substantially affects the likelihood of *eh* usage: for every instance of *eh* uttered by speakers above the age of 40, there are 4.48 instances uttered by speakers between the ages of 16 and 39. The gender of speakers has the smallest effect size as males are only 2.26 times more likely to use *eh* than female speakers.

Table 1: Results of the final minimal adequate mixed-effects binominal logistic regression model.

	Groups	Variance	Std. Dev.	L.R. χ^2	DF	Pr(>|z|)	Significance
Random Effect(s)	file.speaker.id	1.22	1.1	143.44	1	0.0000	p<.001***
Fixed Effect(s)	Estimate[11]	VIF[12]	OddsRatio[13]	Std. Error	z value	Pr(>|z|)	Significance
(Intercept)	-6.97	0	0	0.44	-15.72	0.0000	p<.001***
age:young	1.50	1.24	4.48	0.38	3.91	0.0001	p<.001***
sex:male	0.81	1.02	2.26	0.28	2.96	0.0031	p<.01**
ethnicity:maori	1.51	1.26	4.51	0.35	4.28	0.0000	p<.001***
Model statistics				L.R. χ^2	DF	Pr(>|z|)	Values
Number of Groups							140
Number of Cases in Model							17,770
Observed Misses							17,553
Observed Successes							217
Residual Deviance[14]							2,168.87
R^2 (Nagelkerke)[15]							0.079
C[16]							0.845
Somer's D_{xy}[17]							0.689
AIC[18] (BIC[19])							2,178.87 (2,217.79)
Prediction Accuracy							98.78%
Model Likelihood Ratio Test				174.4	4	0.0000	p<.001***

6 DISCUSSION

The current analysis of the discourse-pragmatic particle *eh* in New Zealand English has shown that the usage of *eh* is socially stratified. The results of the mixed-effects model show that *eh* is used particularly often by young speakers, by male speakers, and by speakers who are Maori, while neither priming nor the occupation type of speakers – a proxy for social class – correlate significantly with the usage of *eh*. To emphasize, the present case study of *eh* in NZE cannot confirm that this syntactically optional and semantically bleached discourse particle is significantly affected by priming effects, i.e. by a psycholinguistic factor. The absence of a significant effect of occupation type is intriguing as both Stubbe & Holmes (1995) and Meyerhoff (1994) reported that *eh* is more common among working class speakers. The present analysis thus raises doubts over previous claims according to which *eh* serves as a marker of working-class speech (Stubbe & Holmes 1995). Similarly, the absence of significant interactions contrasts with what would be expected based on previous research which found interactions between the gender and ethnicity of speakers (Meyerhoff 1994) or between the social class of speakers and their gender (Stubbe & Holmes 1995: 75).

What the analysis shows is that *eh* has a distinct social profile as the sociolinguistic factors – with the exception of occupation type – investigated in this study correlate significantly with the usage of *eh*. The result that ethnicity has the strongest impact on the occurrence of *eh* validates previous research which suggested that *eh* serves as an identity marker among Maoris (Bauer 1997; Bell & Kuiper 1999; Holmes 1997, 2005; Holmes, Stubbe & Marra 2003; Meyerhoff 1994; Stubbe & Holmes 1995). Both male and female Maori display comparatively high frequencies of this particle and the regression model reports ethnicity as the predictor with the strongest effect.

Also, younger speakers make use of this variant substantially more often than older speakers (above age 40) which strongly suggests that younger speakers adopt this form to set themselves apart from older speakers. This appears to be the case, particularly among Pakehas since – although the difference between young and old Maoris is much larger than the difference between young and old Pakehas – *eh* is almost completely absent from the speech of old Pakehas. It seems that young Pakehas adopt this variant quite readily while it is rejected by older Pakehas probably due to overt stigmatization. This pattern suggests ongoing language change as *eh* is entering the Pakeha speech community and young speakers appear to be the early adopters. In addition, it may well be speculated that *eh* carries covert prestige among youngers speakers who use it to distinguish themselves linguistically from their parents and grand-parents.

With the exception of the insignificant impact of occupation type, i.e. the proxy for social class, the stratification of *eh* usage reflects the typical pattern of innovative variants which undergo change and carry covert prestige because male speakers commonly lead changes if the innovation is overtly stigmatized but carries covert prestige as in the present case (Trudgill 1975; Labov 1994). This pattern thus conforms with Labov's 'Gender Paradox' according to which women conform more closely than men to sociolinguistic norms that are overtly prescribed, but conform less than men when they are not (Labov 2001: 292—3). The fact that *eh* conforms to the 'Gender Paradox' in the present analysis is surprising as previous research suggested that *eh* does not exhibit the social profile predicted by the 'Gender Paradox', i.e. that males use stigmatized elements more frequently than females. To elaborate, Meyerhoff (1994: 384) found that Pakeha men lag behind Pakeha women in the usage of *eh* which led her to conclude that *eh* does not exhibit the prototypical behavior of linguistic elements undergoing change. We will return to this point below.

Previous studies have suggested change in progress as *eh* diffuses through the New Zealand speech community with Maoris being the leaders of change. Among Pakehas, young male working-class speakers were deemed the leaders of change (Stubbe & Homes 1995: 84). The present study partially confirms this claim as young, male Pakehas are indeed the cohort among Pakehas who use *eh* most. However, the notion that it is specifically young male working-class Pakehas is not supported, as the occupation type of speakers -- the proxy for social class in this study – is not correlated significantly with the use of *eh*. Furthermore, although the apparent-time[20] distribution is suggestive of ongoing change, an additional real-time analysis would be necessary to confirm this impression as only real time evidence can prevent potential misdiagnoses of apparent-time evidence (Milroy & Gordon 2003: 36). Unfortunately, all spoken private dialogue data of the New Zealand component of the ICE were collected in 1993 and 1994 so that the data do not lend themselves to a real--time analysis.

The absence of significant interactions in the ICE NZ data is particularly intriguing as it conflicts with the results of previous analyses of *eh*. While Meyerhoff (1994) found an interaction between the age and ethnicity of speakers, this study cannot confirm the existence of such an interaction based on the ICE data. The results of the current study, in contrast, suggest that *eh* does indeed match the prediction that male speakers adopt stigmatized variants that carry covert prestige more readily than female speakers. In other words, the results of the current study perfectly match the distribution that would be expected based on Labov's 'Gender Paradox'. In addition, the mixed-effects binomial logistic regression model failed to confirm the findings by Stubbe & Holmes (1995) who argued that *eh* functions as a marker of male working-class identity – in contrast to Meyerhoff (1994) where *eh* is considered a marker of ethnic identity among Maoris. In fact, the occupation type of speakers, which served as a proxy for social class in this study, could neither be confirmed to be significantly correlated with the use of *eh* nor could the analysis validate the claim that social class and gender interact.

There are at least five possible explanations for the conflicting results.

1. The most obvious cause of the difference in the results could be a difference in data sets that were analyzed. However, the kind of data used for the present study and Meyerhoff (1994) as well as Stubbe & Holmes (1995) are very similar. The present analysis relies on the spontaneous private dialogue data that is part of the ICE NZ while Meyerhoff (1994) is based on the Wellington social dialect data and uses transcripts of the free speech/casual conversations (Meyerhoff 1994: 368). The data used by Stubbe & Holmes (1995) were taken from the Wellington Corpus of Spoken New Zealand English (WCSNZE) (Stubbe & Holmes 1995: 66) which does have a substantial overlap with the ICE data Vien (1999: 8), especially the spoken sections are closely linked and share nine categories. Therefore, the difference in the findings is unlikely to stem from differences in the data sets.
2. Meyerhoff herself offers a likely explanation for her finding that an interaction between the sex and ethnicity of speakers is at work with respect to the usage of *eh*. She hypothesized that the working class

Pakeha women in her data might have been more prepared to use *eh* compared to their male peers because of their personal associations with Maori men who they were in personal relationships with (Meyerhoff 1994: 384). If this were true, then the absence of significant interactions in the ICE NZ data would reflect actual language use of the population while the results of Meyerhoff's study reflect an idiosyncrasy of her data.

3. Mixed-effects regression models are more likely than fixed-effect models to commit β-errors (Type II errors), i.e. they are prone to reporting independent variables as being insignificant although they are, in reality, significant predictors. The probability of β-errors in mixed-effect models can be extremely high in cases where the effect size of the predictor is very weak, there is high within-speaker variance, and each speaker utters only a few utterances (Johnson 2009: 368—69).
4. In the present case, the low number of occurrences of *eh* or large within-speaker variance could potentially cause the mixed-effects model to be less reliable than would be desirable. Both the fixed- and the mixed-effects model have low explanatory power (cf. the pseudo-R^2 values in Table 4) which – in addition to the summary in Table 2 – indicates large within-speaker variance which is a contributing factor to the occurrence of β-errors in mixed-effects analyses (Johnson 2009: 368—69). However, mixed-models are generally preferable as they are substantially less likely to commit α-errors (Type I errors), i.e. they are less likely to report predictors as being significant although they are not (Johnson 2009: 368—69).
5. Previous studies have not used multivariate statistics, let alone mixed-effects models. This is very unfortunate, as mono- and bivariate statistics are prone to be misleading as they cannot control for underlying confounding factors, i.e. factors which correlate with both the dependent and the independent variable. In previous analyses of *eh*, it is very likely that the effects of age and ethnicity represent such confounding factors when applying bivariate statistics to investigate the correlation between *eh* and social class as in the case of Stubbe & Holmes (1995).

Since Meyerhoff (1994) based her analysis on a relatively limited, partially atypical data base (remember that three of the five Pakeha women either lived with or were married to Maori men), her results may not lend themselves easily to generalizations. The results of the present study are based on a substantially larger data set and retrieved with the help of multivariate regression modeling. Hence, the outcome of the present analysis is more reliable than the results of previous studies. This stresses the advantages of applying multivariate statistics to avoid overlooking confounding factors and to inspect the relative effect size of factors.

An additional issue of the present analysis, which requires discussion before a conclusion can be drawn, relates to the operationalization of the dependent variables. The procedure of collapsing age groups and, in particular, occupations into binary factors and the application of an 18-word demarcation mark for priming effects can be viewed as problematic.

Collapsing factors into binary divisions is certainly sub-optimal as this causes a loss of vital information. Labov even states that binary divisions are of little value in sociolinguistic analyses and suggests that social class should be represented as a factor with a minimum of four levels or groups: two peripheral groups and two central groups (Labov 1990: 220). This criticism is undoubtedly valid but it does not sufficiently take an existing trade-off between the reliability of the results and the granularity of the analysis into account. In essence, this means that there are two conflicting principles at work here: the need to provide reliable results – which requires cohorts that are large enough to license generalizations about the speech community – and the need for fine-grained analysis, which requires a maximum of detail -- the more-fine grained the data are, the more accurately can we observe the actual interactions within the sample. In previous variationist research, studies have used four divisions at the cost of violating the assumption of independence. This study does not violate the assumption of independence as it is statistically problematic (Johnson 2009) but was consequently forced to apply binary divisions.

As to the decision to apply an 18-word rather than a 10 word demarcation mark for priming: considering that the mean length of speech units is 7.3 words with a standard deviation of 6.7 words, a 10-word span which would have led to the inclusion – if at all(!) – of the speech unit final slot immediately adjacent to the occurrence of an *eh* while the 18-word span encompasses between two and three slots subsequent to a slot filled with *eh* which is more likely to reflect actual re-use.

Whatever the exact causes for the differences between the findings of the present analysis and previous research might be, the current analysis has raised some interesting issues and indicated directions for further research. One issue which has not been addressed in detail above is the lack of a qualitative analysis in this study. It would be intriguing to also consider the discourse-pragmatic functionality of *eh* and investigate whether social stratification correlates with certain functional variants of *eh*.

7 Conclusion

This study shows that extra-linguistic, socio-demographic variables (gender, age, occupation type, and ethnicity) correlate significantly with the usage of the speech unit final discourse particle *eh* in the New Zealand component of the ICE. Speakers who are 39 or younger use *eh* significantly more often than speakers who are 40 or older. Male speakers are significantly more likely than female speakers, and Maoris are significantly more likely than Pakehas to use *eh*. In contrast to what has been suggested by the relevant literature (Meyerhoff 1994), the analysis neither confirmed a significant effect of the occupation type of speakers on their usage of *eh* nor did it confirm significant interactions between gender and ethnicity (Meyerhoff 1994) or between occupation type and gender (Stubbe & Holmes 1995). In addition, the results of a mixed-effects binary logistic regression model show that priming does not significantly impact the occurrence of *eh*.

APPENDIX

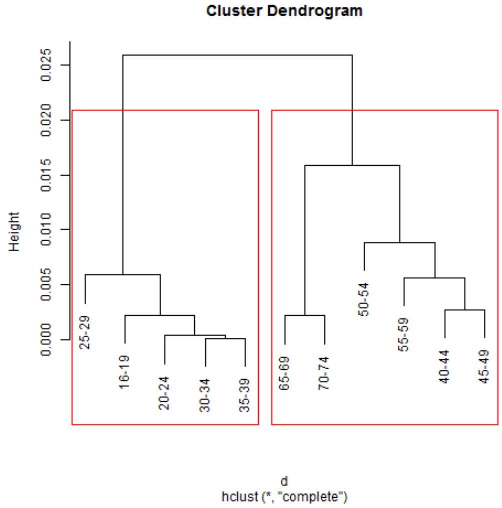

Figure 9: Clustering of age groups based on their mean frequency of eh.

Table 5: Results of the model fitting process (step-wise step-down): mixed-effects binominal logistic regression model.

Full Model	suf.eh~age+sex+ethnicity+priming+soc.stat+age:sex+age:ethnicity+sex:ethnicity+ age:soc.stat+sex:soc.stat+ethnicity:soc.stat+age:sex:ethnicity+(1|le.speaker.id)										
Model	Term Deleted	Compared to…	DF	AIC	BIC	LogLikelihood	Residual Deviance	χ^2	χ^2 DF	p-value	Significance
m12.glmer	age:sex:ethnicity	m11.glmer	14	2,188.3	2,297.3	-1,080.1	2,160.3	0.11	1	0.7437	n.s.
m11.glmer	ethnicity:occ.type	m10.glmer	13	2,186.41	2,287.6	-1,080.2	2,160.4	1.72	1	0.1902	n.s.
m10.glmer	sex:occ.type	m9.glmer	12	2,186.12	2,279.5	-1,081.1	2162.1	3.34	1	0.0674	p<:10(*)
m9.glmer	age:occ.type	m8.glmer	11	2,187.47	2,273.1	-1,082.7	2,165.5	0.79	1	0.3752	n.s.
m8.glmer	sex:ethnicity	m8.glmer	10	2,186.25	2,264.1	-1,083.1	2,166.3	0.02	1	0.8923	n.s.
m7.glmer	age:ethnicity	m8.glmer	9	2,184.2	2,254.3	-1,083.1	2,166.3	0.83	1	0.3614	n.s.
m6.glmer	age:sex	m7.glmer	8	2,183.1	2,245.3	-1,083.6	2,167.1	0.59	1	0.4407	n.s.
m5.glmer	occ.type	m6.glmer	7	2,181.7	2236.2	-1083.9	2,167.7	1.17	1	0.2798	n.s.
m4.glmer	priming	m5.glmer	6	2,180.8	2,227.5	-1,084.4	2,168.9	0.00	1	0.9678	n.s.
m3.glmer	ethnicity	m4.glmer	5	2,178.8	2,217.7	-1,084.4	2,168.9	18.8	1	0.0000	p<:001***
m2.glmer	sex	m1.glmer	4	2,195.6	2,226.7	-1,093.8	2,187.6	5.19	1	0.0227	p<:05*
m1.glmer	age	m0.glmer	3	2,198.8	2,222.1	-1,096.4	2,192.8	7.02	1	0.0080	p<:01**

Table 6: Results of the model fitting process (step-wise step-up): mixed-effects binominal logistic regression model.

Base-Line Model	suf.eh~(1|le.speaker.id)

Model	Term Added	Compared to…	DF	AIC	BIC	LogLikelihood	Residual Deviance	χ^2	χ^2 DF	p-value	Significance
m1.glmer	age	m0.glmer	3	2198.8	2222.2	-1,096.4	2,192.8	7.02	1	0.00804	p<:01**
m2.glmer	sex	m1.glmer	4	2,195.6	2,226.8	-1,093.8	2,187.6	5.19	1	0.02275	p<:05*
m3.glmer	ethnicity	m2.glmer	5	2,178.9	2,217.8	-1,084.4	2,168.9	18.7	1	0.00001	p<:001***
m4.glmer	priming	m3.glmer	6	2,180.9	2,227.6	-1,084.4	2,168.9	0.00	1	0.96782	n.s.
m5.glmer	occ.type	m4.glmer	7	2,181.7	2,236.2	-1,083.9	2,167.7	1.17	1	0.27985	n.s.
m6.glmer	age:sex	m5.glmer	8	2,183.1	2,245.4	-1,083.6	2,167.1	0.59	1	0.44079	n.s.
m7.glmer	age:ethnicity	m6.glmer	9	2,184.3	2,254.3	-1,083.1	2,166.3	0.83	1	0.36146	n.s.
m8.glmer	sex:ethnicity	m7.glmer	10	2,186.3	2,264.1	-1,083.1	2,166.3	0.02	1	0.89231	n.s.
m9.glmer	age:occ.type	m8.glmer	11	2,187.5	2,273.1	-1,082.7	2,165.5	0.79	1	0.37529	n.s.
m10.glmer	sex:occ.type	m9.glmer	12	2,186.1	2,279.6	-1,081.1	2,162.1	3.34	1	0.06748	p<:10(*)
m11.glmer	ethnicity:occ.type	m10.glmer	13	2,186.4	2,287.6	-1,080.2	2,160.4	1.72	1	0.19028	n.s.
m12.glmer	age:sex:ethnicity	m11.glmer	14	2,188.3	2,297.3	-1,080.2	2,160.3	0.11	1	0.74374	n.s.

REFERENCES

Aijmer, Karin. 2002. *English Discourse Particles: Evidence from a Corpus, vol. 10*. Amsterdam & Philadelphia, PA: John Benjamins.

Althaus, Scott L. & Young Mie Kim. 2006. Priming Effects in Complex Information Environments: Reassessing the Impact of News Discourse on Presidential Approval. *Journal of Politics* 68(4), 960-976.

Baayen, R. Harald. 2008. *Analyzing Linguistic Data. A Practical Introduction to Statistics using R*. Cambridge: Cambridge University Press.

Bailey, Guy, Tom Wikle, Jan Tillery & Lori Sand. 1991. The apparent time construct. *Language Variation and Change* 3(3), 241—64.

Bauer, Laurie. 2007. Some Grammatical Features of New Zealand English. *New Zealand English Journal* 21, 1—25.

Bauer, Laurie, Allan Bell, David Britain, Graeme Kennedy, Chris Lane, Miriam Meyerhoff & Maria Stubbe. 1999. *The New Zealand Component of the International Corpus of English (ICE-NZ)*. Wellington, NZ: Victoria University of Wellington.

Bauer, Winifred. 1997. *The Reed Reference Grammar of Maori*. Auckland: Reed.

Bell, Allan & Koenraad Kuiper. 1999. *New Zealand English, vol. 25*. Amsterdam & Philadelphia, PA: John Benjamins.

Cedergren, Henrietta J. & David Sankoff. 1974. Variable Rules: Performance as a Statistical Rejection of Competence. *Language* 50(2), 333—55.

Columbus, Georgy. 2009. 'Ah lovely stuff, eh?' Invariant Tag Meanings and Usage across Three Varieties of English. *Language and Computers* 71(1), 85—102.

Du Bois, John W., Wallace L. Chafe, Charles Meyer, Sandra A. Thompson, Robert Englebretson & Nii Martey. 2000—2005. *Santa Barbara Corpus of Spoken American English, Parts 1-4*. Philadelphia, PA: Linguistic Data Consortium.

Field, Andy, Jeremy Miles & Zoe Field. 2012. *Discovering Statistics using R*. London: Sage.

Gold, Elaine & Mireille Tremblay. 2006. Eh? and hein?: Discourse Particles or National Icons?. *The Canadian Journal of Linguistics/La revue canadienne de linguistique* 51(2), 247—63.

Gries, Stefan Th. 2005. Syntactic priming: A corpus-based approach. *Journal of Psycholinguistic Research* 34(4), 365—99.

Gries, Stefan Th. 2013. Sources of variability relevant to the cognitive sociolinguist, and corpus- as well as psycholinguistic methods and notions to handle them. Journal of Pragmatics 52, 5—16.

Holmes, Janet. 1997. Maori and Pakeha English: Some New Zealand Social Dialect Data. *Language in Society* 26 (1), 65—101.

Holmes, Janet. 2005. Using Maori English in New Zealand. *International Journal of the Sociology of Language* 2005(172), 91—115.

HoImes, Janet, Maria Stubbe & Meredith Marra. (2003). Language, Humour and Ethnic Identity Marking in New Zealand English. In Mair (ed.), 431—55.

Johnson, D. Ezra. 2009. Getting off the Goldvarb Standard: Introducing rbrul for Mixed-Effects Variable Rule Analysis. *Language and Linguistics Compass* 3(1), 359—83.

Johnson, D. Ezra. 2010. Why step-wise isn't so wise. Paper presented at NWAV 39 ('New' New Ways of Analyzing Variation). San Antonio, TX, November 4th-6th, 2010.

Kallen, Jeffery L. & John M. Kirk. 2008. *ICE-Ireland: A User's Guide Documentation to accompany the Ireland Component of the International Corpus of English (ICE-Ireland)*. Belfast, NI: Clo Ollscoil na Banrona, School of English, Queen's University Belfast.

Labov, William. 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2(2), 205—54.

Labov, William. 1994. *Principles of Language Change: Internal Factors, vol. 1*. Oxford: Blackwell.

Labov, William. 2001. *Principles of Linguistic Change: Social Factors, vol. 2*. Blackwell.

Mair, Christian (ed.). 2003. *The Politics of English as a World Language: New Horizons in Postcolonial Cultural Studies*. Amsterdam & New York: Rodopi.

Meyerhoff, Miriam. 1994. Sounds Pretty Ethnic, eh?: A Pragmatic Particle in New Zealand English. *Language in Society* 23, 367—67.

Milroy, Lesley & Matthew Gordon. 2003. Sociolinguistics: Method and Interpretation, vol. 13. Malden, MA: Blackwell.

Stubbe, Maria & Janet Holmes. 1995. You know, eh and other 'exasperating expressions': An Analysis of Social and Stylistic Variation in the Use of Pragmatic Devices in a Sample of New Zealand English. *Language & Communication* 15(1), 63—88.

Szmrecsanyi, Benedikt. 2005. Language Users as Creatures of Habit: A corpus-based Analysis of Persistence in Spoken English. *Corpus Linguistics and Linguistic Theory* 1(1), 113—150.

Szmrecsanyi, Benedikt. 2006. *Morphosyntactic Persistence in Spoken English: A Corpus Study at the Intersection of Variationist Sociolinguistics, Psycholinguistics, and Discourse Analysis*. Berlin & New York: Walter de Gruyter.

Thorne, Barrie & Nancy Henley (eds.). 1975. *Language and Sex: Difference and Dominance*. Rowley, MA: Newbury House.

Trudgill, Peter. 1975. Sex, Covert Prestige and Linguistic Change in the Urban British English of Norwich. In Thorne & Henley (eds.), 88—104.

Tulving, Endel & Daniel L. Schacter. 1990. Priming and Human Memory Systems. *Science* 247 (4940), 301—306.

Turner, George W. 1966. *The English Language in Australia and New Zealand*. London: Longmans, Green & Co.

Vine, Bernadette. 1999. *Guide to the New Zealand Component of the International Corpus of English (ICE-NZ).* Wellington, NZ: School of Linguistics and Applied Language Studies, Victoria University of Wellington.

FOOTNOTES

[1] The text in the examples in (1a—d) was set to lower case and cleared of meta-linguistic information for the sake of readability.

[2] The phrase 'Speech Unit' is adopted from the terminology of the compilers of another data base that is part of the ICE family of corpora (ICE Ireland; cf. Kallen & Kirk 2008). The phrase *Speech Unit Final* can be considered equivalent to the phrases 'clause-final' or 'sentence-final' in literature.

[3] That is to say all files with headers starting with S1A.

[4] For a more extensive overview of the results cf. Table 4.

[5] Model Likelihood Ratio Test: L.R. χ^2: 174.4, DF: 4, Significance: p-value<.001***.

[6] χ^2: 7.02, DF: 1, Significance: p-value<.001**.

[7] χ^2: 5.19, DF: 1, Significance: p-value<.05*.

[8] χ^2: 18.8, DF: 1, Significance: p-value<.001***.

[9] It should be noted that an interaction between gender and occupation type was reported as being marginally significant during model fitting (cf. Tables 5 and 6 in the Appendix). This interaction was added to the final minimal adequate model to test if its inclusion would significantly improve the model fit. A comparison between a model with and a model without that interaction showed that the model fit did not improve significantly if the interaction was included (χ^2: 1.42, DF: 2, p-value: 0.49).

[10] Model Likelihood Ratio Test: L.R χ^2: 143.44, DF: 1, p-value <.001***}

[11] Estimates are not standardized estimates of effect size and indicate whether the predictor correlates positively (positive estimate) or negatively (negative estimate) with the occurrence of *eh*.

[12] VIF (Variance Inflation Factor) values provide information about multi-collinearity in the data, i.e. whether independent variables correlate. Values higher than 10 indicate unacceptable correlations between independent variables, but even values above 4 are usually already problematic and, for large data sets, even values higher than 2 can lead inflated standard errors. Values below 2 signal that multi-collinearity is not a problem for the current model.

[13] OddsRatios (odds ratio) are a normalized measure of effect size. Values greater than 1 indicate a positive correlation between the independent variable and the usage of *eh*

while values below 1 indicate a negative correlation (if the value is exactly 1, then there is no correlation between the independent and the dependent variable).

[14] Residual deviance is a measure for the amount of variance not explained by the current model. Deviances can only be meaningfully compared across models that are fit to the same data.

[15] R^2 (Nagelkerke): this pseudo-R^2 is essentially identical to the Cox & Snell pseudo-R^2 but can – at least mathematically – reach 1 (cf. Field, Miles & Field 2012: 317—18). Cox & Snell pseudo-R^2 depends on the deviance of the baseline model, the current model, and the number of cases in the models. The model fit increases as this pseudo-R^2 approximates 1.

[16] C is an index of concordance between the predicted probability and the observed response. When C takes the value 0.5, the predictions are random, when it is 1, prediction is perfect. A value above 0.8 indicates that the model may have some real predictive capacity (cf. Baayen 2008: 204).

[17] Somers' D_{xy} is a rank correlation between predicted probabilities and observed responses ranges between 0 (randomness) and 1 (perfect prediction) (Baayen 2008: 204).

[18] AIC: Akaike Information Criterion (AlC = -2LL + 2k) is an estimate of parsimony which provides information on the balance between the amount of variance being explained by the model on the one hand and the number of predictors necessary to obtain the amount of variance being explained on the other (cf. Field, Miles & Field 2012: 318). The lower the AIC the better the balance between the amount of variance explained and the number of variables necessary to explain that amount of variance. The AICs of models can be compared only, if the models are fitted to the same data!

[19] BIC: Bayesian Information Criterion (BlC = -2LL + 2k * log(n)) is very similar to the AIC but adjusts the penalty included in the AlC (i.e., 2k) by the number of cases in the model (cf. Field, Miles & Field 2012: 318). The lower the BIC the better the balance between the amount of variance explained and the number of variables necessary to explain that amount of variance. The BICs of models can be compared only, if the models are fitted to the same data!

[20] For an explanation and a detailed discussion of the apparent-time construct (cf. Bailey et al. 1991).