

**ASPECTS OF NEW ZEALAND ENGLISH  
INTONATION AND ITS MEANINGS:**  
*An experimental investigation of forms and contrasts*

**Patricia Vermillion**

**Aspects of New Zealand English intonation and its meanings:  
*An experimental investigation of forms and contrasts***

by

**Patricia Vermillion**

A thesis  
submitted to Victoria University of Wellington  
in fulfilment of the  
requirements for the degree of  
Doctor of Philosophy  
in Linguistics

Victoria University of Wellington

2006

**Author: Patricia Vermillion**

School of Linguistics and Applied Language Studies,  
Victoria University of Wellington,  
New Zealand

**Supervisors: Dr. Paul Warren**

School of Linguistics and Applied Language Studies,  
Victoria University of Wellington,  
New Zealand

**Prof. Janet Holmes**

School of Linguistics and Applied Language Studies,  
Victoria University of Wellington,  
New Zealand

## ABSTRACT

---

The first aim of this thesis was to examine the form of New Zealand English intonation. The results of the first series of experiments illustrated several distinctive features of NZE intonation and the preferential tonal use within this language variety. The results from the first experiment suggests that NZE intonation can be characterised as having a narrow pitch range within the phrase and a wide pitch range at the end of the phrase in relation to British English. The findings in the second analysis illustrate that tonal composition, not sentence type affects the pitch range that NZE speaker uses. In addition to pitch range preferences, NZE speakers were also found to prefer an H\*L-L% nuclear tonal composition on statements and an L\*L-H% for two types of questions when conversational cues were not required by the task type.

The second aim of the thesis was to define the tonal features which may adequately describe the semantic contrasts used in this variety. Five experiments were carried out with this aim in mind. The results revealed that NZE listeners use the height of pitch target values when interpreting the meaning of intonation and that the heights of three tonal constituents would be useful in notating the semantic contrasts in this language variety. First, the pitch accent target is used in this variety to indicate speaker involvement, whereby higher (H\*) or later (L+H\*) pitch accent targets indicate a greater degree of involvement than lower (L\*) or earlier targets (H\*), respectively. This claim was supported by a production experiment (Chapter 5) in which speakers were asked to convey contrasting meanings on identical utterances. The results were such that higher and later pitch peaks were produced to convey concern, emphasis and an impressed attitude, while lower and earlier pitch accent peaks indicated an absence of these three meanings. Further support for this claim was provided in a perception experiment (Chapter 7), which investigated how listeners interpret conversational markers indicating discourse completeness. The results show that NZE listeners interpret higher H\* targets as indicating speaker involvement and, subsequently, listener-oriented turn cues. However, a non-emphatic H\*, or a high pitch accent which is lower in pitch than a preceding high pitch accent, does not convey such cues.

Second, the boundary target is used to contrast continuation with high phrase-final targets and finality with low phrase-final targets in NZE. This assertion was supported by a perception experiment (Chapter 6) which examined categorical boundaries determined by the boundary tone height. The results suggest that there is at least one categorical boundary at the IP-Final position, which is marked by the pitch movement to the boundary target from the preceding H\*. In addition, the semantic contrast of the boundary target height was illustrated in two experiments. First, a production experiment (Chapter 5) illustrated how NZE speakers indicate conversational continuation cues and concern with high boundaries whereas low boundaries indicated conversational cessation cues and a lack of concern. A separate perception experiment (Chapter 7) showed that NZE listeners interpret higher boundary targets as speaker continuation cues and listener-oriented speaking cues whereas lower boundary targets again indicated conversational cessation cues for the speaker and to the listener.

Third, the phrase accent may prove useful in distinguishing a further semantic contrast used in this language variety, with a level pitch movement from H\* to the IP-Final boundary target categorised with the H% stimuli (suggested in Chapter 6) while the distinction between H-L% and L-L% may be best defined as a pitch movement which does not fall to the F0 minimum and a movement which does fall to this low value (Chapter 5). Although the existence of a phrase accent could not be proved in this thesis, the results illustrate support for this tonal feature.



## ACKNOWLEDGEMENTS

---

First, I would like to extend my deepest gratitude to my supervisor, Dr. Paul Warren for so much of his assistance throughout my time at Victoria University; first, for teaching me about not only intonation but experimental design, writing, editing and thought processes; second, for providing interesting conversation and discussions, and for the regular 'mental gymnasium' he provided for me through his many thought-provoking comments; third, for providing a platform for growth. Most of all, I am forever indebted to Dr. Warren for being my biggest supporter and my harshest critic – an ideal supervisor!

I would also like to thank Colleen Kelly for helping with the statistics used in several experiments in this thesis, always thoughtfully considering each of my experiments and detailing each analysis and design.

Many thanks to Paul Warren, Janet Holmes, Laurie Bauer, Liz Pearce and Sandra Shearn for allowing me to recruit participants in their classes. Also, I would like to thank the group of speakers who allowed me to use their voices as the source materials in the experiments. I owe many thanks for the assistance of Chris Peirce, Paul Warren, Laurie Bauer, Jonathan Newton, Des Hurley and Liz Warren. I would also like to thank the numerous students at Victoria University of Wellington for participating in my experiments. Although I cannot thank each participant by name, I want to thank those of you who kindly took part in my studies; your participation was much appreciated. In addition, I am grateful to the staff of the Language Learning Centre at Victoria University, notably, Karen Boxall for allowing me to use lab so frequently and assisting me, as well as Richard Keenan for his technical assistance. I also greatly appreciate the assistance of many different native NZE speakers who helped me with the development of my experiments.

Finally, I would like to thank my financial supporters, including Victoria University Wellington for awarding me a Vice Chancellor's Strategic Research Scholarship which provided the financial backing to carry out such a project, and to School of Linguistics and Applied Language Studies for the financial assistance required to carry out and report the experiments detailed in this thesis.

## TABLE OF CONTENTS

---

<b>1.0. INTRODUCTION AND BACKGROUND</b>	<b>1</b>
1.0. General Introduction	1
1.1. New Zealand English intonation	3
1.1.1. The High Rising Terminal in NZ	3
1.1.2. Social factors and intonation in NZE	5
1.2. Methodological issues	11
1.2.1. Unit of Measurement	11
1.2.2. Notation system	12
1.3. Outline of Thesis	18
<b>2.0. INTONATION: FORMS AND MEANINGS</b>	<b>19</b>
2.1. Attitudinal Meaning	21
2.1.1. Universal Meanings	22
2.1.2. Language Specific Meanings	23
2.2. Discoursal and Grammatical Meaning	26
2.2.1. Utterance Information	26
2.2.2. Sentence Type	30
2.2.3. Conversational Cues	31
2.3. Summary and Discussion	34
2.3.1. Pitch range	34
2.3.2. Pitch contour shape	36
2.3.3. Pitch height	39
2.4. Research questions	41
2.5. Summary of findings	42
<b>3.0. AN EXPERIMENT INVESTIGATING TONAL CUES USED TO IDENTIFY NEW ZEALAND ENGLISH</b>	<b>45</b>
3.0. Chapter abstract	45
3.1. Introduction	46
3.2. Perception of Tonal Cues	47
3.2.1. Materials	47
3.2.2. Participants	54
3.2.3. Procedure	55
3.2.4. Scoring	56
3.2.5. Results	57
3.3. Attitude Assessment	59
3.3.1. Materials	60
3.3.2. Procedure	61
3.3.3. Participants	61
3.3.4. Scoring	62
3.3.5. Results	62
3.4. Discussion and Conclusions	62
<b>4.0 THE PRODUCTION OF INTONATION IN READ ALOUD SPEECH</b>	<b>65</b>
4.0. Chapter abstract	65
4.1. Introduction	65
4.2. Methodology	68
4.2.1. Materials	68
4.2.2. Participants	70
4.2.3. Procedure	71
4.2.4. Analysis	72
4.3. Results and Discussion	76
4.3.1. Pitch level	76

4.3.2. Preferential tonal composition	76
4.3.3. Determinants of nuclear pitch range	79
4.4. Summary and conclusions	82
<b>5.0 THE PRODUCTION OF INTONATIONAL MEANING</b>	<b>84</b>
5.0. Chapter abstract	84
5.1. Introduction	85
5.2. Methodology	86
5.2.1. Materials	86
5.2.2. Participants	91
5.2.3. Procedure	92
5.2.4. Auditory and Acoustic Analysis	93
5.3. Results and Discussion	94
5.3.1. Conversational Continuation vs. Cessation cues	95
5.3.2. Authoritative vs. Submissive	97
5.3.3. Emphasis vs. Lack of Emphasis	102
5.3.4. Concern vs. Lack of Concern	105
5.3.5. Impressed vs. Unimpressed	109
5.4. Summary and Conclusions	112
<b>6.0 CATEGORIES AND MEANINGS CONVEYED BY VARIATION IN INTONATION PHRASE-FINAL HEIGHT</b>	<b>116</b>
6.0. Chapter abstract	116
6.1. Introduction	117
6.2. Discrete-Choice Task	118
6.2.1. Materials	118
6.2.2. Participants	122
6.2.3. Procedure	122
6.2.4. Results and Discussion	123
6.3. Gradient-Response Task	126
6.3.1. Materials	126
6.3.2. Participants	128
6.3.3. Procedure	128
6.3.4. Results	130
6.3.4.1. Liveliness	130
6.3.4.2. Confidence	131
6.3.4.3. Insistence	133
6.4. Summary and Conclusions	135
<b>7.0 INTONATIONAL PHRASE BOUNDARY TONES AND NUCLEAR H* AS MARKERS OF DISCOURSE COMPLETENESS</b>	<b>137</b>
7.0. Chapter abstract	137
7.1. Introduction	138
7.2. Methodology	140
7.2.1. Materials	140
7.2.2. Participants	144
7.2.3. Procedure	144
7.3. Results and Discussion	146
7.3.1. Listener Turn Cue	147
7.3.2. Speaker Turn Completion Cue	151
7.4. Summary and Conclusions	153
<b>8.0 SUMMARY AND CONCLUSIONS</b>	<b>155</b>
8.1. Summary and Discussion	156
8.2. Conclusions	163
8.3. Further considerations	166

<b>REFERENCES</b>	<b>168</b>
<b>Appendices</b>	<b>178</b>
<b>A. Chapter 3 experimental materials</b>	<b>179</b>
1. Recording materials used for source utterance and instructions	179
2. Consent and feedback forms	180
3. Instructions	182
4. Questionnaire and answersheet (Perception experiment)	183
5. Questionnaire and answersheet (Attitude experiment)	187
<b>B. Chapter 4 and 5 experimental materials</b>	<b>194</b>
1. Consent forms	194
2. Questionnaire	196
3. Instructions (Chapter 4)	197
4. Instructions (Chapter 5)	197
5. Recording material used for scene	197
<b>C. Chapter 6 experimental materials</b>	<b>198</b>
1. Consent and feedback forms	198
2. Recording material used for source utterance	200
3. Instructions (Task 1)	200
4. Instructions (Task 2)	200
5. Questionnaire and answersheet	202
<b>D. Chapter 7 experimental materials</b>	<b>208</b>
1. Consent and feedback forms	208
1. Recording material used for source utterance	210
2. Instructions (Task 1)	210
3. Instructions (Task 2)	211
4. Questionnaire and answersheet	213

## LIST OF FIGURES

---

<b>Chapter 1</b>		
1.1	F0 trace exemplifying H* and !H* pitch accents, and L-L% boundary	15
1.2	F0 trace exemplifying L*+H pitch accents and H-H% boundary.	15
<b>Chapter 2</b>		
2.1	F0 traces of NZE speaker saying "yeah"	19
2.2	Schematic representation of two rises identified in British English	25
2.3	Schematic representation of a fall-rise	26
<b>Chapter 3</b>		
3.1	Schematic representation of source utterance	47
3.2	Schematic representations of the 14 manipulations performed on the Pitch Points (PPT) in the source utterance (c.f. 3.1) used in the experiment	51
3.3	Mean scores for low and high L values	58
3.4	Mean scores for low and high H% values	58
3.5	Mean scores for low and high L% and H% values	59
<b>Chapter 4</b>		
4.1	Reading task used to investigate pitch range	70
4.2	F0 trace of NZE speaker reading aloud first two sentences in the monologue.	75
<b>Chapter 5</b>		
5.1	Prosodic Completion Tasks (PCTs) contrasting 8 pairs of meanings	90
5.2	Six F0 contours produced on the utterance 'no' by a female NZE speaker.	93
<b>Chapter 6</b>		
6.1	Schematic representation of source utterance	119
6.2	Schematic representation of manipulations carried out on the source utterance	120
6.3	Discrete-choice response, illustrating the percentage of participants who correctly identified the 10 test-stimulus pairs as "different"	124
6.4	Discrete-choice response, illustrating the percentage of participants who incorrectly identified the 11 control stimulus pairs as "different"	126
6.5	Average response score illustrating the degree of "liveliness" across 2 boundary tones (L% and H%) and 5 pitch widths from H* to %	131
6.6	Average response score illustrating the degree of "confidence" across 2 boundary tones (L% and H%) and 5 pitch widths from H* to %	132
6.7	Average response score illustrating the degree of "insistence" across 2 boundary tones (L% and H%) and 5 pitch widths from H* to %	133
<b>Chapter 7</b>		
7.1	Schematic representation of source utterance	141
7.2	Schematic representations of H* (3 levels) and boundary (7 levels) manipulations performed on the lexeme UNION	144
7.3	Listener-oriented turn cue response over 7 IP-final heights and 3 H* pitch accent heights	147
7.4	Speaker turn completion cue over 7 IP-final heights and 3 H* pitch accent heights	151

## LIST OF TABLES

---

### Chapter 1

- 1.1 Correspondences between Pierrehumbert 1980 and British-style nuclear tones 17  
(adapted from Ladd 1996: 82).

### Chapter 2

- 2.1 Meanings associated with pitch range 35
- 2.2 Meanings associated with the shape of the pitch movement 37
- 2.3 Meanings associated with pitch height 40

### Chapter 3

- 3.1 Source pitch points (in ERB) for 'male' and 'female' stimuli 48

### Chapter 4

- 4.1 Sentences listed according to Sentence Type which were included in the current investigation 72
- 4.2 Tonal Composition (7 categories) and Sentence Type (3 categories) produced at the nuclear location by NZE speakers 77
- 4.3 Average pitch range (in ERB) produced by 30 NZE speakers on 7 Tonal Compositions 80

### Chapter 5

- 5.1 Distribution of tonal compositions produced by 30 NZE speakers when conveying conversational continuation vs. cessation cues 96
- 5.2 Distribution of tonal compositions produced by 30 NZE speakers when conveying Authority vs. Submission 97
- 5.3 Distribution of tonal compositions produced by 30 NZE speakers when conveying Emphasis vs. an absence of Emphasis on "no" 102
- 5.4 Distribution of tonal compositions produced by 30 NZE speakers when conveying Emphasis vs. an absence of Emphasis on "yeah" 103
- 5.5 Distribution of tonal compositions produced by 30 NZE speakers when conveying Concern vs. an absence of Concern on "no" 106
- 5.6 Distribution of tonal compositions produced by 30 NZE speakers when conveying Concern vs. an absence of Concern on "yeah" 106

5.7	Distribution of tonal compositions produced by 30 NZE speakers when conveying an Impressed vs. Unimpressed attitude on "no"	110
5.8	Distribution of tonal compositions produced by 30 NZE speakers when conveying an Impressed vs. Unimpressed attitude on "yeah"	110
5.9	Distribution of pitch accents produced by 30 NZE speakers when conveying Meaning (Emphasis, Concern and Impressed) vs. a Lack of these same meanings (Lack of Emphasis, Lack of Concern and Unimpressed)	112

## **Chapter 6**

6.1	Pitch point values for the source utterance	120
-----	---	-----

## **Chapter 7**

7.1	Pitch point values for the source utterance	142
-----	---	-----

## Abbreviations

---

AmE	American English
ANOVA	Analysis of Variance
AusE	Australian English
BrE	British English
ERB	Equivalent Rectangular Bandwidth
H	High F0 target
HRT	High Rising Terminal
Hz	Hertz
IDS	Infant Directed Speech
IP	Intonation Phrase
IViE	Intonation Variation in English
L	Low F0 target
NZE	New Zealand English
PCT	Prosodic Completion Task
PL	Pitch Level
PPt	Pitch Point
RP	Received Pronunciation
ST	Semitones
SBrE	Southern British English
ToBI	Tone and Break Indices
ToDI	Transcription of Dutch Intonation



**For Chris**

### Introduction and Background

---

#### 1.0. General Introduction

Intonation, or the rising and falling of pitch, conveys meaning. Scholars have claimed that several aspects of intonational meaning are *Universal* (Ohala 1984; Ohala 1994; Gussenhoven and Chen 2000; Gussenhoven 2004). Although the term *Universal* implies that a similar tonal distinction is made in *all* languages, this is rarely the case. Ladd (1996:113) similarly notes:

“There are enough widely shared properties of intonation systems for many investigators to have hypothesised a universal common core of intonational signalling [...]. At the same time, however, there are many obvious differences between languages [...].”

For example, one universal form-function distinction is the phrase-final rise, which often signals an interrogative whereas a fall frequently identifies a statement. Although it has been regarded as a universal distinction, falling tonal patterns indicate an interrogative in Chickasaw (Gordon 1999, as cited in Gussenhoven 2004: 92) and rising patterns do not always specify a question in languages such as Australian English (AusE) and New Zealand English (NZE) (Britain and Newman 1992; Britain 1992; Warren and Britain 2000). Following Greenberg (1966), the term *Universal* is used here to refer to the tendency for *most* languages to follow a certain pattern.

As indicated by Ladd in the above quote, the meanings conveyed by intonation may also be specific to a language. For instance, an identical pitch range is interpreted as more excited in Dutch than in British English (BrE) (Chen, Rietveld and Gussenhoven 2001). Different languages have also been found to employ intonation for different functions; intonation has been shown to distinguish an utterance as either a command or an instruction in

Italian (Firenzuoli 1999). This function is language specific as few languages are known to indicate this grammatical division with intonation. Similar variations in meaning may be evident in different varieties of a language, such as English, as well.

It has been claimed that many different meanings are conveyed by intonation. The meanings conveyed by one variety of intonation has received notable attention, namely the standard British variety as spoken in the Southern region of England, which is sometimes referred to as Received Pronunciation (RP). These meanings are of various types, and may coincide in one utterance. For instance, intonation choices can reflect the attitude of the speaker (e.g. enthusiastic, interested), but might also depend on the status of the spoken unit of information (e.g. new information vs. old information; see Chapter 2). It is unclear, at this point, whether NZE speakers employ identical tonal techniques to convey meaning as those ascribed to other varieties of English, such as RP, or whether they employ variety-specific form-function associations.

In this thesis, comparisons are made between NZE and British English, most commonly, RP. RP, rather than American English or Australian English, was selected for comparison with NZE as it is sometimes assumed that "New Zealand English is derived from the English of South-East England. This assumption is most generally made with reference to New Zealand English pronunciation" (Bauer 2000: 41). Although Bauer (2000: 51-52) notes that this claim is not a clear and accurate account of the dialectal origins of NZE, I examine how accurate this assumption is with regards to intonation.

As this chapter will show, the study of the intonation of New Zealand English (NZE) has been limited largely to one form of intonation, the High Rising Terminal, and the meanings associated with the intonational form. It will become quite clear, however, that there is much more that can be said about both the forms and meanings of NZE intonation. This thesis will present new data and analyses that contribute to our understanding of the particular characteristics of this variety.

## **1.1 New Zealand English Intonation**

### **1.1.1. The High Rising Terminal in NZE**

Although it has not been established whether the meanings conveyed by NZE intonation are similar to those discussed for RP, several studies have indicated that this variety has several characteristic realisations of particular intonational forms. In particular, the High Rising Terminal contour produced on statements (HRT) is often cited as one distinguishing feature of NZE intonation (Allan 1990; Britain 1992; Britain and Newman 1992; Ainsworth 1994; Warren and Britain 2000; Warren 2005b), although other varieties of English do realise a similar feature (for Australian English, see Allan 1984, Guy and Vonwiller 1984; for general American English, see Ching 1982).

In the meanings outlined by O'Connor and Arnold (1961) for RP, two meanings conveyed by the rising tune are a sense of "incompleteness" (1961: 50) and "tentativeness" (1961:57). Other scholars have indicated similar perceptions of this feature when used in declaratives: the meaning of this feature has often been interpreted as indicative of the speaker's uncertainty (Lakoff 1975:17) or "helpless[ness]" (Eble 1975:246).

As noted above, the HRT is often noted by RP speakers as indicating uncertainty. If uncertainty was the sole function, Britain and Newman (1992) claim speakers would use this feature more frequently in speech acts "...where the truthfulness of a speaker's statement can be more readily challenged" (1992: 9), such as in opinions. Britain (1992) compared how HRTs are used within narrative texts and opinion texts in NZE. Narrative texts are described as "consist[ing] of at least two temporally ordered clauses using preterit verb forms, and in most cases additional elements which frame the story, provide orientational detail and evaluate the dramatic highlights of the plot" (Warren and Britain 2000: 159). Opinion texts may be described as "contain[ing] initiating opinion statements, possibly with supporting evidence and a resolution with the hearer to agree or disagree" (Warren and Britain 2000: 159). When comparing the use of high rising terminal contours in narratives and opinions, Britain (1992) found that HRTs were three times more prominent in narratives than in opinions. This evidence illustrates that

HRTs are used less frequently in opinions and thus supports the claim that – in NZE at least - this tonal feature indicates more than simply the speaker's uncertainty.

Warren and Britain (2000) examined where HRTs were used in 968 narrative tone groups and 1341 opinion tone groups by young speakers in order to examine how NZE speakers use HRTs within specific discourse genres. First, they confirmed previous research in that HRTs were much more prevalent in narrative text than opinion-giving text. However, they also found a great deal of distributional variation within these texts. In narratives, they found HRTs to occur most frequently in the abstract, orientation and evaluation, which are specifically segments of narratives "in which there is a heavy burden on the speaker to both instrumentally and affectively address the positive face of the hearer [...], suggesting that the speakers use [HRTs] as a tool with which to perform such functions" (ibid: 168-9). In opinion giving texts, they found HRTs to occur most frequently in evidence and pre-sequences, which are specifically segments of opinion giving texts in which "Both [speakers] seek to overcome potential misunderstandings, and both seek to foreground common knowledge and misunderstanding" (ibid: 169). Based on this evidence in NZE, Warren and Britain assert that "HRTs function as positive politeness markers, serving to overcome interspeaker hurdles and to build and maintain speaker-hearer solidarity" (ibid: 169).

The different use of this tonal feature across language varieties is shown by Fletcher, Grabe and Warren (2004), who report a comparative study of the form-function relationships of rising tunes in four varieties of English. The varieties which they compared included NZE, AusE and those varieties spoken in Glasgow and Belfast. Specifically they examined how these four varieties use rising tunes according to a *typology of contrasts* (based on Wells 1982, and developed for the discussion of intonational differences by Ladd 1996); the contrasts included (1) semantic differences, (2) systemic differences, (3) realisational differences and (4) phonotactic differences. Fletcher et al. (2004) assert that the high rising tune illustrates both a semantic and a realisational difference amongst these varieties. In other

words, the high rising tune conveys a different meaning in NZE and AusE than in Glasgow or Belfast English as well as realisational differences in the detailed shape of the contour between each variety.

Given that this research has established some semantic and realisational differences of intonation between varieties of English, any observed difference between NZE intonation and other varieties is ambiguous; we do not know, for instance, whether these observed differences are realisational, systemic, phonotactic or semantic in nature. If HRTs were semantically similar in NZE and BrE, the more frequent use of HRTs by NZE speakers than British speakers may indicate that NZE speakers are more uncertain (Lakoff 1975; O'Connor and Arnold 1961) or encouraging (O'Connor and Arnold 1961) than British speakers (for further discussion, see Warren 2005b). However, the evidence suggests that the meanings ascribed to BrE phrase-final rises cannot be ascribed to the NZE HRT. Other semantic and realisational differences may also exist between NZE and other varieties of English.

### **1.1.2. Social factors and intonation in NZE**

There have been a number of studies which examine how social groups within New Zealand produce intonation differently. One such study illustrated how the form of the phrase-final rise differs between men and women in NZ. Specifically, Warren and Daly (2000) examined how five male and six female speakers produced pitch range, pitch dynamism and alignment differently on rising phrase-final intonation. They examined the speech in a sentence reading task, which contained three echo questions, and a dialogue task where the participants were to instruct moves on a given map. They found that females produced wider pitch ranges than men in both tasks, and that the difference was more profound in the map task. They also found that females produced slightly more rapid rises than men in the sentence list but not in the map task data, where they produced equal rates of change. Their rise alignment results illustrated that females align their rises later than men in both tasks.

Warren and Daly hypothesised that the different results for their two tasks (reading task vs. map task) were possibly due to “a sex difference in approach to the task, rather than a difference in more direct sociophonetic markers of speaker sex” (Warren and Daly 2000: 107). For instance, the later rises produced by female speakers may be used to convey “polite insistence” whereas the earlier rises produced by male speakers may be used to indicate “contrast” (ibid: 107). In order to confirm such a supposition, they carried out a further analysis to explore the semantic differences of these different intonation forms (pitch range, pitch dynamism and rise alignment) as used by NZE men and women. They examined pairs of intonation-only questions in which one question was asking for information and the other question with the same wording was put into a context that should indicate surprised or contrastive questioning. Their results showed that females produced more rapid rises and started their rises later than their male counterparts, confirming their initial results (see above). They also did not find evidence to suggest that men were conveying contrast with their earlier rises in the first experiment. Therefore, they assert that realisational differences exist in the production of rises by NZE men and women, whereby “females used larger rises, they rose more rapidly and they started rising later” (ibid: 111).

Another study carried out by Daly and Warren (2001) compared men and women’s pitch range and pitch dynamism in sentence list reading and story reading. The tasks were selected in an attempt to “provide a context in which we expected speaker sex differences to be salient” (ibid: 89). Haan and van Heuven (1999) and Graddol and Swann (1983) found that pitch range variation occurred in data which included both statements and questions. With this in mind, two sets of materials were used. The first set was a sentence list, which included a number of both statements and questions. The second set of reading materials were taken from a children’s story, namely *Cinderella*, which also contained a number of statements and questions. They analysed the data using both Semitone and ERB (Equivalent Rectangular Bandwidth) scales (for further discussion, see Section 1.2.1) as their unit of measurement, and found that female speakers used a greater pitch range and more dynamic pitch using both units of measurement for each

task type. They also found an effect for sentence type (sentence vs. question), whereby the speakers produced significantly more dynamic pitch movements in questions than statements, and task type, as the difference between male and female pitch range and dynamism was greater in the dialogue task than in the sentence reading task. From these results, they assert that NZE men and women use different ranges and dynamisms and that the choice of task type is important in finding such sex salient features of intonation.

While the research findings have generally agreed that women use more dynamic pitch than men, one study found that NZE females do not employ more pitch movements than men. Specifically, Ainsworth (2003) examined whether five female dairy farmers from the Taranaki region of NZ employed more or less dynamic pitch than five male dairy farmers from the same region. She analysed pitch dynamism as the number of changes in pitch direction for each intonation group and accent unit, which is different from the method used by Warren and Daly (2001) to assess pitch dynamism. Accent units included the intonationally prominent syllable and any "syllables up to, but not including the next pitch accent" (Ainsworth 2004: 107). Intonation groups were separated at boundaries which were identified by a number of prosodic features, such as pause, final syllable lengthening, volume adjustments and anacrusis. Ainsworth analysed fact reporting and narrative style samples of speech, which were taken from her recordings of a conversational interview with the participants. The results were that men produced more changes in pitch direction than women. In addition, she found that women used more level pitch rather than upward or downward pitch movements.

The result was surprising in light of stereotypes often made about men and women's speech; Henton noted that female intonation is "swoopy" (Henton 1989: 299) and McConnell-Ginet (1983:555) noted women use "more rapid shifts in pitch" than men. Ainsworth notes that this difference between the sexes may be evidence of a language change in progress, where women are adopting the new feature (a "flat intonation") and men are expressing their



“pride in being from the Naki via their dynamic pitch movement” (Ainsworth 2003: 13). At present, it appears that Taranaki men and women do not conform to the stereotypes of men and women’s intonation use with regards to the number of pitch movements in a phrase (e.g. “swoopiness”).

Ainsworth (2003) did, however, find that Taranaki women used more non-falling terminal pitch patterns than men, which supports other studies. Specifically, a series of studies have also found that the High Rising Terminal is used preferentially by certain social groups in New Zealand. As noted in Warren and Britain (2000:155), the majority of studies have shown that HRTs are used more by women than by men, more by Māori (indigenous New Zealander) than by Pākehā (New Zealander of European descent), and more by younger speakers than by older speakers (Allan 1990; Britain 1992; Britain and Newman 1992). One study, however, has shown no significant speaker sex difference in the use of HRTs. Ainsworth (1994) conducted a comparative study of 4- and 9-year-old boys and girls, using narrative-style speech. The initial results showed no significant difference of HRT use between the sexes within either age group. Ainsworth admits that the collected scores were dramatically skewed by one boy, who used “an extraordinary high frequency of HRTs, namely 39.4%”, (1994: 11). Therefore, she notes that if the data from this boy were removed, the remaining data would show a significant sex difference, with female speakers using more rising phrase-final pitch movements than male speakers.

Therefore social variation between men and women has been shown to exist in NZE intonation. In addition to speaker sex, Ainsworth (2004) analysed a number of other social factors as variables in NZE intonation variation. Specifically, she examined the speech of young (20-35 years) and old (70+ years) men and women from three different regions in New Zealand: South Taranaki, New Plymouth and Wellington. She analysed the recordings of conversational interviews with the participants, and assessed a number of intonational features to determine how age, sex and regional variation affected such variables. She explicitly examined seven dependent variables, which are focused around pitch movement over different domains (i.e. across

the accent unit or the complete intonation phrase), within different areas within an intonation phrase (nuclear and non-nuclear) and movement direction (falling and level). Her thesis results are too numerous to list here (c.f. Ainsworth 2004: 117-138). On the whole, she found (2004: 141): "there is evidence of regional variation in New Zealand English intonation, but the picture is complicated by the interaction of regional differences with factors such as gender, age and rural or urban orientation".

Thus the research has shown that various social groups in New Zealand use intonation differently. Other studies have examined the unique forms of NZE intonation. Warren and Daly (2005) explicitly examined the form of rises in NZE. Specifically, they examined the properties of rising accent units formed by 6 male and 6 female NZE speakers completing a map task in same-sex pairs. The rise onset location was determined as its position (in syllables) within the accent unit, which was represented as the proportional value of the syllables. To use Warren and Daly's example, a rise starting on the third syllable in a four syllable accent unit would be 0.75 (i.e.  $\frac{3}{4}$ ). The results confirmed their earlier finding reported above, i.e. that male speakers produce earlier rise onsets than female speakers, and additionally showed that question rises begin earlier than statement rises for both sexes, although this difference between the two utterance types is greater for males.

Warren (2005a) later reanalysed this data, noting whether the rise onset was located at one of three positions (ibid: 214): the nuclear syllable, a metrically strong post nuclear syllable (MSPNS), or a metrically weak post-nuclear syllable. In addition to confirming the sex differences with rise alignment, he found that the NZE speakers were more likely to start their rise on the MSPNS than on the nuclear syllable and that speakers were more likely to align their rise on the MSPNS for statements and on the nuclear syllable for questions. He asserts that a categorical distinction exists between a statement rise and a question rise in NZE which can be differentiated as a "concave later rise" pattern and a "convex early rise pattern" (ibid: 225), respectively. These two patterns are said to be realisations of different anchoring point of the phrase accent ('-'); a single anchoring of H- is located

on a MSPNS and is used to convey a statement whereas a double anchoring of H- to both the nuclear syllable and a MSPNS is used to indicate a question.

In addition to examining the rise onset location of NZE rises, Warren and Daly (2005) also examined whether NZE speakers compress their pitch movement, i.e. whether they would “increase the rate of change of fundamental frequency as the segmental material became shorter, apparently in an effort to reach a similar end point for the pitch movement” (ibid. 230). Alternatively they may truncate their pitch movement, which would employ “the same pitch trajectory regardless of the amount of segmental material” (ibid. 230). Truncation would result in the speaker not reaching their F0 endpoint (maximum or minimum) in shorter accent units as there would not be enough segmental material for them to reach the maximum or minimum target. They assert that NZE uses a “mix of both compression and truncation” (ibid. 233); it appears that the NZE speakers employ truncation when increasing the rate of change any further is too difficult. They then further explain that the greater rate of change in their NZE data than that found in the four varieties of British English (Grabe, Post, Nolan and Farrar 2000) in addition to the alignment results discussed in the previous paragraph “are two aspects of the same phenomenon – a tendency for young NZE speakers to produce relatively late and relatively large pitch movements (Warren and Daly 2000)” (ibid: 234). They assert that their collective results illustrate that NZE is more dramatic than BrE, as NZE appears to use more complex tunes, and to realise later rising tunes thus producing a more dynamic pitch movement on rising and falling tunes.

Although there appear to be a number of particular intonational forms used in NZE and functional uses of such forms, there are still large gaps in our knowledge concerning the form and meaning of New Zealand English intonation. For instance, we currently know little about what makes NZE intonation distinctive. In addition, there is even less data indicating how speakers use intonation as a meaningful linguistic device in this language variety. It is currently unknown whether the intonational means of signalling meanings differ between NZE and other varieties of English. This appears to

be especially pertinent as the forms of several features have been found to indicate meanings other than those claimed for other varieties. In addition, it is currently unknown how NZE speakers employ specific tonal features in order to indicate meanings which are identified as universal. For instance, NZE speakers may raise their high pitch peaks or both their pitch peaks and low values in order to convey meanings ascribed to the Frequency Code. The aim of this thesis is to explore first which tonal features are used in NZE and second, how intonation is used as a meaningful linguistic device in this language variety.

## **1.2. Methodological issues**

The motivations for the current examination of NZE intonation were provided in the previous section. At this point, I must outline several methodological considerations, which are important to the description of intonation of a language variety. This section outlines the necessary background information with this focus in mind. First I examine the different units of measurements used by various researchers in intonation research. I then discuss the different notation systems adopted by researchers when describing intonation.

### **1.2.1. Unit of Measurement**

Intonation is often described in terms of fundamental frequency (F0), which is the acoustic correlate of pitch. One type of measurement which has been traditionally used to measure F0 is a linear scale such as Hertz (Hz). However perception experiments have found that a linear Hz scale does not accurately reflect our perceptions of pitch. One such study found that men's and women's voices are not perceived in a linear fashion; Hudson and Holbrook (1981) claimed that the Hz scale does not make the appropriate normalisation which corresponds to our perceptions.

Another two units of measurement measure the perceptual property of F0, namely pitch, which can be derived from the Hz values. Typical units of measurement on a logarithmic scale are musical scales. One such unit of measurement is Semitones (ST), which have been claimed to provide a more representative pitch excursion size. Graddol (1986) supported this claim,

stating "The ear is known to perceive pitch in a roughly logarithmic fashion..." (ibid: 228). A logarithmic scale makes comparisons between male and female speakers more perceptually valid than the linear Hz scale.

Other scholars have asserted that an Equivalent Rectangular Bandwidth scale, or ERB, is more accurate than ST as it is a closer measure to our perceptual interpretation of pitch. A characteristic of the ERB scale is that it is linear for part of its range and logarithmic for another part of its range, reflecting perceptual salience at different frequencies. ERB values are derived from the Hertz units for instance using the formula  $ERB = 16.7 * \log_{10}(1 + Hz/165.4)$ .

Hermes and van Gestel (1991) conducted a study on the perception of speech, comparing the three scales. Specifically, the subjects adjusted excursion sizes to match synthesised speech in different pitch registers. They found that "...the average excursion sizes in different registers are equal on the psychoacoustic scales", (1991: 97). For example, if a male speaker's pitch jumps from 120 Hz to 180 Hz, a linear scale would illustrate an equal excursion in a female voice as 240 Hz to 300 Hz. However, using an ERB-rate scale, the equivalent female pitch excursion would convert as 240-325 Hz, which is closer to our perceptions than the linear scale. The ERB scale was chosen as the unit of measurement in this thesis as it is closer to listeners' perceptions of pitch excursion in speech than either a simple linear scale, such as Hz, or a purely logarithmic scale such as ST.

### **1.2.2. Notation System**

There are two main approaches to the analysis and therefore notation of intonational patterns. The first approach, such as that outlined by the IPO (Institute of Perception Research) grammar of intonation (Collier & 't Hart 1981; 't Hart, Collier & Cohen, 1990) or the British traditional approach to intonation (O'Connor & Arnold 1961) is based on the pitch movement, such as rises and falls. The contrasting pitch contours defined for Southern British English (SBrE) by O'Connor and Arnold (1961) are discussed in detail in Chapter 2 (Section 2.1.2).

The second school of thought is dissimilar from the traditional British approach since it defines intonation as individual targets, such as high 'H' pitch targets and low 'L' pitch targets, rather than pitch movement. One recent example of this approach is the Autosegmental Metrical framework (Pierrehumbert 1980). Several transcription systems follow this framework (c.f. Jun 2005), including Tone and Break Indices, or ToBI (Beckman & Ayers 1994) and IViE or Intonation Variation in English (Grabe, Nolan & Farrar 1998). Each of these transcription systems uses a hierarchy of separate tiers in order to transcribe intonation, including but not limited to the orthographic tier (segmental information) and the phonological tier (H and L tones). More importantly, this approach separates the features of intonation into pitch accents and edge tones (Bruce 1977), which are both discussed below.

Regarding the phonological categorisation of intonation, there is no one system which is better than another; a follower of any one system may find disadvantages or "flaws" in another system. In addition, the transcription system can vary within one phonological approach. For instance, IViE uses 0% to indicate a lack of pitch movement from the penultimate target onto the boundary target whereas ToBI does not use this symbol. In this thesis, I adopt the Autosegmental Metrical (AM) framework in order to notate the intonation of NZE rather than the traditional British approach, which has traditionally focussed on the researcher's auditory assessment of pitch movement. In an attempt to determine the phonological status of pitch in NZE, it is necessary to incorporate both an acoustic and auditory analysis of the data in order to avoid the risk of over relying on either my auditory assessment, which may be influenced by my first language (AmE), or the acoustic pitch trace, which may be affected by various segmental factors and pitch processing errors.

In addition, having a comparable notation system is vital in order to compare the phonological inventory of languages and language varieties. The notation system adopted for this thesis is the Tone and Break Indices (ToBI) system (Beckman & Ayers 1994). Adopting a ToBI-style analysis will prove especially useful as there have been a number of such analyses carried



out on a range of languages and language varieties. In short, the phonological contrasts found for NZE may be directly compared with other varieties, such as British English or American English. Such comparisons may show the uniqueness of NZE intonational phonology in comparison to other varieties of English.

ToBI includes three features of intonation. First, the ToBI system identifies pitch accents, or pitch targets associated with a perceptually salient tone, which are indicated with a '\*' (e.g. H\*). For the purpose of describing American English (AmE), ToBI allows for five different pitch accents (Beckman and Ayers 1994); a high pitch accent (H\*), a low pitch accent (L\*), a low tone target followed by a sharp rise (L\*+H), a sharp rise in pitch onto a high target located on the accented syllable (L+H\*), and a step down in pitch from a high unaccented syllable onto a high accented syllable (H+!H\*). Second, boundary specifications indicate the end of an intonation phrase (IP) with the symbol '%' appearing after the tone target. ToBI allows for two possible boundary tones: L% indicates a low IP-Final boundary specification and H% indicates a high IP-Final boundary specification. The third type of accent noted in ToBI is the intermediate phrasal accent, which occurs at intermediate phrase (ip) boundaries and is marked with the symbol '-'. Beckman and Ayers (1994) claim that this second level of phrasing is denoted by a lesser degree of disjuncture within the prosodic string than the Intonation Phrase boundary. There are two possible intermediate phrase accents in ToBI; a Low phrase accent is indicated as L- and H- indicates a High phrase accent. The intermediate phrase symbol also serves to notate the range in pitch at the end of an intonation phrase. For the description of AmE, ToBI allows for four intermediate phrase and intonation phrase boundary combinations at the end of an intonation phrase (Beckman and Ayers 1994); an IP target falling to a low value within a speakers pitch range (L-L%), a low intermediate phrase accent followed by a high boundary target (L-H%), a high intermediate phrase accent, which upsteps the following high boundary target (H-H%), a high intermediate phrase accent, which upsteps the following low boundary target and forms a level or "plateau" pitch (H-L%). Also for AmE, a downstepped phrase accent (!H-) is permitted. The current analysis of

intonation does not employ the downstepped phrase accent feature and uses the four boundary combinations outlined above.

For example, Figure 1.1 illustrates an F0 trace which includes ToBI labels. The NZE speaker of Figure 1.1 produced three High pitch accents, specifically on *MANny* *WON* and *RUMmy*. The three pitch accents are followed by a low phrase-final target which reaches the bottom of the speaker's range, or in ToBI terms, a low intermediate phrasal accent (L-) and a low boundary target (L%).

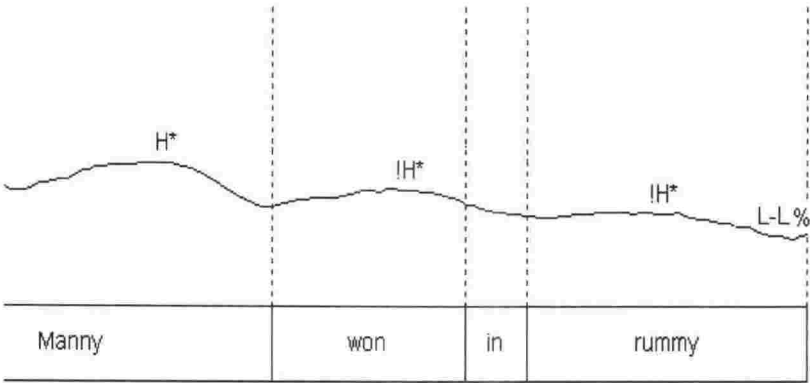


Figure 1.1. F0 trace exemplifying H\* and !H\* pitch accents, and L-L% boundary.

Figure 1.2 illustrates the same sentence produced by a different NZE speaker with a different intonation.

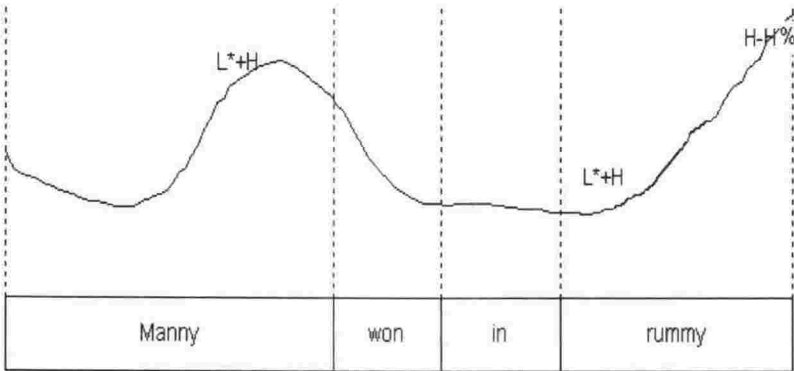


Figure 1.2. F0 trace exemplifying L\*+H pitch accents and H-H% boundary.

As shown, a low tone is followed by a sharp rise on the utterances *MANny* and *RUMmy*, and each pitch accent is labelled as L\*+H. The pitch trace also



depicts a high boundary target which exceeds the height of previous targets. In ToBI, this would be claimed to be a high intermediate phrase accent, which upsteps the following high boundary target (e.g. H-H%). Other examples of the notations used to describe F0 traces are provided throughout this thesis.

In using ToBI, I will be applying a transcription system which was not specifically developed for NZE. As noted by Beckman and Ayers (1994), "The ToBI labelling system was originally developed to cover the three most widely used varieties of spoken English -- namely, general American, standard Australian, and southern British English". Warren (2005b) discusses the issues of applying an existing framework based on one language variety to a separate language or language variety. Adopting an existing framework to describe a new variety has the advantage of applying an already established set of phonological contrasts. However, adoption may also result in losing possibly distinctive, phonological contrasts in the new variety by attempting to assemble and label the new language's data according to an established phonology of another language variety. He concludes that adapting a framework to the new variety would be the best solution (ibid: 348):

"In practice, the best approach to establishing a new analysis may well be a compromise solution, in which the researcher is informed by an existing analysis, while remaining ready to abandon it if closer investigation of the data suggest that this might be prudent (see similar conclusions reached by Knowles et al., 1996)".

Although the ToBI transcription system is used in the current thesis, I examine the adequacy of this system in the description of NZE intonation throughout this thesis. The benefit of using ToBI in this thesis is that it may permit a possible future comparison of the phonology of NZE with other varieties of English, for which there already exist established intonational analyses within the ToBI framework.

Although the ToBI framework is used in this thesis, a number of analyses, including several that are described in Chapter 2, have employed

the traditional British approach to intonation. Therefore, the following table (Ladd 1996: 82) provides comparisons between the two notation systems: ToBI, which is based on Pierrehumbert (1980), and the traditional British-style.

Pierrehumbert	British-style
H* L L%	fall
H* L H%	fall-rise
H* H L%	stylised high rise
H* H H%	high rise
L* L L%	low fall
L* L H%	low rise (narrow pitch range)
L* H L%	stylised low rise
L* H H%	low rise
L+H* L L%	rise-fall
L+H* L H%	rise-fall-rise
L+H* H L%	stylised high rise (with low head)
L+H* H H%	high rise (with low head)
L*+H L L%	rise-fall (emphatic)
L*+H L H%	rise-fall-rise (emphatic)
L*+H H L%	stylised low rise
L*+H H H%	low rise
H+L* L L%	low fall (with high head)
H+L* L H%	low fall-rise (with high head)
H+L* H L%	stylised high rise (low rise?) with high head
H+L* H H%	low rise (high range)
H*+L H L%	stylised fall-rise ('calling contour')
H*+L H H%	fall-rise (high range)

*Table 1.1. Correspondences between Pierrehumbert 1980 and British-style nuclear tones (Ladd 1996: 82).*

As shown above, a high target followed by a low target in Pierrehumbert's system (e.g. H\*LL%) would generally correspond to a fall in the British style approach whereas a low followed by a high (L\*HH%) would be noted as a rise. As noted above, this is not always the case; for instance, a high phrase accent upsteps a following low boundary target and forms a level pitch (H-L%). Other factors, such as pitch range, are also identified in both the AM and traditional British approach; for example, a "low rise" with a narrow pitch range in the British approach may be labelled as L\*LH% in ToBI whereas an L\*HH% in ToBI may refer to a low rise using the traditional British labels.

Issues regarding intonational analysis are considered in Chapter 2.

### **1.3. Outline of thesis**

The current chapter provided the motivations for examining the forms and contrasts in NZE intonation. In addition, the necessary background for an examination of intonation was presented. Chapter 2 serves as a review of a number of meanings ascribed to intonation. Chapters 3 through 7 (inclusive) are data chapters, each detailing a different experiment carried out for this thesis. Specifically, Chapter 3 reports the evidence from a perception experiment, which examined which tonal cues are used by NZE listeners in distinguishing NZE intonation from British English. The experiment reported in Chapter 4 examines the forms of intonation produced by NZE speakers in read-aloud speech, specifically examining preferential use of tonal compositions on statements and questions, in addition to pitch range and pitch level. In Chapter 5, the experiment focuses on how NZE speakers convey meaning using their intonation. The interpretation of the IP-Final boundary tone is examined in Chapter 6. The perception experiment detailed in Chapter 7 assesses intonation as a conversational cue in NZE. In Chapter 8, the main findings are summarized and discussed.

## Intonation: forms and meanings

Figure 2.1a illustrates a wide pitch movement onto a low boundary target on the utterance *yeah*. This same utterance is produced with a similar pitch movement although with a narrower pitch range in Figure 2.1b. The two examples were selected from a number of utterances spoken aloud by a male NZE speaker (39 years) saying *yeah* in a number of different ways.

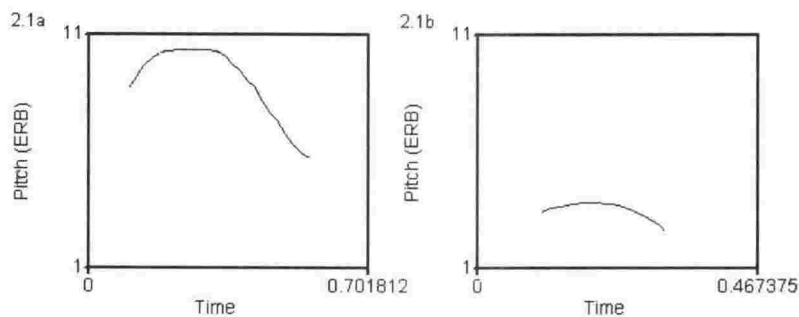


Figure 2.1. F0 traces of NZE speaker saying “yeah”

Some scholars believe that the *shape* of the pitch movement determines the meaning. If 2.1b were perceived as a low falling pitch movement, several scholars in this tradition claim that this pitch movement would indicate the speaker’s attitude as being unexcited. In contrast, if 2.1a were perceived as a “high fall”, it may indicate the speaker’s excitement (O’Connor and Arnold 1961). Similar meanings may also be conveyed with pitch *range*, which can be measured as the difference between the highest pitch value and the lowest value; wide pitch excursions (Figure 2.1a) are said to indicate “authority” and “enthusiasm” whereas narrow pitch ranges (Figure 2.1b) are said to indicate the speaker’s lack of interest (Gussenhoven 2004: 88). Pitch *height*, or the value of specific pitch points or targets, may also convey meaning, with higher pitch height, such as in Figure 2.1a, indicating “submission” and “appeasement” and lower pitch height (Figure 2.1b) indicating “authority” or “aggression” (Ohala 1994).

Assuming the multiple meanings provided for Figure 2.1 are correct, the examples illustrate how two intonation patterns may contrast several meanings. However, problems arise when analysing linguistic features with multiple meanings. This issue becomes evident when two or more contradictory meanings are said to be conveyed by one intonation. For instance, Figure 2.1a may be said to convey “authority” with a wide pitch movement in addition to “submission” with its higher pitch, therefore illustrating two contradictory meanings on one intonation. Distinguishing such meanings may become clear within the context that *yeah* was spoken, such as the previous utterances and the relationship between the speakers. Other factors may also play a role. For instance, phonetic features such as duration, alignment and amplitude may also assist in distinguishing such meanings. The appropriate meaning may also be different for separate languages. As briefly noted in Chapter 1, meanings acclaimed to be universal in all languages may be employed differently in separate languages or language varieties. For instance, listeners of one language variety may interpret the intonation illustrated in Figure 2.1a as authoritative whereas listeners of another language variety may interpret this same intonation as submissive.

The example discussed above also illustrates how a number of tonal features, such as pitch height, F0 shape or pitch range, are used to express meaningful distinctions. Identifying semantic contrasts in a language or language variety is considerably more problematic if it is unclear which component of intonation is conveying the meaning. For instance, if we assume that pitch range conveys the same degree of “enthusiasm” independent of pitch register (average height of utterance), then two intonation patterns equal in pitch range but in different pitch registers would indicate identical degrees of enthusiasm. On the other hand, if enthusiasm is conveyed not by pitch width but by the height of the pitch target, the intonation produced in the higher pitch register would be said to convey notably more enthusiasm. It may be true that both tonal descriptors adequately describe the semantic contrasts, and determining which meaning is being conveyed is dependent on the context. It may also be the case that the former tonal descriptor (i.e. pitch range) may adequately describe the meanings in one

language variety whereas the later descriptor (pitch height) may more appropriately describe another variety.

In order to identify semantic contrasts in intonation, the possible contrasting units of intonation which identify such contrasts need to be considered. In segmental phonology, this task is considerably easier; /l/ and /r/ can be confirmed as phonemes in English by establishing that these segments produce a meaningful difference, such as in the minimal pair *led* and *red*. In intonational phonology, however, separating the tonal components and identifying the contrasting units are challenging tasks due to the inherent relativity of the tonal constituents. For instance, raising an F0 peak would create a wider pitch range, and determining whether the higher peak or the wider pitch range is indicating the meaning associated with this intonation is, at very best, problematic. Nevertheless the tonal descriptors used to define meaningful contrasts in intonation must be considered.

In order to explore the phonology of NZE intonation, we must first examine several possible semantic contrasts which may exist in the intonation of this language variety. Three different types of meaning ascribed to intonation are outlined in the following two sections: attitudinal (section 2.1), and discursual and grammatical (Section 2.2). The meanings conveyed by three tonal descriptors, namely pitch height, pitch range and shape of the pitch movement are summarised in tables and subsequently discussed in Section 2.3. The review of the literature provided in Chapters 1 and 2 identifies gaps in the available data concerning NZE intonational phonology. The research questions arising from the detailed examinations of intonational forms and contrasts are provided in Section 2.4.

## **2.1. Attitudinal Meaning**

Traditionally, scholars have attributed a number of attitudinal meanings to intonation. For some of these meanings it has been claimed that they are conveyed similarly in all languages, while for others it has been claimed that they are conveyed differently in separate languages. The meanings claimed as *universal* and *language-specific* are outlined in the sections 2.1.1 and 2.1.2, respectively.

### 2.1.1. Universal meanings

Pitch height has been claimed to indicate a number of similar attitudinal meanings in all languages. Specifically, the *Frequency Code* (Ohala, 1984: 340-341) suggests that a high pitch signals meanings associated with smallness whereas a low pitch signals meanings associated with largeness. Ohala (1984) proposes that the meanings associated with pitch are signalled similarly in all languages because of our evolutionary development; at about the same time as the sexual organs in humans begin to develop, the vocal anatomy of boys and girls begin to develop differently, creating distinctly different vocal anatomy size and location - the male larynx is lower in the throat as well as 50% larger from front to back than the female larynx. This physical disparity, he suggests, is brought about because of sexual dimorphism. The reason for the male's growth, similar to growth in other species, is "...to enhance the acoustic component of aggressive displays" (ibid: 342). Therefore, Ohala claims that low pitch conveys "largeness" and secondary meanings such as "threat, intention to prevail in a contest, dominance, self-sufficiency" (1994: 341). Conversely, high pitch conveys "smallness" and subsequently "nonthreat, submission, appeasement, desirous of goodwill and cooperation of the receiver" (1994: 341). For example, the tonal targets and the overall pitch register are significantly higher in Figure 2.1a than in Figure 2.1b. If the utterances (i.e. *yeah*) were answers to the question *Can I borrow your car tonight?*, the listener may interpret the speaker of 2.1b as threatening (i.e. *Yes, but you had better bring it back without any dents!*) whereas 2.1a may be interpreted as non-threatening (i.e. *Yes, of course you can borrow my car!*).

In addition to pitch height, pitch range has also been claimed to convey universal meanings. Gussenhoven (2004:85) claims that increases in pitch production effort via a wider pitch range convey meanings which can be attributed to this greater effort. Thus he calls this biological code of intonational meaning the "Effort Code". There are a number of such meanings conveyed with the greater expenditure of effort. Meanings of this type that are conveyed with a wide pitch span include "[the] speaker's concern that his [sic] message should come across" (ibid: 85), providing "focus" on



specific units within the utterance, as well as affective meanings such as “‘helpfulness’, ‘authoritative’, ‘insistent’ and ‘enthusiastic’” (ibid: 88). Conversely, a narrow pitch range may be used to indicate that “the speaker is withdrawing information” or to “express a gentle negation” (ibid: 88), in addition to “‘lacking commitment’ and ‘uninterested’”. Applying the Effort Code meanings to the examples in Figure 2.1, the speaker of 2.1a may appear to be enthusiastic about lending the car to a friend whereas the speaker of 2.1b may seem to be unenthusiastic.

A third universal meaning, referred to as the Production Code (Gussenhoven 2004), is detailed in Section 2.2.1.

### **2.1.2. Language specific meanings**

In addition to both pitch height and pitch range, the traditional British approach to intonation assumes that the meaningful units of intonation are based on the shape of the pitch contour across a tone unit. The structure of the tone unit is comprised of several components. O'Connor and Arnold (1961) isolated the semantic contrasts in this language variety that are conveyed by different *nuclear tones*. These tones are carried by the “...the stressed syllable of the last prominent word” (O'Connor & Arnold 1961: 12) and by any additional syllables which follow the nucleus, referred to as the *tail*. They also identify the prenuclear pattern as important, including the head, which “begins with the stressed syllable of the first prominent word... and ends with the syllable immediately preceding the nucleus” (ibid: 15) and the prehead, which “consists of any syllables before the stressed syllable of the first prominent word” (ibid: 15). O'Connor and Arnold maintain that the nucleus is the defining part of the tone unit.

They identify four types of nuclear tunes, each conveying unique types of meaning: falling, rising, compound and complex tunes. The meanings attached to these tunes differ for sentence types, such as statements and questions. As the purpose here is to provide a general overview of several language-specific meanings ascribed to BrE intonation, I will outline the meanings ascribed to statements rather than attempting to summarise the numerous meanings assigned to other sentence types (e.g. questions, commands and interjections). In addition, I will limit the discussion to the



meanings attributed to nuclear tunes rather than the meanings described to the nucleus and the many variations in the head, pre-head and tail.<sup>1</sup>

First, the falling tunes are said to convey a sense of being "complete" or "definite" (O'Connor and Arnold 1961: 43) in "Southern British English" (SBrE) (O'Connor and Arnold 1961: vii). For this variety of British English, O'Connor and Arnold describe three types of nuclear falls<sup>2</sup>, each of which conveys additional attitudinal meanings (O'Connor and Arnold 1961: 33). The first fall begins in the lower part of the speaker's range and drops to the bottom of the range (i.e. a "low fall"). The low fall is similar to that pictured in Figure 2.1b. This intonation is claimed to convey a "calm", "reserved" or "dull" attitude (1961: 33) in this variety. For instance, the speaker of 2.1b may appear to be calm or reserved about lending his car to a friend.

Another falling pattern is a wide pitch movement on the nuclear fall (i.e. "high fall"), such as that depicted in Figure 2.1a. O'Connor and Arnold claim this tune conveys "*personal concern or involvement* in the situation; [...] more lively and interested, sometimes surprised, always more airy and lighter in mood than with Tone Group 1 [low-fall] or 2 [a low fall with a high prehead]" (1961: 39). Similarly, Cruttenden states this tune is "more interested, more excited and more involved" (1986: 100). For example, the speaker producing the pattern in Figure 2.1a might be interpreted as being more involved in the conversation, livelier and more interested in lending his car to a friend than the speaker in Figure 2.1b.

The third phrase-final falling nuclear pitch movement is a *rise-fall*. This tune begins at a low pitch onset, then rises and completes with a fall. O'Connor and Arnold claim that the attitudes associated with this tone group in SBrE include "impressed, awed, complacent, self satisfied, smug, challenging, censorious, disclaiming responsibility" (O'Connor and Arnold 1961: 147).

---

<sup>1</sup> For a complete assessment of intonational meanings ascribed to tunes in SBrE, I refer the reader to O'Connor and Arnold (1961).

<sup>2</sup> The three types of nuclear falls identified here do not include variations in the head. Including pitch variations in the head, O'Connor and Arnold (1961: 30) identify five types of phrase-final falls: low fall with a low head, a low fall with a stepping or high head, a high fall with a low head, a high fall with a high or stepping head and a rise fall.

Second, the rising tunes in SBrE are said to convey a sense of “incompleteness” (1961:50), “tentative[ness]” (1961:57) and are “used for continuative purposes” (1961:50). The following Figure is adapted from O’Connor and Arnold (1961: 6) for illustrative purposes here:

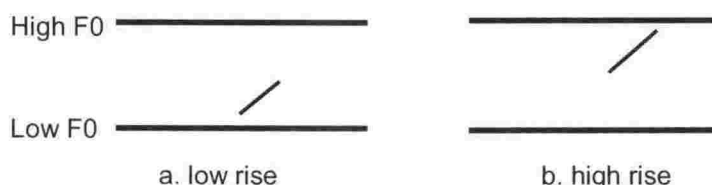


Figure 2.2. Schematic representation of two rises identified in British English (adapted from O’Connor and Arnold 1961: 6)

O’Connor and Arnold identify three types of nuclear rises in SBrE<sup>3</sup>, each conveying additional attitudinal meanings. First is a rising tune in the lower part of the speaker’s pitch range. This *low rise* has a low pitch onset which rises to its completion before the middle of the speaker’s range, as depicted in Figure 2.2a. The meanings ascribed to this intonation in SBrE include “inviting further conversation”, “...guarded, reserving judgement, reprovably critical, resentful, deprecatory” (1961: 169). If the speaker produced an intonation similar to that in Figure 2.2a on the utterance *yeah*, the speaker may appear to be guarded about lending his car to a friend.

The second type of rise is placed in the higher part of the speaker’s range; the *high rise* begins in the middle of the speaker’s range and rises until its completion point. This tonal description appears to denote a pattern similar to that depicted in Figure 2.2b. This high rise is said to indicate a number of meanings in SBrE: “questioning, trying to elicit a repetition, but lacking any suggestion of disapproval or puzzlement; [and] (in non-final [intonation phrases]) tentative”, (1961: 210). The speaker of 2.2b may appear to be questioning why his friend wants to borrow the car.

A tune which begins high and falls from this position and completes with a final rise is referred to as a *fall-rise*, such as that illustrated here:

<sup>3</sup> The three types of nuclear rises identified here do not include variations in the head. Including variations in the head, O’Connor and Arnold (1961: 30-31) outline five different types of phrase-final rises: a low rise with a low head, a low rise with a high or stepping head, a high rise with a low and stepping head, a fall-rise and a high fall followed by a low rise.

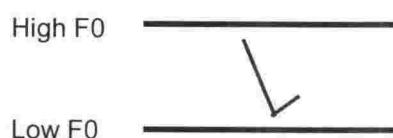


Figure 2.3. Schematic representation of a fall-rise  
(adapted from O'Connor and Arnold 1961: 6)

O'Connor and Arnold claim that this tune “draws particular attention to one element for the purpose of contrast and at the same time shows an intention to continue the utterance” (1961: 59-60) in SBrE statements. They also claim that this tune can be “grudgingly admitting, reluctantly or defensively dissenting, concerned, reproachful, hurt, reserved, tentatively suggesting” (1961: 226).

The meanings outlined above are by no means an exhaustive account of attitudinal meaning assigned to SBrE intonation. However, they do illustrate a number of attitudinal meanings ascribed to the shape of the pitch movement in SBrE. In the next section, two other types of meaning ascribed to intonation are outlined, namely discoursal and grammatical meanings.

## 2.2. Discoursal and Grammatical Meaning

The intonation of an utterance has also been shown to convey discoursal and grammatical information. Three types of such meanings are discussed here. First, the intonation produced on an utterance provides information about that individual segment of speech. For instance, intonation has been claimed to indicate whether the utterance is new to the discourse or already known. Such meanings are discussed in Section 2.2.1. Second, intonation has been claimed to indicate grammatical distinctions, such as the sentence type (e.g. question vs. statement). These distinctions are outlined in Section 2.2.2. Finally, the intonation that is used to indicate conversation turn-taking within discourse is discussed in Section 2.2.3.

### 2.2.1. Utterance Information

There are a number of tonal features that have been claimed to indicate new information and given information within the discourse. *New* information is, as

its name implies, new to the discourse "in the sense that the speaker presents it as not being recoverable from preceding discourse" (Halliday 1967: 204). Conversely, given information is already known within the conversation. For instance, Halliday (1967) and Chafe (1974) note that new information is accented and emphasised whereas given information is unaccented and unemphasised.

It is interesting to note that a similar new vs. old information distinction is evident in paragraph structure. Specifically, a speaker's pitch range has been claimed to be narrowed over the duration of the paragraph; Sluijter and Terken (1993) assert that pitch range is significantly wider in phrases at the beginning of a paragraph than in phrases occurring later in the paragraph. Similarly, Ladd (1988) found that the topline, or sequences of pitch accent peaks, produced in three consecutive clauses decline within and across clauses. It appears that newness of topics and utterances, as well as the newness of the speaker turn are similarly signalled with higher and wider pitch ranges while given information and non-introductions are identified with a narrower pitch range and lower pitch.

Similar assertions have been made by other scholars. Gussenhoven (2004: 89-92) claims that speakers use the natural reduction of subglottal air pressure over time to convey meaning. There is a gradual lowering of F0 over time as a consequence of air pressure reduction. As a result, listeners anticipate that pitch levels at the beginning of a phrase or breath group will be high and that those at the end of such a unit will be low. Gussenhoven claims that a communicative effect of this is that "high beginnings signal new topics, low beginnings continuations of topics, while a reverse relation holds for the utterance end, where high endings signal continuation, low endings finality and end of turn" (ibid: 89). He refers to this universal meaning of intonation as the Production Code.

In earlier work, Gussenhoven (1983) also claims that the shape of the pitch contour over the nucleus indicates information about the utterance in BrE. He identifies three primary tunes and four modifications of these tunes, each of which indicates the status of the spoken information within the

conversation. He asserts that the speaker produces a fall to add the *variable* (e.g. information provided by the lexeme) to the background as it is new information. He refers to this option as *addition*. The utterance used to exemplify the meaning is *Howard Johnson's* as a response to the question *Where shall we take Aunt Annie for lunch?* (Gussenhoven 1983: 206). If a fall is realised on this response, he claims that this might indicate a "speaker who is merely concerned to commit the variable to the background" (ibid: 206). That is to say, the speaker is introducing new information into the discourse, which in this case is the name of the restaurant (i.e. *Howard Johnson's*).

The second primary tune, a fall-rise tune, selects the variable from the background, which he refers to as *selection*. A speaker who uses a fall-rise would be said to be "tak[ing] note of the fact that this is (was) part of the background" or "remind[ing] you of the fact that this is background" (ibid: 204). Using the same example (e.g. *Howard Johnson's*), a fall-rise tune might indicate "You know that in our background there only is one restaurant I can fill your slot with, and therefore I am telling you it was already there: Where else could we take her?" (ibid: 206). Alternatively, the fall-rise might signal that the question had already been asked and that an answer had been given.

The third primary tune that Gussenhoven identifies is the rise, which he refers to as *relevance testing*; a speaker producing this tune is said to be testing whether the variable belongs in the background (ibid: 202). The rise on the variable *Howard Johnson's* might indicate that the speaker "chooses not to add it to the background, possibly because he wishes to allow [the other speaker] to give an opinion on his choice first" (ibid: 206).

The four modifications identified by Gussenhoven are as follows. The first modification is referred to as *delay*, which is a pitch movement realised after the onset of the stressed syllable. Delay modifies the three primary tones: (1) a fall with a delay would be realised as a rise-fall, (2) a delayed fall-rise would create a rise-fall-rise, and (3) a delayed rise would be realised as a later occurring onset of the rising pitch movement. Gussenhoven claims that *delay* signals "non-routineness" and indicates that the information being conveyed is "very significant" (ibid: 218). Contrary to delay, *stylisation* is said

to convey the discursial meaning of routineness (Gussenhoven 1983: 220). Stylisation is realised by flattening and extending the tones on the word, such as the vocative chant (Liberman 1975), or as a level tone, and leads to an interpretation that the variable is a common occurrence. For example, Gussenhoven (1983: 220) claims that a stylised production of "*John--- ny!---*" "does not signal that the speaker is calling the hearer, as had been widely assumed (cf Gibbon 1976: 276), but that the act of initiating a conversation with the hearer is presented as a matter of routine by the speaker". Another modification of the original three tones is *half-completion*. This tone would realise a short 'trip' from its onset point which completes before it crosses the middle of the speaker's pitch range. This tune is claimed to display the speaker's "non-involvement" in the conversation.

In addition to the shape of the pitch movement, Gussenhoven claims that a wide pitch range conveys emphasis on the given utterance and a narrow pitch range conveys a lack of emphasis, as noted in the Effort Code (Section 2.1.1). For example, if a speaker produced a wide pitch movement on the utterance *I LOVED the film!*, it would be said to emphasise this utterance and its meaning: how much the speaker had enjoyed the film. Ladd and Morton (1997) confirmed that pitch range indicates emphasis. Specifically, they employed three separate categorical perception experiments in order to determine whether there is a categorical distinction defined by pitch range. In the first categorical perception experiment, they asked the participants to identify whether the stimulus was either an "everyday occurrence" or an "unusual experience". They also asked participants whether pairs of stimuli were the "same" or "different". They manipulated the F0 peak in the source utterance *The ALARM went off* in 11 steps of 6 Hz. The results from this perception task indicated that a categorical distinction defined by pitch range exists, with a wider range indicating an "unusual experience" and a narrower range indicating an "everyday occurrence". However, the discrimination task indicated that participants heard pitch differences within these two categories and did not discriminate well enough to distinguish an exact categorical boundary. Although their follow-up experiments illustrate a number of interesting findings, these will not be

discussed here for the sake of brevity. I will note, however, that their collective results confirm that the distinction between normal, or an “everyday occurrence” and emphatic, or an “unusual experience” intonations are interpreted categorically. However, listeners can perceive pitch variations along a gradient scale between these categories.

### **2.2.2. Sentence Type**

In addition to indicating discourse information, intonation has also been claimed to indicate the type of sentence. Traditionally, scholars have associated statements with falling pitch movements whereas rising pitch movements have been claimed to indicate interrogativity in yes/no questions (see Gussenhoven 2002 for a review of intonation in questions). Other scholars believe that the statement vs. question distinction is made with pitch height; Hirst and Di Cristo (1998: 1) state that “...in a vast majority of languages some sort of raised pitch (final or non-final) can be used in contrast with lower pitch to indicate that an utterance is intended as a question rather than as a statement.” As noted in Chapter 1, this may not always be the case. For instance, Gordon (1999:53) found that in Chickasaw, a Muskogean language, “an IP ends in a boundary tone, a H% in the case of statements and a L% in Wh- and yes/no questions”. In English as well, there have been a number of varieties which employ high rising terminal contours on statements (HRTs), as noted in Chapter 1.

In a study comparing varieties of English spoken in the British Isles, it was found that falling vs. rising intonation or low vs. high pitch is not always employed to differentiate statements and yes/no questions. Grabe, Kochanski and Coleman (2005) have shown that a variety of intonation patterns may be used by speakers of Cambridge, Newcastle and Belfast English on different sentence types (referred to as “utterance types”). However, they also note that there are clear preferences in each of these three varieties of English for specific tonal patterns on a specified utterance type. Specifically, they analysed the speech of six speakers from the three varieties of English named above. The speech was read aloud from a sentence list and included eight statements (e.g. *We live in Ealing.*), three wh-questions (e.g. *Where is the manual?*), three polar questions (e.g. *May I lean*



*on the railings?*) and three echo questions (e.g. *He is on the lilo?*). They show that a high pitch accent followed by a low boundary target, labelled as H\*L% in IViE (for a review of the IViE transcription system, see Grabe, Nolan and Farrar 1998) is the preferred nuclear pattern for wh- questions and statements by Cambridge and Newcastle English speakers. However, a low pitch accent followed by a high boundary target (L\*H%, in IViE) is preferred in Belfast English. For echo questions, they found that Belfast and Newcastle English speakers preferred an L\*H% whereas a low pitch accent followed by an extended high boundary target (L\*H H% in IViE) was the preferred nuclear pattern in Cambridge English. For polar questions, there was quite a bit more variability; Cambridge English speakers produced a nuclear fall most frequently although there were a number of nuclear rises realized as well. Newcastle and Belfast speakers shared L\*H % as the most popular pattern for polar questions, although Newcastle speakers also produced a significant number of nuclear falls (23.5%).

### **2.2.3. Conversational Cues**

On a parallel to the patterns discussed above for sentence type, a universal meaning that has been attributed to phrase-final intonation is that rising intonation indicates openness and continuative meanings whereas falling intonation conveys closed and non-continuative meanings (Cruttenden 1997). In other words, a rise would be said to indicate continuation cues to the listener whereas the fall would indicate cessation cues. Similar assertions have been made by other scholars. As noted above (Section 2.2.1), the Production Code stipulates that “high endings signal continuation, low endings finality and end of turn” (Gussenhoven 2004: 89).

Other research has similarly maintained that intonation is used to distinguish turn-taking cues in discourse. Sacks, Schegloff and Jefferson (1974) introduced the term ‘transition relevance place’ (TRP) to describe the place where a speaker turn change may occur. For instance, a person may relinquish their turn-at-talk and provide the listener with a cue to take the speaking floor. Conversely, speakers may also indicate that they would like to reserve the speaking floor for themselves and continue speaking. Studies



have indicated that the TRP is identified by the syntax and prosody of the utterance.

One specific prosodic feature which has been attributed to the TRP is the pitch at Intonation Phrase (IP) boundaries. A number of studies have illustrated how the IP boundary conveys conversational cues in several languages and language varieties. For instance, conversational cues, such as speaker-turn continuation and completion cues, are signalled by a speaker's intonation at IP-Final locations in Dutch (Caspers 2000; Caspers 2001), Ulster English (Wells and Peppé 1996), Tyneside English (Local, Kelly and Wells 1986), London Jamaican English (Local, Wells and Sebba 1985) and Southern British English, or SBrE (Wichmann and Caspers 2001).

In the study carried out on SBrE, Wichmann and Caspers used two perception studies to investigate how syntactic completeness and intonation at phrase-final locations are used to interpret speaker-turn completion by SBrE speakers. The materials were derived from two complete Map Task dialogues, which were separated into interpausal units (IPUs) and labelled according to syntactic completeness ("complete" vs. "incomplete"), whether the IPUs occurred within the speaker's turn or not ("hold" vs. "change") and tonal contour. Listeners were then played fragments of the dialogue, each fragment finishing at the end of an IPU, and asked "to judge what the first speaker had intended – to continue or to cede the turn" (ibid: 181). The results illustrated that "the high level tone...appears to operate as a turn holding device" (ibid: 180) even when the utterance was considered syntactically complete. Different levels of low pitch at this location (L%) were also found to indicate speaking turn cues; a "truncated fall", or a fall which does not reach the F0-minimum, "is much more likely to cue a turn hold (71% of responses compared with 41% for the complete fall)" (ibid: 182).

Wichmann and Caspers found two tonal features which are interpreted as speaker turn holds in SBrE, including a high level tone and a low tone which does not reach the speaker's F0 minimum. Both phrase-final turn-holding cues are pitch movements with a reduced pitch range (e.g. a "high level tone" and a "truncated fall"). Conversely, contours with a wider pitch span, a "high rise" and a "complete fall", were interpreted as a speaking turn change by the majority of SBrE listeners. Post (2000:134) notes similar findings for French,

stating "This type of reduction of the range of the falling movement [a L% not reaching F0-minimum] is therefore comparable to that in final rises, where a reduced range has been claimed to be consistently associated with continuation, as opposed to rises with a wide range which are predominantly associated with questions or exclamations".

Wells and Peppé (1996) examined Ulster English, and similarly explored how intonation conveys that the speaker has finished his turn-at-talk in this language variety. They qualitatively examined one conversation, outlining the prosodic features which speakers regularly employed at turn-endings. They found clusters of phonetic features which were used to designate the end of speakers' turns. One such feature was a rise in pitch which reached or surpassed the pitch height of a previous peak, thus illustrating that a wide rising pitch movement indicates speaker turn completions in both Standard British English and Ulster English. Speakers of Tyneside English have also been found to use intonation in the same way; Wells and Peppé (1996: 123) note that the end of a turn in Tyneside English (Local, Kelly and Wells 1986) is indicated similarly with a "pitch step up at the end of a turn, which is usually greater than any other pitch step up in the turn and which always attains a higher pitch than any other step up in the turn". Thus wide pitch spans on phrase-final rising pitch patterns appears to denote turn-finality in three varieties of English: Tyneside, Ulster and Southern British English.

For Ulster English, however, it was found that in addition to wide pitch range, narrow pitch spans were also found to indicate turn completion. Wells and Peppé identified another two pitch-related features at turn-final locations, which they refer to as "spouts". At several turn-completion points, phrase-final rises were followed by subsequent syllables with either a sustained high pitch level or a stepping down in pitch "without pronounced movement" (1996: 110). These rises with post-peak pitch movements may be a high pitch accent (H\*) followed by a level or truncated fall onto the boundary tone, which is similar in form to those pitch movements described by Wichmann and Caspers as a "high level tone" or a "truncated fall". Such pitch movements have been identified as speaker turn-holds by SBrE listeners whereas in Ulster English, speakers appeared to use such a "sustained high pitch level" to convey speaker turn completions. The results may be illustrating a similar

phonetic form conveying two opposing meanings in two varieties of English: Ulster English and Southern British English.

## **2.3. Summary and Discussion**

In sum, there are a number of meanings conveyed by intonation; the types of meaning that have been discussed are the attitude of the speaker, discoursal and grammatical meaning. To date, we know little about how NZE intonation is used to convey such meanings. Also noted earlier, identifying the tonal component used to contrast the many types of meaning is problematic due to the inherent relativity of the tonal features. For instance, if a speaker raises their high pitch targets without adjusting their low values, the pitch span is subsequently widened, and it may not be obvious whether this corresponds to a phonological feature of pitch target height or to one of pitch range. Although such issues exist, we must consider the tonal descriptors used to define meaningful contrasts in order to explore semantic contrasts in a language variety. This section summarises those meanings which are said to be contrasted by three tonal descriptors, namely pitch range, the shape of the pitch movement and pitch height, and discusses these features in the context of NZE intonation.

### **2.3.1. Pitch range**

As discussed, pitch range has been claimed to indicate a number of meanings. Table 2.1 outlines the meanings associated with a wide pitch range and a narrow pitch range.

Pitch Range	Attitudinal	Utterance Information	Conversational cues
Wide range	Concern, helpfulness, authority, insistent, enthusiastic, speaker involvement	Emphasis	Speaker turn finality
Narrow range	Gentle negation, lacking commitment, uninterested, not involved		Speaker turn continuation

Table 2.1. Meanings associated with pitch range

Although a number of these meanings have been claimed to be “universal”, there may be language-specific uses of such meanings. For instance, Chen, Gussenhoven and Rietveld (2002, 2004) examined how universal meanings may be realised differently in two languages, namely Dutch and British English. Specifically, they examined the perceived level of “emphasis” and “surprise” on a number of manipulated utterances. The test utterances were dissimilar in the peak height of the High pitch accent (H\*) and in peak alignment (3 levels) which both combined with three IP-Final targets and with different pitch registers (5 levels). What is noteworthy here is that pitch register, or the raising of low and high targets in steps of 20 and 30 Hz, respectively, was interpreted differently by listeners in the two languages; the Dutch listeners believed that the higher the pitch register, the more emphatic the stimulus, whereas the opposite effect was true for BrE speakers, as lower pitch registers were interpreted as more emphatic than higher pitch registers. They assert that the Dutch listeners employ the Effort Code in interpreting meaning on the wider and higher “pitch registers” whereas British English speakers employ the Frequency Code on the same intonation (2004: 337). The evidence from this study illustrates that “universal” meanings may be employed differently in two separate languages.

Little is known about how NZE speakers employ pitch range. Bauer (1986) noted the lack of research carried out on New Zealand English intonation at that time. However, he did indicate his impressions of NZE intonation as being “rather flat” (ibid: 68), which may be suggestive of a narrow pitch range. If it is true that NZE speakers produce notably narrow pitch movements, it may be the case that NZE speakers and listeners have a heightened sensitivity to small deviations from a narrow range. It may also be true that speakers use alternative tonal constituents to convey meanings that are conveyed with pitch range in other languages and language varieties. For instance, NZE speakers may chose to raise the pitch heights of both their high and their low values within an intonation phrase rather than widening their pitch range by raising the high values and lowering the low values. In doing so, they may retain their “natural” narrow pitch range whilst using higher high targets to convey the meanings ascribed to this added effort.

### **2.3.2. Pitch Contour Shape**

The shape of the pitch contour has been claimed to convey attitudinal meaning and three types of discorsal meanings: utterance information, sentence type and conversational cues. These meanings are summarised in Table 2.2.

It is not completely clear how NZE speakers use the shape of the pitch contour to convey meaning. At this point, there is no textbook outlining the meanings conveyed by various shapes of pitch movements in NZE. However, research examining the shape of the pitch movement in NZE intonation has shown the proportional use of the pitch shapes in NZE and SBrE are dissimilar. Warren and Daly (2005) undertook a partial characterisation of NZE intonation in order to define a transcription system. They used corpus material consisting of scripted news broadcasts and unscripted lectures in order to gain “a general picture of the intonation system” (ibid: 218). The first set of data was transcribed using the traditional British school of notation by

<b>F0 shape</b>	<b>Attitudinal</b>	<b>Utterance Information</b>	<b>Conversational cues</b>	<b>Sentence type</b>
<b>Falls</b>	Complete, definite	New information	Finality, closed, non-continuative	Statement
	<b>Low fall</b>	Calm, reserved, dull		
	<b>High fall</b>	Personal concern, involvement, lively, interested, excited		
<b>Rises</b>	Incompleteness, tentative, continuative	Testing the relevance of the utterance	Continuation, open	yes/no Question
	<b>Low rise</b>	Guarded, critical		
	<b>High rise</b>	Questioning, eliciting repetition		
<b>Complex</b>				
	<b>Rise-fall</b>	Impressed, challenging, censorious	Non-routine	
	<b>Fall-rise</b>	Contrast, concern, hurt, reluctant	Reminder, focus	
	<b>Rise-fall-rise</b>		Non-routine	

Table 2.2. Meanings associated with the shape of the pitch movement

noting pitch movements (e.g. falls, rises, rise-fall, fall-rise) on the final accented syllable in an intonation unit (i.e. the nucleus). They found that SBrE and NZE differ in their proportional use of level and complex (e.g. rise-fall, fall-rise, rise-fall-rise and fall-rise-fall) tones. Warren and Daly also explored the shape of the rises, which is discussed in further detail in Chapter 1, Section 1.1.

Although the heavy use of complex pitch movements appears to be indicative of NZE intonation, it is currently unclear what meanings they are conveying. They may of course be carrying the same meanings as those ascribed to complex tunes in SBrE, and thus indicating that these speakers are “impressed”, etc. On the other hand, the more frequent use of complex tunes by NZE speakers may not be indicating that they are more “impressed” than SBrE speakers, but might simply be a stylistic difference in newsreader speech in the United Kingdom and New Zealand. Also, while it may be true that NZE speakers produce a greater number of complex pitch movements, it may also be the case that NZE speakers use other tonal cues to convey meanings which are conveyed with complex pitch movements in SBrE. For instance, NZE speakers may use a wider pitch range to indicate an impressed attitude or non-routineness rather than complex tonal movements.

In addition to the proportional use of pitch contours, it has been suggested that the function of the pitch contours may be different across varieties of English. A study carried out by Warren and Daly (2000; detailed in Section 1.1.2) found that the majority of contrastive questions (as opposed to questions used to gather information) were produced in NZE with a rising-falling tune. What is interesting is that “the use of an early rise to mark contrast... is not a typical contrastive question intonation” (ibid: 109), which is dissimilar from American English contrastive questions (cf. Pierrehumbert and Hirschberg 1990).

Other studies have shown that the location of the rise onset, which also affects the pitch contour shape, distinguishes between a question rise and a statement rise in NZE. Zwartz and Warren 2003 (reported in more detail in

Warren 2005a) carried out a perception experiment with this aim in mind. Specifically, Zwartz and Warren (2003) adjusted the phrase-final rise onset location produced by a male speaker and female speaker reading aloud one utterance. There were two sets of temporal rise alignment adjustments made on the source utterances. The first set of manipulations produced the gradient stimuli, with five rise onset locations after the accent unit which rose from this adjusted location directly onto the IP-Final target. The second set of manipulations produced another five stimuli which had a “rise-and-plateau” (Zwartz and Warren 2003: 55) after the rise onset. For the plateau stimuli, the pitch rose from the onset location and levelled out at the same height as the IP-Final target and remained level until the IP-Final boundary target was reached. The participants were provided with a reading dialogue illustrating an example of an identical utterance being used as a question and a statement. They were then played aloud the stimuli and asked to determine whether the utterance was a statement or a question. The results showed that early rises were more likely to be interpreted as questions whereas later rises were more likely to be interpreted as statements. Zwartz and Warren also found that the shape of the rise is important in distinguishing between statement and question rises in NZE; “The main effect of continuum type reflects the higher ‘statement’ response for the gradient stimuli [...] than for the plateau stimuli” (Zwartz and Warren 2003: 59).

It is possible that there are a number of pitch contour shapes which have a distinctive form-function relationship in NZE. At this point, however, it is unclear which pitch contour shapes, if any, are distinctive to the phonology of NZE intonation.

### **2.3.3. Pitch height**

Finally, pitch height was shown to convey attitudinal, conversational and grammatical meanings. Table 2.3 outlines the meanings associated with high and low pitch.



Pitch Height	Attitudinal	Conversational cues	Sentence Type
High Pitch	Smallness, submission, nonthreat, submission, appeasement	High endings = continuation; high beginnings = new topics	Question
Low pitch	Largeness, authority, threat, intention to prevail in a contest, dominance, self-sufficiency	Low beginnings = continuation of topic; low endings = speaker turn finality	Statement

*Table 2.3. Meanings associated with pitch height*

In order to convey meanings signalled by pitch height, speakers may adjust several tonal features. For instance, speakers may adjust the height of the pitch accent to convey the appropriate meaning, or the same speakers may raise or lower the overall pitch register of the utterance. At this point, it is unclear how NZE speakers employ pitch height variations to convey meanings in their variety of English, as there have been no studies, to my knowledge, which examine pitch height as a meaningful contributor in NZE intonation.

It may be the case that NZE speakers employ the heights of H% targets, for example, to differentiate between question rises and statement rises whereby the former may be indicated with lower targets than the later. Scholars have noted, however, that the HRT has been found to take various forms. Warren and Britain (2000:154) note that “[d]espite the fact that HRTs are relatively easily recognised, their physical realisations can vary dramatically, corresponding to rises in fundamental frequency (F0) that may be less than 50Hz or more than 100Hz, and stretching over a period of speech that may be as small as a single syllable or greater than an entire word (Britain and Newman 1992:4-8)”. It is currently unknown whether the

height of the IP-Final H% target, in addition to the pitch height of other tonal targets, differentiates meanings in NZE.

## **2.4. Research Questions**

Two strands of research have been reviewed in the first two chapters. The first strand reported in Chapter 1 concerns previous research on NZE intonation. As noted by Warren (2005b: 345), New Zealand English is a relatively under-studied variety of English. Nevertheless, there have been several studies carried out, most of which have examined how social groups within New Zealand produce different patterns of intonation. The second strand (Chapter 2) concerns how meaning is conveyed with intonation. As discussed above, scholars have traditionally assigned meanings to individual pitch target height, as well as to overall pitch register, pitch range, pitch movements and pitch trajectory. Taking these two strands together, it is clear that there is still much scope for understanding how speakers of NZE use intonation as a communicative device. In order to further explore this question, two questions are examined in this thesis. The first question concerns the forms of NZE intonation; although there may be a number of similar intonation patterns shared between language varieties, it is important to determine which tonal features distinguish NZE intonation from the intonation employed in other varieties of English. This question is specifically examined in Chapters 3 and 4. Second, there have been few studies carried out on the meaning of NZE intonation. It is unclear, for instance, whether an identical intonation pattern conveys identical meanings to those outlined for other varieties of English. Given that previous research has exhibited meaningful differences between language varieties, it would be unwise to assume that the semantic contrasts of one variety are realised in the same way in a separate variety.

Therefore, this thesis explicitly examines the intonational forms of NZE and the ways in which intonational contrasts are used to signal several semantic contrasts in this NZ variety. Examining these research questions naturally resulted in an examination of several phonological categories used

to label the intonation of NZE. With these research questions in mind, the following section outlines the findings from this thesis.

## **2.5. Summary of Findings**

The results of this investigation are divided into five sections. In the first two chapters, I explore the intonation use that defines New Zealand English. First, I examined which tonal features are used to identify NZE intonation in comparison to British English. Specifically, the results revealed:

- The lower the phrase-final low target (L%), the more likely the listener is to interpret the intonation as indicative of NZE speech;
- The higher the phrase-final high target (H%), the more likely the listener is to interpret the intonation as indicative of NZE speech; and
- Higher phrase-medial low values are perceived as more typical of NZE than lower values.

Second, I explored how NZE speakers use pitch width within an intonation phrase and the tonal composition they use on specified sentence types in a reading passage. The findings were that:

- NZE speakers prefer a H\*L-L% on statements and an L\*L-H% for questions on a reading task which does not require conversational cues;
- Tonal category, not sentence type, affects the pitch range produced by NZE speakers; and
- Male and female NZE speakers produce similar relative pitch ranges to one another.

In addition to defining the characteristic features of NZE intonation, I then examined how intonation is used to contrast meaning in this language variety. Specifically, I examined how speakers contrasted 5 pairs of meanings. The results indicated the following:

- NZE speakers prefer a high pitch accent (H\*) rather than a low pitch accent (L\*), or a low pitch target followed by a high target (L+H\*) rather than a high pitch accent (H\*) when conveying Emphasis, Concern and an Impressed attitude. Conversely, these same speakers prefer a L\*

rather than a H\*, or a H\* rather than a L+H\* pitch accent to convey an absence of Emphasis, and an Unconcerned and Unimpressed attitude;

- NZE speakers prefer a high IP-Final target (H%) to indicate Continuation cues and Concern for the speaker whereas a low IP-Final target (L%) was used to indicate Cessation cues and an absence of Concern; and
- NZE speakers prefer higher L% targets to indicate a Submissive refusal to an invitation whereas an Authoritative refusal to the same invitation is indicated with lower L% targets.

In addition to the production experiment, I also employed perception experiments in order to examine what tones are interpreted as meaningful by NZE listeners. I first investigated whether the height of the IP-Final target differentiates meaning in NZE. There findings were as follows:

- One categorical boundary appears to exist at the IP-Final boundary, evidenced by the finding that NZE listeners were more easily able to distinguish pitch levels at the boundary when the direction of pitch movement from H\* to the IP-Final target was different (i.e. rising from a H\* onto an H% or falling from a H\* onto an L%);
- The level pitch movement from the H\* to the boundary appears to be categorised with the H% in NZE rather than with L% (e.g. H\*H-L%), as traditionally labelled using the ToBI framework;
- NZE listeners interpret greater degrees of confidence and insistence on lower IP-Final targets than on higher IP-Final targets; and
- Higher targets within the range of H% values are interpreted as conveying greater degrees of liveliness, whereas all other lower targets, including within the range of H% and targets within L% were interpreted as less lively.

Finally, I examined whether IP-Final targets and nuclear H\* are markers of discourse completeness in NZE. Specifically, the findings were:

- The pitch accent and the IP-Final target are used to identify listener-oriented turn cues;
- The IP-Final target is used to identify speaker-oriented turn-cues;

After the results are presented in this thesis, the overall assessment of NZE intonational phonology is provided in Chapter 8. Specifically, this thesis

shows that three tonal characteristics may be necessary to accurately describe the intonational phonology of NZE: the pitch accent, the phrase accent and the boundary tone.

### **An experiment investigating tonal cues used to identify NZE intonation**

---

#### **3.0. Chapter abstract**

The aim of the experiment reported in this chapter was to investigate which tonal cues are indicative of NZE intonation in comparison to BrE. Fourteen potential tonal cues were chosen for manipulation. Results from a rating task demonstrated that listeners used several tonal characteristics when differentiating NZE from BrE. The principal cues were the relative height of the final intonation phrase boundary target and the height of the phrase-medial low values. Specifically, the lower of the two different values of a low phrase boundary (L%) was interpreted as sounding more like NZE; conversely if the final target was high (H%), the higher target was believed to sound more like a NZE speaker. For phrase-medial low values (L), it was found that higher L values indicated NZE. From these results, I conclude that native NZE listeners use several tonal cues, including the height of H%, L% and L values, when determining what makes NZE intonation distinctive from that of BrE. This indicates that NZE listeners believe NZE speech is characterised as having relatively high pitch values within an intonation phrase with marked departures from this to extreme low L% or high H% phrase-final boundaries. In order to determine whether the opinions of the participants about NZE people in comparison to British English people influenced the results, these same participants completed an attitude questionnaire. The results revealed that there were no significant correlations between the participants' opinions about the two speaker groups and their perception of which tonal cues are "indicative" of NZE intonation.

### 3.1. Introduction

A number of phonetic cues help us determine which variety an English speaker is using. These include the phonetic realisations of different segments, such as the use of a postvocalic /r/ in *car*, as a marker (for instance) of General American English in comparison to the English spoken in many parts of New Zealand. There is also semantic variation of certain lexical items; *bush*, for example, refers to a “native forest” in NZE and a “shrub” in British English (BrE). In addition, listeners may also use the intonational pattern of a word or phrase, such as a rising terminal contour on a statement which has been claimed to be used in NZE (Ainsworth 1994; Allan 1990; Britain 1992; Warren and Britain 2000; see also Chapter 1) as well as other varieties, rather than a falling tune, which has been claimed to be a “universal” marker of declaratives (see Chapter 2).

Past research conducted on NZE intonation has shown several prosodic features to be indicative of this variety. As discussed in Chapter 1, the most noted feature is the High Rising Terminal Contour (HRT). Timing also appears to be a factor; Ainsworth (1993) and Warren (1998, 1999) indicate that there is a rhythmic distinction between NZE and BrE. Ainsworth examined strong vs. weak vowel productions in function words, comparing news readings from several radio stations. She found that Māori NZE newsreaders produced more full vowels than Pākehā NZE newsreaders, who produced more full vowels than BBC newsreaders. Warren (1999) confirmed these findings in acoustic measurements of the same newsreader data.

The aim of the current experiment is to add to the studies discussed above and in Chapter 1 and to serve as a starting point for future examinations, looking at various features of intonation as possible factors which may be indicative of NZE intonation. Using native speakers as expert judges, this study examines explicitly fourteen different tonal cues which listeners might use when distinguishing NZE intonation from the non-regional variety of BrE, or RP (Section 3.2). The study also examines whether the perceptions of such indicative tonal targets are related to listeners’ opinions about NZE people and BrE people (Section 3.3).

3.2. Perception of Tonal Cues

3.2.1. Materials

The source utterance used in this study was taken from an adult male speaker of Southern British English. He was instructed to read aloud from a provided script (Appendix A1), “as naturally as possible” whilst being recorded onto a Maxell 74 minidisk using a Sony MDS-JB930 minidisk recorder and a Sony unidirectional microphone. A schematic representation of the recorded source utterance is provided in Figure 3.1.

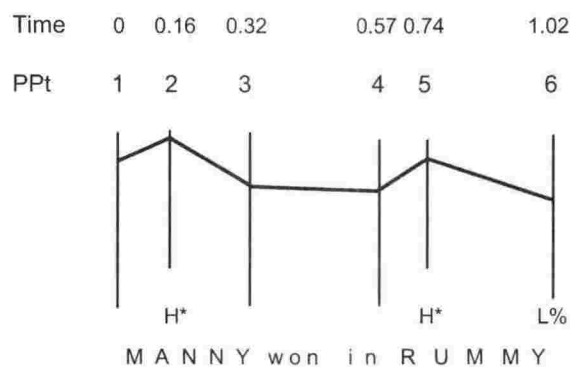


Figure 3.1. Schematic representation of source utterance (Time = sec.)

The speaker was recorded in a recording room in the Language Learning Centre at Victoria University of Wellington.

The segmental makeup of an utterance invariably affects the F0 trace. Voiceless obstruents, for instance, are not evident on the F0 trace due to the absence of voicing whereas plosives cause the F0 trace to rise after the stop is released. Such segmental effects often prove problematic for synthesising speech though they are easily and unconsciously normalised by listeners’ perceptual system. The sonorant utterance *Manny won in rummy* was selected as the base utterance for the experimental materials, in order to avoid such complications. This utterance was digitised at a sampling rate of 16 kHz in mono. Using the signal processing system Praat (Boersma & Weenink 1996), the source utterance was stylised and six key pitch points (Ppt) were located as shown in Figure 3.1.



To create male and female stimuli and thus allow a comparison of male and female pitch levels, the Total Rescaling method (Ladd & Morton 1997), or an up/down-scaling of all pitch points within the contour, was adopted. The six pitch points from the source utterance were adjusted to one of two levels using pitch point values as shown in Table 3.1. This produced two pitch levels (PLs) reflecting the average values measured from ten male (3.98 ERB) and ten female (5.83 ERB) NZE speakers reading aloud a story.

Pitch Points	1	2	3	4	5	6
Male	3.80	4.52	3.90	3.84	4.15	3.11
Female	5.69	6.41	5.78	5.73	6.04	4.99

Table 3.1. Source Pitch Points (in ERB) for 'male' and 'female' stimuli

The intonation of the source utterance was further manipulated in order to assess which intonational variations may be used by listeners to identify NZE. Since this is an exploratory experiment, a wide range of different tonal characteristics was assessed. The factors examined can be described with reference to the example pitch contour in Figure 3.1. The assessed potential indicators are as follows: [1] high pitch accent values (pitch points 2 and 5), [2] pitch range, [3] low pitch values between pitch accent peaks (pitch points 3 and 4), phrase-initial intonation phrase boundaries (pitch point 1), including [4] low values, and [5] high values, phrase-final phrase boundaries (pitch point 6), including [6] low values, and [7] high values, [8] pitch level, [9] number of peaks, [10] peak alignment on earlier peaks (i.e. high peak occurring on the first syllable of the utterance, at pitch point 2), [11] peak alignment on later peaks (i.e. high peak occurring on the penultimate syllable in the utterance, at pitch point 5). A final set of manipulations involve the relative steepness of the pitch slope, or *dynamism* at three locations within the phrase: [12] dynamism on the high phrase-final boundary (pitch point 6), [13] dynamism on the high phrase-initial boundary (pitch point 1) and [14] dynamism on a high phrase-internal target. The following paragraphs describe the manipulation of each of these tonal characteristics.

First, eight sets of manipulations involved setting pitch targets to high (represented as  $H$ ) or low (represented as  $L$ ) values, relative to the source Pitch Points (PPTs) indicated in Table 3.1. The range between these high and low values was set at 1.25 ERB. As noted earlier, studies examining and defining NZE intonation are relatively sparse. Due to this lack of information, the two contrasting high and low values were not based on actual values used by NZE speakers; they were selected according to what appeared natural to the author, which was later confirmed by two native NZE speakers.

Specifically, the adjustments were as follows (illustrated in Figure 3.2). First, the height of the  $H^*$  targets were compared (Fig 3.2a). The two  $H^*$  targets (PPT 2 and PPT 5, Table 3.1) were adjusted to levels 0.25 ERB below and 1.00 ERB above the source PPT values, creating low and high  $H^*$  targets, respectively. I also examined whether the values to which pitch is allowed to fall between two peaks are used as indicators of NZE; thus  $L$  values at PPT 3 and PPT 4 were lowered by 0.75 ERB or raised by 0.50 ERB to create low and high  $L$  values (Fig 3.2b). The height of the final boundary target was also assessed as a possible determiner of NZE (PPT 6 in Fig 3.1). A final boundary tone can be either  $L$  or  $H$ . The low final boundary tone ( $L\%$ ) was set to 0.75 ERB below or 0.50 ERB above the source PPT values, creating a low  $L\%$  and a high  $L\%$  (Fig 3.2c). In addition, the relative height of a high final boundary tone was assessed (Fig 3.2d) by setting the boundary tone (now designated  $H\%$ ) to 1.50 ERB or 2.75 ERB above the source value, creating a low and a high  $H\%$ , respectively. Similar to the phrase-final boundary, the height of the phrase-initial boundary target (PPT 1 in Fig 3.1) was examined as a possible factor. A low phrase-initial boundary ( $\%L$ ) was adjusted to one of two values relative to the source PPT: creating a high and low  $\%L$ : the PPT was set 0.25 ERB above and 1.0 ERB below the source PPT 1, respectively (Fig 3.2e). An initial high boundary tone ( $\%H$ ) was set at 0.75 and 2.00 ERB above the source PPT 1, creating both a relatively low and a high  $\%H$  (Fig 3.2f).

Adjusting the heights of tonal values inevitably affects the mean  $F_0$  over the utterance as a whole. Therefore, I also examined explicitly how pitch level affects judgments of the stimuli as being perceived as NZE or BrE (Fig 3.2g).

Pitch levels were adjusted either to a low level by lowering all PPTs by 0.625 ERB or to a high level by raising the PPTs values 0.625 ERB above the source PPTs. The resulting average PLs for 'males' and 'females' were 3.35 and 5.23 ERB for  $PL_L$ , and 4.57 and 6.48 ERB for  $PL_H$ , respectively.

Pitch range may also be an intonational feature which listeners use when identifying NZE. Effects of pitch range can be explored while keeping pitch level (relatively) constant if high and low values are adjusted around a midpoint. The high and low values were raised and lowered, respectively, by 0.625 ERB from the source utterance PPTs, creating a wide range (Figure 3.2h). It was determined by the experimenter that further raising and lowering of both low and high values, respectively, made pitch movements difficult to hear. Therefore, high and low values in the source utterance were not adjusted for a narrow range in order to maintain perceivable pitch movement in the contour.

In addition to adjusting PPT values, the timing of pitch movements and the number of peaks were also manipulated. Previous studies have suggested that alignment of the pitch movement may affect the meaning of an utterance. One characterization of the difference between a fall and a rise fall is as a delay in the peak of a nuclear pitch accent, conveying non-routineness (Gussenhoven 1983, Chapter 2). Similarly a delayed pitch movement on a rise, contrasting an early rise and a late rise, may indicate whether a rising phrase final contour is interpreted as a statement or a question (Zwartz and Warren 2003, Warren 2005). In the current study, peak alignment was also assessed as to whether the NZE variety of English is distinguishable from BrE according to the alignment of tonal material with the text. In an attempt to make variations in peak alignment easier to perceive, the original pitch movement positions (Figure 3.1) were adjusted for each manipulation (c.f. Fig 3.2i and 3.2j). The location of each PPT is described here in terms of how many seconds they occur after the onset of the source utterance voicing (0 seconds).

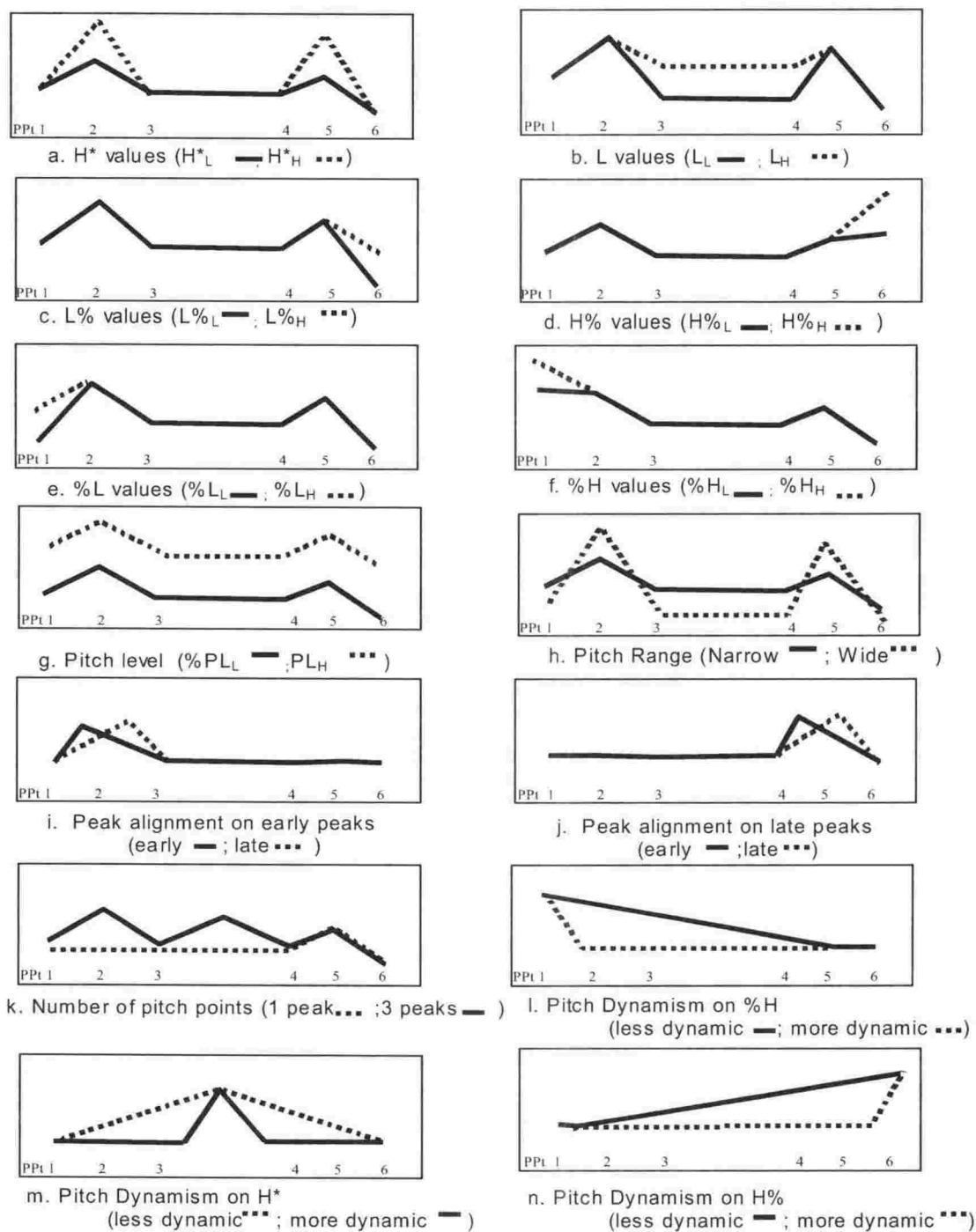


Figure 3.2. Schematic representations of the 14 manipulations performed on the Pitch Points (PPt) in the source utterance (c.f. Figure 3.1) and used in the experiment.

Peak alignment was assessed in two locations. First, the F0 peak occurring early in the phrase was examined (Fig 3.2i). Specifically on the first vowel in the utterance *Manny* (Ppt 2, Figure 3.1), the F0 peak was adjusted to either the onset of the vowel, located at 0.06 sec after the voicing onset of the utterance, or the end of the vowel, found at 0.16 sec after the voicing onset. The F0 peak occurring later in the IP, specifically on the first vowel of the utterance *rummy* (Ppt 5, Figure 3.1), was also assessed (Fig 3.2j); the peak was aligned to the onset of the vowel, located at 0.76 sec. after the onset of voicing of the utterance, or at the end of the vowel, found at 0.83 sec. In total, the manipulation of peak alignment created 4 separate stimuli.

The number of pitch movements in an intonational phrase may also be a tonal feature distinguishing NZE and BrE. For example, if more contrasting high and low targets are produced in an utterance, the utterance may appear to be more dynamic and musical than an utterance with fewer contrasting targets. In order to assess such an assumption, the incidences of peaks were examined as a possible indicator of NZE (Fig 3.2k). Accordingly, the IP was manipulated to include either one or three pitch peaks. First, one F0 peak was created by lowering the first high value (Ppt 2) to the following low value (Ppt 3), which reduced the number of H\* targets from two in the source to one. Relative to the source Ppts given in Table 3.1, an utterance with three high pitch peaks was established by creating an additional high peak located in between Ppt3 and Ppt 4, specifically situated at 0.45 sec. The natural effects of declination (the progressive lowering of a sequence of high values that results from reduction in subglottal air pressure over a breath unit) were considered when determining the height of this further high peak; therefore the height of this peak was set to a value half-way between the two original H\* values given in Table 3.1 (Ppt 2 and Ppt 5), creating a progressive lowering effect across the three peaks.

Dynamic pitch has been noted as a marker of social dialects, in particular women's speech (Austin 1965; Brend 1975; Ainsworth 2003; Daly and Warren 2001; Warren and Daly 2005) and homosexual male speech (Barrett 1997). In the current study, pitch dynamism is examined as a possible distinguishing

feature of NZE. Dynamism at three locations within the phrase was examined, including phrase-initial, phrase-internal and phrase-final locations (Figures 3.2l, 3.2m and 3.2n, respectively). In an attempt to make variations in dynamism greater and subsequently easier to perceive, the original pitch movement positions (Figure 3.1) were adjusted for each manipulation and PPT locations different from Table 3.1 were adopted. Two PPT ERB values were used for all high and low targets in assessing dynamism. Specifically, the height of any high target PPT was set to 4.52 and 6.41 ERB for male and female speakers, respectively, and the height of any low target PPT was set to 3.80 and 5.69 ERB for male and female speakers. The location of each PPT is explained here as seconds after the onset of the utterance voicing (0 seconds).

To assess dynamism at IP-initial position (Fig 3.2l), three PPTs were used; the first PPT (high target) was set at 0 sec., followed by a low target producing either a more dynamic pitch movement (the target was at 0.3 sec.) or a less dynamic pitch movement (the target was at 1.0 sec.), which was trailed by a final low target located at 1.02 sec. Three PPTs were also used to create more and less dynamic stimuli at phrase-final locations (Fig 3.2n); the first PPT (low target) was set at 0 sec., the second low target was set to create either a less dynamic pitch movement (starting at 0.02 sec.) or a more dynamic pitch movement (starting at 0.72 sec.) and the final high target was set at 1.02 sec. For more and less dynamic pitch movements at IP-internal position, five PPTs were used (Fig 3.2m). For a less dynamic H\*, two low targets, located at 0 sec. and 0.09 sec. (respectively), were followed by a high target set at 0.49 sec., and two low targets located at 0.89 sec. and 1.02 sec. after the onset of the utterance. As a contrast with the less dynamic H\*, a more dynamic H\* was created: two low targets, located at 0 sec. and 0.39 sec., were followed by the high target (again located at 0.49 sec), and two low targets set at 0.59 sec. and 1.02 sec. Collectively, dynamism was assessed as relatively more dynamic and less dynamic in three locations within an IP, including phrase-final, phrase-initial and phrase internal.

Each of the 14 manipulation types shown in Figure 3.2 has two values, giving 28 stimuli for each of the two voices (male and female), creating 56 stimuli in all. Two native NZE speakers listened to each manipulation and confirmed that they created “natural” sounding utterances. A Low-Pass Filter (LPF) was used on the manipulated stimuli using Cool Edit (cut-off frequency 110 Hz, filter order 3.5). The LPF was used to mask any British English production of segmentals in the source utterance, which may hinder the interpretation of any stimulus as “NZE-like”. The filter used in the current experiment was previously used in an experiment by Rietveld and Vermillion (2001 & 2003), who confirmed that listeners are able to distinguish the height of the H values at the different pitch levels associated with male and female voices.<sup>4</sup> After filtering, the stimuli were amplified by 500%, to ensure the pitch information was audible.

The “female” and “male” stimuli were randomised separately; separating the sexes and indicating the sex of the speaker at the top of each answer sheet ensures that the listeners are rating the appropriate sex (in case they perceive a different sex than that one intended). One half of the completed recordings had “female” speakers played first while the other played “male” speakers first. Prior to every fifth stimuli, a 0.5 second beep was inserted in order for the participants to determine whether they were keeping pace with the recording. Each stimulus was separated by an interval of 3.5 seconds. The first five stimuli were played a second time within the experiment and the answers given for the first set were discarded from the included data; this was done to discount erroneous answers often made at the beginning of experiments. In total, the experiment extended over a period of approximately 9 minutes, plus a 5-minute introduction.

### **3.2.2. Participants**

56 participants took part in this experiment on a voluntary basis. A questionnaire was used to determine social variation amongst the participants

---

<sup>4</sup> Although the filter has a “cut-off” of 110Hz, the slope of a filter used ensures that pitch-relevant information above the cut-off frequency, including the range for female voices, is still present.



(Appendix A4). The data provided by 30 native NZE speakers were included in this experiment. Participants' datasheets were discarded if they were not a native NZE speaker (5). In addition, data were not included if the participant did not complete both this experiment and a related attitude experiment (Section 3.3). Six participants in the selected group were male and the other twenty-four were female. All were students at Victoria University of Wellington. 27 of the selected group were aged between 18-22 years and three were 23-29 years old. None of the listeners had any known hearing impairments.

### **3.2.3. Procedure**

Participants were told that they would hear a number of speakers saying one sentence, which was played over a loudspeaker. The participants were told to mark on a given scale "how much the speaker sounds like a New Zealand English speaker" with the left side of the scale indicating that "the speaker does not sound like a NZE speaker" and the right side of the scale indicating that "the speaker sounds very much like a NZE speaker" (see Appendix A3).

To assess the listeners' responses, I used a Magnitude Estimation Scale. This was provided by a 100mm line on the answer sheets, between the two labels for the scale ends. This type of scale does not assume equal intervals of perception (Zraick and Liss 2000). The listeners were asked to mark on the line their impressions of each stimulus. The introduction to the experiment, which was read aloud by the author, identified the variety of NZE that I was investigating. Specifically, the listeners were asked to identify a variety of NZE that is spoken by Pākehā speakers with a 3-year University degree and from the Wellington area. I asked them to compare this variety to one other variety of English (the "standard variety of British English" or RP) in order to give all of the listeners a similar comparative model. That is to say, I wanted the listeners to indicate what makes NZE distinct from one specific variety rather than having different listeners judging NZE against different comparison varieties. I avoided indicating a specific city or town for the BrE comparator in the introduction, as listeners may know a person from this place who may speak a different variety from the one which I am attempting to locate. The



listeners were provided with samples of speech from a SBrE speaker (educated speaker from South England) and a NZE speaker (educated speaker from Wellington, NZ) in order to illustrate the varieties of English I was referring to (Appendix A1b). Both samples of speech were taken from female speakers, however speech from neither one was used as source data for this experiment. These example speech samples were not altered or filtered in any way, as I wished to indicate clear examples of both varieties of English.

The introduction was also used to provide motivation for the listeners, by telling them that I wanted to determine what a NZE accent is so that I could synthesise a realistic sounding NZE accent on the computer. The participants were told that each utterance had been read aloud by one of two RP British English speakers: one male and one female. The purpose of using RP speakers, rather than NZE speakers was to reduce the danger of all stimuli being regarded as having a “NZE accent”, which might have resulted had I said they had been derived from the speech of a NZE speaker. By using a British speaker as our source utterance rather than a speaker of some other variety, I am also priming a comparison to be made between NZE and BrE, which is the comparative aim of this experiment. The participants were also told that each utterance was then put onto a computer and changed in some way. Their task was to determine which of these manipulated stimuli sounded like a NZE speaker and which did not.

#### **3.2.4. Scoring**

Scores were measured to the nearest whole millimetre from the 100mm response scale. Measurements were taken as the centre point on any marks which were a millimetre or so wide. Scores could range from 0 (for “not at all like a NZE speaker”) to 100 (for “very much like a NZE speaker”). The answers (marks along the given lines) were measured again before the statistical analyses were carried out in order to ensure their reliability. These scores were then entered as the dependent variable into separate ANOVAs, or Analyses of Variance (repeated measures design, Huyhn-Feldt corrected) for each pair of stimuli (as determined by the manipulations listed above, e.g.

the  $H_H$  and  $H_L$  versions of the  $H^*$  height manipulation formed one such pair). In each case the independent variables were the two within-item factors: Sex of Voice (2 levels: male and female, based on the initial manipulation of pitch level of the source utterance, as indicated in Table 3.1) and Manipulation (2 levels: high and low, as described in the Method section for each manipulation).

Note also that there are three points of comparison that involve similar manipulations. Two contrasting pairs are manipulations at the same location within the phrase; they are the phrase-initial boundary (high and low values of %L and %H) and the phrase-final boundary (high and low values of L% and H%). Another contrasting pair is a mirrored change in pitch movement at two opposite ends of the phrase, specifically pitch dynamism at %H and H% (less and more dynamic pitch movements on %H and H%). For each of these points of comparison an additional three-way ANOVA (Huyhn-Feldt corrected) was carried out, with Sex of Voice (2 levels), Manipulation (2 levels) and Tune (2 levels). Specifically, the two compared levels of Tune were L vs. H for the initial and final boundary and %H and H% for pitch dynamism. It should be noted that statistics were not performed repeatedly on all of the responses. Specifically, statistics were performed first on the contrasted pairs (once per pair), and secondly on three *a priori*, or planned comparisons.

### 3.2.5. Results

When analysing the fourteen separate phonetic variables, two factors turned out to have a significant effect on the 'NZE-like' score: H% values and L values. Of these two sets of results, those for the height of the L values ( $F_{1,29} = 5.489$ ,  $p = 0.026$ ) showed the larger effect size ( $\eta^2 = 0.053$ ). Comparing the values for the low and high L values, the response averages were 36.75 and 44.82 respectively (on a scale from 0 to 100), illustrating that the listeners, on average, believed that higher L values are more indicative of NZE than lower L values (see Figure 3.3). The second largest effect size ( $\eta^2 = 0.013$ ) was the height of the H% values ( $F_{1,29} = 5.075$ ,  $p = 0.032$ ), illustrating a response average of 41.73 for low H% and 52.27 for high H% (Figure 3.4). In other

words, the higher the H% values, the more likely the listener would be to interpret the speaker as a NZE speaker.

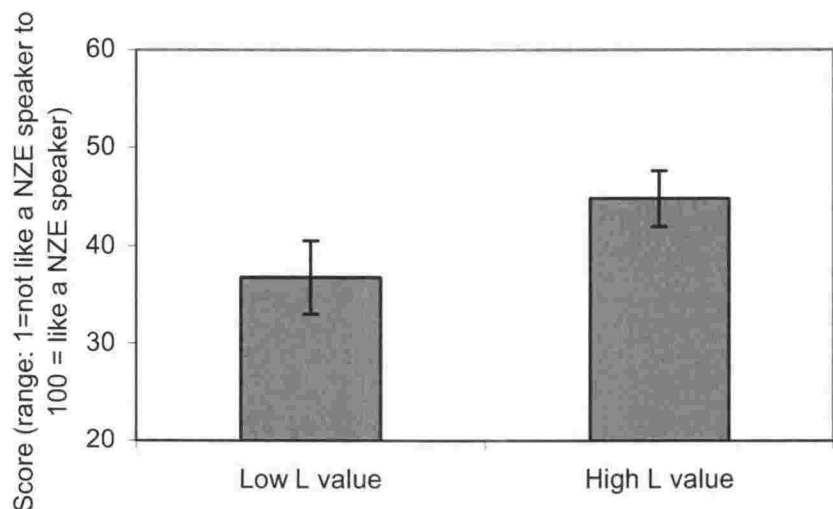


Figure 3.3. Mean scores for low and high L values (higher score = more NZE-like)

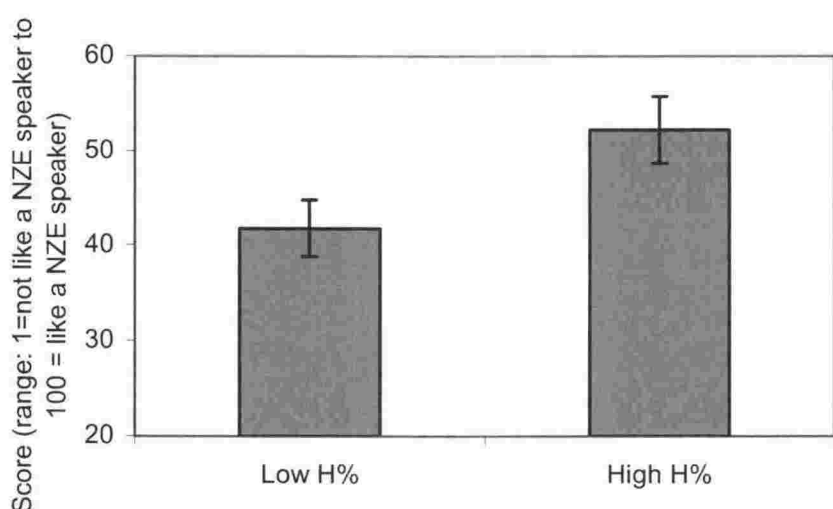


Figure 3.4. Mean scores for low and high H% values (higher score = more NZE-like)

In the three separate “contrasting pairs” comparisons, there was one significant interaction, which involved Tune and Manipulation for phrase final targets, with an effect size of  $\eta^2 = 0.043$  ( $F_{1,29} = 6.190$ ,  $p = 0.019$ , see Figure 3.5). As shown in Figure 3.5, listeners were more likely to believe the stimulus sounded like one from a NZE speaker if the phrase final low target

(L%) was lower and conversely, if the phrase final high target (H%) was higher.

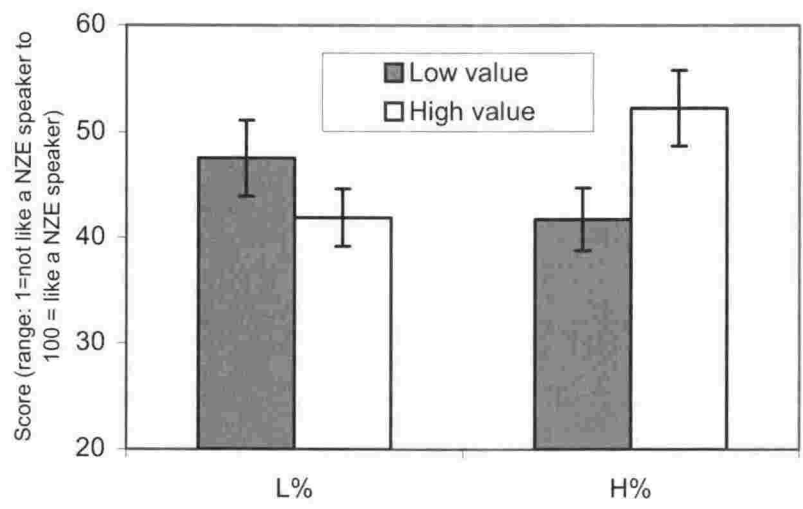


Figure 3.5. Mean scores for low and high L% and H% values (higher score = more NZE-like)

3.3. Attitude Assessment

Participants' interpretations of which tonal cues are indicative of NZE may have been influenced by their attitudes about NZE and BrE speakers. That is, the results presented above may not accurately illustrate which tonal cues are indicative of NZE but may instead reflect the intonational meanings which were expected to be conveyed by NZE speakers. For example, listeners may interpret higher phrase-medial values as indicative of NZE if they expect NZE speakers to be more lively than RP speakers, as variation in the height of such tonal targets may correlate positively, for example, with the degree of perceived "liveliness" in this variety of English. It was for this reason that the participants who took part in the above listening task were also asked to take part in an attitude assessment task, so that we could determine whether there is a relationship between the three tonal features found to be indicative of NZE and participants' opinions about the two speaker groups (New Zealand people and British people).

### 3.3.1. Materials

A questionnaire was used to assess the same NZ listeners' attitudes about NZE speakers with regards to sixteen attributes. As an exploratory experiment, a range of possible attitudinal factors was examined, each of which has been claimed to be conveyed by pitch. Chapter 2 reviewed a number of such meanings. The factors assessed here are as follows: [1] Lively (O'Connor and Arnold 1961: 40, 71), [2] Dull (O'Connor and Arnold 1961: 79), [3] Encouraging (O'Connor and Arnold 1961: 70), [4] Gloomy, or "morose" (Addington 1968: 495), [5] Resentful (O'Connor and Arnold 1961: 68, 70), [6] Appreciative (O'Connor and Arnold 1961: 70), [7] Apologetic (O'Connor and Arnold 1961: 249), [8] Strong willed (Addington 1968:495), [9] Confident (Ohala 1983:341; O'Connor and Arnold 1961: 193), [10] Self-doubting, or "questioning" (O'Connor and Arnold 1961: 210), [11] Dominant (Ohala 1994:341), [12] Submissive (Ohala 1994:341), [13] Threatening (Ohala 1994:341), [14] Welcoming, or "desirous of goodwill" (Ohala 1994: 341), [15] Insistent (O'Connor and Arnold 1961: 92), and [16] Compliant, or "Servile" (Addington 1968: 495).

To assess the listeners' attitudes about NZE speakers in relation to BrE speakers, I used a Magnitude Estimation Scale. This was provided by a 100mm line on the answer sheets (see Appendix A5). The listeners were asked to use this scale in their assessment of both males and females, as follows. The participants were told to mark their beliefs on a given scale, with each scale measuring a different attitude (listed above). Specifically, the left side of the scale indicated that "Kiwi men are not more [attitude] than British men", the right side of the scale indicated that "Kiwi men are much more [attitude] than British men" and the mid-point denoted "Kiwi men are somewhat more [attitude] than British men". For comparability, an attempt was made to keep the attitude assessment scales, which were used here, similar to the scales used determine the distinguishing features of NZE intonation (Section 3.2). I also wanted to avoid assessing two separate meanings on one scale, such as a polar-adjective scale contrasting Submission and Dominance, which may not accurately describe the participants' feelings about the speaker groups with regards to both meanings.

Therefore, rather than being presented as polar-adjective scales (e.g. contrasting Dominant and Submissive on the one scale), each scale contrasted degrees of one individual meaning, with the left side of the scales indicating a lack of the assessed characteristic (e.g. the “not more Dominant than British men”) and the right side of the scales indicating very much of the assessed characteristic (e.g., for the same example “much more Dominant than British men”). This scale, however, failed to assess “less than” the assessed characteristic on the given scale. Although keeping the scales similar in the first (section 3.2) and second (section 3.3) experiment was important for comparability (i.e. both scales contrast “not more than” with “very much more/like”), it may be possible that assessing “less than” on the scale would have provided different results.

The same attitude assessment scales, with appropriate changes in the wording of “men” to “women”, were used to assess attitudes about New Zealand women in comparison to British women. Thus the attitude assessment included 32 attitude scales (16 attitudes for both sexes).

### **3.3.2. Procedure**

The questionnaire was provided to the participants with a delay of one week after the tonal cues experiment (Section 3.2) in order to avoid the possible effects of one experiment influencing the results of another. An introduction was provided, which was read aloud by the author, and which identified the groups of speakers under investigation. Just as above (see Section 3.2.3), the participants were asked to indicate their impressions of NZE speakers (Pākehā speakers with a 3-year university degree and from the Wellington area) in comparison to BrE speakers (university educated and from the south of England).

### **3.3.3. Participants**

The same participants completed the Attitude questionnaire as those who took part in the tonal cues experiment reported above (see Section 3.2.2).

### **3.3.4. Scoring**

Scores were measured to the nearest whole millimetre from the 100mm response scale. Measurements were taken as the centre point on any marks which were a millimetre or so wide. Scores could range from 0 (for “not more [attitude]”) to 100 (for “much more [attitude]”). The measurements were reassessed before the statistical analyses were carried out in order to ensure their reliability. A Bivariate Correlation analysis was carried out for each sex (male and female) looking for correlations of each of the three tonal cues which were found to be indicative of NZE (high phrase-medial L, high H% and low L%) with each of the sixteen meanings assessed in the questionnaires. The significance level was adjusted to account for the number of statistical tests carried out on the data; thus, in order to obtain a significance level of 0.05, I used the Bonferroni correction (Rice 1995:323-324) and tested each hypothesis test at a level of 0.05 divided by the number of tests (48). Thus the significance level was set at 0.001.

### **3.3.5. Results**

There were no significant correlations found between the 16 attitudes and any of the three tonal cues (i.e. low L value, high H% and low L%) for either speaker sex (male and female). The results suggest that the assessed attitudes were not related to the listeners’ interpretations of which tonal cues are indicative of NZE, at least when using the sixteen scales used here.

## **3.4. Discussion and Conclusions**

From these results, I conclude that native NZE listeners use several tonal cues, including the heights of H%, L% and L values when determining NZE as distinctive from BrE. Specifically, a higher value for a high final boundary target (H%) was more characteristic of a NZE speaker. In addition, the lower value for a final low boundary target (L%) was interpreted as sounding more like NZE. These two results may be due to the rate of change at IP-Final positions; the steepness of the F0 change is greater the higher the H% and the lower the L%. However, pitch dynamism as measured separately (see

Figures 3.2l, 3.2m and 3.2n) was found not to be a significant factor in this experiment.

The results may also be illustrating that pitch range may be a factor. That is, rather than the steepness of the pitch change, the important factor may be the pitch span of the final target, which is considerably larger for lower L% and higher H% than the alternatives. However, pitch range was examined as a possible main effect in this experiment by adjusting all of the pitch values in the phrase to either a wide or a narrow position. The results indicated that global pitch range was not used by listeners as a distinguishing feature of NZE. Second, the height of the phrase-medial low values (L) was also a significant factor; higher L values were found to indicate NZE. As evident in Figure 3.2b, higher L values illustrate a more shallow pitch movement within the phrase. The two results show that a wide global pitch range and a wide pitch range within the phrase are not indicative of NZE. However, these results don't nullify the possibility that localised pitch range is a factor. It may be that the two contrasting variables, specifically wide and narrow pitch range, are bound to the location within a phrase; although wide pitch movements at phrase-final locations were found to be indicative of a NZE speaker, within the phrase high L values —



Nonetheless, the results illustrate which features of intonation native NZE speakers believe to be indicative of NZE, when contrasted with BrE. Examining these intonation features further may prove beneficial to the understanding of NZE intonation and the meanings conveyed by such use. In order to assess the validity of these perceptions, it would be necessary to compare BrE and NZE intonation use, specifically examining whether NZE speakers realise narrow pitch ranges within the phrase and wider pitch movements onto the IP-Final boundary targets. Further examination of the features which speakers believe are indicative of their variety may also be valuable in understanding the perceived meanings of intonation, investigating specifically what is interpreted by such use of intonation. Such research questions are examined further in Chapters 6 and 7.

### NZE intonation patterns in read-aloud speech

---

#### 4.0. Chapter abstract

The aim of the experiment reported in this chapter was to examine the forms of NZE intonation in order to better describe the characteristic properties of this language variety's tonal use. The speech data for this analysis were taken from a reading task completed by 30 speakers. For each speaker, I examined the same set of 12 intonation phrases within the reading passage in order to further explore the nature of nuclear accents and boundary targets in NZE. In the first analysis, I examined the preferential use of tonal composition for two examined utterance types: NZE speakers preferred an H\*L-L% on statements and an L\*L-H% for both wh-questions and questions without the lexical question marker (e.g. wh). In the second analysis, I explored the pitch range produced by NZE speakers on the nuclear pitch movement, comparing 7 different tonal compositions and 2 sentence types. This was done in order to determine whether NZE may be characterised not only by localised features of pitch range, as shown in Chapter 3, but also by pitch range features that might be specific to sentence type or to tonal composition. The results of the pitch range analysis revealed that tonal composition, not sentence type, was a significant factor. Finally, this chapter also presents average pitch level results for male and female speakers in order to better describe the form of NZE intonation.

#### 4.1. Introduction

The two principal characteristics of NZE intonation demonstrated in Chapter 3, namely higher medial L values and more extreme final L% and H% values, reflect two independent observations made about NZE intonation. On the one hand, Bauer (1994: 391) makes the general claim that NZE is “generally

speaking rather flat", which is compatible with the experimental observation in Chapter 3 that higher medial L values and subsequently a narrower pitch width was used within the phrase. On the other hand, Warren and Daly (2005: 233) make a more specific claim that "...the extent of pitch movement is considerably larger for NZE speakers [than for either Cambridge or Leeds]" on the nuclear pitch movement, which is reflected in the finding in Chapter 3 that more extreme L% and H% values and consequently a wider pitch range at the end of a phrase were more indicative of NZE.

The results paint a general picture of NZE intonation (as interpreted by NZE listeners) as having narrower pitch movements within the phrase and wider pitch movements onto the boundary tone. However, the results presented in Chapter 3 fail to describe the specific forms which make up NZE intonation. For instance, we do not know the preferential use of tonal structure which may characterise this variety of English. This question is especially important when previous results are considered which have found that speakers of different varieties of English produce distinctive forms of intonation on separate sentence types (e.g. statement vs. question).

As noted in Chapter 2, Grabe, Kochanski and Coleman (2005) have shown that a variety of intonation patterns may be used by speakers of Cambridge, Newcastle and Belfast English on different sentence types. However, they also note that there are clear preferences in each of these three varieties of English for specific tonal patterns on a specified type of sentence. They found that a high pitch accent followed by a low boundary target is the preferred nuclear pattern for wh- questions and statements amongst Cambridge and Newcastle English speakers but that a low pitch accent followed by a high boundary target is preferred by Belfast English speakers. For echo questions, they found that Belfast and Newcastle English speakers preferred a low pitch accent followed by a high boundary target whereas a low pitch accent followed by an extended high boundary target was the preferred nuclear pattern in Cambridge English.

To date, we do not know the preferential use of intonation for specific sentence types in NZE. Given that research has established some different

preferential use of intonation for sentence types between varieties of English, it would be unwise to assume that NZE is similar to any one other variety. Thus the first aim of the experiment reported here is to examine the preferential forms of intonation produced at the nuclear location by NZE speakers.

In addition to the preferential use of certain intonational forms on specific sentence types, the pitch range produced on these linguistic features (tonal composition and sentence type) may also be a characteristic feature of NZE intonation. As noted above, a wide pitch movement from the final pitch accent onto the boundary tone is indicative of NZE. It may also be the case that NZE speakers use the size of the pitch movement to distinguish, for example, questions and statements. For instance, if it is the case that speakers produce the same tonal pattern (e.g. a high pitch accent followed by a low boundary tone) on both statements and questions, they may nevertheless distinguish these two sentence types by differences in the width of the pitch movement. If this turns out to be the case, then a fuller description of pitch range effects in NZE should include reference not only to the location of such effects within the phrase but also to the sentence type.

Previous research has illustrated that sentence type influences the rise onset location on phrase-final rising intonation. As noted in Chapter 1, there have been several studies examining how rises are produced differently in questions and statements in NZE. For instance, Zwartz and Warren (2003) carried out a perception experiment in which they examined the phrase-final rise onset location on a set of gradient stimuli and another set of "rise-and-plateau" stimuli (Zwartz and Warren 2003: 55). The results showed that the shape of the rise is important and that, in general, listeners were more likely to interpret early rises as questions and later rises as statements. Similar findings were reported in a separate analysis of connected speech; one finding presented by Warren and Daly (2005) was that both men and women start their question rises earlier than their statement rises.

Other than rise onset location, Daly and Warren (2001) found that speakers produced significantly more dynamic pitch movements on phrase-

final rising pitch movements on questions than statements. We do not currently know if other tonal features on different pitch movements (e.g. phrase-final falling pitch movements) are used to differentiate between sentence types in NZE. Pitch range may contribute to such distinctions. It is also possible that NZE speakers produce characteristic pitch ranges on different tonal structures; for instance, NZE speakers may produce significantly wider pitch ranges on H\*L-L% in comparison to L\*H-H%. It is unclear at this point whether pitch range is important in differentiating sentence type in NZE or whether it is used differently on specific tonal structures in this language variety. Such features may illustrate further characteristic properties of NZE intonation.

In sum, the overall aim of this chapter is to add further detail to the description of the forms of NZE intonation started in Chapter 3. First, I examine the preferential use of intonation at the nuclear location for statements and questions. Next, I examine whether or not NZE may be characterised by pitch range which is specific to either the sentence type or the tonal composition of an utterance. In order to better define NZE intonation, I also present the average pitch levels produced by NZE men and women.

## **4.2. Methodology**

### **4.2.1. Materials**

The data reported here are taken from a reading task, in which a paragraph was read aloud by participants. It should be noted here that the material was not designed specifically for each of the research questions set out in this chapter. The monologue used here was originally intended to elicit the speakers' average F0 for the benefit of future examinations. A paragraph was read aloud rather than a word or sentence in order to minimise the effects of tonal composition which may affect the average pitch level in shorter utterances. In addition, different punctuation marks were used to elicit different tonal structures; full stops and question marks were used to elicit statement and question intonation (respectively) and commas were placed

within a number of sentences in order to initiate a sentence medial phrase break and possibly another intonation. Although it was believed at the time of carrying out this experiment that the punctuation marks would elicit the desired sentence types, it may have been better to ask each participant if the punctuation was successful in eliciting the different types of sentences.

For the current analysis, the monologue also provided the materials used to examine preferential tonal use and pitch range. Additional materials were not developed for these additional experimental aims for two reasons. First, tonal features other than those examined in Chapter 3, such as a wider pitch range than that created, could not be explored due to the controlled nature of the experiment. In the current experiment, I did not want to define more specific investigative aims and subsequently more controlled experiments before analysing several other tonal features produced by NZE speakers in speech. Thus the current experiment is exploratory. Further controlled experiments were carried out later and reported in Chapters 5, 6 and 7.

Second, the read-aloud paragraph was deemed appropriate for the current analysis. First, the materials were designed to elicit two sentence types, namely question and statements, and as such, the materials could be used to address the additional research questions in this chapter. In addition, using a paragraph within a reading task, such as that used here, rather than sentence lists was preferred in order to control for variation in the participants' interpretations of the utterances. Such evidence was provided by Cauldwell (1999) who demonstrated how context can change the interpretation of an utterance. Specifically, he showed how speakers' interpretations of an identical utterance can change from "irritated/angry" to "not irritated/angry" when additional information was provided, such as surrounding utterances and the speakers' relationship with one another. Similar observations have been noted by others (Cruttenden 1997; Vermillion 2000). In the current analysis, the utterances are in the context of a conversation rather than a list of unrelated sentences. Additional context was provided in the form of the participant's discourse partner, who the participant was told to imagine was

the person that they are in a relationship with.<sup>5</sup> Finally, using the read paragraph rather than using spontaneous speech minimises the effects of lexical and syntactic variation of the utterances whilst also controlling for the segmental effects on the F0 trace. It should be noted, however, that the results from this reading task may illustrate how intonation is used on read aloud utterance types rather than that produced in a genuine question or statement in a dialogue.

The monologue incorporated 12 sentences, varying in length and structure. There were 7 statements, which were prompted with a full stop ('.'), and 5 questions, which were indicated by a question mark ('?'). As noted above, the intention of the commas was to initiate phrase-medial breaks. The monologue is illustrated in Figure 4.1.

George was sweet, you know? He had that thing. What is that thing where you, fall asleep in the middle of a sentence? You know? What is it? Narcolepsy. Right. So anyway, George, uh, went to the union to get his free turkey, because the union always gives George this free turkey at Christmas time. Because he was shell-shocked. You know what I mean? In the first world war. So anyway, George was standing in line, getting his free turkey, but the thing is, he falls asleep and never wakes up.

*Figure 4.1 Reading task used to investigate pitch range*

#### **4.2.2. Participants**

The data used in this production study consist of recordings of 30 students from Victoria University of Wellington. Social variation between the participants, which has been found to be a factor influencing linguistic variation, was assessed using a questionnaire (Appendix B2). Thirteen participants in the selected group were male and the other seventeen were

---

<sup>5</sup> It was believed at the time of preparing the experiment that an imaginary discourse partner would be sufficient in providing context. It may be better to have a person present in the room during the experiment or to have recorded a dialogue between two people to elicit a more natural speech style.



female. All speakers were full-time University students, with an average age of 21.2 years, in a range of 18-29 years. There were not enough speakers to justify an analysis of age as a social factor; with 25 speakers within the age range of 18-22 years and 5 older speakers (23-29 years). All listeners were either Pākehā (New Zealander of European descent) or mostly Pākehā with some Māori ethnicity. All of the participants claimed to be native New Zealand English speakers, and none reported having any known hearing impairments. Speech impediments were not heard in any of the speakers who were included in this experiment. The analyses to be reported below examined speaker sex as the only between-subject factor, as participants were within a narrow range of other social categories.

#### **4.2.3. Procedure**

The recording session started with an introduction (Appendix B3), which was prerecorded by the author onto a Sony Prism 74 minidisk using a MDS-JB90 minidisk recorder and AKG D19DCS unidirectional microphone. The introduction was played aloud to each participant and served to reduce possible contextual variation, such as the relationship between the speaking partners. Specifically, the paragraph was introduced as a monologue in which one speaker was telling a story to their discourse partner, who they were in a relationship with. The participants were first instructed to read the paragraph silently in order to become familiar with the text. They then read the paragraph aloud from a 125 x 75 mm index card, as “naturally as possible” as if “you were speaking to your partner” whilst being recorded using the same equipment as described above. Each experiment was conducted individually in a quiet room in the Language Learning Centre at Victoria University of Wellington.

The recording ended with an opportunity for speakers to ask questions and to provide feedback about the task. In this feedback session, the participants were asked whether or not they were familiar with the monologue. As the monologue was taken from an American movie, I wanted to ensure that they were not “acting” American. No participants noted any familiarity with the monologue.



4.2.4. Analysis

The recorded speech material was later resampled using Praat (Boersma and Weenink 1996) with a sampling rate of 16 kHz in mono. The entire monologue was used to determine average pitch level for male and female NZE speakers in order to describe the intonation of NZE. In order to obtain the average pitch level for male and female NZE speakers, the pitch point values at every 10 milliseconds were first extracted from the read paragraphs. The data were manually examined for sampling errors. The remaining pitch point values were used to calculate the mean pitch level for each speaker, which was then used to calculate the average pitch levels of NZE male (13) and female (17) speakers reading aloud a paragraph (Section 4.3.1).

The speech materials analysed in the second (Section 4.3.2) and third (Section 4.3.3) analyses included both statements and questions, as illustrated in Table 4.1.

Statements	Questions
1. He had that thing.	1. George was sweet, you know?
2. Narcolepsy.	2. What is that thing where you, fall asleep in the middle of the sentence?
3. Right.	3. You know?
4. because the union always gives George this free turkey at Christmas time.	4. What is it?
5. Because he was shell-shocked.	5. You know what I mean?
6. In the first world war.	
7. ...he falls asleep and never wakes up.	

Table 4.1. Sentences listed according to sentence type which were included in the current investigation

As shown in Table 4.1, the questions included samples of different types of interrogatives. For this analysis, we examine two question types: wh-questions (Question numbers 2 and 4 in Table 4.1) and questions without the wh- question marker (Question number 1, 3 and 5 in Table 4.1). As noted

above, the material was not designed specifically for each of the research questions set out in this chapter, and as a consequence, there were not enough data to examine the different question types further. The 12 specified utterances were included in the data if the speaker produced an intonation phrase boundary at the end of the listed utterance, and if there were no reading errors at this location.

In order to analyse pitch range at a specific location, it is necessary to consider tonal phrasing in addition to paragraph structure. It has been claimed that pitch naturally declines over the course of an intonation unit as the subglottal air pressure decreases (Lieberman 1967; Cohen and Hart 1967). Although F0 reduction across an intonation unit is not based solely on air pressure reduction, there appears to be a similar tendency in all languages for such effects to occur, as noted for the Production Code (see Chapter 2). As a result, pitch range becomes narrower later in the intonation phrase until the final low value is reached and the speaker resets their pitch in a new intonation phrase. As a consequence, a speaker who utters two sentences as two separate intonation phrases will have a wider pitch range on the second sentence than a speaker who utters the same two sentences as a single intonation phrase. As noted above (Section 4.2.1.), commas were placed within sentences in order to prompt phrase boundaries and different intonations. Although there was a general tendency to produce phrase breaks at places designated with a comma in the text and intonation breaks at places designated with either a full stop or a question mark, there was some additional phrasing variation. The speakers, however, did not appear to produce different phrasings in order to convey different meanings but rather as a result of false starts, reading errors and hesitation with certain segments of the reading task. These appeared more readily in longer sentences than in shorter ones. Although it is unclear why this effect occurred, it may be due to the speakers not familiarising themselves with the text enough to read the monologue without errors and pauses. As there was an even distribution of long and short sentences for both statements and questions, I do not anticipate the phrasing variation to affect the pitch range results on the sentence type. Similarly, there was an even distribution of questions and

statements across the reading text, which means the effects of paragraph structure should be minimal (e.g. wider pitch ranges produced early in the paragraph and more narrow pitch ranges produced at the end of the paragraph, c.f. Sluijter and Terken 1993).

The experiment reported here examines the intonation produced from the final pitch accent in the phrase onto the boundary tone, which we will refer to as the nucleus, following the traditional British-style labels. The assessment of this nuclear location was motivated by two factors. First, previous assertions made about the meaning of intonation indicate that the nucleus contributes the most meaning in an intonation phrase. O'Connor and Arnold (1961:12) state: "Clearly the stressed syllable of the last prominent word is a landmark of the highest importance, and it is on this syllable that the whole tune centres". The second motivation for examining the nuclear tonal movement is to compare how NZE speakers use phrase-final intonation on specific sentence types. As noted in Chapter 2, statements are often associated with phrase-final falling pitch in many languages and language varieties, whereas a phrase-final rising pitch often indicates a question in these languages. However, rising intonation patterns occurring at the end of statements "are often cited as typical of this variety" (Warren and Britain 2000: 169). In this examination, I want to examine whether NZE speakers produce phrase-final low boundaries (L%) on statements and phrase-final high boundaries (H%) on questions, reflecting a universal tendency in the use of intonation, or whether the preferential use is more varied, such as that shown for three other varieties of British English (Grabe, Kochanski and Coleman 2005).

In using the read aloud paragraph, however, I did not expect many High Rising Terminals to be produced. Warren and Britain (2000) have noted that speakers use HRTs to reduce the distance between speakers. Specifically, they note that HRTs would appear more frequently in speech that contains "those structural aspects which especially function to seek understanding and enhance common ground" (2000: 168). Although the participants were told to imagine that they were speaking to their intimate

partner, there was no other speaker present other than the participant. In imagining a discourse partner rather than having one present, the text type and experimental procedure may have made it difficult for the speakers to produce linguistic features required for conversation between speakers. Therefore the text type and procedures may not prompt a feature which aims at increasing understanding and reducing distance between speakers. The current analysis is not an assessment of HRTs, as has been carried out by various researchers (Allan 1990; Britain 1992; Britain and Newman 1992; Ainsworth 1994; Warren and Britain 2000; Warren 2005b), but rather an assessment of intonational forms of NZE when inter-speaker solidarity cues are not required.

The list of the utterances included in the current analysis were provided in Table 4.1. The utterances were transcribed using a ToBI framework. Figure 4.2 illustrates an F0 trace derived from one NZE speaker reading aloud the first two sentences in the monologue.

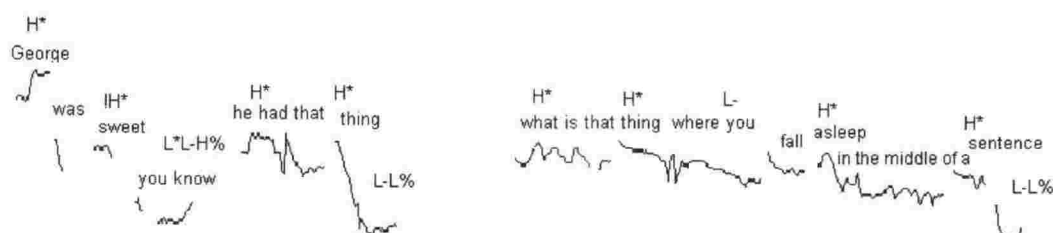


Figure 4.2 F0 trace of NZE speaker reading aloud the first two sentences in the monologue.

High and Low pitch was noted on pitch accents (\*), intermediate phrase accents (-) and boundary tones (%), as detailed in Chapter 1. The final pitch accent, the intermediate phrase accent and intonation phrase boundary tone in an intonational phrase (i.e. together constituting the nucleus) were the only assessed components in this examination. Downstep was not examined in this analysis. All transcriptions were reassessed after the completion of labels being assigned to all data, and any differences between the first and second transcriptions were assessed again before statistical analyses were carried out. High and low pitch point values were also noted (in ERB) within the nucleus and were used to calculate pitch range at this location.

### 4.3. Results and Discussion

#### 4.3.1. Pitch Level

The average pitch level for NZE male speakers reading aloud a monologue was 3.82 ERB, with a standard deviation of 0.48 ERB. The range for NZE male speakers' pitch level was from the lowest pitch level of 3.22 ERB to the highest pitch level of 4.65. As expected, the average pitch level for female speakers reading aloud the same monologue is significantly higher due to the physiological differences between men and women (see Chapter 3). The average pitch level for NZE female speakers was 6.11 ERB with a standard deviation of 0.31 ERB, with the lowest level female producing a pitch level of 5.71 ERB and the highest female speaker produced a pitch level of 6.80 ERB.

#### 4.3.2. Preferential tonal composition

In order to investigate how NZE speakers produce intonation on specified sentence types, each of the 12 examined locations was categorised according to Tonal Composition (7 categories: H\*LL%, L\*HH%, L\*LH%, H\*HH%, L\*LL%, H\*LH% and H\*HL) and Sentence Type (3 categories: statement, wh- question and other question). A 7 x 3 chi-square analysis was performed on the number of productions of each Tonal Composition produced for each Sentence Type. The results revealed a highly significant difference in the proportional use of Tonal Composition for the three Sentence Types ( $\chi^2 = 231.630$ ,  $p < 0.0001$ ). Table 4.2 outlines the proportional use of each Tonal Composition (rows) in each Sentence Type (column).

	Statement	Wh- Question	Other Question
Tonal Composition	Per cent (Total)	Per cent (Total)	Percent (Total)
H*L-L%	<b>77.44%</b> <b>(151)</b>	24.44% (11)	-
L*L-L%	1.02% (2)	-	2.74% (2)
H*H-L%	0.51% (1)	-	9.59% (7)
H*L-H%	0.51% (1)	2.22% (1)	6.85% (5)
L*H-H%	8.72% (17)	2.22% (1)	-
L*L-H%	8.72% (17)	<b>48.89%</b> <b>(22)</b>	<b>80.82%</b> <b>(59)</b>
H*H-H%	3.08% (6)	22.22% (10)	-
Total	(195)	(45)	(73)

*Table 4.2. Tonal Composition (7 categories) and Sentence Type (3 categories) produced at the nuclear location by NZE speakers*

As illustrated in Table 4.2, although NZE speakers use a variety of intonation patterns on the two Sentence Types, for each one there is a clear preference. The most frequently produced Tonal Composition on statements was H\*L-L% (77.44%). The distribution of Tonal Composition dramatically shifts when the speaker is asking a question; specifically, the most frequently produced intonation for both types of questions was L\*L-H% (48.89% and 80.82%).

The evidence presented here suggests that New Zealand English speakers employ a similar tonal preference for statements as Cambridge English speakers (Grabe et al. to appear). Specifically, speakers of both language varieties chose a high pitch accent followed by a low boundary target. A low pitch accent followed by a high boundary tone was the preferred intonation for statements in Belfast and Newcastle English. For questions, it

appears that NZE speakers preferred a low pitch accent followed by a high phrase-final target (L\*L-H%). This tonal preference is similar, although not identical, to that used in Belfast for wh-questions, who produce a low pitch accent followed by a rise-plateau (labelled as L\*H % in IViE) and Cambridge English for echo questions, who preferred a low target on a prominent syllable followed by a notably high target (L\*HH%).

The data presented in Table 4.2 illustrates an equal number of L\*L-H% and L\*H-H% intonation types produced on statements. For questions, however, there was only one L\*H-H% intonation and 81 L\*L-H% intonations. This result was surprising in light of previous findings; although difficult to compare results due to different categories of rises assessed (e.g. H\*H-H%, L\*L-H%, or simply “rises”), previous studies have shown that the NZE question “rises” start later than statement “rises” (Zwartz and Warren 2003; Warren 2005b; Warren and Daly 2005). An L\*H-H% intonation is produced with a continual rise from the low pitch accent (L\*) onto the phrase-final boundary target (H%). Conversely, an L\*L-H%, which is often noted as a continuation rise, anchors the low phrase accent (L-) toward the end of the phrase after the low pitch accent and rises to its completion point. Thus the results presented here indicate that NZE speakers more readily produce a later rise onset than an earlier onset for question rises. The difference in pitch range between the intonation patterns is discussed in Section 4.3.3.

It is tempting to interpret the use of L-H% (rather than the use of H-H%) as functioning as a continuation rise and indicating that the speaker’s turn at talk was not yet completed, as may be expected in a monologue. If the speaker were successful at imagining their discourse partner being present for the conversation, such an explanation may be accurate. However, the speakers were reading aloud a paragraph whilst sitting alone in a quiet room. In such a situation, asking questions or indicating a speaking turn cue may appear unnatural without a discourse partner present, even though they were prompted to imagine having a conversation with their discourse partner. The preferential use of L\*L-H% on questions may be illustrating the reluctance of the speakers to use such conversational cues without a speaking partner present. It may be that speakers typically produce an early rise onset on



questions, such as with an L\*H-H% intonation. However, their reluctance may have produced a delayed rise, in which “the L spreads to the end of the tail” (Gussenhoven 1983: 217; see also Chapter 2, Section 2.1.2). Warren (2005a:225) describes this convex rising pattern as a double-anchoring of the phrase accent: “a H- phrase accent is anchored to both the nuclear syllable and the later anchor point”. This “delay” may result in an intonation similar to that described as L\*L-H% above. Although this interpretation is speculative, the results may be showing that NZE speakers produce delayed rises, such as an L\*L-H%, when asked to produce conversational cues when such cues are not required in the context.

#### **4.3.3. Determinants of Nuclear Pitch Range**

In order to examine whether nuclear pitch range is related to the tonal composition of an utterance or to the sentence type in NZE, an Analysis of Variance was performed, with pitch range within the nucleus as the dependent variable. Nuclear pitch range was simply calculated as the difference between minimum and maximum values between the intonation-final pitch accent and the final intonational boundary target of the intonation phrase. The independent variables included Sentence Type (2 levels: statement and question), the Tonal Composition produced on the nucleus (7 levels: H\*LL%, L\*HH%, L\*LH%, H\*HH%, L\*LL%, H\*LH% and H\*HL), Speaker Sex (2 levels), and Speaker was the random factor. Sentence Type was not found to be a significant factor here, which suggests that NZE speakers produce equal or nearly equal pitch ranges on both the statements (0.939 ERB, s.e. 0.126) and questions (1.119 ERB, s.e. 0.113) examined here. There were two significant main effects. First, there was a significant main effect of Speaker ( $F_{28,260} = 2.658$ ,  $p < 0.01$ ) indicating that pitch range is speaker specific. There was also a significant main effect of Tonal Composition ( $F_{6,280} = 2.545$ ,  $p < 0.05$ ). No other significant main effects or interactions were found. Table 4.3 lists the average pitch range (represented as ERB values) produced in each Tonal Composition.



Tonal Composition	Pitch Range	
	in ERB	Std. Error
H*L-L%	1.075	0.106
L*L-L%	0.348	0.048
H*H-L%	0.792	0.274
H*L-H%	0.879	0.275
L*H-H%	1.598	0.237
L*L-H%	1.001	0.100
H*H-H%	1.166	0.241

Table 4.3. Average pitch range (in ERB) produced by 30 NZE speakers on 7 Tonal Compositions

The widest pitch range (1.598 ERB) is produced on an L\*H-H% nuclear intonation pattern, followed by the H\*H-H% (1.166 ERB), H\*L-L% (1.075 ERB), L\*L-H% (1.001 ERB), H\*L-H% (0.879 ERB), H\*H-L% (0.792 ERB) and L\*L-L% (0.348 ERB). This result is largely unsurprising as certain tonal compositions restrict the pitch range that a speaker may produce. For instance, a wide pitch movement on an H\*H-L% tonal structure is not possible, as the tonal combination produces a narrow pitch movement.

This result was anticipated and as such, two *a priori*, or planned comparisons were designed. First, I wanted to compare how speakers use pitch range within the lower end of their range (L\*L-L%) and the higher end of their range (H\*H-H%). However, an unanticipated result was that there were few speakers who produced these two intonations on the nucleus when reading aloud the paragraph; specifically, there were only 4 productions of an L\*L-L% intonation (4 speakers who each produced an L\*L-L% intonation once in the paragraph) and 16 productions of an H\*H-H% intonation (14 speakers who produced an H\*H-H% once in the read paragraph and one speaker who produced this intonation twice). The second planned comparison aimed to examine how speakers produce pitch range from a high pitch accent onto a low boundary target (H\*L-L%) and a low pitch accent onto a high boundary target (L\*H-H%). There was considerably more data for this comparison, but while there were 162 occurrences of H\*L-L% (all 30 speakers), there were only 18 productions of L\*H-H% (14 different speakers). As a consequence of the few available data, an analysis of the planned comparisons would not accurately illustrate significant differences in pitch range due to the tonal composition of specific utterances, and as such, were not assessed for

statistical significance. Chapters 6 and 7 further examine pitch range at this location, specifically exploring variations in pitch height as a meaningful linguistic device in this language variety.

Speaker Sex was not found to be a significant factor in any of the analyses done here, indicating that male and female NZE speakers produced similar pitch ranges at the nuclear location in the analysed reading task. This result is surprising in light of previous findings. As noted in Chapter 1, Daly and Warren (2001) analysed the speech of six female and five male NZE speakers in two reading tasks: a sentence list and a children's story, namely *Cinderella*. The use of the materials is based on the tasks used in the study of Intonation Variation in English (IViE) in the British Isles (Grabe, Post and Nolan 2001). Daly and Warren found that female speakers produced wider pitch ranges than male speakers in both reading tasks. They also found that the pitch range differentiation was greater for the story telling task, which provoked much wider pitch ranges for women than men.

Daly and Warren note that "differences between studies reporting speaker sex differences in pitch patterns and those failing to find them are principally to do with the speech task and its analysis" (2001:93). It may be true that the two text types used by Daly and Warren instigated a sex difference in pitch range whereas the current analysis did not. Evidence indicates that men and women do, in fact, employ pitch differently. For instance, van Leeuwen (1999: 109, as cited in Chen et al. 2004: 343) asserts that "where men raise their voice to a higher pitch, women may do the reverse – for instance in newsreading". In addition, Vermillion (2001:69) found "male speakers used a significantly lower [pitch] register to convey support and non-support whereas female speakers realised significantly later peaks to convey authority and submission". The evidence suggests that men and women use pitch differently in separate contexts. It may be that female speakers used a wider pitch range in Daly and Warren's tasks than male speakers as they may have been using a speech style in which men and women use different pitch ranges: reading aloud a children's story or isolated utterances. It may be the case that male and female speakers produced similar pitch ranges in the current study because of the context and the reading task. Specifically, the reading task used in the current study

promoted a different style of speech which was a read aloud monologue, which was directed toward an adult audience.

#### **4.4. Summary and Conclusions**

Although this study does not show how NZE speakers use intonation in spontaneous speech, the results presented above illustrate some of the forms of intonation produced by NZE speakers in a read aloud monologue. The examination of the preferential use of intonational forms revealed that NZE speakers prefer a low pitch accent followed by a high IP-Final boundary target to indicate two different types of questions whereas a high pitch accent followed by a low IP-Final boundary target is preferred to indicate statements in a scripted monologue. It is interesting to note that universal meanings assigned to phrase-final rises as questions and phrase-final falls as statements appear to be evident in this variety of English. Specifically, speakers preferred a low phrase-final target on statements and conversely, a high phrase-final target on questions.

The results also illustrate few high boundary targets (H%) in the narrative style speech, which has previously been shown to be a strong candidate for High Rising Terminals (Allan 1990; Britain 1992; Warren and Britain 2000). As noted above, however, HRTs may be used to "overcome interspeaker hurdles and to build and maintain speaker-hearer solidarity" (Warren and Britain 2000: 169). The tonal preferences described above may be indicative of an intonation used when conversational cues between speakers are not necessary. Specifically, the data are derived from a reading task, which the participants were told to "act out" as if they were talking to their discourse partner. However, as the discourse partner was imaginary, the speaker did not need to signal "positive politeness markers" (ibid) to their discourse partner. In addition, the speaker would not need to check for the listener's understanding (Guy and Vonwiller 1984: 4). Thus the data may illustrate the intonation used in NZE when such conversational cues are not required, such as in the scripted monologue used here. The intonation used to indicate conversational cues are examined further in Chapters 5 and 7.

It was also interesting that both male and female speakers produced equal pitch ranges at the nuclear location throughout the paragraph. This result was surprising in light of previous research carried out on men's and women's pitch range differences at this location (Warren and Daly 2000). As noted by Warren and Daly, however, the task type is important in uncovering gender salient speech. They note from Holmes (1993) that "New Zealand women are highly attentive to the needs of their conversational partners" (Warren and Daly 2000: 112). They also state that "Elsewhere (Daly and Warren 2001) suggest that the use of a wider pitch range and more dynamic pitch is evidence of female speakers making an effort to attract and maintain the interest of their conversational partners". From these results, it was suggested the task type used in the current analysis (reading aloud a paragraph) did not prompt sex-salient speech, as there was no discourse partner present for the female speaker to be "attentive to" (at least if the participant was unsuccessful in imagining their discourse partner).

This chapter also investigated whether nuclear pitch range in NZE is determined by the tonal compositions being produced by the speaker and/or by sentence type. In addition to localised pitch range being a characteristic property of NZE, as shown in Chapter 3, the results presented here illustrate that such ranges may be affected by the tonal composition of the utterance, but not the sentence type. The meanings associated to these different intonations and variations in pitch range are examined further in Chapters 5, 6 and 7.

## **The production of intonational meaning**

---

### **5.0. Chapter abstract**

In the two previous chapters, I examined the characteristic forms of NZE intonation. The aim of the experiment reported here was to explicitly examine how NZE speakers use intonation as a meaningful linguistic device. A production experiment using Prosodic Completion Tasks (PCT) was designed with this aim in mind. In each PCT, listeners were prompted to convey a specific meaning on a given utterance. The intended meaning was elicited with the help of an audio prompt, which the participants had to understand as being their “discourse partner’s” utterance. This study illustrates how intonation may be used to convey five contrasting pairs of meaning in this variety of English. First, significantly different choices in pitch accent were made for three pairs of contrasting meanings: Emphatic vs. Unemphatic, Concerned vs. Unconcerned and Impressed vs. Unimpressed. These tonal contrasts are claimed to be due to the Effort Code (Gussenhoven 2004; Chapter 2); higher pitch accent peaks, such as H\* rather than L\*, and later occurring pitch accent peaks, such as L+H\* rather than H\*, were used by the majority of speakers to indicate meanings that might be associated with greater effort, such as Emphasis, Concern and an Impressed attitude. Second, significantly different choices of Intonation Phrase (IP) boundary tones were made for two pairs of contrasting meanings: Conversational Continuation vs. Cessation and Concerned vs. Unconcerned. Although a low boundary tone (L%) may be used to signal both Conversational Continuation and Cessation, a high boundary tone (H%) was used to indicate Conversational Continuation. Also, the majority of speakers preferred to produce L% to signal an absence of Concern and H% to indicate Concern. The results for the PCT examining Submission vs. Authority illustrated no significant differences. That is, H\*L-L% was used by the majority of speakers to indicate both Submission and Authority. It would seem that different tonal

target categories, such as high vs. low, are not used to contrast this meaning pair at the specific locations within the intonation phrase being investigated here, i.e. the pitch accent and the boundary tone. However, a further analysis of the pitch heights of H\* and L% in these contours illustrates that the L% is higher when conveying Submission than when conveying Authority, possibly indicating a categorical distinction between levels of L tone.

### 5.1. Introduction

If speakers want to convey a change in meaning without changing the lexical composition of an utterance, they may manipulate the tonal characteristics of the utterance. For example, Ladd and Morton (1997) found that higher peak heights convey emphasis whereas lower peak heights do not convey this meaning. Thus, if speakers want to emphasise how many sandwiches they have just eaten, they may realise a pitch-extended High pitch accent (H\*), for example, on the utterance *two*. They may also want to portray the non-routineness of having eaten two sandwiches, which they may convey by realising a later pitch peak (Gussenhoven 1983). As discussed in Chapter 2, there are a number of such universal and language-specific meanings ascribed to intonation. However, there is a lack of research examining how NZE speakers specifically employ pitch to convey such meanings. For instance, if NZE speakers want to convey the meanings listed above, would they use similar tonal techniques to those noted above?

Speakers readily produce intonation using specific form-function relationships which are typical of their language or language variety. For instance, Zwartz and Warren (2003, reported in more detail in Warren 2005a) found that NZE listeners are more likely to interpret early rise onsets as interrogatives and later rise onsets as declaratives. This evidence supports production experiment results (Warren and Daly 2004; Warren 2005a), which showed that “there is a stronger tendency in statements than in questions for late rise onsets rather than onsets on the accented syllable” (Warren 2005a: 214). Therefore the evidence suggests that rise onset location distinguishes statement rises from question rises in NZE. Such distinctions have not been found in other varieties of English.

Given that the form-function relationship of intonation may differ between varieties of English, the objective of the current experiment was to examine explicitly how NZE speakers use intonation to convey meaning. An additional aim of this experiment was to provide a starting point for future examinations, which explore possible intonational contrasts as used in this variety (Chapters 6 and 7).

## 5.2. Methodology

### 5.2.1. Materials

A number of methodologies have been employed in order to examine the meaning of a speaker's intonation, dating back to Uldall (1960). One such method, developed in Vermillion (2001), was used here to examine how NZE speakers produce intonation to convey meaning. Similar to a Discourse Completion Task (see Blum-Kulka 1983; Brown 2001), a *Prosodic Completion Task* (PCT) elicits a linguistic variable which illustrates how a speaker conveys a given meaning. In the Discourse Completion Task, the speakers are given background information, such as what the speaker's discourse partner had said, the setting and the speakers' relationship with one another. A specific feeling or emotion is also prompted, which the speaker is then supposed to convey through his/her response. The aim of a Discourse Completion task is to elicit lexical and/or syntactic forms that convey the prompted meaning. Therefore the lexical responses are provided spontaneously by the speaker rather than read aloud from a prepared script. The aim of a PCT, however, is to elicit the speaker's prosodic technique for conveying the prompted meaning. For this purpose, every PCT includes a lexical response that the speaker is required to use, thus limiting the speaker's own contribution of meaning to intonation.

In the current experiment, five meaning contrasts, each claimed to be distinguished by a speaker's intonation in BrE, were elicited using the following PCT design. Each PCT includes one *scene*, two *conditions*, and two *test items*, as shown in each of the panels of Figure 5.1. First, the purpose of the *scene* is to reduce contextual variation, which has been suggested as an influence on the meaning of tunes (Cruttenden 1997; Cauldwell 1999;



Vermillion 2000). The context provided by the *scene* constitutes the “discourse partner’s” utterance, to which the participant responds. Since participants’ interpretation of their discourse partner’s utterance is likely to vary according to the emotive use of intonation in that utterance, it is important to provide the same utterance for each participant. Therefore, each scene was provided by using a single recording of a native speaker of New Zealand English (who did not otherwise take part in the PCTs), who had been instructed to read aloud the sentences from a provided script “as naturally as possible” and “without emphasis” (see Appendix B5). In the PCT, the participants were asked to imagine that this speaker was their “discourse partner”. In playing aloud the previous speaker’s utterance, it would create a more realistic discourse environment than that the experiment reported in Chapter 4, where no “discourse partner” was provided.

The speech was recorded using a MDS-JB90 minidisk recorder and AKG D19DCS unidirectional microphone. The recorded speech was downsampled into a signal processing system, namely Praat (Boersma and Weenink 1996) at a sampling rate of 16 kHz. All of the pitch points in each utterance were moved up or down, creating one of two pitch levels based on the average level of NZE male (3.98 ERB) and female (5.83 ERB) speakers reading aloud a story (Vermillion 2003). As the current experiment was done at the same time as the experiment reported in Chapter 4, I could not use the average pitch levels found for 30 speakers reading aloud a paragraph. Nevertheless, evidence provided in Chapter 4 also showed that these two pitch levels were within the range of NZE male (3.22 - 4.65 ERB) and female (5.71-6.80 ERB) pitch levels reading aloud a monologue. The participants later confirmed that the utterances with lower and higher average pitch levels were “believable” examples of a male and a female speaker, respectively.

The *conditions* in the PCTs provided the meanings which the participant was to convey. The five meaning contrasts elicited in this experiment were (1) Conversational Continuation vs. Cessation, (2) Authority vs. Submission, (3) Emphatic vs. Unemphatic, (4) Concerned vs. Unconcerned, and (5) Impressed vs. Unimpressed (See Figure 5.1). These



conditions were presented along with the *test items*, which were the lexical responses for each condition. It is important to note that a single test item was employed as the response in the two contrasting conditions. As indicated above, an identical response inhibits the lexical expression of these types of meaning which, therefore, compels the speaker to convey these meanings prosodically. There were two test items, which were used in separate PCTs; these were the words *yeah* and *no*. These words were chosen because they consist entirely of voiced sonorant segments, which tend not to perturb the F0 trace.

My first prediction is that the first meaning contrast – Continuation vs. Cessation of the conversation – will be indicated by using a high boundary tone (H%) to show Continuation and a low boundary tone (L%) to show Cessation. This prediction is based on previous claims made about the meaning of rising vs. falling phrase-final pitch, which are often equated with the autosegmental metrical notations, H% and L% respectively (Ladd 1996: 82; this thesis Table 1.1). It should also be noted, as noted by Ladd, that there are exceptions to this association between ToBI and the British-style transcriptions; L% and H% do not always mean falling and rising, respectively. For instance, an H\*H-L% tonal composition would represent a level pitch movement from the high pitch accent onto the low boundary target.

Other scholars have associations between conversational cues and the intonation produced with the phrase-final pitch movement or the boundary tone (see Chapter 2). Specifically, Hirst and Di Cristo (1998:27) claimed that rises often denote unfinished utterances and similarly Gussenhoven (2004: 89) asserts that high pitch at the end of a phrase denotes continuation. Thus rises may indicate speaker-oriented speaking turn cues whereby a speaker may produce a rise to indicate that their speaking turn is not yet completed. I examined how NZE speakers use their intonation to indicate a continuation request to their discourse partner and how these same speakers cued that they wanted the conversation to cease (see Figure 5.1a).

Figure 5.1b illustrates the PCT used to elicit the second meaning contrast, namely Authority and Submission. Ohala (1994) claimed that low pitch signals “largeness”, which consequently conveys secondary meanings such as “threat, intention to prevail in a contest, dominance [and] self-sufficiency” (1994: 341), while high pitch conveys a sense of “smallness” and consequently conveys “nonthreat, submission, appeasement [and] desirous of goodwill...” (1994:341). Based on the meanings associated with the “Frequency Code” (Ohala 1994; Chapter 2, Section 2.1.1), I predict that higher pitch will be contrasted with lower pitch when distinguishing Submission and Authority in NZE.

These first two PCTs (Figures 5.1a and 5.1b) each used only one test item (*yeah* and *no* respectively) for reasons to do with contextual appropriateness. For instance, the participants’ “discourse partner” offers an invitation to dinner in both PCTs. In the first PCT, speakers may have difficulty conveying conversational continuation using a response which indicates a rejection (i.e. *no*). In the second PCT they may similarly struggle to indicate Authority using a response that indicates an acceptance to the dinner invitation (i.e. *yeah*). Such issues did not arise in the remaining PCTs, and so the following three PCTs use two semantically-contradictory test items (i.e. *yeah* and *no*) in order to examine how identical meaning contrasts may vary according to the lexical response.

The third meaning contrast explored how speakers convey Emphasis and an absence of Emphasis. The “Effort Code” (Chapter 2, Section 2.1.1) stipulates that the greater the pitch movement, the greater the degree of perceived insistence and emphasis, in addition to other meanings that can be derived from the greater expenditure of effort (Gussenhoven 2004: 85-89). Speakers may manipulate several tonal features in order to create greater pitch movement. For instance, they may raise their high targets, lower their low targets or use a combination of both tonal manipulations in order to produce greater pitch movement. I examine here which tonal features are manipulated in NZE, if any, in order to contrast Emphasis and an absence of Emphasis (see Figure 5.1c and 5.1d). Two lexical items, *yeah* and *no*, were

Scene: Your partner tells you "I want to go out to dinner this weekend. Would you like to join me?"

Condition 1: You are interested in this idea and signal to him/her to continue:

Test item 1: *Yeah*

Condition 2: You are not interested in this idea and signal to him/her to stop talking!

Test item 2: *Yeah*

a. Conversational Continuation vs. Cessation Cues

Scene: Your partner tells you "I want to go out to dinner this weekend. Would you like to join me?"

Condition 1: You say, with authority:

Test item 1: *No*

Condition 2: You say, submissively:

Test item 2: *No*

b. Authority vs. Submission

Your partner asks you "Did you enjoy the film?"

Condition 1: You think it was a good film. You tell your partner without emphasising:

Test item 1: *Yeah*

Condition 2: You think it was a good film and you want to emphasise to your partner how good you thought the film was:

Test item 2: *Yeah*

c. Emphatic vs. Unemphatic on *yeah*

Your partner asks you "Did you enjoy the film?"

Condition 1: You think it was a bad film. You tell your partner without emphasising:

Test item 1: *No*

Condition 2: You think it was a bad film and you want to emphasise to your partner how bad you thought the film was:

Test item 2: *No*

d. Emphatic vs. Unemphatic on *no*

Scene: Your partner tells you "I told my boss off at work the other day. Did I tell you?"

Condition 1: You are concerned and say:

Test item 1: *No*

Condition 2: You are not concerned and say:

Test item 2: *No*

e. Concerned vs. Unconcerned on *no*

Scene: Your partner tells you "I told my boss off at work the other day. Did I tell you?"

Condition 1: You are concerned and say:

Test item 1: *Yeah*

Condition 2: You are not concerned and say:

Test item 2: *Yeah*

f. Concern vs. Unconcerned on *yeah*

Scene: Your partner tells you "I told my boss off at work the other day. Did I tell you?"

Condition 1: You are impressed and say:

Test item 1: *No*

Condition 2: You are not impressed and say:

Test item 2: *No*

g. Impressed vs. Unimpressed on *no*

Scene: Your partner tells you "I told my boss off at work the other day. Did I tell you?"

Condition 1: You are impressed and say:

Test item 1: *Yeah*

Condition 2: You are not impressed and say:

Test item 2: *Yeah*

h. Impressed vs. Unimpressed on *yeah*

Figure 5.1: Prosodic Completion Tasks (PCTs) contrasting 8 pairs of meanings.

used for this PCT in order to determine whether a strong negative assertion, such as that conveyed with the response *no* (i.e. I did NOT like the film!), is emphasised differently to a strong positive assertion, such as *yeah* (i.e. I DID like the film!).

Two test items (Figures 5.1e and 5.1f) were also used in the PCTs investigating a Concerned vs. Unconcerned attitude, and also (Figures 5.1g and 5.1h) in the PCTs looking at the contrast of Impressed and Unimpressed. In each contrast, the test item *yeah* conveys that the information provided by the discourse partner is already known, and the test item *no* indicates that this same information is new to the conversation. The motivation for eliciting two test items was to examine how speakers indicate that the information is new or given whilst also contrasting meaning pairs (see Chapter 2, Section 2.2.1). Concern is claimed to be conveyed by a fall-rise intonation pattern on statements in British English (O'Connor and Arnold 1961:226), while a rise-fall pattern indicates that the speakers is awed and Impressed (O'Connor and Arnold 1961:147; Cruttenden 1986:100). The two meaning contrasts, Concerned vs. Unconcerned and Impressed vs. Unimpressed, were elicited in order to examine specifically how NZE speakers convey these two meaning contrasts with their intonation.

### 5.2.2. Participants

The speakers who took part in the PCTs were the same participants who provided the paragraph data reported in Chapter 4. Specifically, the data used in this production study consist of recordings of 30 students from Victoria University of Wellington. Thirteen participants in the selected group were male and the other seventeen were female. All speakers were full-time University students, with an average age of 21.2 years, in a range of 18-29 years. All listeners were either Pākehā or mostly Pākehā with some Māori ethnicity. All of the participants claimed to be native New Zealand English speakers, and none reported having any known hearing impairments or speech impediments. The analyses to be reported below examined speaker sex as the only between-subject factor.

### 5.2.3. Procedure

The recording session began with the participant reading aloud the monologue paragraph described in Chapter 4. They then listened to a prerecorded introduction (Appendix B4), which had been made by the author onto a Sony Prism 74 minidisk using a MDS-JB90 minidisk recorder and AKG D19DCS unidirectional microphone. The introduction was played aloud to each participant and served to reduce additional contextual variation which could not easily be controlled by the individual scenes for each PCT. The introduction established the participant's "discourse partner" (i.e. the recorded speaker) and their relationship to the speaker (i.e. they were to be imagined as the speaker's intimate partner). After the introduction was played, the participants were asked to read each PCT card to themselves to ensure that they understood the task. Any questions that they had were addressed at this point.

At the beginning of each PCT, the scene was played aloud to the speaker. Each male and female participant was played aloud a "female" and "male" scene, respectively. Thus the results represent how speakers respond to a discourse partner of the opposite sex. The motivation for using the voice which was the opposite sex to the participant rather than the same sex voice was to potentially make comparisons between the data reported here and the data obtained from a similar study on BrE (Vermillion 2001). As the purpose of the current investigation is to examine how NZE speakers contrast specific meanings, the data between the two studies will not be compared here. After each scene was played to the participant, they read aloud the test item from a provided 125 x 75 mm index card, as naturally as possible, whilst being recorded onto a Sony Prism 74 minidisk using a MDS-JB90 minidisk recorder and AKG D19DCS unidirectional microphone. At the end of the task, the participants were asked for any comments or questions regarding the experiment. Although it was not verified whether or not the participants followed the instructions of the PCTs, no one noted any difficulty with the task. Each experiment was conducted individually in a recording room in the Language Learning Centre in Victoria University of Wellington.

Although all speakers participated on a voluntary basis, they were provided with a music/book voucher worth 5 New Zealand dollars for their time.

#### 5.2.4. Auditory and Acoustic Analyses

With the help of the *Praat* signal processing system, the utterances were transcribed using ToBI (Beckman and Ayers 1994; Pierrehumbert 1980), as detailed in Chapter 1, with pitch accents, phrase accents and boundary tones all being marked (see Chapter 1). Figure 5.2 is provided here in order to illustrate several different intonations produced by one female speaker when conveying different meanings on the utterance *no*.

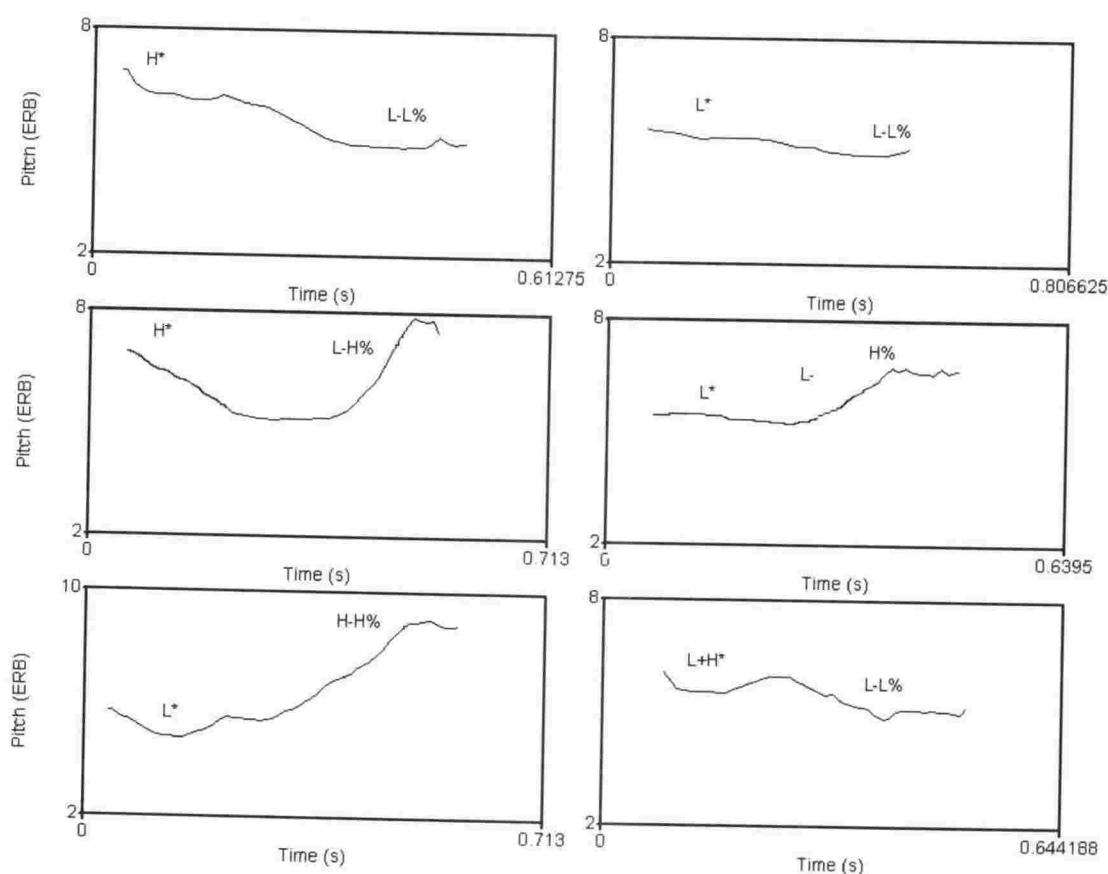


Figure 5.2. Six F0 contours produced on the utterance 'no' by a female NZE speaker.

In ToBI, the  $L^*+H$  pitch accent, which is referred to as a “scooped accent”, is realised as a low tone target focused on an accented syllable which rises. An  $L+H^*$  pitch accent, which is referred to as a “rising peak accent”, is realised as a high pitch tone target focused on the accented syllable which is preceded

by a sharp rise (Beckman and Ayers 1994). As the data in this analysis were monosyllabic, it was impossible to ascertain whether the Low or High tone in L+ H was accented as both tones were produced on one syllable. Therefore, the utterances were always labelled L+ H\*, which appeared as a sharp rising pitch movement from the Low to the High tone. Thus the current analysis fails to distinguish the possible differences between L+H\* and L\*+H.

All transcriptions were done "blind" in that I did not know the intended meaning while labelling the speech data. This was done so that the judgements were not influenced by knowing the intended meaning. The transcriptions were reassessed after the completion of labels being assigned to all data, and then again before statistical analyses were carried out. Any questionable labels assigned to the data were reassessed by a graduate student who was experienced in intonation research and in the ToBI labelling system. For each PCT, the pitch targets at the pitch accent and the intonation phrase boundary were recorded (in ERB) for later analysis. The pitch target height was not recorded for the phrase accent as the specific tonal location and subsequent pitch height proved difficult to pinpoint within the phrase.

### **5.3. Results and Discussion**

Two sets of statistical analyses were performed on the PCT data. First, a test of marginal homogeneity for paired data (Agresti 2002: 422) was carried out in order to investigate the effect of the independent variable (meaning) on the tonal composition of the phrase, and to determine whether there were any interactions between the pitch accent, phrase accent and the boundary tone. That is to say, the use of these three intonational features under the two conditions were examined, with consideration of the repeated measures design of the experiment. A second test for marginal homogeneity was used to determine the effect of the independent variable (meaning) on each individual tonal target rather than the combined tonal composition of the utterance. Specifically, I wanted to examine whether or not speakers produced different (1) pitch accents, (2) phrase accent or (3) boundary tones when indicating the pairs of elicited meanings. A McNemar's test (2 x 2 matrix) was used to assess the effects of the independent variable on both the phrase accent and boundary tones, comparing high (H) and low (L) targets. A



3 x 3 matrix was required to test the effects of meaning on the pitch accent, comparing the three pitch accent types produced by the participants in the PCTs: high (H), low (L) and a low plus high (L+ H\*).

In total, 24 tests were carried out on the individual targets, examining three tonal targets in eight pairs of contrasted meanings. The significance level was adjusted to account for the number of statistical tests carried out on the data; in order to obtain a significance level of 0.05 for the tests carried out on the individual targets, I used the Bonferroni correction (Rice 1995:323-324) and tested each hypothesis test at a level of 0.05 divided by the number of tests (24). Thus the significance level was set at 0.002 for the tests on the individual targets.

### 5.3.1. Conversational Continuation and Cessation

I start here by examining how speakers use their intonation to indicate a *continuation* request to their discourse partner, and how these same speakers indicate that they want the conversation to *stop* (Figure 5.1a). The test of marginal homogeneity showed that the speakers did not produce different intonations to convey the two contrasted meanings, as the differences were not significant in light of the Bonferroni correction ( $p < 0.005$ ). Table 5.1 outlines the distribution of the tonal compositions used by the NZE speakers when indicating this meaning pair. The table depicts the number of participants (each box) who produced two specific tonal compositions to indicate Continuation (column) and Cessation (row). The highlighted cells identify the number of participants who produced identical tonal compositions when indicating each of the two meanings.



<b>Cessation</b>	H*H-L%	H*L-L%	L*H-L%	L*L-L%	TOTAL
<b>Continuation</b>					
H*H-L%	-	-	-	1	1
H*L-H%	-	-	-	2	2
H*L-L%	-	5	-	3	8
L*H-L%	1	1	-	3	5
L*L-H%	-	4	3	3	10
L*L-L%	-	-	-	1	1
L*H-H%	-	1	-	1	2
L+H*L-L%	-	1	-	-	1
TOTAL	1	12	3	14	30

Table 5.1. Distribution of tonal compositions produced by 30 NZE speakers when conveying Conversational Continuation vs. Cessation.

Nevertheless, the second set of tests for marginal homogeneity performed on the individual targets identified that one tonal target was found to be significantly different when indicating Conversational Continuation and Cessation – the boundary tone ( $p < 0.002$ ). Specifically, a low pitch final target was used by all speakers (30) when conveying cessation cues to their discourse partner. 53% of these speakers (16) also used L% when indicating continuation, while the others used H%. No significant differences were found for the pitch accent or the phrase accent when indicating Conversational Continuation and Cessation on the individual target matrices. Thus the individual tests of marginal homogeneity indicate that there is an effect of condition (Conversational Continuation vs. Cessation) on the production of the boundary tones, but no effect of condition was found for the pitch accent or the phrase accent.

Although one boundary tone (i.e. L%) may be used to indicate both contrasted meanings, speakers may also contrast Conversational Continuation and Cessation with the use of a high boundary tone and a low boundary tone, respectively. The results were largely unsurprising, as one universal meaning of intonation is that a rising pitch pattern indicates “open” or that there is “more to come” whereas falling pitch patterns indicates “closed”, and “no such meaning” (Cruttenden 1997:57). As noted in section 5.2.1, falling pitch at the phrase-final location has often been equated with Low boundary tones (L%) and rising pitch with High phrase-final targets (H%) (c.f. Ladd 1996: 82; this thesis Table 1.1). NZE speakers appear to use their phrase-final pitch target to signal conversational cues, whereby H% targets

indicate a conversational continuation request and L% targets indicate that the conversation should cease.

### 5.3.2. Authoritative vs. Submissive

As noted in Section 5.2.1, Ohala (1994) suggests that a universal meaning of pitch (e.g. the Frequency Code) is that low pitch indicates Authority and high pitch indicates Submission. Speakers may choose different tonal targets in order to convey such universal meanings. For instance, speakers may produce high targets, such as a high pitch accent or a high boundary tone, to indicate Submission rather than Authority, which may be signalled with a low pitch accent or a low boundary tone. The second meaning contrast investigated here – Authority vs. Submission - examines how NZE speakers use intonation to convey such meanings.

The test for marginal homogeneity illustrated no significant differences for tonal composition or any of the three individual targets. Table 5.2 outlines the distribution of the tonal compositions used by the NZE speakers when indicating Authority and Submission. As illustrated in Table 5.2, speakers preferred to produce one tonal composition, namely H\*L-L%, in order to convey both meanings.

Submission	H*H-L%	H*L-H%	H*L-L%	L*L-H%	L*L-L%	TOTAL
Authority						
H*L-L%	-	1	22	1	1	25
H*H-H%	-	-	1	-	-	1
L*L-H%	1	-	2	-	1	4
TOTAL	1	1	25	1	2	30

Table 5.2. Distribution of tonal compositions produced by 30 NZE speakers when conveying Authority vs. Submission.

The Frequency Code may also be signalled by absolute pitch target values rather than tonal target categories (H vs. L), as stipulated above. Although the two are related, since high tonal targets would often have higher absolute pitch values than low tonal targets produced by the same speaker, this is not always the case. For instance, speakers may raise the overall pitch level on an H\*L-L% in order to convey Submission and subsequently produce higher H\* and L% values than on an H\*L-L% that is produced with a lower pitch level in order to indicate Authority. In this example, the tonal composition of both would be identical if we were to adopt an intonation

labelling system such as ToBI. However, the absolute pitch target values may distinguish the two meanings. Thus, an analysis was made of the pitch heights of two tonal targets in order to determine whether absolute pitch height (in ERB) rather than category (e.g. High vs. Low) was used contrast Authority and Submission. Specifically, an ANOVA was performed on data from the 30 speakers in order to determine whether contrasting target heights, rather than category (i.e. H vs. L), were used to convey the two opposing meanings. The ANOVA had (1) the height of boundary tone, and (2) the height of the preceding pitch accent (in ERB) as the dependent variables. Considering the pitch accent target was not always easy to locate, the highest point preceding an L% or the lowest point preceding an H% was taken as the pitch accent value. The opposing meanings (2 levels: Authority and Submission) and Speaker Sex (2 levels: male and female) were the independent variables. As before, phrase accent was not assessed as a dependent variable because of the difficulty in locating such boundaries within the intonation phrase. Speaker Sex proved to be a significant factor in both Analyses; female speakers realised significantly higher pitch targets (both pitch accents and boundary tones) than their male counterparts. Such differences are unsurprising given the physiological differences between female and male speakers. The analyses did not show that meaning was a significant factor affecting the heights of the pitch accent or the boundary tone.

It is interesting that neither intonation category nor pitch height distinguished these two meanings, conflicting with assertions made about pitch height. Specifically, Ohala (1994) claimed that low pitch signals "largeness", which consequently conveys secondary meanings such as "threat, intention to prevail in a contest, dominance [and] self-sufficiency" (1994: 341), while high pitch conveys a sense of "smallness" and consequently conveys "nonthreat, submission, appeasement [and] desirous of goodwill..." (1994:341). The results of the two ANOVAs performed on the data illustrate that NZE speakers do not appear to use the height of their pitch accent or of the boundary tone to contrast Authority and Submission. In addition, the tests of homogeneity performed on the individual targets within the intonation phrase showed that speakers did not produce different

categories on specific tonal targets, such as the pitch accent, phrase accent or the boundary tone when indicating this meaning pair. The lack of significant results on individual targets may have been due to the communicative intention of the condition not being understood. The condition "Submission" could have been interpreted several ways. For instance, the speaker could have been attempting to convey an apology with their declined invitation to attend dinner: "I am *sorry*, but I cannot go to dinner". Alternatively, they may have intended to convey uncertainty, such as "I *may* have something else to do, so I cannot go to dinner".

It should also be noted that the overall analysis of pitch height did not examine whether the pitch height within categories, such as H\* or L\*, was different but rather compared the pitch height at two locations, regardless of H or L category. Such analyses may reduce the effects of possible intra-category distinctions made by the speakers. For instance, the speakers may have used higher H\* to indicate Submission whereas Authority may have been signalled with lower H\* targets. As I have used a transcription system adopted for another variety of English, namely AmE, I may have neglected a categorical distinction between levels of High pitch accents that may exist in NZE. An ANOVA was carried out on data from the preferred tonal category for both meanings, namely H\*L-L%, in order to determine whether there were tonal distinctions made within the categories or whether the intonation was indeed identical, as the previous analyses suggest. This analysis is therefore of data from 73% (22) of the speakers who realized this H\*L-L% pattern when conveying both Authority and Submission. The dependent variable for the analysis was the Pitch height (in ERB). The independent variables for the analysis were Pitch location (2 levels: the pitch accent and the boundary tone), Meaning (a within-subject factor with 2 levels: Authority and Submission) and Speaker Sex (a between-subject factor with 2 levels).

As the experiment reported here is exploratory and there were few data for this analysis (22 speakers), I will briefly comment on the general pattern of this result in order to examine possible tonal contrasts in NZE. The interaction of Meaning and Pitch Target Location turned out to be nearly significant ( $F_{1,20} = 10.167$ ,  $p < 0.01$ ); the speakers produced slightly higher pitch accents for Authority (5.402 ERB, s.e. 0.131) than for Submission

(5.224 ERB, s.e. 0.125), but lower boundary targets for Authority (4.090 ERB, s.e. 0.099) than for Submission (4.402 ERB, s.e. 0.129). There were no other significant interactions. Pitch Target Location ( $F_{1,20} = 48.346$ ,  $p < 0.01$ ) turned out to be nearly significant as a main effect. As expected, Speaker Sex proved to be a significant main effect, which again reflects the physiological differences between male and female speakers. Thus Speaker Sex as a main effect will not be discussed further in this chapter.

A further two ANOVAs were carried out on the data in order to confirm whether the height of the L% and the height of the preceding pitch accent were significantly different when indicating Authority and Submission. The dependent variables for each analysis were the (1) height of L%, and (2) the height of the preceding pitch accent (in ERB). The independent variables for both analyses were the same as described above. A comparison of the heights of L% and the preceding peak in these contours reveals that the height of the L%, but not that of H\*, was different when conveying Authority and Submission ( $F_{1,20} = 7.335$ ,  $p < 0.05$ ); the speakers realized significantly higher L% values when conveying Submission (4.402 ERB, s.e. 0.129) than Authority (4.090 ERB, s.e. 0.099). There were no significant interactions between the variables.

The difference in the height of L% may also indicate discourse completeness and finality. One meaning claimed for phrase final falls is that the utterance is a statement rather than a question, the latter being indicated by a rise (or at least high terminals rather than low terminals). Rather than simply conveying that the statement is complete, differing heights of L% may also convey differing degrees of finality (as claimed in traditional analyses by e.g. Crystal 1969). When signalling Authority, the speaker may also be indicating that their assertion is final and that they do not want to discuss the topic further, and so they are more likely to use a fall to a low pitch value. Alternatively, the speaker may not want to assert such finality if their intention is for instance to indicate a submissive refusal, and their pitch may therefore not fall to such a low value.

Although it is possible that the communicative intention of the PCT may not have been understood, I would nevertheless expect the speakers to convey tentativeness, uncertainty, or hesitancy when submissively refusing an

invitation to dinner (see Figure 5.1b). Although recent research has disputed the following association, the use of High Rising Terminal contours which occur on statements (HRTs) has been traditionally associated as markers of “tentativeness, uncertainty and hesitancy” (Warren and Britain 2000:164). However, it is interesting to note that it is not the phrase-final rise, generally occurring with a H% boundary tone (Ladd 1996:82), which is used to indicate a submissive refusal to an invitation, since only 10% of speakers produced H% to convey Submission, compared with 17% to convey Authority. The results illustrate that H% produced by NZE speakers on statements is not used to indicate meanings traditionally ascribed to this intonation use in some varieties of English, namely BrE (see Chapter 1), including tentativeness, uncertainty or hesitancy.

It should be noted that the H% produced on statements, which traditionally has been noted as a marker of tentativeness, uncertainty and hesitancy for other varieties of English (*ibid*), may be produced with a different shape than the H% described here. In this experiment, I adopted a framework developed for American English, namely ToBI, in order to describe the intonation of NZE. However, NZE may employ different tonal contrasts than American English. Warren (2005b: 355) examines such issues, and states that “An analysis which assumes (implicitly or explicitly) that categories are easily transferred between the varieties may lead to inadequate or erroneous conclusions”. At this point, we do not know whether the HRT is dissimilar in form to the H% described here. Such issues are explored further in Chapter 6.



### 5.3.3. Emphatic vs. Unemphatic

The next meaning contrast – Emphatic and Unemphatic – was examined to further explore how NZE speakers employ the Effort Code. Two PCTs examined how the same 30 NZE speakers indicate Emphasis or an absence of Emphasis on two different test items, and explored whether or not a negative assertion (i.e. *I did NOT like the film!*) on the response *no* is emphasised differently from a positive assertion (i.e. *I liked the film!*) on the response *yeah*. The two PCTs were illustrated above in Figures 5.1c and 5.1d.

On the response *no*, the test of marginal homogeneity showed a significant proportional difference of tonal composition when conveying Emphasis and an absence of Emphasis ( $p < 0.01$ ). Table 5.3 outlines the distribution of the tonal compositions used by the NZE speakers when indicating Emphasis vs. an absence of Emphasis on this response. What is observed in the matrix is that all differences can be explained by the use of the pitch accent. With consideration of the Bonferroni correction used to determine statistical significance ( $p < 0.002$ ), the 3 x 3 test of homogeneity performed on the pitch accent did not produce significant results ( $p = 0.011$ ). However, I will briefly comment on the general pattern of this result in order to examine possible tonal contrasts in NZE. Although the majority of speakers used a high pitch accent ( $H^*$ ) to indicate both meanings, this pitch accent was used by more speakers to convey Emphasis (77%) than an absence of Emphasis (63%).

Unemphatic	H*L-L%	L*L-H%	L*L-L%	TOTAL
Emphatic				
H*L-H%	1	-	-	1
H*L-L%	14	1	7	22
L*L-H%	1	-	2	3
L+H*L-L%	3	-	1	4
TOTAL	19	1	10	30

Table 5.3. Distribution of tonal compositions produced by 30 NZE speakers when conveying Emphasis vs. and absence of Emphasis on “no”.

Conversely, a low pitch accent ( $L^*$ ) was produced by more participants when conveying an absence of Emphasis (37%) than Emphasis (10%). Although both tonal targets may be used to convey both meanings, 9 speakers used

different tones for Emphasis and for an absence of Emphasis. Eight of these speakers used a high pitch accent when conveying Emphasis to their discourse partner, but L\* to show an absence of Emphasis, while there was one speaker who produced a low pitch accent to indicate Emphasis and a high pitch accent to indicate an absence of Emphasis. In addition, one pitch accent, namely L+H\*, was never produced to indicate an absence of Emphasis although 13% of speakers (4) used this pitch accent to indicate Emphasis. The intermediate and intonation phrase targets were found to be used similarly for both conditions ( $p = 1$ ).

Similar results were found for the productions of Emphatic vs. Unemphatic on the response *yeah* ( $p < 0.005$ ). Table 5.4 illustrates the tonal compositions produced by 30 NZE speakers when conveying these two meanings on this response.

Unemphatic	H*H-L%	H*L-L%	L*L-H%	L*L-L%	TOTAL
Emphatic					
H*L-H%	2	3	1	1	7
H*L-L%	2	7	-	6	15
L*H-L%	-	1	-	-	1
L+H*L-L%	-	4	-	3	7
TOTAL	4	15	1	10	30

Table 5.4. Distribution of tonal compositions produced by 30 NZE speakers when conveying Emphasis vs. an absence of Emphasis on "yeah".

The differences in proportions observed can be attributed to the types of pitch accent that were used. The 3 x 3 homogeneity test, which compared the pitch accents produced when indicating Emphasis and an absence of Emphasis on *yeah*, found near significant results ( $p = 0.003$ ). The preferred pitch accent for both meanings was a high pitch accent (H\*), which was produced by the majority of speakers when conveying an absence of Emphasis (63%) and Emphasis (73%). For 27% of the speakers (8), however, different pitch accents were used when contrasting these two meanings, with a low and a high pitch accent indicating an absence of Emphasis and Emphasis, respectively. Conversely, there was only one speaker who produced a low pitch accent to indicate Emphasis and a high pitch accent to indicate an absence of Emphasis. Although there is a similar number of speakers who contrast this meaning pair with their pitch accent on the two words, it should be noted that they are not the same 8 speakers. It is also interesting that 23%



of Emphatic utterances were produced with an L+H\* pitch accent, whereas no such target was employed to convey an absence of Emphasis. It should be noted again, however, that these results were only nearly significant in light of the Bonferroni correction. The intermediate phrase accent and the boundary tone were not used differently to contrast these two meanings.

The tonal productions on both responses (*yeah* and *no*) show a similar trend for pitch accent; although a high pitch accent (H\*) can be used to signal both meanings, a sizeable group of speakers use a high pitch accent (H\*) and a low pitch accent (L\*) to indicate Emphasis and an absence of Emphasis, respectively. In addition, L+H\* was used to signal Emphasis on both responses, but was never used to convey an absence of Emphasis. It should be noted here that L+H\* may be a substitute feature for higher pitch height; Gussenhoven (2004: 90) notes the theory that “peak delay may replace greater peak height”. Thus both higher pitch targets (i.e. H\*) and what some may refer to as later occurring high pitch targets (i.e. L+H\* or L\*+H) were used to indicate Emphasis. Both tonal manipulations produce a greater expenditure of Effort put forth by the speaker than lower or earlier respective pitch targets. Gussenhoven (2004: 85) claims that speakers use such pitch variations “to signal meanings that can be derived from the expenditure of Effort”, which he identifies as the “Effort Code”. The results presented in Tables 5.3 and 5.4 support such suppositions, illustrating how NZE speakers use a high pitch accent rather than a low pitch accent, in addition to L+H\*, to indicate Emphasis with the production of greater expenditure of Effort.

What is observed in the two matrices (Table 5.3 and 5.4) is that the pitch accent and the phrase accent appear to be markedly similar for both responses, *yeah* and *no*. However, it appeared that the speakers used marginally different boundary targets for the two responses; more speakers produced a phrase-final H% when conveying Emphasis on the lexeme *yeah* (23% of the speakers) than when conveying the same prompted meaning on *no* (10% of the speakers). This result may have been due to the speakers’ desire to indicate supplementary meanings, such as conversational continuation cues, and the contextual appropriateness of such meanings on

each test item. For instance, both responses were opinions, either indicating that the speaker did not like the film (i.e. *no*) or that the speaker did like the film (i.e. *yeah*). The speakers who produced H% on *yeah* may be using such a tonal target to indicate their "openness" (Cruttenden 1997) and a desire to continue the conversation and tell their discourse partner all about the film. Alternatively, the speaker may use L% to emphasise their dislike of the film by indicating finality in their negative opinion of the same film *no*. Therefore more speakers may choose to indicate openness with H% on a positive response, such as *yeah*, than a negative response, such as *no*. Such suppositions would necessitate further studies, as the results here merely indicate marginal differences in boundary tone preference for the two test items.

#### **5.3.4. Concerned vs. Unconcerned**

Next, I examine how speakers use their intonation to indicate Concern to their discourse partner, and how these same speakers indicate that they are Unconcerned. Two different responses, *yeah* and *no*, were used to assess this meaning contrast, illustrated in Figures 5.1e and 5.1f. On the response *no*, the test of marginal homogeneity showed a significant proportional difference of tonal composition when conveying Concern and an absence of Concern ( $p < 0.0003$ ), which can be explained by the different proportions in the pitch accent types ( $p < 0.002$ ) and the boundary tone types ( $p < 0.002$ ). Table 5.5 outlines the distribution of the tonal compositions when indicating Concern and an absence of Concern on the response *no*. As can be seen from the table, the most frequently produced boundary tone which was used to convey an absence of Concern was L% (77%) whereas Concern was signalled with H% by the majority of speakers (73%). In addition to the overall proportions, more than half of the speakers used different boundary tones to indicate these two meanings; specifically, 53% of speakers produced an L% to convey an absence of Concern and all of these speakers changed their target to an H% when conveying Concern on the response *no*. Thus, although both L% and H% may be used to indicate both meanings, many speakers make different choices for indicating Concern and an absence of Concern using a high boundary tone and a low boundary tone, respectively.

Unconcerned	H*H-L%	H*L-L%	L*H-L%	L*L-H%	L*L-L%	TOTAL
Concern						
H*L-H%	3	3	-	6	3	15
H*L-L%	-	1	-	1	-	2
L*L-H%	1	-	1	-	5	7
L+H*L-L%	-	3	-	-	3	6
TOTAL	4	7	1	7	11	30

Table 5.5. Distribution of tonal compositions produced by 30 NZE speakers when conveying Concern vs. an absence Concern on "no".

The proportional use of the pitch accent was also different when comparing realisations of these two meanings on the response *no*. The majority of speakers (63%) produced a low pitch accent (L\*) when conveying an absence of Concern, while a small majority (57%) of speakers signalled Concern with a high pitch accent (H\*). There was also an overall tendency of individual speakers to use different pitch accents for the two meanings; 58% of the speakers who produced H\* to convey Concern changed their pitch accent to L\* when conveying an absence of Concern. There was only 1 speaker who produced the opposite pattern, with an L\* to indicate Concern and an H\* to convey an absence of Concern. Similarly Concern was signalled with L+H\* by 6 speakers, all of whom changed to L\* or H\* to indicate an absence of Concern. There were no significant differences found for the phrase accent when indicating these two meanings.

Similar results were found on the response *yeah*, which showed a proportional difference on the productions of Concerned and Unconcerned ( $p < 0.001$ ). Table 5.6 summarises this distribution of tonal compositions.

Unconcerned	H*H-L%	H*L-L%	H*H-H%	L*H-L%	L*L-L%	TOTAL
Concern						
H*H-L%	-	-	-	-	1	1
H*L-H%	-	-	-	-	4	4
H*L-L%	1	1	-	-	1	3
L*L-H%	2	-	1	1	3	7
L*L-L%	-	1	-	-	-	1
L+H*L-L%	7	4	-	-	3	14
TOTAL	10	6	1	1	12	30

Table 5.6. Distribution of tonal compositions produced by 30 NZE speakers when conveying Concern vs. an absence of Concern on "yeah".

What can be seen from the table is that all differences can be explained by the proportions of two tonal targets. First, the boundary tone was proportionately different for the two meanings, which was confirmed by the 2 x 2 McNemar's test ( $p < 0.002$ ) performed on this specific tonal target. As evidenced in Table 5.6, the majority of speakers (29/30) preferred to produce an L% for an absence of Concern. In order to convey Concern, however, there was a more mixed result; 11 speakers produced an H% and 19 produced an L%.

The second tonal target that had a different distribution for these two was the pitch accent, as confirmed by the 3 x 3 test of homogeneity ( $p < 0.002$ ) performed on this target. Table 5.6 shows that L+H\* was the most frequently produced pitch accent to convey Concern (46%) in comparison to H\* (27%) and L\* (27%). However L+H\* was not produced by any speaker when conveying an absence of Concern. There were no significant differences found for the phrase accent.

In sum, the intonation patterns produced on both responses (*yeah* and *no*) show a similar trend for both pitch accent and boundary tone. In general, speakers appeared to use high pitch accents (i.e. H\*) and later occurring peaks (i.e. L+H\*) in order to convey Concern whereas these same speakers tended to use low pitch accents (L\*) or high pitch accents (H\*) to indicate an absence of Concern. Similar to the production of Emphasis (subsection 5.3.3), speakers may be producing higher pitch accents, such as H\* in comparison to L\*, and later pitch peaks, such as L+H\* rather than H\*, in order to signal extra articulatory effort. Consequently, speakers may indicate meanings which are derived from this added effort. One affective meaning of a greater pitch movement is "more surprised" (Gussenhoven 2004: 95) whereas "uninterested" (ibid: 88) is indicated with a more narrow pitch movement. NZE speakers may be employing the Effort Code by producing lower and earlier pitch peaks to indicate that they are uninterested, as well as an absence of concern for what the speaker had said (i.e. *I told off my boss at work today*). Conversely, speakers may produce a wider pitch movement to indicate interest in the conversation, their surprise in their discourse partner's assertion and subsequently their concern for the speaker.

The results also illustrate that NZE speakers use a high boundary tone more readily to indicate Concern than an absence of Concern, and that absence of Concern is almost exclusively marked by L%. Such tonal use may be used to signal conversational continuation and cessation requests in order to indicate an attitudinal meaning, such as concern for the speaker. In subsection 5.3.1, a high boundary tone was used to indicate conversational continuation whereas a low boundary tone was used to indicate cessation. The speakers may be adopting this form-function relationship in order to indicate that they are interested in hearing more from their discourse partner. Such meanings would subsequently imply the speaker's concern.

What is observed in Tables 5.5 and 5.6 is that the tonal preferences are different for the two lexical responses. It appears that the speakers preferred to produce a high pitch accent when conveying Concern on the test item *no* (57% of speakers) and more speakers preferred a L+H\* than any other pitch accent on the same prompted meaning on the response *yeah* (47% of speakers). Such preferential use may be due the compensatory nature of intonational meaning. That is to say that the intonation is produced as a meaningful device in a manner that complements the meaning of the lexeme which carries the tune. For instance, on the response *yeah*, the speakers are indicating that they already know that their discourse partner has "*told off the boss*" and thus any information which would follow is already known. In order to convey concern on such a response, speakers may need to compensate for a response which may bring the conversation to an end by conveying "*yes, I already know that you told off your boss today*". Conversely, the speakers may not need to convey so much concern on the *no*, as such a response indicates that the information is new to the discourse and as such, would not end the conversation. These results may be illustrating that L+H\* indicates greater degrees of concern than H\*. Similar assertions have been made concerning other research results; as argued in Gussenhoven (2004: 91), the results presented in Ladd and Morton (1997) show the "listeners apparently perceive late peaks as if they were higher". Therefore L+H\* may be used to indicate greater degrees of emphasis and concern than H\*. The results presented here suggest that intonation is a meaningful compensatory

tool for the lexeme; tonal devices may be used to convey meanings which supplement those meanings conveyed by the lexeme.

The majority of speakers also conveyed Concern with a high boundary tone on the response *no* (73% of speakers) and a low boundary tone on the response *yeah* (63% of speakers). Once again, the prompted response may have been used to indicate conversational continuation. As discussed above, the response *no* indicates that the information is new in the discourse (i.e. “*no, you did not tell me that you told off your boss today*”). Speakers may also want to indicate their desire to continue the conversation (i.e. “*please tell me more*”) with a conversational continuation marker, such as H%. The meaning conveyed with the response *yeah*, however, is such that the information is already known within the discourse (i.e.: “*yes, you have already told me that you told off the boss*”). Such conversational continuation cues may not be as appropriate when the discourse topic is already known.

#### **5.3.4. Impressed vs. Unimpressed**

Next, I examine how speakers use their intonation to convey an Impressed attitude and how these same speakers indicate that they are Unimpressed on two different lexical items (i.e. *no* and *yeah*). The PCTs were illustrated above in Figures 5.1g and 5.1h. The test of marginal homogeneity performed on the response *no* showed that the speakers produced different intonations to convey the two meanings ( $p = 0.003$ ). Although there were 14 speakers who produced identical pitch accents for both meanings, the test of marginal homogeneity performed on the individual targets identified the pitch accent as significantly different when indicating an Impressed and Unimpressed attitude ( $p = 0.0053$ ). Table 5.7 outlines the distribution of the tonal compositions when indicating an Impressed or an Unimpressed attitude on the response *no*.

Unimpressed	H*H-L%	H*L-L%	L*L-H%	L*L-L%	L+H*L-L%	TOTAL
Impressed						
H*L-H%	1	1	-	-	-	2
H*L-L%	-	-	2	-	-	2
L*L-H%	1	3	5	6	-	15
L+H*L-L%	-	5	2	3	1	11
TOTAL	2	9	9	9	1	30

Table 5.7. Distribution of tonal compositions produced by 30 NZE speakers when conveying an Impressed vs Unimpressed attitude on “no”.

The matrix above shows that the production of L+H\* is proportionally different for the two meanings. Specifically, 37% of the speakers (11) produced this pitch accent when conveying an Impressed attitude but only 1 of these speakers produced an L+H\* when conveying an Unimpressed attitude. The remaining 10 used either H\* or L\* when conveying the Unimpressed meaning. There were no significant differences found for the phrase accent or the intonational phrase boundary targets.

The test of marginal homogeneity performed on the response *yeah* showed similar results; it was found that speakers produced different intonations to convey the two meanings ( $p < 0.003$ ), which again can be explained by the proportional use of the pitch accent ( $p < 0.002$ ). Table 5.8 outlines the distribution of the tonal compositions when indicating an Impressed or Unimpressed attitude on the response *yeah*.

Unimpressed	H*H-L%	H*L-L%	L*L-H%	L*L-L%	TOTAL
Impressed					
H*L-H%	-	-	-	1	1
H*L-L%	-	-	1	-	1
L*L-H%	1	2	4	2	9
L+H*L-L%	3	4	2	10	19
TOTAL	4	6	7	13	30

Table 5.8. Distribution of tonal compositions produced by 30 NZE speakers when conveying an Impressed vs. Unimpressed attitude on “yeah”.

Table 5.8 illustrates that, once again, L+H\* was used to convey an Impressed attitude; 63% of speakers produced an L+H\* pitch accent when conveying an Impressed attitude and all of these speakers used a different accent, either



H\* (33%) or L\* (67%), when convey an Unimpressed attitude. There were no significant differences found for the phrase accent or the intonational phrase boundary targets.

Once again, the intonations produced on both responses (*yeah* and *no*) show a similar trend. What was observed in Tables 5.7 and 5.8 was that the difference in proportions could be explained by the pitch accent rather than phrase accent or intonational phrase boundary targets. Similar to the patterns discussed above for Concern (subsection 5.3.4), the data illustrate how NZE speakers produce higher and later pitch peaks in order to indicate an Impressed attitude with greater articulatory effort.

Tables 5.7 and 5.8 also show that the tonal preferences are different for the two lexical responses. Speakers preferred a low pitch accent when conveying an Impressed attitude on the test item *no* (50% of speakers), as evidenced in Table 5.7, and an L+H\* on the same prompted meaning on the response *yeah* (63% of speakers), as shown in Table 5.8. Once again, speakers may have used their intonation in order to compensate for the meanings conveyed by the response word. For instance, *yeah* indicates that the information is already known within the discourse, specifically conveying “yes, I already know that you told off your boss today”. Such a meaning may bring a conversation to an end as further information may be repeated information. In order to indicate an impressed attitude, speakers may need to employ intonation which conveys surprise, such as later occurring pitch accent peaks (i.e. L+H\*). Conversely, the response *no* indicates that the information is new to the conversation, and thus would not stop the storyteller from continuing their story about having *told off the boss*. That is, such lexical meanings do not require additional support, such as with the use of L+H\*, in order to convey that the speaker is impressed.

It is also interesting to note another similarity between the Concern vs. Unconcerned PCTs and the Impressed vs. Unimpressed PCTs; once again, the majority of speakers produced a phrase-final high target (H%) on the response *no* (57%) and conversely, a phrase-final low target on the response *yeah* (67%). Once again, speakers may also want to indicate their desire to continue the conversation (i.e. “please tell me more”) when the topic is new to



the discourse (i.e. on the response *no*) rather than indicating continuation cues when the discourse topic is already known (i.e. on the response *yeah*). A conversational continuation marker, such as H%, may be used to contrast such meanings. However, such suppositions would need to be explored further in supplementary studies.

5.4. Summary and Conclusions

Although these results merely illustrate how NZE speakers produce intonation in a laboratory setting, it is clear that NZE speakers may use different pitch accent and boundary tone types when indicating opposite meanings. First, the pitch accent was found to be significantly different in the comparisons of three pairs of meaning: Emphatic vs. Unemphatic, Concerned vs. Unconcerned and Impressed vs. Unimpressed. Table 5.9 summarises the distribution of the pitch accent when indicating Emphasis, Concern and an Impressed attitude, compared with an absence of these three meanings.

Absence of Meaning	L*	H*	L+H*	TOTAL
Meaning				
L*	29	14	-	43
H*	36	40	-	76
L*+H*	27	33	1	61
TOTAL	92	87	1	180

Table 5.9. Distribution of pitch accents produced by 30 NZE speakers when conveying Meaning (Emphatic, Concerned and Impressed) vs. an absence of the same meaning (Unemphatic, Unconcerned and Unimpressed).

In sum, speakers appeared to prefer a high pitch accent (H\*) or a low pitch target followed by a high target (L+H\*) when conveying Emphasis, Concern and an Impressed attitude whereas speakers appeared to prefer a low (L\*) or a high pitch accent (H\*) to convey an absence of these three meanings. These contrasts were said to be due to the Effort Code, since greater pitch movements may be produced in order to convey meanings derived from this added effort, which appears to include Emphasis, Concern and an Impressed attitude in NZE. Second, the speakers also employed variations in the boundary tone when distinguishing Continuation and Cessation and Concerned vs. Unconcerned. The use of the high boundary tone (H%)

appears to indicate “openness” and continuation whereas the low boundary target indicates cessation and finality.

The results found for both pitch accent and boundary tone were largely unsurprising as they illustrate how NZE speakers use two universal codes. First, the Effort Code was signalled with larger pitch excursions on the pitch accent, thus conveying more emphasis, concern and an impressed attitude. In addition, high boundary tones were used to convey “openness”. It is interesting to note, however, that H% was not used to indicate a submissive refusal to an invitation. Although not explicitly examined, this result suggests that H% is not used to signal secondary meanings such as uncertainty, hesitancy or tentativeness in this variety of English. In addition, speakers did not produce L% more frequently to convey an Authoritative refusal to the same invitation. As noted above, H% was rarely produced to indicate either meaning. The results of this PCT illustrate that NZE speakers do not contrast Submission with high boundary tones (H%) and Authority with low boundary tones (L%). This offers supporting evidence that H% produced on statements in NZE are not used to indicate meanings traditionally ascribed to the H% in other varieties of English. Such meanings include tentativeness, uncertainty or hesitancy. Such attitudinal associations to the phrase-final H% are further explored in Chapter 6.

Contextual appropriateness may also offer an explanation for the lack of H% in specific discourse genres, such as opinions (Guy, Horvath, Vonwiller, Daisley and Rogers 1986; Allan 1990; Britain 1992). It was noted that few speakers produced high boundary targets (i.e. H%) when asserting Emphasis and an absence of Emphasis on their opinion about whether or not they enjoyed a film (Figures 5.1c and 5.1d). Similar evidence was noted by scholars who investigated the meaning of phrase-final rises on statements; specifically scholars noted that High Rising Terminal contours are much less present in opinion-giving texts than narratives (see Guy et al 1986; Allan 1990; Britain 1992; Warren and Britain 2000). Such scholars have used the proportional distribution of HRTs as evidence against the claim that HRTs indicate “uncertainty” in Antipodean varieties. It may also be the case that speakers in these studies may have avoided such an intonational form in

opinion-giving texts because it may convey the speakers' uncertainty about their opinion, for instance, of whether they enjoyed the movie or not. Speakers may be able to assert themselves confidently when affirming their opinion and as such, opinion-giving text is not always an appropriate discourse genre for H%.

Also, the phrase accent was not found to be proportionally different in any of the meaning contrasts examined here. There are several descriptions of intonation which do not recognise the phrase accent, questioning its validity in intonational phonology (for German and English, see Grabe 1998; for Dutch, see Gussenhoven, Rietveld and Terken 1999). The phrase accent was proposed by Beckman and Pierrehumbert (1986) in order to describe different levels of intonational phrasing. In ToBI descriptions of intonation, such as used in this dissertation, the phrase accent is described as less complete than the intonational phrase boundary. It is also used to "up-step" a phrase final low target which produces a level pitch movement from a high pitch accent to the boundary tone (see Chapter 1 for a review). A phrase accent may be helpful in distinguishing tonal contrasts such as an Authoritative refusal to an invitation, with an L% which reaches the F0 minimum (e.g. H\*L-L%), and a Submissive refusal to the same invitation, which is conveyed with higher levels of L%, or an "upstepped" L% (e.g. H\*H-L%). ToBI uses the later notation (i.e. H\*H-L%) for level pitch movements from H\* to L% in American English. Rather than a level pitch movement from a high pitch accent to the boundary tone, NZE speakers may distinguish between Authority, which may be conveyed with a wider pitch movement from H\* to L%, and Submission, which may be signalled with a narrow pitch movement from these two tonal targets. For future investigations, it may prove useful to further explore the necessity of the phrase accent, and the possible categorical distinction between a wide pitch movement from H\* to L% (H\*L-L%) and a more narrow pitch movement between these two tonal targets (H\*H-L%) in NZE.

Intonologists quickly discover that it is difficult to separate the many-layered meanings of intonation. PCTs do not evade such issues. For example, NZE speakers may use H% to signal conversational continuation

and at the same time to imply other meanings, such as the speaker's interest in the conversation as well as concern for their discourse partner. Conversely, L% may signal a lack of interest and conversational cessation, which may imply a speaker's finality in his opinion. As shown in the PCTs above, this methodology does not unveil all of the meaningful tonal contrasts used by speakers of a language variety. The PCT results illustrate possible contrasts used in NZE intonation and illustrate how tonal techniques may be used to convey a number of different meanings in this variety of English. As such, the conclusions derived from the results would necessitate further examination. Variations in meaning of such tonal categories are explored further in Chapters 6 and 7.

### Categories and meanings conveyed by variation in boundary tone height

---

#### 6.0. Chapter abstract

The previous chapter reported that boundary tone height was important in distinguishing between Conversational Continuation vs. Cessation and Concerned vs. Unconcerned. In addition, it was also found that there may be a categorical distinction between levels of L tone when NZE speakers contrast Submission and Authority. The aim of the experiment reported in this chapter was to determine whether the height of the boundary tone differentiates categories of meaning in New Zealand English. With this in mind, two experiments were carried out. The first was a perception experiment using a discrete-choice task to test for the presence of categorical boundaries within the range of boundary tone height. 21 test stimuli were created from the same source utterance and each was played in conjunction with a second test stimulus with a boundary target that was 0.50 ERB higher. 36 listeners were asked to specify whether the two stimuli were the same or different. The results indicate that there is at least one categorical boundary, which is determined by boundary height. It is argued that different pitch levels of the boundary height can be categorized as either L% or H% in NZE.

Second, a gradient-response task was used to further explore how three separate meanings are interpreted according to the height of the boundary tone. 10 test stimuli with different boundary heights were created from a single utterance and played in conjunction with an anchor stimulus. 37 participants were instructed to judge each stimulus against the Anchor stimulus according to three different meaning scales: the comparative levels of LIVELINESS, CONFIDENCE and INSISTENCE. The participants showed a general tendency to interpret stimuli with higher boundaries as livelier, less confident and less insistent. The results are interpreted as demonstrating that NZE listeners employ the Frequency Code (Ohala 1994) when interpreting

attitudinal meaning conveyed by the height of the boundary tone. Specifically, listeners' responses show a linear relationship with the height of the boundary tone, with the degree of perceived insistence and confidence correlating negatively with the boundary, and the degree of perceived liveliness correlating positively with the same boundary.

## 6.1. Introduction

As discussed in Chapter 2, speakers of different languages and language varieties often employ both universal and language-specific meanings when using intonation as a meaningful linguistic device. For example, one universal form-function distinction is between a phrase-final rise, which often signals an interrogative and a phrase-final fall, which frequently identifies a statement (see Chapter 2, Section 2.2.2). Language-specific meanings have also been associated with the phrase-final rising intonation. For RP, for instance, O'Connor and Arnold (1961) describe several attitudes conveyed by rising phrase-final pitch movements when used on declarative statements; they assert that different types of rises on statements communicate a sense of "incompleteness", are "used for continuative purposes" (1961:50) and indicate "tentative[ness]" (1961:57). O'Connor and Arnold also divide up the meanings associated with rises according to the range of pitch. For instance, a low rise, which has a low F0 onset which rises to its completion before the middle of the speaker's range, invites further conversation, as well as being "...guarded, reserving judgment, reprovably critical, resentful, deprecatory" (1961:169). A second type of phrase-final rise begins in the middle of the speaker's range and rises until its completion point, or a *high rise*. As well as being incomplete, O'Connor and Arnold claim the high rise is "questioning, trying to elicit a repetition, but lacking any suggestion of disapproval or puzzlement; [and] (in non-final [intonation phrases]) tentative", (1961:210).

It is clear that NZE phrase-final high targets may convey meanings other than those ascribed to phrase-final rises in RP, or those meanings attributed as "universal". Phrase-final intonation, notably the High Rising Terminal (HRT), may have an additional form and function in NZE from those rises noted above; as noted in Chapter 1, HRTs function as "positive politeness

markers" (Meyerhoff 1991) rather than indicators of uncertainty: "our analyses showed that those structural elements that function strongly to appeal to hearers' understanding and empathy were those most likely to carry HRTs" (Warren and Britain 2000: 170).

It is unclear, however, whether the NZE HRT is somehow "different" from other contours with high boundary targets (H%). One possibility is that high H%, such as an HRT, and low H% may be meaningfully distinct in NZE, just as a low rise and a high rise are claimed to be meaningfully distinct in BrE. It may be that NZE speakers employ high H% to assess listener comprehension, for example, whereas low H% may indicate uncertainty and a lack of confidence. A further unresearched question is whether NZE speakers use different levels of low boundary tones (L%) to convey unique categories of meaning, as suggested in the previous chapter. Thus the aim of the two experiments reported in this chapter is to determine whether the height of the boundary tone differentiates unique categories of meaning in NZE. Two experimental tasks were carried out for this purpose.

## **6.2. Discrete-choice task**

### **6.2.1. Materials**

The source utterance used in this study was taken from a male speaker of New Zealand English (Appendix C2). He was instructed to read aloud from a provided script, "as naturally as possible" whilst being recorded onto a Maxell 74 Minidisk using a Sony MDS-JB930 minidisk recorder and a Sony unidirectional microphone. The recorded speech material was later resampled using Praat (Boersma and Weenink 1996), with a sampling rate of 16 kHz in mono. The speaker was recorded in a quiet room at Victoria University of Wellington.

The utterance *Manny won in rummy* was selected as the source utterance in this perception experiment for several reasons. As discussed in Chapter 3, the segmental makeup of an utterance affects the F0. An utterance consisting entirely of voiced sonorants was selected in order to avoid such complications. In addition, the source utterance was selected because it could potentially convey each of the three attitudinal meanings being



assessed in the second part of this study (see Section 6.3.3). For example, speakers may convey different degrees of LIVELINESS on the sentence *Manny won in rummy*, as the speaker may want to show excitement (or lack of excitement) that Manny had won the card game. A speaker's CONFIDENCE may also be conveyed through intonation; a speaker may appear both confident and unconfident in his assertion that Manny had won in rummy. A speaker's intonation may also be employed in order to convey INSISTENCE, asserting his certainty that Manny has indeed won in rummy. Finally, the utterance is durationally short (1.11 sec), limiting the total duration of the experiment and the subsequent demand on the participants' time and concentration.

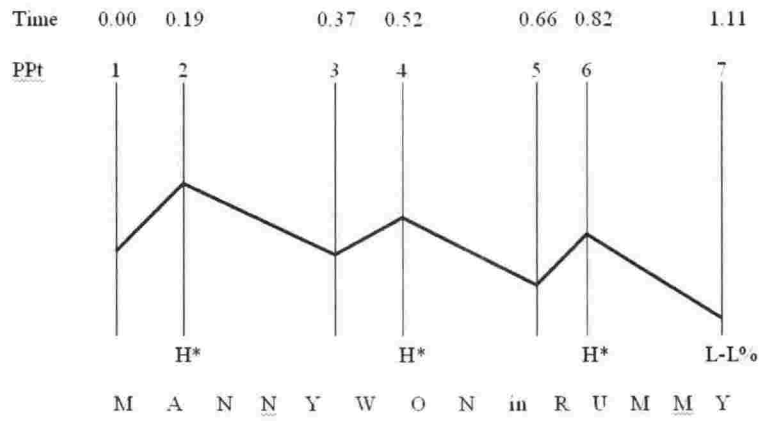


Figure 6.1: Schematic representation of source utterance (Time = sec, PPt = Pitch Point).

Using Praat, the pitch contour of the recorded utterance was stylised and seven pitch points (PPt) were located at tonal movement positions (PPts 2-6, Figure 6.1) and phrase boundary locations (PPts 1 and 7, Figure 6.1). The average pitch level for the recorded utterance was 3.73 ERB, which is similar to the average pitch levels of 13 NZE male speakers reading aloud a story (mean 3.82, s.d. 0.48, see Chapter 4). Following Ladd and Morton (1997: 324), a male voice was used rather than a female voice. Ladd and Morton wanted to avoid the possibility of having to explain unanticipated problems due to the sex of the speaker. As the primary focus of the current thesis is not an examination of social variation within NZE, a similar approach to that



used in Ladd and Morton (1997) was used here. The pitch values at each of the seven tonal positions for the source utterance are provided in Table 6.1.

	PPt 1	PPt 2	PPt 3	PPt 4	PPt 5	PPt 6	PPt 7
ERB	4.19	4.67	4.00	4.30	3.77	4.09	3.09

Table 6.1: Pitch Point (PPt) values (in ERB) for the source utterance

In order to examine possible boundaries at the IP-Final boundary position, the height of the boundary tone located at PPt 7 (Figure 6.1 and Table 6.1) was manipulated. The height of the boundary target was adjusted to one of eleven pitch levels. Figure 6.2 illustrates the manipulations carried out on the source utterance.

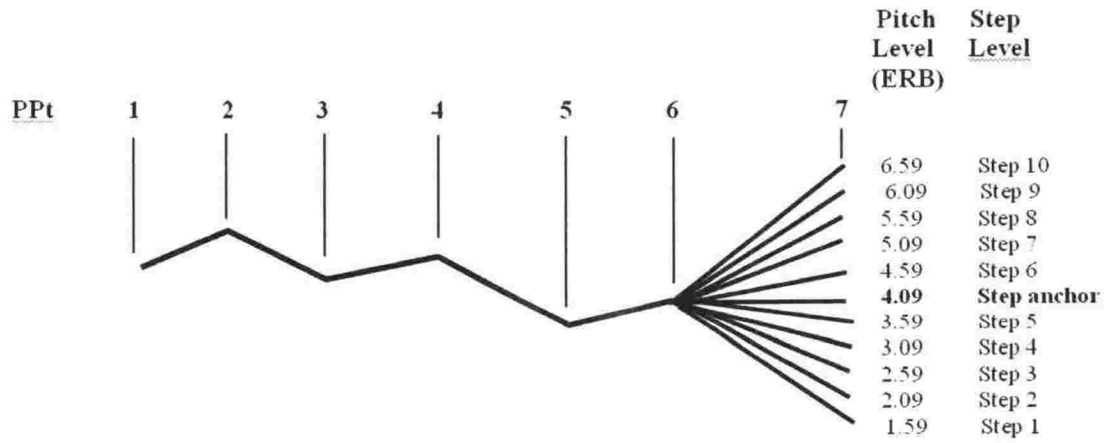


Figure 6.2: Schematic representation of manipulations carried out on source utterance, (PPt = Pitch Point).

The Anchor stimulus, referred to as Step Anchor in Figure 6.2, was a level pitch movement from H\* (PPt 6, Figure 6.1) onto the IP-Final boundary target (PPt 7, Figure 6.1). Five lower targets (Step 1 to Step 5 in Figure 6.2) and five higher targets (Step 6 to Step 10) were adjusted relative to the Anchor stimulus pitch level (4.09 ERB, Figure 6.2). The difference between each adjacent step was 0.50 ERB. The motivation for adopting pitch targets in relation to the boundary tone target on the Anchor stimulus was theoretical, and is primarily of relevance to the second task (see section 6.3).

A pilot study, which was performed in a Linguistics class at Victoria University, helped establish the step size and the experimental procedure. The primary aim of the pilot study was to determine whether naïve listeners could distinguish a step size of 0.40 ERB, which was determined as audible by the author, or whether a larger step size (e.g. 0.50 ERB) was required. The pilot study revealed that a step size of 0.40 ERB at IP-Final boundary was nearly undifferentiable to the participants; the students remarked that members of each stimulus pair sounded alike while other students noted that they indicated that the stimuli were “different” even though they believed the stimuli sounded alike. These results were surprising in light of Ladd and Morton’s assertions based on a pilot study: “two-step pairs (with peak levels 12 Hz apart) might be too easily discriminated” (1997:325). Although Ladd and Morton did not specify the exact pitch range (in ERB) of the utterance used in their pilot study, it appears unlikely that a step size as large as 0.40 ERB was used; a difference in 12 HZ may represent, for example, 0.156 on an ERB scale at the high end of the voice pitch range (e.g. from 400 to 412 Hz), or 0.39 ERB at an exceedingly low voice pitch range (e.g. 52 to 64 Hz). There may be several explanations for this discrepancy in discernable pitch levels. First and foremost, the pitch steps were adjusted at the boundary position in this experiment rather than a high pitch accent ( $H^*$ ), as done in Ladd and Morton’s experiment. It may be easier to distinguish pitch variation at a perceptually more salient stressed point (such as the pitch accent) than an unstressed point (such as the boundary tone). Other possible explanations concern the experimental procedure. Specifically, the pilot study stimuli here were played over a loudspeaker rather than individual headsets. Based on the evidence of the pilot study, steps of 0.50 ERB for boundary targets were adopted for this study, which were played over individual headsets.

Pairs of stimuli, each with an interstimulus interval of 0.35 seconds, were used as the test stimulus pairs. For these discrete-choice task stimulus pairs, each stimulus step (Figure 6.2) was combined with a higher step (e.g. Step 1 was combined with Step 2) forming the test stimulus pairs (following Ladd and Morton 1997). In addition, pairs of identical stimuli formed the control stimulus pairs (following Ladd and Morton 1997). The stimulus

combinations, resulting in 21 pairs of stimuli, formed the discrete-choice task. It should be noted that reverse pairs of stimuli, which combine a higher boundary target stimulus with a lower boundary target stimulus (i.e. Step 2 + Step 1), were not created and tested. This was because Ladd and Morton found (1997:331) that "it appears that the subject cannot discriminate stimulus pairs presented in BA [higher-then-lower] order....". They posited that the order of presentation is important when differentiating pitch levels as "different" due to the effects of declination, which mean that a later pitch peak that is physically lower in F0 than an earlier pitch peak may nevertheless sound at least as high as the earlier peak (Gussenhoven and Rietveld 1994). Thus listeners may interpret a stimulus pair which combines a stimulus with a higher boundary target and a stimulus with a pitch target lower in pitch level as equal in pitch target height.

### **6.2.2. Participants**

36 participants took part in this experiment on a voluntary basis. Social variation was assessed using a questionnaire (Appendix C5). 16 were male and the other twenty were female. All were students at Victoria University of Wellington. 27 participants were aged between 18-22 years, seven were 23-27 years old and two were 28-35. All of the participants claimed to be native New Zealand English speakers, and none reported having any known hearing impairments. This analysis examined listener sex as the only between-subject factor. There were not enough speakers to analyse age as a factor. The participants represented a single native language category (NZE) and occupation (student).

### **6.2.3. Procedure**

One task typically employed in categorical perception experiments asks listeners to indicate whether pairs of stimuli are identical or different. Discrimination between pairs of stimuli is predicted to be better if the stimuli are perceived to be in different categories than if the stimuli are interpreted as within the same category (Ladd and Morton 1997:12). In order to determine whether or not NZE listeners differentiate pitch levels at the boundary as

categorically distinct, participants were asked to tick one of two boxes, indicating whether the pairs of sentences were the “Same” or “Different”.

In addition to providing instructions for each task, the introduction identified the speaker of the test stimuli as a “Kiwi” man. Prior to every eighth stimulus, a 0.5 second beep was inserted in order for the participants to determine whether they were keeping pace with the playback. Each stimulus pair was separated by 3.5 seconds. The pilot study confirmed that participants were able to complete the task using this interstimulus interval (ISI). The first five stimuli were played a second time within the experiment and the answers given for the first set were discarded from the included data; this was done to discount erroneous answers often made at the beginning of experiments.

The introduction (Appendix C3), which was prerecorded using the author’s voice, and stimuli were played over individual headsets in an audio classroom in the Language Learning Centre, Victoria University. Groups of students took part in a single listening exercise. The participants were provided with a \$10 book/music voucher as recompense for their time.

#### **6.2.4. Results and Discussion**

Figure 6.3 illustrates the percentage of listeners who correctly identified each of the 10 test stimulus pairs as “different”. The number of participants who indicate an accurate response (i.e.: indicating that pairs of different stimuli are “different”) increases as stimulus pairs become closer in pitch height to the Step Anchor stimulus. Nearly all listeners (91.89%) correctly identified the Step 5-Anchor stimulus pair as “different”, and 83.78% correctly identified the Step Anchor-Step 6 stimulus pair as “different”. It is unclear at this point why there is a sharp decrease in discriminability between the Anchor stimulus - Step 6 stimulus pair and the Step 6-Step 7 stimulus pair. It may be due to the fact that the trajectory of the pitch for the later stimulus pair was more similar (both rising from the H\* to the IP-Final target) and therefore, less easy to discriminate than the former stimulus pair, which may have been perceived as less similar (a level pitch movement as opposed to a rising pitch movement from the H\* to the IP-Final target). However, such assumptions cannot be confirmed using the given data as I did not ascertain whether or not the

participants perceived the stimuli as rising or falling pitch movements at this location.

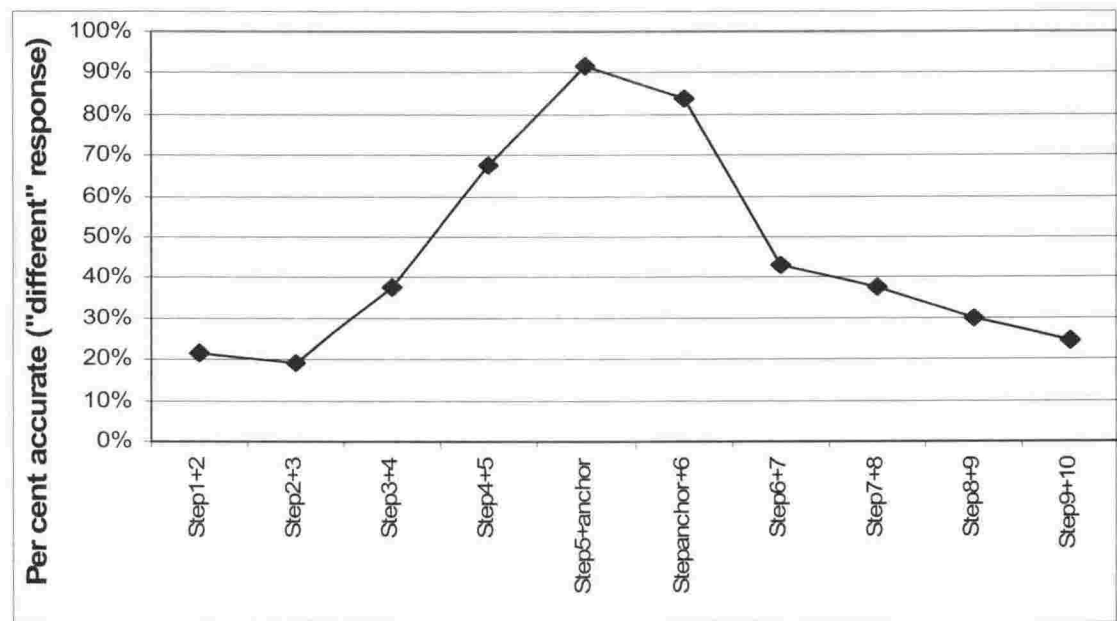


Figure 6.3.: Discrete-choice response, illustrating the percentage of participants who correctly identified the 10 Test-Stimulus pairs as "different".

The results indicate that there is at least one categorical boundary which is determined by boundary tone height in this experiment and categorised by the pitch movement to the boundary tone from the preceding H\* value used in this experiment. The phrase-final targets in these stimuli would be labelled as either L% (Step 5) or H% (Step 6), using the ToBI labelling system (Beckman and Ayers 1994; Pierrehumbert 1980). The results also illustrate that the level pitch movement from H\* to % (i.e. Step Anchor) would be categorised along with Steps 6 - 10, on the basis that the Step 5 - Step Anchor stimulus pair is more discriminable than the Step Anchor – Step 6 stimulus pair. Therefore the results indicate that a two-level phrase-final categorisation, such as that used in ToBI, may be accurate for NZE. It should be noted, however, the level pitch movement from H\* to the boundary tone would be categorised with the H% stimuli rather than as an upstepped L% (i.e. it would not appear to be "H-L%").

However, it should be noted here that such a two tone (high vs. low) dichotomy at this location cannot be confirmed by the results here. This is due to the high level of distinguishability of stimuli which are shallow in pitch movement from H\* to the boundary tone (Step 4+Step5 and StepAnchor + Step 6 stimulus pairs). It may be the case that an additional categorical distinction exists within the shallow pitch movements at this location.

The results also suggest that higher and lower H% are not in different categories in NZE, as previously hypothesised in this chapter. That is, within what we might now characterise as the H% range, it appears to be the case that the higher the boundary target, the smaller the number of participants who recognise the test stimulus pair as “different”. The finding that there is not a second peak of discrimination within this H% range strongly suggests that a raised H% does not indicate a separate category in NZE.

It should also be noted that, as claimed by O'Connor and Arnold (1961), a “low rise” (i.e. commencing from a low pitch point in the speaker's pitch range) and a “high rise” (i.e. beginning in the middle of the speaker's pitch range) may still be phonologically distinct in NZE. That is to say, the height of the nuclear pitch accent, such as the H\* used here, may differentiate categories of meaning in NZE which are similar to those ascribed by O'Connor and Arnold (see section 6.1). As the height of the pitch accents were not examined here, supplementary studies would be necessary to determine such a supposition.

Figure 6.4 illustrates the percentage of listeners who incorrectly identified the 11 identical control stimulus pairs as “different”. Note the peak in errors across the stimulus pairs Step Anchor+Anchor and Step 6+6; 36.11% and 33.33% of listeners, respectively, believed that the stimuli in these two pairs of identical items were different. The general tendency for such responses (i.e.: incorrect identification of identical stimulus pairs as “different”) decreases as the Step levels become increasingly distant (higher and lower) from the H\* at the IP-Final boundary position. As noted by Ladd and Morton (1997: 328), categorical perception studies show an increase in

the number of “false alarms” at a categorical boundary. This is therefore further evidence that there may be a categorical boundary at this location.

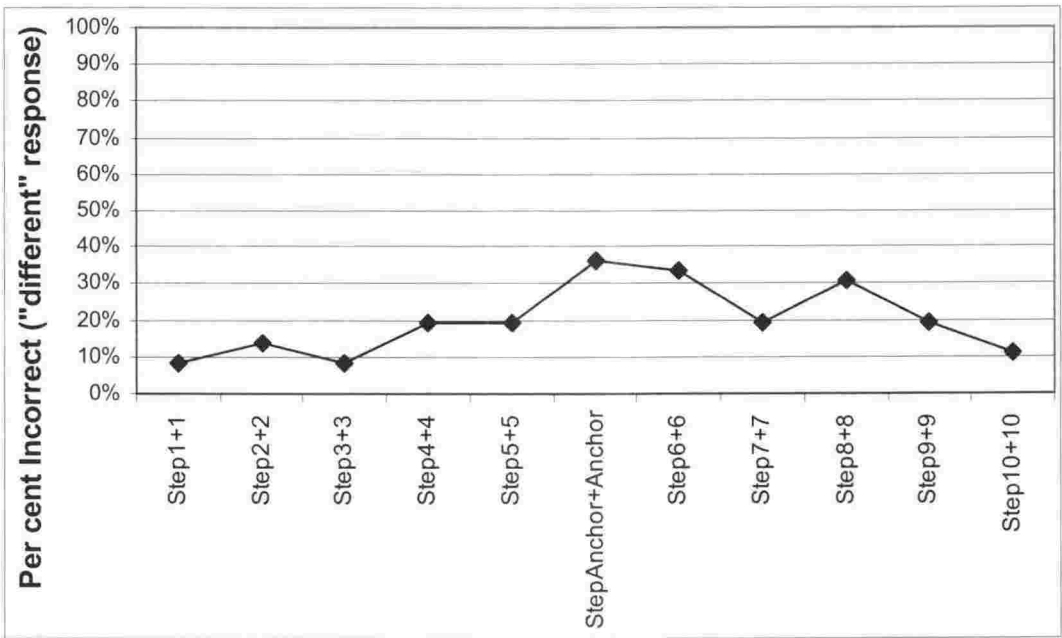


Figure 6.4. Discrete-choice response, illustrating the percentage of participants who incorrectly identified the 11 control stimulus pairs as “different”.

### 6.3. Gradient-response task

#### 6.3.1. Materials

The stimuli used for the discrete-choice task reported above (Section 6.2) were also used for the gradient-response task (see Figure 6.2). However, different stimulus pairs were used in the two tasks. In the discrete-choice task, pairs of stimuli with adjacent step levels were played in conjunction with one another in order to examine where the level of discriminability increased and thus, where a categorical boundary may exist. The aim of the gradient response task, however, was to examine how a range of boundary targets convey different meanings. Thus, the stimulus pairs for the gradient-response task were combinations of one stimulus, namely the Anchor stimulus (referred to as Step Anchor in Figure 6.2), with the stimulus from each Step (Steps 1 - Step 10, Figure 6.2).

The motivation for the use of a level pitch contour as the anchor was based on common assertions about the intonational phonology of English. As



noted above, scholars have traditionally categorised the boundary tone into two groups. Researchers who adopt the traditional British approach to intonation, for example, often categorise phrase final intonation into falling pitch patterns vs. rising pitch contours (c.f. O'Connor and Arnold 1961) whereas a similar dichotomy is often used in Autosegmental Metrical approaches, such as ToBI, which categorizes the boundary tone as either low (L%) or high (H%), (c.f. Beckman and Ayers 1994). As noted earlier, L% does not always equate to a falling pitch movement. For instance, an H\*H-L% tonal composition would represent a level pitch movement at the nuclear position. This experiment explores whether a categorical distinction exists between falling and rising movements (here corresponding to L% and H% boundary tones) when interpreting three meanings. It does this by comparing meanings ascribed to various falling pitch movements (i.e. movements to L% targets) and rising pitch movements (to H% targets) when compared to a level pitch movement from H\* to the boundary tone.

Another motivation for selecting the level pitch from H\* to % as the Anchor stimulus was to examine whether or not phrase-final pitch width (regardless of boundary tone) is a factor in interpreting the three assessed meanings. The Anchor stimulus, with equal pitch target heights on H\* and the IP-Final boundary position, was directly compared with the test stimuli, which produced subsequent steps of wider pitch widths from H\* to each of the L% and H% targets. Such a direct comparison permitted a 2 by 5 Analysis of Variance, in which the two independent factors were boundary tone (2 levels, H% and L%) and pitch widths (5 levels, either above (H%) or below (L%) the H\*).

A further 11 filler stimuli were created from the source utterance in an attempt to disguise the investigative aim of the experiment. Thus the pitch level and pitch movement position of tonal targets other than those under examination were adjusted (PPTs 1, 2, 3, 4, 5 and 6, Figure 6.2). The filler and test stimuli produced a total of 21 stimulus pairs, each of which consisted of a combination of the anchor stimulus with a different stimulus. Stimulus pairs were separated by an interval of 3.5 seconds.



### 6.3.2. Participants

The same listeners participated in the gradient-response task as the discrete-choice task (see Section 6.2.2).

### 6.3.3. Procedure

This task started with an introduction, which was prerecorded by the author, as in the first task (Appendix C4). A second task typically employed in traditional categorical perception experiments asks listeners to assign stimuli from a phonetic continuum to one of two categories. In intonation studies, such categories are aligned with meaning. For instance, listeners may be asked to assign stimuli from a phonetic continuum, such as pitch height, to categories such as emphatic vs. not emphatic (Ladd and Morton 1997). A task involving only one meaning was difficult to apply to this experiment due to the number of meanings associated with H% and L% in English. For instance, H% has been claimed to convey conversational continuation (see Chapter 2 for discussion), a speaker's uncertainty (Lakoff 1975) and an indication that a question has been asked (see Chapter 2 for discussion), whereas conversational finality and a speaker's certainty are identified with L%. If the perception task were limited to assessing one meaning, it could not be determined whether or not other meanings are categorised into either H% or L% in NZE (as hypothesised in Section 6.1). Instead, three meanings were selected for this assessment. The perceived levels of CONFIDENCE, LIVELINESS and INSISTENCE were examined in order to ascertain whether or not NZE listeners interpret the same meanings on rising vs. falling phrase final stimuli as have been ascribed to these tonal contrasts in another variety of English (e.g. SBrE, c.f. O'Connor and Arnold 1961). It should be noted that intonation may convey several meanings simultaneously. For instance, varying degrees of liveliness, confidence and insistence may be conveyed by the same intonation pattern. In this task, I wanted to explore how these three meanings are interpreted with respect to differences in the pitch height of the boundary target.

In addition to several meanings being assessed rather than one, this experiment differed from typical experiments on the interpretation of intonation

in that a gradient response task was used rather than a forced choice task. The aim of this type of task was to determine whether these three meanings are interpreted along a gradient scale or whether these meanings are more obviously aligned into different categories (for instance, corresponding to the distinction between L% and H% demonstrated above). The three meanings in the gradient-response task were assessed separately. For example, the participants were told to mark on the given scale "how lively the second sentence sounded in comparison to the first sentence" with the left side of the scale indicating "much less lively" and the right side of the scale indicating "much more lively", and with the mid-point denoting that the test stimulus conveys the "same amount" of liveliness as the Anchor stimulus. The degree of INSISTENCE and the level of CONFIDENCE were assessed using similar scales (Appendix C5).

The pairs of stimuli were randomised in three separate orders so as to reduce the interpretational effects that preceding pairs of stimuli may have on subsequent pairs of stimuli. It should also be noted that each stimulus pair was judged once according to each meaning, allowing participants to indicate their perceptions of one meaning at a time for each stimulus pair. Prior to every fifth stimulus, a 0.5 second beep was inserted in order for the participants to determine whether they were keeping pace with the presentation. The first five stimuli were played a second time within the experiment and the answers given for the first set were discarded from the included data; this was done to discount erroneous answers often made at the beginning of experiments.

For the gradient-response task, scores were measured to the nearest whole millimetre on the meaning scales, measuring 100 mm in length. Measurements were taken as the centre point on any marks which were a millimetre or so wide. Scores ranged from 0 (for "much less lively", "much less confident" and "much less insistent") to 100 (for "much more lively", "much more confident" and "much more insistent"), with 50 denoting that the second stimulus conveys the "same amount" of the assessed meaning as the Anchor stimulus. These scores were then entered into three separate

Analyses of Variance (repeated measures design, Huyhn-Feldt corrected), one for each meaning: LIVELINESS, CONFIDENCE and INSISTENCE. In each analysis, there was one between subject factor: Listener Sex (2 levels, male and female) and two within subject factors: Boundary Tone (2 levels, H% and L%) and Pitch Width from H\* to the boundary target (5 pitch levels, differing in subsequent steps of 0.50 ERB). Thus the analysis paired responses for comparisons which were equal in Pitch Width from the H\* to %, such as Step Level 1 (to a L% target) and Step Level 10 (to a H% target) which represented a pitch width of 2.50 ERB (see Figure 6.2). The results are provided in the following section.

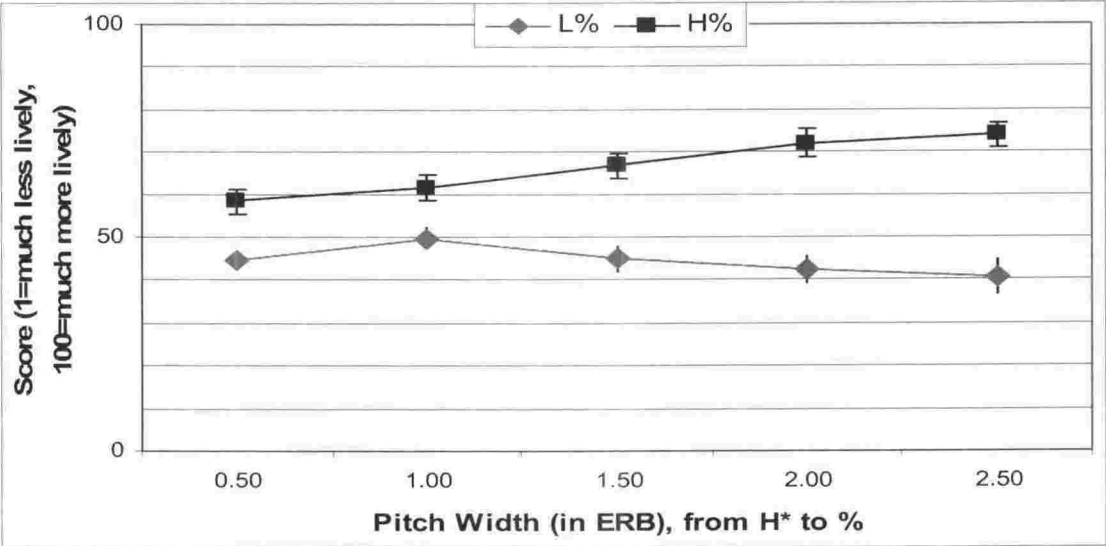
### **6.3.4. Results**

#### **6.3.4.1. Liveliness**

The interaction of Pitch Width and Boundary Tone turned out to be highly significant for judgements of LIVELINESS ( $F_{4,136} = 8.191$ ,  $p < 0.01$ ). For high boundary tones, wider pitch movements from H\* to H% were interpreted as more lively, while for low boundary tones, narrower movements from H\* to L% were interpreted as more lively, as shown in Figure 6.5. There were no other significant interactions. The type of Boundary Tone turned out to be significant as a main effect ( $F_{1,34} = 41.424$ ,  $p < 0.01$ ) but Pitch Width was not a significant main effect.

On the basis of the significant interaction, and in order to assess differences between the 5 pitch levels within each category (i.e. within H% and L%), two separate Analyses of Variance were carried out for H% stimuli and for L% stimuli. In each analysis, there was one between subject factor: Listener Sex (2 levels, male and female) and one within subject factor: Pitch Level (5 levels). There was a significant effect of Pitch Level within H% ( $F_{4,136} = 9.098$ ,  $p < 0.01$ ), with higher H% targets indicating more liveliness. On the other hand, there was no significant effect of Pitch Level within L%, although there was a linear relationship across the five pitch levels ( $F_{1,34} = 5.504$ ,  $p < 0.05$ ), illustrating a positive correlation between pitch level and degree of perceived liveliness. No quadratic relationship was found. That is to say, the

participants' responses for any of the five steps were not significantly skewed, or out of line with the other four pitch levels within L%. A similar linear effect was found for pitch level within H% ( $F_{1,34} = 32.654$ ,  $p < 0.01$ ). There were no significant interactions between the variables.



H%	Pitch width from H* to H%	0.50 ERB	1.00 ERB	1.50 ERB	2.00 ERB	2.50 ERB
	Mean	58.32	61.62	66.76	72.00	73.95
	Standard Error	2.16	2.77	2.94	2.84	3.42
L%	Pitch width from H* to L%	0.50 ERB	1.00 ERB	1.50 ERB	2.00 ERB	2.50 ERB
	Mean	44.65	49.73	45.03	42.49	40.69
	Standard Error	2.16	2.31	2.94	3.23	3.98

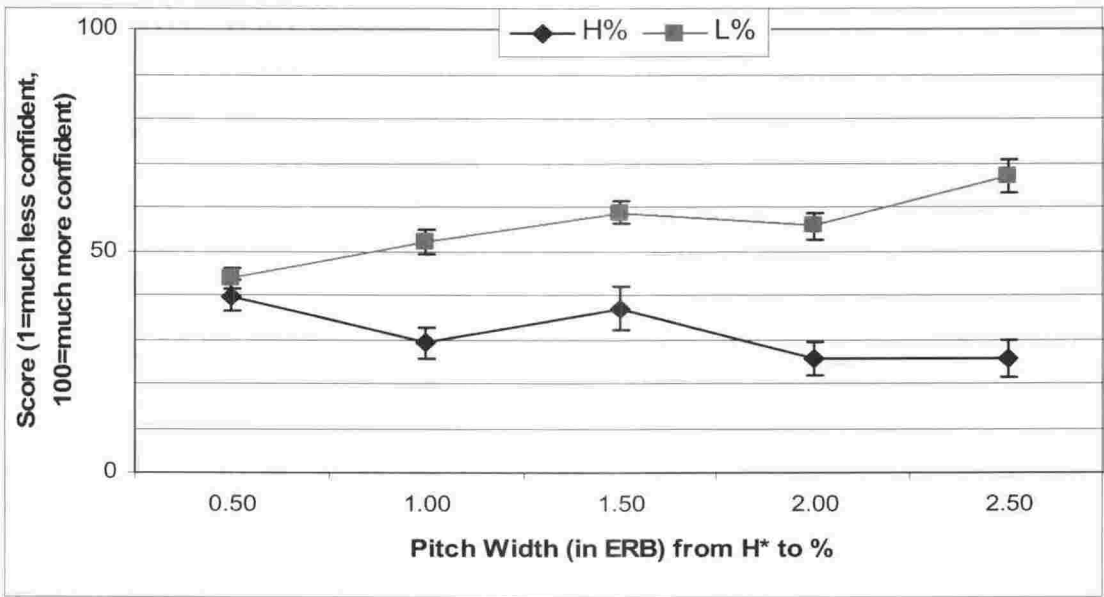
Figure 6.5.: Average response score illustrating the degree of "Liveliness" across 2 boundary tones (L% and H%) and 5 pitch widths from H\* to %.

### 6.3.4.2. Confidence

or Pitch Width and Boundary Tone turned out to be highly significant for the degree of perceived CONFIDENCE ( $F_{4,136} = 7.971$ ,  $p < 0.01$ ). Listeners interpreted wider pitch movements from H\* to L% and narrower pitch movements from H\* to H% as showing greater confidence. Figure 6.6 illustrates this result. Although there were no other significant interactions, the

The interaction between Pitch Width and Boundary Tone was highly significant for the degree of perceived CONFIDENCE. Listeners interpreted wider pitch movements from H\* to L% and narrower pitch movements from H\* to H% as showing greater confidence. Figure 6.6 illustrates this result.

two main effects were also significant: Boundary Tone ( $F_{1,34} = 33.208$ ,  $p < 0.01$ ) and Pitch Width ( $F_{1,34} = 3.279$ ,  $p < 0.05$ ).



L%	Pitch width from H* to L%	0.50 ERB	1.00 ERB	1.50 ERB	2.00 ERB	2.50 ERB
	Mean	43.65	51.95	58.81	55.65	66.94
	Standard Error	2.44	2.86	2.64	3.08	3.70
H%	Pitch width from H* to H%	0.50 ERB	1.00 ERB	1.50 ERB	2.00 ERB	2.50 ERB
	Mean	39.73	29.16	36.92	25.59	25.62
	Standard Error	3.46	3.51	4.96	3.93	4.35

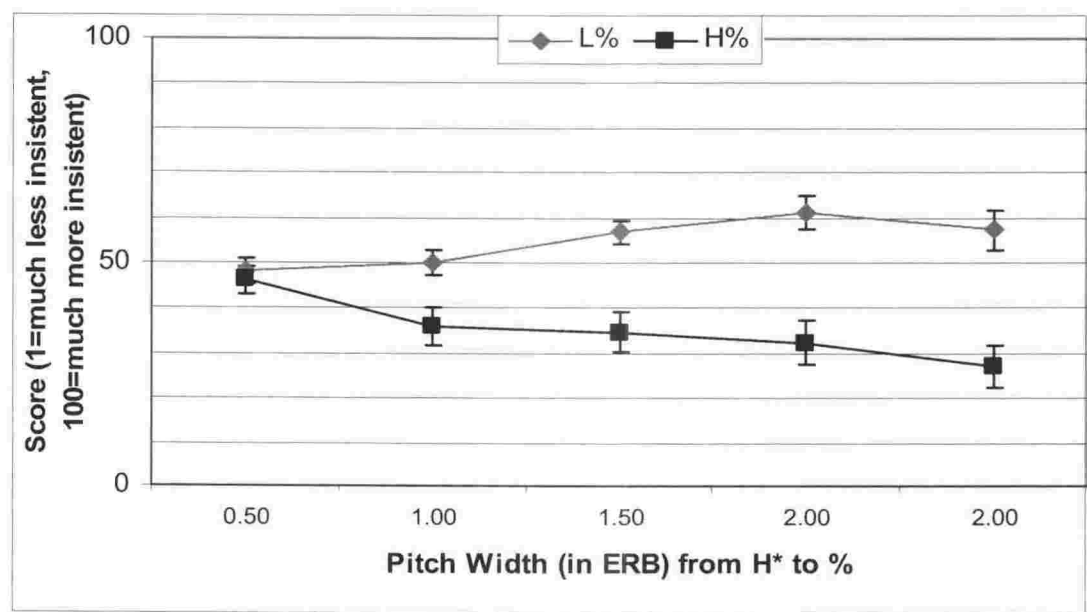
Figure 6.6.: Average response score illustrating the degree of “Confidence” across 2 boundary tones (L% and H%) and 5 pitch widths from H\* to %.

In order to examine whether the 5 pitch levels within each category (i.e. H% and L%) were interpreted differently, the scores for H% and L% were entered into separate Analyses of Variance. In each analysis, there was one between subject factor: Listener Sex (2 levels, male and female) and one within subject factor: Pitch Level (5 levels). The results illustrated a significant effect of Pitch Level within both L% ( $F_{4,136} = 10.482$ ,  $p < 0.01$ ) and H% ( $F_{4,136} = 6.836$ ,  $p < 0.01$ ). Specifically, listeners interpreted higher pitch levels on both boundary tones as more confident. In addition, a linear effect was found for pitch level within both L% ( $F_{1,34} = 27.075$ ,  $p < 0.01$ ) and H% ( $F_{1,34} = 13.042$ ,  $p < 0.01$ ) across the five pitch levels, illustrating a negative correlation between

pitch level and degree of confidence for both boundary tones. There were no significant main effects for the pitch levels within H% and no significant interactions between the factors.

### 6.3.4.3. Insistence

The interaction of Pitch Width and Boundary Tone turned out to be highly significant for the degree of perceived INSISTENCE ( $F_{4,136} = 6.229, p < 0.01$ ); listeners interpreted wider pitch movements from H\* to L% and narrower pitch movements from H\* to H% as more insistent, as illustrated in Figure 6.7.



L%	Pitch width from H* to L%	0.50 ERB	1.00 ERB	1.50 ERB	2.00 ERB	2.50 ERB
	Mean	48.08	49.81	56.84	61.27	57.33
	Standard Error	2.68	2.76	2.63	3.66	4.33
H%	Pitch width from H* to H%	0.50 ERB	1.00 ERB	1.50 ERB	2.00 ERB	2.50 ERB
	Mean	46.19	36.03	34.78	32.41	27.03
	Standard Error	2.98	4.03	4.42	4.91	4.83

Figure 6.7.: Average response score illustrating the degree of “Insistence” 2 boundary tones (L% and H%) and 5 pitch widths from H\* to %.

Boundary Target ( $F_{1,34} = 20.683$ ,  $p < 0.01$ ) turned out to be a significant main effect although Pitch Width was not found to be significantly different. There were no other significant interactions.

In order to examine whether the 5 pitch levels within each category (i.e. H% and L%) were interpreted differently, the scores for H% and L% were entered separate Analyses of Variance. In each analysis, there was one between subject factor: Listener Sex (2 levels, male and female) and one within subject factor: Pitch Level (5 levels). Pitch Level had a significant effect on the degree of perceived INSISTENCE for both L% ( $F_{4,136} = 3.190$ ,  $p < 0.05$ ) and H% ( $F_{4,136} = 5.960$ ,  $p < 0.01$ ) stimuli; the results are similar to the degree of perceived confidence discussed above, illustrating that listeners perceive greater confidence and insistence with higher pitch levels within the two targets. There were no significant interactions between the factors. In addition, a linear effect was found for pitch level within both L% ( $F_{1,34} = 6.929$ ,  $p < 0.05$ ) and H% ( $F_{1,34} = 15.833$ ,  $p < 0.01$ ) across the five pitch levels, illustrating a negative correlation between pitch level and degree of confidence for both boundary tones. There were no significant interactions between Listener Sex and the other factors.

The consistent findings of linear effects of pitch level on response scores indicate that the height of the boundary target is positively correlated with the degree of perceived LIVELINESS on the boundary target and negatively correlated with the level of perceived CONFIDENCE and INSISTENCE. These results appear to support the Frequency Code, which stipulates low F0 signals "largeness" which consequently conveys secondary meanings such as "threat, intention to prevail in a contest, dominance [and] self-sufficiency" (Ohala 1994: 341) while a high F0 indicates "smallness" and consequently conveys "nonthreat, submission, appeasement [and] desirous of goodwill..." (Ohala 1994:341). The meanings associated with the Frequency Code may offer reasons why NZE listeners perceived increased levels of LIVELINESS with higher pitch targets on H% and increased levels of CONFIDENCE and INSISTENCE with lower pitch targets, regardless of boundary category (L% or H%).



## 6.4. Summary and Conclusions

NZE listeners were more easily able to distinguish different pitch levels at the IP-Final boundary position when the direction of pitch movement from H\* to the boundary tone was different (i.e. rising to an H% target vs. falling to L%). This result provides support for a categorical boundary in NZE which is determined by the direction of pitch movement from the H\* nuclear pitch accent to the boundary position, such as L% and H%. The results provided above did not indicate that there is more than one category within the higher pitch targets of H% in NZE, as speculated earlier in this chapter. Thus the evidence suggests the use of a phonological system at boundary based on pitch height at this location relative to the preceding H\* (either L% or H%), just as that used in transcription systems such as ToBI (see Chapter 1). The categorisation is dissimilar from ToBI, however, in that the level pitch movement from H\* to the boundary target appears to be grouped with the high boundary tones, rather than the low boundary tone (e.g. H\*H-L%).

It should also be noted that the results merely illustrate that listeners may perceive two boundary categories (H% and L%) when the preceding pitch accent is high (H\*). Although this thesis describes intonation as autonomous tonal targets, the results presented above do not confirm that a third boundary category does not exist when the pitch accent is low (L\*), or even with a High pitch accent (H\*) with a different pitch height than that used here.

The gradient-response task illustrated that NZE listeners use the pitch height at the boundary position to interpret different degrees of three attitudinal meanings in NZE. Lower boundary targets are perceived as more confident and insistent than higher boundary targets, and higher targets within the range of H% values convey more liveliness. These meanings can be explained with the Frequency Code (Chapter 2, Section 2.1.1), as lower pitch signals authority and meanings associated with self-sufficiency (Ohala 1994: 341) and higher pitch indicates submission and meanings such as nonthreat and desirous of goodwill (ibid.). The results also illustrate that the sometimes stigmatised meanings attached to this intonation use, such as "uncertainty"

(see Warren and Britain 2000:164), a “questioning” attitude (Fletcher and Harrington 2001: 215) and a “lack of confidence” (Lakoff 1975), manage to survive in the perception of NZE listeners even though NZE speakers use this phrase-final rise frequently in speech. Such meanings may not be the only ones arrived at by listeners when they hear H%; they may also interpret other meanings not assessed here such as conversational continuation cues. It should also be noted that listeners may prefer one meaning, such as verifying listener comprehension, over another meaning, such as lack of confidence, once the intonation is put into context (i.e. surrounding utterances, speaker’s relationship, etc.). Furthermore, it appears that liveliness, confidence and insistence are interpreted along a gradient “Frequency Code” scale, with the height of the boundary target positively correlated with the degree of perceived liveliness and negatively correlated with the degree of perceived insistence and confidence.

As shown in Chapter 5, two pairs of meaning were contrasted with the boundary tone; the first meaning contrast was attitudinal (concerned vs. unconcerned) and the second meaning contrast was discorsal (conversational continuation vs. cessation). In addition to examining categorical boundaries designated with the height of the IP-final boundary target, this chapter explored how variations in the height of this pitch target affects the interpretation of three attitudinal meanings. In the following chapter, I specifically investigate how NZE listeners interpret conversation cues on varying heights of the boundary target and the preceding pitch accent.

**Boundary tones and nuclear H\* as conversational markers indicating discourse completeness**

---

**7.0. Chapter abstract**

In Chapter 5, boundary tones were found to distinguish between conversational continuation and cessation with high (H%) and low (L%) tones, respectively, in NZE. In addition, it was suggested that different degrees of finality may be signalled by different heights of L%. The aim of this chapter is to further explore how both boundary tones and the preceding pitch accent indicate conversational cues in New Zealand English. A perception experiment was designed with this aim in mind. 21 test stimuli were created from a single utterance, combining 7 IP-Final boundary target heights with 3 pitch heights on H\* targets occurring on the penultimate syllable. 30 participants were instructed to complete two separate discrete-choice tasks; first, they were asked to specify whether the utterance indicates a conversational turn to the listener (i.e. implying that it is their turn to speak), and second, they were instructed to specify whether the utterance indicates that the speaker's turn is complete (i.e. implying that they have finished speaking, although not necessarily indicating a conversational continuation cue). For the listener-oriented turn cue response, the number of participants indicating a listener turn cue positively correlated with the height of the high pitch accents (H\*) and boundary targets (%); firstly listeners may have interpreted the higher H\* targets as indicating enthusiasm and interest in the conversation and secondly, listeners may have interpreted higher boundary targets as indications of questions and openness, and subsequently a listener turn cue. For the lower H\* target, however, listeners were more likely to interpret the "widest" pitch movements from H\* to H% (as opposed to narrower pitch movements between these same two targets) as a speaking turn cue, but much less likely to give this interpretation to all other phrase-final pitch movements. For the speaker-oriented turn cue response,

judgements about whether the speaker had completed their speaking turn appeared to be dependent primarily on the height of the boundary target; the data showed that the higher the boundary target, the more likely the listener would interpret the utterance as indicating that the speaker wished to continue their turn at talk, while lower boundary targets were more likely to be interpreted as indicating that the speaker had finished their turn at talk. Collectively the results show that high boundary targets indicate conversational continuation (either for the speaker or to the listener) and low boundary targets convey finality, in addition to higher pitch accents indicate the speakers' enthusiasm and thus, cues to the listener to continue the conversation.

### **7.1. Introduction**

As detailed in Chapter 2, listeners may arrive at various meanings from the intonation pattern of an utterance. For instance, a listener may interpret a higher level of overall pitch as indicating a speaker's "liveliness" (Gussenhoven and Rietveld 1994) whereas a lower level of pitch may indicate a "dull" attitude (O'Connor and Arnold 1961: 33). The informational status of the spoken information within the conversation may also be conveyed by intonation; falling pitch patterns may demonstrate that the lexeme is new information in the context of the spoken discourse whereas a rising pitch pattern on this same lexeme may be the speaker's way of checking the relevance of the utterance to the discourse (Gussenhoven 1983: 193-266). Intonation may also distinguish a declarative from an interrogative with low and high phrase-final pitch targets, respectively. This chapter explores yet another function of intonation, namely how the intonation patterns occurring at the phrase-final location function as a conversational cues and subsequently indicate the management of speaking turns in NZE.

Research has shown that pitch at Intonation Phrase boundaries conveys conversational cues in several languages and language varieties. Conversational cues, such as speaker-turn continuation and completion cues, are signalled by a speaker's intonation at IP-Final locations in Dutch (Caspers 2000; Caspers 2001), Ulster English (Wells and Peppé 1996), Tyneside English (Local, Kelly and Wells 1986), London Jamaican English (Local,

Wells and Sebba 1985) and Southern British English (Wichmann and Caspers 2001). As noted in Chapter 2, Wichmann and Caspers found two tonal features which are interpreted as speaker turn holds in Southern British English (SBrE), including a high level tone and a low tone which does not reach the speaker's F0 minimum. Both phrase-final turn-holding cues are pitch movements with a reduced pitch range (e.g. a "high level tone" and a "truncated fall"). Conversely, contours with a wider pitch span, a "high rise" and a "complete fall", were interpreted by the majority of SBrE listeners as signalling a change in speaking turn. Similarly, a wide pitch span on phrase-final rising pitch patterns appears to denote turn-finality in three other varieties of English: Tyneside, Ulster and Southern British English (See Chapter 2).

There has been scarce research investigating explicitly how intonation indicates conversational turn-taking cues in NZE. One of the earlier experiments in this thesis, reported in Chapter 5, examined how 30 NZE speakers produced intonation to indicate conversational continuation and cessation to the listener (section 5.3.1). The results showed that the boundary tone may be used to differentiate these two meanings; although an L% was used to indicate both conversational continuation and cessation, an H% was only produced to indicate conversational continuation. The results are largely unsurprising, as one universal meaning of intonation is that a rising pitch pattern indicates that there is "more to come" whereas falling pitch patterns indicate "closed" meanings, such as finality (Cruttenden 2001:57). The speakers may have used a high phrase-final target (H%) to indicate their willingness to continue the conversation. The results of the Prosodic Completion Task in Chapter 5, however, failed to show how such a contrast between high vs. low boundary tones may be affected by other tonal factors, such as pitch span or the pitch height of specific tonal targets. For instance, a wide pitch movement from H\* to H% may indicate speaker turn-completion cues whereas speaker turn-holds may be indicated with narrow pitch movement between the same tonal targets, just as in SBrE (Wichmann and Caspers 2001). It is interesting to note that neither the pitch accent nor the phrase accent categories were used to contrast continuation and cessation in the Prosodic Completion Task. It may be the case that absolute pitch height

value of the pitch accent, rather than simply the pitch accent category (i.e. H\*, L\* or L+H\*), contributes to the interpretation of such conversational cues in NZE. In order to examine such suppositions, the following perception experiment was carried out. Specifically, a discrete-choice task was used to investigate how turn-taking in NZE conversation is marked by the height of both the boundary tone and of the preceding H\* pitch accent.

## **7.2. Methodology**

### **7.2.1. Materials**

The source utterance used in this study was taken from a male speaker of NZE (age 38 years). He was instructed to read aloud from a provided script, “as naturally as possible” and “without emphasis” whilst being recorded using Praat (Boersma and Weenink 1996) with a sampling rate of 16 kHz in mono, using a Audio-Technica Cardioid microphone (Appendix D2). The speaker was recorded in a quiet room at Victoria University of Wellington.

The utterance *Manny went to the union* was used as the source utterance in this perception experiment. The motivation for selecting this sentence was three-fold. First, the meaning of the sentence would not interfere with the potential for such conversational cues to be conveyed. For instance, if a speaker utters the sentence “*I don’t want to go to the ball with you*”, a listener may not understand the speaker to have conveyed conversational continuation regardless of the intonation employed. The semantics of the uttered words limits the possible interpretation of such meanings. The source utterance was selected due to its potential to convey all of the meanings being assessed in this experiment. For instance, speakers may convey conversational turn-holds on the sentence *Manny went to the union*, as such an utterance could occur at the beginning of a speaking monologue. Conversely, a speaking-turn completion cue may be realised on this utterance as it could potentially occur at the end of a speaking turn or monologue. Speaking turn cues may be conveyed to the listeners by either indicating or not indicating that the listener should say something following the utterance.

In addition to the lexical content of the sentence, the segmental makeup of an utterance invariably affects the F0 trace. Such methodological considerations were discussed in Chapter 3 (section 3.2.1). Although in this case the utterance was not made up entirely of voiced sonorants, the portions carrying the tonal targets which would be raised or lowered were made up of such segments. That is, both the L% and the preceding H\* were anchored to the lexeme *union*. Finally, the utterance is durationally short (1.079 sec), limiting the total duration of the experiment and the subsequent demand on the participants' time and concentration.

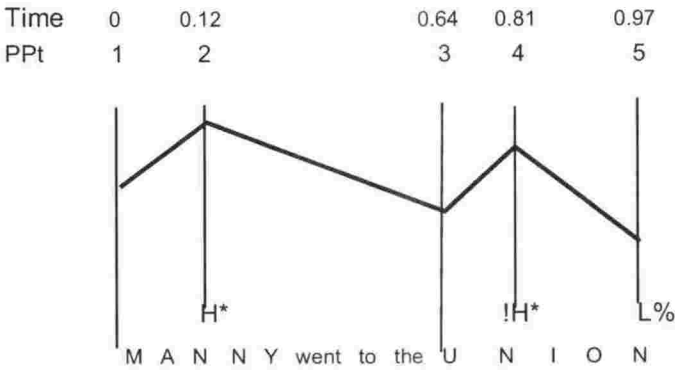


Figure 7.1. Schematic representation of source utterance (Time = sec)

Using Praat, the recorded utterance was stylised and five pitch points (PPT) were located at tonal movement positions (see Figure 7.1). The average pitch level for the recorded utterance was 3.25 ERB. The speaker's average pitch level was relatively low in comparison to an average male pitch level of 3.82 ERB, which was derived from 13 male NZE speakers reading aloud a monologue (see Chapter 4). Low pitch targets and overall pitch levels may convey several meanings. As discussed in Chapter 2, the Frequency Code stipulates that low pitch signals authority or aggression and high pitch signals submission (Ohala 1984: 340). Such meanings may affect the listeners' interpretations of conversational cues such as, "Please continue the conversation" or conversely, "Shut up!", as well as "I am not interested". I decided that raising the pitch level of a naturally produced utterance was more appropriate than prompting the same speaker to produce the utterance in a higher pitch register or prompting another speaker to produce the



utterance as instructed (e.g. using a higher or lower pitch). In order to avoid the possibility of such interpretational effects, the source utterance was adjusted to a level which was more similar in pitch level to the average male NZE pitch level. The Total Rescaling Method was employed for this purpose; all of the pitch points in the utterance were adjusted to a higher level in an attempt to avoid the possible interpretational effects of stimuli with relatively low pitch targets or with a low overall pitch level.

Raising the source utterance pitch level to that of the NZE male speaker average (3.82 ERB) appeared unnatural and “Mickey Mouse-like”. The cause for such vocal distortions is uncertain. The unnatural effects may be due to the distortions introduced by Praat’s resampling routine. It may also be true that such a high pitch level is too high for an isolated utterance, whereas a story or monologue context will also require higher pitch in order to allow for declination over longer intonation phrases and paragraph structure, and will result in a raised overall average pitch level. In order to elicit responses from the participants which were natural interpretations of the utterances, it was necessary for the stimuli to appear natural. The average F0 of the resulting stimulus was therefore set at 3.50 ERB. The pitch level was selected as it raised the overall pitch level of the utterance closer to that of the NZE male average whilst retaining the “natural” sound of the utterance; an auditory assessment by the author, which was corroborated by a Linguistics student, confirmed that the resulting utterance sounded natural (i.e. not computationally manipulated). This pitch level also falls within the range of NZE male speakers (3.22 – 4.65 ERB, see Chapter 4). The pitch values at each of the five tonal positions for this source utterance are provided in Table 7.1.

	PPt 1	PPt 2	PPt 3	PPt 4	PPt 5
ERB	4.18	4.54	3.35	4.28	2.78

Table 7.1: Pitch Point (PPt) values (in ERB) for the source utterance

Two sets of manipulations involved setting the values of two pitch targets relative to the source PPts indicated in Table 7.1. First, the height of the boundary target (PPt 5, Figure 7.1) was adjusted to one of seven values. The lowest of the seven values, illustrated as IP-Final target (IPF) 1 in Figure



7.2 was equal to value of the source utterance (PPt 5, Table 7.1). L% targets were not shifted below the source utterance L% target because, in the author's view, lower values were not audibly distinct from the source target value. Six higher targets (IPF2 to IPF7, Figure 7.1) were adjusted relative to PPt 5 (Table 7.1), in steps of +0.50 ERB.

I also examined whether the high pitch target on the penultimate syllable (PPt 4, Figure 7.1) helps listeners distinguish conversation cues. Thus PPt 4 was manipulated to one of three pitch levels relative to the source PPt height (PPt 4, Table 7.1). The PPt value from the source utterance was used as the lowest H\*, illustrated as H\*1 in Figure 6.6, as it was a natural, unemphasised (as instructed) production of the utterance. To create a mid and high H\*, the PPt was set 0.50 and 1.0 ERB above the source PPt, respectively (H\*2 and H\*3, in Figure 7.2).

The motivation for selecting this step size was that I did not want to unduly burden the participants with a long experiment containing numerous similar stimuli resulting from smaller step sizes. Steps of 0.50 ERB for both H\* and boundary targets were determined as audibly distinct steps in pitch by the author. In addition, the highest boundary targets and pitch accent targets were judged as natural-sounding stimuli, while stimuli with even higher targets appeared "unnatural". Both the step size and the upper pitch values for both pitch targets (boundary target and pitch accent) were later confirmed as "natural" by two native NZE speakers.

The three H\* and seven boundary manipulations were combined, resulting in 21 test stimuli (Figure 7.2). A further 20 filler stimuli were created from the source utterance in an attempt to mask the experimental manipulations and to therefore make the goal of the experiment less obvious to the participants. The filler and test stimuli produced a total of 41 stimuli.

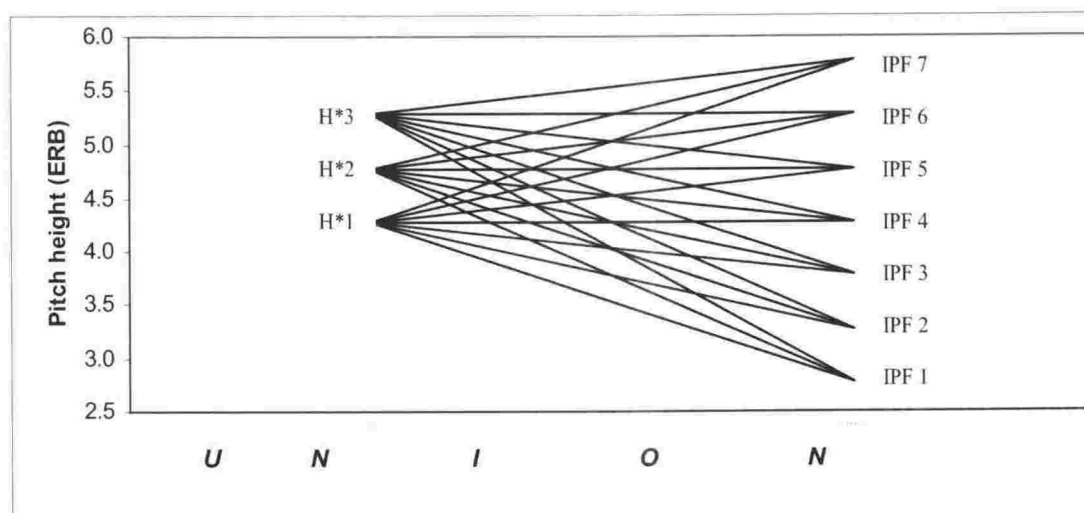


Figure 7.2: Schematic representations of H\* (3 levels) and boundary (7 levels) manipulations performed on the lexeme **UNION** (PPt 4 and 5 in Figure 7.1).

## 7.2.2. Participants

30 participants took part in this experiment on a voluntary basis. A questionnaire was used to determine social variation amongst the participants (see Appendix D5). 13 men and 17 women took part in this experiment; all were students at Victoria University of Wellington. 21 of the selected group were aged between 18-22 years, five were 23-29 years old and four were 30-35. 25 in the group claimed to be Pākehā, 3 participants claimed to be mostly Pākehā with some Māori, and 2 participants stated that they were New Zealanders of Asian descent. All of the participants claimed to be native New Zealand English speakers, and none reported having any known hearing impairments.

## 7.2.3. Procedure

The experiment began with an introduction (Appendix D3), which was a recording of the author using the computer program Praat and an AKG D19DCS unidirectional microphone. The introduction informed the participants that they would hear a male native speaker of New Zealand English, who they should assume was their discourse partner. They would hear this voice say one sentence in a number of ways. In an attempt to

record interpretations which occur naturally in conversation, they were also instructed to imagine that the speaker was talking directly to them.

In order to examine whether different heights of the boundary targets as well as the preceding high target on an H\* pitch accent distinguish different types of conversational continuation, the participants were given two separate discrete-choice tasks. First they were asked whether the speaker of each utterance was prompting a listener-oriented turn cue. Specifically, they were instructed to tick the "INDICATING: say something" box if they believed the utterance was indicating "it is your turn to say something". If they did not believe the utterance was prompting their turn to talk, the "NOT INDICATING" (say something) box was to be used. The participants were not provided a separate box, indicating an "Other" meaning that they may have interpreted on the stimulus. The purpose of this experimental limitation was to allow the participants to concentrate on one meaning at a time, rather than offering the possibility for a vague answer.

After hearing all of the stimuli once and indicating their response for each utterance, they were told they would hear the utterances once again (Appendix D4). In the next task, I wanted to determine whether these same pitch variations indicate the speaker's turn completion. In order to examine this speaker-oriented turn cue, the participants were asked to tick one of two boxes; the "FINISHED turn-at-talk" box was used to indicate that the utterance communicated the speaker's turn completion, and, ticking the "MORE to say" box indicated that utterance was communicating a conversational turn-hold.

The listener turn cue and the speaker turn cue are interrelated. For instance, if a speaker is indicating that he has "more to say", he would not be simultaneously indicating a listener-oriented speaking turn cue. It should be noted, however, that the meanings are not always mutually exclusive. For example, the speaker may indicate that he is finished speaking whilst also indicating either a listener turn cue or no such cue. Although the two sets of meanings are not entirely independent, I did not want to make the assumption that listener turn cues and speaker turn cues are co-dependent. Thus the two conversational cues are assessed separately. In order to avoid responses being influenced by the preceding task, the two separate tasks were

reordered; one experiment had “speaker completion cues” task played first while the other played “listener turn cue” task first. At the end of the experiment, participants were provided with the opportunity to ask questions or comment on the experiment. None reported any difficulty with differentiating these two meanings.

In order to reduce the interpretational effects that preceding stimuli may have on subsequent stimuli, the stimuli were randomised in three separate orders. The stimulus orders were examined before the experiments in order to ensure that one stimulus was not always followed by another specific stimulus, as may sometimes occur in random orders. For instance, I did not want H\*2 stimuli to always follow H\*1 stimuli in the task, as the former would be interpreted as higher than if these same stimuli were followed by H\*3 stimuli. Prior to every fifth stimulus, a 0.5 second beep was inserted in order for the participants to determine whether they were keeping pace with the presentation. Each stimulus was separated by an interval of 3.5 seconds. The first five stimuli were played a second time within the experiment and the answers given for the first set were discarded from the included data; this was done to discount erroneous answers often made at the beginning of experiments. The experiment, including the introduction and the tasks, took approximately 15 minutes.

The introduction and stimuli were played over a Sony CF5-W420 radio cassette recorder in a small classroom in Victoria University of Wellington for groups of 4 people or more, or in the author's office for smaller groups. Small groups ranging from two to ten participants took part in a single listening exercise. Each listening exercise was attended by the author, who set the volume to an appropriate level which was then confirmed by the participants' feedback. The participants were provided with a \$5 book/music voucher as recompense for their time.

### **7.3. Results and Discussion**

The dependent variables for the two separate analyses were the “listener turn cue” response and the “speaker turn completion” response. The independent variables for both analyses were (1) the height of the boundary target (3 levels) and (2) the height of H\* on the penultimate syllable (7 levels). Counts

of (1) “INDICATING: Say something”, which assessed listener-oriented turn cues and (2) “More to say”, which assessed speaker-oriented turn cues were entered into two separate Hierarchical Loglinear Analyses (Hays 1981: 561-566), each with H\* height and IP-Final boundary height as factors. For each analysis a parsimonious model was built by backwards elimination of factor combinations. I did not include speaker sex as an independent variable in the current analysis as it was not the aim of this chapter to examine how male and females interpret cues differently. Thus, the results represent how NZE listeners as a whole interpret conversational cues.

7.3.1. Listener- oriented Turn Cue

For the listener-oriented turn cue response, there was a significant interaction between boundary height and H\* height ( $G^2 = 7.823$ ,  $p < 0.05$ ), as well as significant main effects of both H\* height ( $G^2 = 52.108$ ,  $p < 0.05$ ) and boundary height ( $G^2 = 11.908$ ,  $p < 0.01$ ). The turn cue response trend across the seven boundary heights and three H\* heights is illustrated in Figure 7.3.

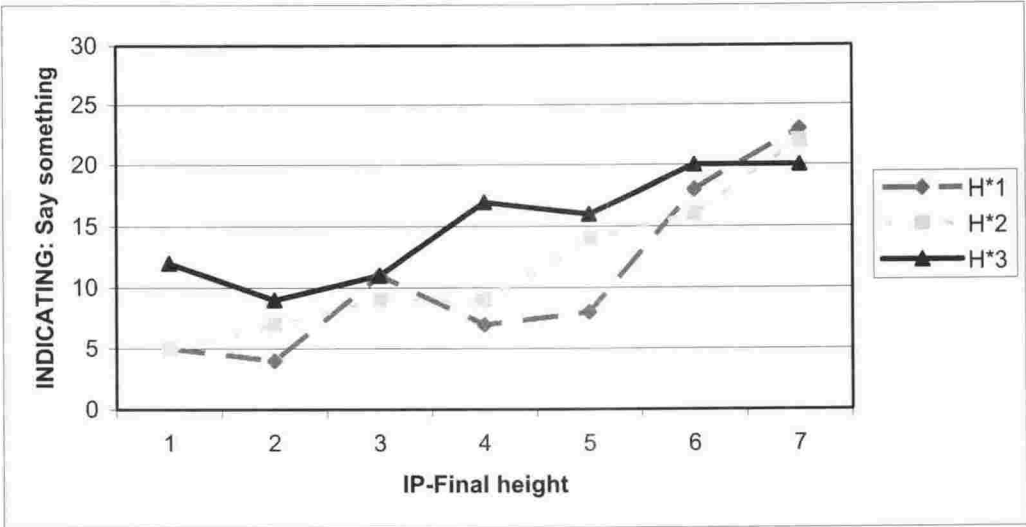


Figure 7.3: Listener-oriented turn cue response over 7 IP-Final boundary heights and 3 H\* pitch accent heights, with the vertical axis showing the number of participants making this response (N = 30).

For the two higher H\* pitch accents, namely H\*2 and H\*3, there was a near linear response trend for boundary height; the results show that the higher the boundary, the more likely it is that the stimulus indicated that the

listener should now have the conversational floor. The results for H\*2 and H\*3 stimuli are unexceptional as higher boundary targets are more likely to be interpreted as either questions or High Rising Terminal contours (HRTs). Hirst and Di Cristo (1998: 1) state that "...in a vast majority of languages some sort of raised pitch (final or non-final) can be used in contrast with lower pitch to indicate that an utterance is intended as a question rather than as a statement." As questions prompt responses from the speaker's discourse partner, a listener turn cue would be a likely interpretation if the participant understood that a question had been asked. It is also possible that high boundary tones may be interpreted as statement rises. High Rising Terminals (HRTs) are said to function as "positive politeness markers, serving to overcome inter-speaker hurdles and to build and maintain speaker-hearer solidarity" (Warren and Britain 2000:169). If listeners interpreted the stimuli as statement rises, they may also interpret the speaker using the high boundary tone to be checking for the listener's understanding. In such cases, the listener may also want to respond to such a cue and indicate (in some way) whether or not they understand the given statement and thus, interpret a listener-oriented turn cue.

A comparable tendency was evident for the height of the high pitch accent, at least for the H\*2 vs. H\*3 stimuli; the results revealed that the higher H\* target (H\*3) was more likely to be interpreted as a listener-oriented turn cue than the lower H\* target (H\*2). This result may be due to the listeners interpreting higher pitch peaks according to the Effort Code; higher pitch peaks, and the resulting wider pitch excursion size may be interpreted as "enthusiastic" whereas lower peaks with a narrower pitch excursion may be interpreted as "uninterested" (Gussenhoven 2004: 88). The results may suggest that such affective meanings are important when interpreting listener-oriented turn cues. Listeners may be more likely to interpret cues to continue the conversation when their discourse partner is enthusiastic and interested in the conversation.

For H\*1 and H\*2 stimuli, there is an increase in listener continuation cues as the low boundary target rises (IP-Final heights 1-3 for H\*1 and 1-4 for H\*2, see Figure 7.3). This response trend, however, was not always the case;

the lowest pitch accent height ( $H^*1$ ) did not appear to be interpreted on a similar gradient meaning scale across the 7 boundary targets. A distinction appears to exist between  $H^*1$  stimuli with "wide" pitch span from  $H^*$  to  $H\%$  and other IP-Final boundary targets for  $H^*1$ ; the majority of listeners indicated that a wide pitch span from the lowest  $H^*$  target onto a high phrase-final pitch target (level 6 and 7, Figure 7.3) conveyed listener oriented turn cues whereas for lower boundary targets (levels 1 to 5, Figure 7.3), the majority of listeners believed the stimuli were not conveying such conversational cues.

The different response trend for  $H^*1$  stimuli in comparison to  $H^*2$  and  $H^*3$  stimuli may have been due to the degree of perceived emphasis on the different levels of the high pitch accents ( $H^*$ ). Ladd and Morton (1997) have shown that pitch peak height positively correlates with the degree of perceived emphasis, although listeners categorically interpret the pitch contour as belonging to one of the two categories (emphatic vs. unemphatic). The listeners may have also categorised the height of  $H^*$  into two separate categories: unemphatic stimuli ( $H^*1$ ) and emphatic stimuli ( $H^*2$  and  $H^*3$ ). It should be noted here that such a categorisation is merely speculation, as the participants were not asked whether they interpreted the stimuli as "emphatic" or "unemphatic". Nevertheless, the categorisation of pitch height into two separate categories is not new. O'Connor and Arnold (1961), for instance, categorised pitch range of falling and rising tunes in British English into two categories: the high fall vs. low fall and the high-rise vs. low-rise. The separation of the two categories between  $H^*1$  and  $H^*2$  may be due to the perceptual effects of declination; the heights of  $H^*2$  and  $H^*3$  were greater than that the preceding high pitch accent in the utterance (Ppt 2, Figure 7.1), whereas the height of  $H^*1$  was lower than this preceding peak. Declination results in a lowering of pitch targets in an utterance. Such a lowering effect appears unmarked due the natural decrease in air pressure over a breath unit. The  $H^*1$  stimuli, which formed a lowering of subsequent pitch peaks across the intonation phrase, would have appeared to the listener as unmarked. Conversely listeners may interpret pitch peaks which are later and measurably higher than earlier pitch peaks (e.g.  $H^*2$  and  $H^*3$ ) as emphatic. Similar findings have resulted from other experiments (Pierrehumbert 1979; Gussenhoven and Rietveld 1988).



The possible categorization of H\* heights according to perceived emphasis provides one explanation of why the two groups (H\*1 on the one hand and H\*2 and H\*3 on the other) were responded to differently. However, it does not explain why a gradient response trend occurred for the supposed “emphatic” stimuli (H\*2 and H\*3 contours) in comparison to a possible categorical distinction made for the “unemphatic” stimuli (H\*1 contours). One such explanation may be that the additional meanings interpreted on “emphatic” vs. “unemphatic” stimuli could have caused the different response trends. As discussed in the preceding paragraph, later peaks which are measurably higher in pitch than peaks occurring earlier in the utterance would appear marked to the listener. This is due to the greater articulatory effort made by the speaker; although there is lower subglottal air pressure present, the speaker produces a pitch peak which exceeds the pitch height of a previous peak. In addition to emphasis, other meanings may also be conveyed by such effort, such as a speaker’s enthusiasm (Gussenhoven 2004: 88). Such supplementary meanings may be aligned with stimuli which are perceived as emphatic due to the greater expenditure of pitch production effort (e.g. H\*2 and H\*3). Therefore all of the perceived emphatic stimuli would be signalling conversational continuation cues and subsequently, listener turn cues. Accordingly, the number of respondents interpreting a listener turn cue increases as the height of H\* and the IP-Final boundary targets are raised.

It is evident from Figure 7.3 that a dramatic shift in judgment (from 8 to 18 of the 30 listeners) occurs between IP-Final 5 and IP-Final 6 target heights on H\*1 stimuli. The form of the intonation containing IP-Final 5 and IP-Final 6 targets is in both cases a rising pitch movement from the high pitch accent (H\*) onto the boundary target (illustrated in Figure 7.2). The shallower pitch movement from H\* to the H% in the case of the IP-Final 5 target may have indicated that the speaker was continuing their turn at talk. As discussed in Section 7.1, turn-holding cues are indicated by phrase-final boundaries with a reduced pitch range on a “high level tone” and a “truncated fall” in SBrE (Wichmann & Caspers 2001). It may be the case that shallow pitch movements from H\* to H% indicate a turn-hold rather than a listener speaking-turn cue in NZE, as in SBrE.

It is also interesting to note that listeners do not categorise the listener-oriented turn cue response into two traditional categories on the H\*1 stimuli. That is, listeners did not separate the stimuli into two categories differing in whether they have rising pitch movements (boundary targets 5 - 7, Figure 7.3) or falling pitch movements (boundary targets 1 - 3, Figure 7.3), nor did they make a categorical distinction between high boundary tones (boundary targets 5 – 7, Figure 7.3) and low boundary tones (boundary targets 1 – 4, Figure 7.3). As discussed above, the data illustrate a categorical distinction made within the traditionally labelled H% category, specifically between a shallow pitch movement and a wide pitch movement.

**7.3.2. Speaker turn completion cue**

For speaker turn completion responses, the boundary height was found to be a significant factor ( $G^2 = 18.956, p < 0.01$ ); generally, the higher the boundary target, the more likely the respondent would choose “more to say” (Figure 7.4).

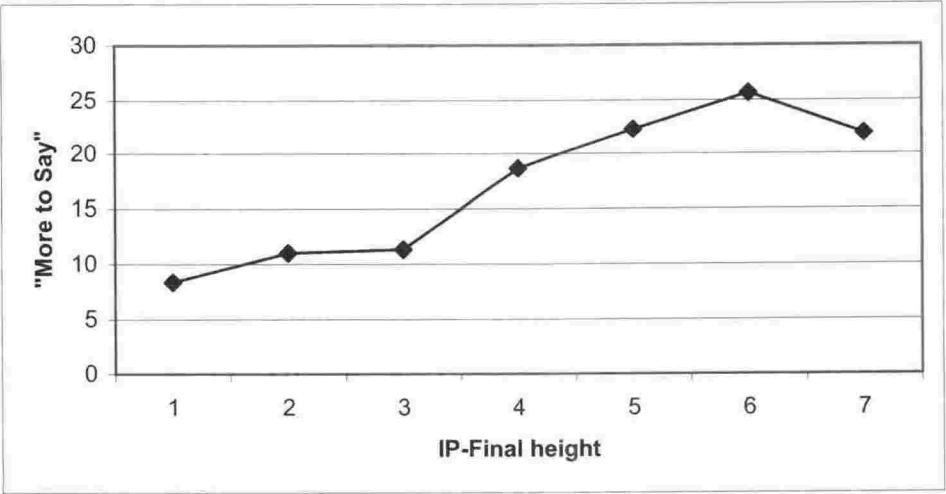


Figure 7.4: Speaker turn completion cue over 7 IP-Final boundary heights for 3 H\* pitch accent heights, with the vertical axis showing the number of participants making this response (N = 30).

H\* height was not found to be a significant factor in determining turn completion, nor was there a significant interaction of H\* height and boundary height.

Cruttenden (2001) comments that a universal meaning of rising pitch movements is that they indicate openness and/or that there is “more to come”

whereas falling movements indicate no such meanings. The results presented here support a similar form-function relationship in NZE intonation; although not depicting a simple rise/fall dichotomy, the results illustrate higher pitch targets, which form rising pitch movements from H\* to H% (see Figure 7.2), as more likely to indicate that the speaker has “more to say” than lower targets. The results here do not suggest a categorical distinction as indicated by Cruttenden, but rather a gradient response which roughly corresponds with pitch target height.

Although this experiment did not explicitly test whether or not listeners heard the stimuli as statements or questions, it appears that they interpreted the higher boundary targets as statement rises rather than question rises. If the listeners had interpreted the utterances as questions, they should also have expected that the speaker had ceded the conversational speaking floor to the listener, which would indicate a speaker turn completion. The interpretation of the stimuli as declaratives rather than as interrogatives may be due to the language specific use of tonal distinctions between declaratives and interrogatives in NZE. Previous scholars have found that the rise onset on phrase-final high targets differentiates phrase-final H% produced on interrogatives from those produced on statements. Zwartz and Warren (2003) and Warren (2005a) found that NZE listeners are more likely to interpret early and later rise onsets as interrogatives and declaratives, respectively. This supports production experiment results (reported in Warren and Daly 2005; Warren 2005a), which showed that “there is a stronger tendency in statements than in questions for late rise onsets rather than onsets on the accented syllable” (Warren 2005a: 214). The rise onset onto H% is located relatively late (within the final two syllables of the stimuli used here). Listeners may have used this late rise onset in interpreting a declarative H%. Nevertheless, supplementary experiments will need to be carried out in order to examine whether earlier rise onsets affect the interpretation of speaker-oriented turn cues. In addition, it cannot be proved whether or not listeners interpreted the stimuli as statements or questions.

#### 7.4. Summary and Conclusions

The results presented in this chapter illustrate that NZE listeners perceive both the listener-oriented turn cues and the speaker-oriented turn cues on the intonation of an utterance, as would be expected. For listener-oriented turn cues, the pitch accent and the boundary tone were distinguishing factors. The use of the pitch accent height may be illustrating that identifying the speaker's enthusiasm on higher pitch accents is an important aspect of identifying cues to the listener to continue the conversation. Listener-oriented turn cues may require the speakers to indicate their interest in continuing the conversation. Speaker-oriented turn cues may not require such meanings. Accordingly, speaker-oriented turn cues were indicated with the boundary target alone and not with the height of the pitch accent. Although we may expect enthusiasm to indicate that the speaker wants to continue the topic by him/herself, this does not appear to be the case here; pitch accents were not used to indicate speaker-oriented turn cues. From these results, I conclude that listeners interpret both speaker-oriented and listener-oriented conversational turn cues on boundary tones. Listener-oriented turn cues may also be interpreted on the height of the pitch accent whereby higher pitch accent targets are more likely to induce a listener-oriented response than lower targets.

It is unsurprising that NZE signals openness (e.g. there is more to come from either the speaker or as a cue to the listener to continue the conversation) with higher pitch targets, and closure (e.g. speaking turn finality with no continuation cues) is signalled with lower pitch target. However, the results illustrate that this distinction is not categorised according to rising vs. falling pitch movements, as asserted by Cruttenden (2001), but rather boundary tone height. Listeners do not align the listener-oriented turn cue response into two traditional categories on the H\*1 stimuli. For instance, the listeners did not interpret the stimuli as prompting a listener speaking-turn cue on rising pitch movements (or H%). The listeners appear to make a distinction within the traditionally labelled H% category, specifically between a shallow pitch movement and a wide pitch movement. Thus the use of a shallow pitch movement from a lower level H\* (in comparison to higher H\* pitch targets) to H% indicates a turn-hold rather than a listener speaking-turn

cue in NZE, which is similar to the conversational cues indicated on intonation in Southern British English. It is unclear, however, whether such conversational cues are moderated by the height of the pre-final H\* (e.g. any High pitch accents which precede the phrase-final pitch accent) in Southern British English as in NZE.

The results are also interesting in that they show that it is not the pitch movement that is important in determining speaking cues, but rather the pitch target height. For the listener-oriented turn cue response, higher H\* targets were found to indicate continuation, even though L% did not differ. This result was unanticipated as it indicates that the bigger fall, or the larger pitch movement from H\* to the low boundary tone, shows continuation rather than finality. Other research carried out on BrE has suggested that a wide pitch movement onto a L% target indicates finality whereas a narrow pitch movement indicates non-finality in this variety (see Chapter 2, Section 2.2.3). It appears that listeners are using the height of the H\* pitch accent to identify “enthusiasm” and “interest in the conversation” rather than interpreting the consequent larger pitch movements from H\* to L% to indicate finality.

**Summary, Discussion and Conclusions**

---

Intonation is claimed to convey a number of meanings. As outlined in Chapter 2, higher pitch is alleged to signal submission and appeasement whereas low pitch is claimed to identify aggression (Ohala 1994). In addition to pitch height, a wider pitch span has been claimed to indicate meanings associated with the greater expenditure of effort, such as emphasis, insistence and enthusiasm (Gussenhoven 2004). Another meaning assigned to intonation is that of the phrase-final rise which is often said to signal a question whereas a fall frequently identifies a statement. However, these meanings are not always conveyed in all languages and language varieties. A case in point given was the High Rising Terminal contour occurring in statements (HRT), which is used frequently in NZE statements.

At present, there is no detailed or complete phonological analysis of the intonation of NZE, since in this respect it is a “relatively under-studied variety of English” (Warren 2005b: 345). Several studies have shown how social groups within NZE use their intonation differently (Allan 1990; Britain 1992; Britain and Newman 1992; Warren and Britain 2000; Warren and Daly 2000; Daly and Warren 2001; Ainsworth 2003; Ainsworth 2004). A smaller number of studies have examined the characteristic form of NZE intonation (Fletcher, Grabe and Warren 2004; Warren and Daly 2005). Others have explored the form-function distinctions in this language variety, such as the rise onset location distinguishing between a statement and a question rise in NZE (Zwartz and Warren 2003; Warren 2005a; Warren and Daly 2005). However, there are many more features of intonation which make up this language variety. One of the key under-explored issues is whether these various forms of NZE intonation are phonetic variants on other English intonational forms or whether they represent distinct phonological categories in this variety.

The overall aim of this thesis was to explicitly examine the phonology of NZE intonation. With this aim in mind, the first two studies investigated which

features of intonation make NZE intonation unique. A number of subsequent studies explored how speakers (chapter 5) and listeners (chapters 6 and 7) employ intonation as a meaningful communicative device in this language variety. The following section summarises these experiments and discusses them in relation to the research questions raised in Chapter 2. The main findings from this thesis are summarised in Section 8.2, which highlights several characteristic properties of NZE intonation that have been found. Finally, further considerations are given attention in Section 8.3, such as the limitations of the methodological techniques applied in this thesis.

### **8.1. Summary and Discussion**

In the Introduction (Chapter 1), I reviewed the relatively few studies carried out on NZE intonation. The review illustrated that intonational variation occurs between social groups in New Zealand, reflecting the particular differences in speakers' age, sex and regional origin. The review of other researchers' experiments which have been carried out on NZE also highlighted two tonal features that have been found to be functionally distinctive in this variety. First, statements appear to have a later rise onset than questions in NZE and second, rising phrase-final pitch movements on statements (HRTs) function as positive politeness markers in this variety of English. This review of research highlighted the paucity of information concerning the meaning of NZE intonation. It was unclear, for instance, whether an identical intonation pattern conveys identical meanings to those outlined for other varieties of English. Given that previous research has exhibited meaningful differences between language varieties, it would be unwise to apply the semantic contrasts of one variety to a separate variety.

It was argued that the phonological status of any observed difference between NZE intonation and other varieties is ambiguous as it can often be unclear what the linguistic function of such differences is within the variety. The HRT was discussed to illustrate this point. Although two meanings assigned to this form of rising intonation in RP are "tentativeness" and "incompleteness", NZE speakers readily employ such an intonation when such meanings are not likely to be the intention of the speaker. Thus applying



a form-function relationship based on analysis of the intonation of one language variety does not necessarily give an accurate description of the intonation in a different variety (Warren 2005b). The aim of this thesis was provided an explicit examination of NZE intonation with special attention given to the meaning of such intonation use.

Chapter 2 outlined a number of attitudinal, discoursal and grammatical meanings ascribed to intonation, or more specifically, to individual pitch target heights and to the overall pitch register, pitch width, pitch movements and pitch trajectory. The questions raised by the review of literature provided the research questions examined in this thesis. Specifically, the aim of this thesis was to examine the intonational forms and contrasts of NZE intonation. The first question concerns the forms of NZE intonation; although there may be a number of similar intonation patterns shared between language varieties, it is important to determine which tonal features distinguish NZE intonation from the intonation employed in other varieties of English. The second question concerns the meaning of NZE intonation. It is unclear, for instance, whether an identical intonation pattern conveys identical meanings to those outlined for other varieties of English.

The first two data chapters examined the characteristic forms of NZE intonation. Specifically, Chapter 3 presented a perception experiment which was used to examine which tonal cues NZE listeners use to identify NZE intonation as distinctive from BrE. Three such cues were identified. First, the lower of two different values of a low phrase boundary (L%) is interpreted as sounding more like NZE; second, the higher value of a final high boundary target (H%) is believed to sound more like a NZE speaker; and third, higher phrase-medial L values are also interpreted as more NZE-like. Taken together, these results suggest that localized pitch range may be a factor, since higher phrase-medial low values produce a narrow pitch range within the phrase and higher H% and lower L% values produce a wider pitch range at the edge of the intonation phrase.

The results of the perception experiment were compared statistically with the same listeners' opinions about NZE and BrE speakers. This

correlation analysis revealed that there were no significant correlations between the participant's opinions about the two speaker groups and their perception of which tonal cues are "indicative" of NZE intonation. Thus the results of the perception study do not appear to be influenced by the attitudes that the listener may have had concerning NZE and BrE speakers.

As noted in Chapter 2, Bauer (1986: 68) indicated his impressions of NZE intonation as being "rather flat", which may be suggestive of a narrow pitch range. Others have asserted that young NZE speakers produce relatively large pitch movements on phrase final rising pitch movements. The evidence appears to support both of these claims; NZE listeners believe that NZE speech is characterised by a narrow pitch range within the phrase and a wide pitch movement at the end of the phrase.

Although the results paint a general picture of NZE intonation, they do not describe the specific forms which make up NZE intonation. With this in mind, Chapter 4 described an analysis of a read paragraph in order to further explore other characteristic properties of NZE intonation. 12 nuclear pitch movements within the reading passage were examined in order to further explore how NZE speakers use pitch at this location. The results also revealed a preferential use of particular tonal compositions for the three examined sentence types. Although there was variability in nuclear tonal composition, illustrating that NZE speakers may employ a number of intonation patterns on statements, wh- questions and other questions, NZE speakers preferred an H\*L-L% on statements and an L\*L-H% for questions in the scripted monologue. It is interesting to note that High Rising Terminals were not frequently found in this narrative style text, even though others have claimed that this intonation is used frequently in narratives in NZE. However, the preferential use of particular intonation patterns may simply illustrate that H\*L-L% is produced on statements and L\*L-H% is produced on two types of questions when conversational cues are not required, such as in a scripted monologue.

Another analysis examined how NZE speakers produce pitch range on the nuclear pitch movement, comparing 7 different tonal compositions and 2

sentence types. The aim of this analysis was to determine whether NZE may be characterised not only by localised pitch range features, as shown in Chapter 3, but also by pitch range that might be specific to sentence type or tonal composition. The results showed that pitch range is affected by the tonal composition of the utterance, but not by sentence type, as both statements and questions are produced with similar pitch ranges at the nuclear location in NZE.

What is also interesting to note is that male and female speakers produced similar pitch ranges at the nuclear position throughout the reading, which was surprising in light of previous research (Warren and Daly 2000). Other studies have noted that men and women employ intonation as a communicative device differently. For instance, female speakers produced a wider pitch range when talking to infants, but male speakers did not (Fernald, Taeschner, Dunn, Papousek, de Boysson-Bardies and Fukui 1989). Leeuwen (1999: 109, as cited in Chen et al. 2004: 343) similarly asserts that men raise their voice in news reading whereas women do not. In addition, Vermillion (2001:69) found that women adopt a raised pitch register to indicate support whereas men do not. The results presented in Chapter 4 suggest that men and women employ similar pitch ranges, at least in a text type which does not require conversational cues which may assist their discourse partner.

The remaining data chapters examined the meaning of NZE intonation. Specifically, Chapter 5 investigated how NZE speakers produce meaning using intonation as a communicative device. These speakers tended to use two locations within the intonation phrase to contrast meaning on a monosyllabic utterance. First, the nuclear pitch accent was found to be significantly different when three pairs of meanings were contrasted: Emphatic vs. Unemphatic, Concerned vs. Unconcerned and Impressed vs. Unimpressed. The results of Chapter 5 provided the first evidence that either higher ( $H^*$  as opposed to  $L^*$ ) or later ( $L+H^*$  as opposed to  $H^*$ ) occurring pitch accent targets indicate a number of meanings in NZE, each of which convey a greater degree of speaker involvement within the conversation. The tonal contrasts are claimed to be due to the Effort Code (Gussenhoven 2004).

Higher pitch accent peaks, such as H\* rather than L\*, and later occurring pitch accent peaks (a “substitute phonetic feature” for peak height, as claimed by Gussenhoven 2000: 90), such as L+H\* rather than H\*, were used by the majority of speakers to indicate meanings reflecting this added effort, such as Emphasis, Concern and an Impressed attitude.

Second, the boundary tone at the end of the Intonation Phrase (IP) was found to be significantly different in the Conversation Continuation vs. Cessation PCT and also in the Concern vs. Lack of Concern PCT. Although a low boundary target (L%) was observed in both Conversation Continuation and Conversation Cessation, a high boundary target (H%) was used to indicate Conversation Continuation. Also, the majority of speakers preferred to produce L% to signal a Lack of Concern and H% to indicate Concern. The results suggest that NZE speakers use the phrase-final high target to indicate openness and continuation but phrase-final low targets to indicate closure or non-continuation (as claimed in traditional analyses by e.g. Cruttenden 1997). However, the results presented in later experiments (Chapter 6) suggested that the NZE boundary tone may not be categorised as other varieties of English (see section 8.2).

Other evidence in this chapter suggested that the height of the low boundary target (L%) conveys varying shades of closed meanings. Specifically, higher L% values on an H\*L-L% intonation pattern were used to indicate Submission whereas lower L% values on the same pattern were used to signal Authority. Rather than simply conveying that the statement is complete, the results were said to suggest that differing heights of L% convey differing degrees of finality; NZE speakers may want to assert that their authoritative refusal to an invitation is final with lower L% whereas they may not want to assert such finality if their intention is to indicate a submissive refusal to the same invitation, and their pitch may not fall to such a low value.

These results provided in Chapter 5 provided the initial evidence that NZE speakers use the height of target values to convey meaning rather than pitch width, as supposed by the Effort Code. Gussenhoven (2004:88) states: “Increases in effort expended on speech production will lead to greater

articulatory precision (de Jong 1995), including a wider excursion of the pitch movement. Speakers exploit this fact by using pitch span variation to signal meanings that can be derived from the expenditure of effort". Although the heights of both the pitch accent and the boundary target may be manipulated in order to create a wider pitch span, speakers of NZE appear to prefer to use the height of individual targets to convey meaning. The results show that the greater expenditure of effort conveyed with higher and later pitch accent peaks convey speaker involvement and lower L% targets indicate greater degrees of finality and authority.

In Chapter 6, two perception experiments were used to explore whether the height of the boundary target differentiates categories of meaning in New Zealand English. The first perception experiment was a discrete-choice task which tested for the presence of categorical boundaries within the range of IP-Final boundary height. The results indicate that there is at least one categorical boundary, which is determined by boundary height and categorized by the pitch movement to the boundary target from the preceding H\*. It is argued that different pitch levels of the boundary tone can be categorized as either L% or H% in NZE, which roughly corresponds to falling vs. rising pitch movements. It appears that the level pitch movement from H\* onto the boundary tone would be categorised along with H%. The lack of a second peak in the categorization response function within the H% range strongly suggests that a raised H%, such as that supposed for the HRT, does not indicate a separate category in NZE, at least not in the combinations with the preceding tonal material used here. Other researchers have found that alignment of the rise onset may be used to distinguish such meaningful distinctions (Zwartz and Warren 2003; Warren 2005).

However, it should be noted here that a simple two level categorisation of the boundary (i.e. either L% or H%) was not unambiguously demonstrated by the experiment in Chapter 6, as the results suggested that there may be an additional categorical boundary on shallow pitch movements from H\* to the IP-Final boundary target. This is reflected in the high level of distinguishability of stimuli with shallow pitch movements from H\* to the IP-Final boundary

targets. It may be the case that an additional categorical distinction exists within the shallow pitch movements at this location. Such distinctions may be evidence that both a phrase accent and a boundary tone exist here, as a High phrase accent (H-) “upsteps” the low boundary target L%, as used in ToBI. However, such a supposition would require further examinations with this explicit experimental aim in mind.

The second perception experiment in Chapter 6 was a gradient-response task, assessing listeners’ interpretations of Liveliness, Confidence and Insistence on varying heights of the boundary target. The participants showed a general tendency to interpret stimuli with higher boundary targets as livelier, less confident and less insistent. The results, although speculative, are interpreted as suggesting that the sometimes stigmatised meanings attached to this intonation use, such as “uncertainty” (see Warren and Britain 2000:164), a “questioning” attitude (Fletcher and Harrington 2001: 215) and a “lack of confidence” (Lakoff 1975), manage to survive in the perception of NZE listeners even though NZE speakers use the high boundary tone frequently in their speech. It is interesting to note that although the rising tune at the end of a statement has been claimed to indicate a number of meanings in this language variety, the meanings ascribed to the rise in RP or by non-native speakers of the NZE variety can still be perceived in NZE. However, the interpretation of a specific meaning, such as either “lack of confidence” or “positive politeness marker”, may be aided by the context of the speaking situation or the surrounding utterances.

Chapter 7 described a perception experiment used to assess intonation as a conversational cue in New Zealand English. One task found that the number of participants indicating that they detected a listener turn cue correlated positively with the height of the high pitch accents (H\*) and boundary positions (%), at least for stimuli containing the highest two H\*. The results suggest that listeners interpret the higher H\* values as signals of enthusiasm and interest as well as markers that a question is being asked. It was suggested that such enthusiasm is important when signalling to the listener that they should continue the conversation. This evidence supports



the claims made in Chapter 5 that higher pitch accent peaks indicate speaker involvement. It is perhaps unsurprising that speakers did not respond to pitch accent height in interpreting speaker-oriented conversation cues; variation in pitch accent height, which indicate speaker involvement, is not required in interpreting whether the speaker is finished talking or whether s/he is going to continue his/her speaking turn.

For the lowest H\* target, however, listeners were more likely to interpret the widest subsequent pitch movements from H\* to H% as a speaking turn cue than any other phrase-final pitch movements. It was suggested that a shallow pitch movement from H\* to H% indicates a turn-hold rather than a listener speaking-turn cue in NZE, at least when the pitch accent is not conveying a great amount of speaker involvement in the conversation.

The other set of responses in this experiment showed that the height of the boundary target indicated whether or not the speaker had completed their speaking turn. Listeners were more likely to interpret the higher boundary target as indicating that the speaker wants to reserve the speaking floor and lower IP-Final boundary targets as indicating that the speaker has finished their turn at talk. The findings offer further evidence that NZE speakers employ the open/closed dichotomy, one of the universal intonational meanings suggested by Cruttenden (1997). However, the results do not suggest a categorical distinction between rises and falls, as indicated by Cruttenden, but rather a gradient response which roughly corresponds with pitch target height.

## **8.2. Conclusions**

This thesis showed that three tonal characteristics may be necessary to accurately describe the intonational phonology of NZE. First, the pitch accent is used in this variety to indicate speaker involvement, with higher or later pitch accent targets indicating a greater degree of involvement than lower or earlier targets. This notion of involvement collectively accommodates the several specific meanings found to be conveyed by these types of target. Higher and later pitch peaks were found to indicate Concern, Emphasis and an Impressed attitude, while lower and earlier pitch accent peaks were taken



to indicate an absence of these three meanings (Chapter 5). In addition, higher H\* targets indicated listener-oriented turn cues. Such cues indicate cues to the speaker's discourse partner to continue with the conversational turn and topic, which I suggested as indicative of speaker interest in the conversation (Chapter 7).

Next, there appears to be at least one categorical boundary at the IP-Final position, which is marked by the pitch movement to the IP-Final boundary target from the preceding H\*. The different pitch levels of IP-Final boundary height can be categorized as either L% or H% in NZE (Chapter 6). The height of the boundary target is used to distinguish openness and continuation cues with a high target and finality with a low target. Once again, the categorisation of H% and L% as open vs. closed, respectively, accommodates the meanings found for each target. High boundaries indicate conversational continuation and concern whereas low boundaries are used to convey conversational cessation cues and an absence of concern (Chapter 5). In addition, levels within these two categories indicate degrees of these meanings. It was suggested that lower levels of L% convey greater degrees of finality than higher L% targets. Higher boundary targets identified speaker continuation cues and listener-oriented speaking cues (Chapter 7).

The third tonal characteristic which may be used to describe shallow pitch movements from the nuclear pitch accent onto the boundary target is the phrase accent. Although this thesis does not provide unequivocal proof that this phonological category exists in this variety, two chapters highlighted the importance of its use, including Chapters 5 and 6 (see below). The finding supporting the use of a phonological unit in the intonational hierarchy corresponding to the intermediate phrase is important due to the contentiousness of the intermediate phrase (and its boundary marker, the phrase accent) in intonational phonology. Several descriptions of intonation do not recognise the intermediate phrase and thus question its validity (for German and English, see Grabe 1998; for Dutch, see Gussenhoven, Rietveld and Terken 1999). In ToBI descriptions of intonation, however, the phrase accent is used to describe a second level of phrasing. The findings presented

in this thesis suggest that the intermediate phrase may exist, at least within NZE.

In ToBI, a H\*H-L% would depict a level pitch movement from a High pitch accent to the “low” boundary target (c.f. Pierrehumbert 1998). In NZE, however, a tonal composition of H\*H-L% would describe a shallow falling pitch movement, which would probably best be treated as a realisational variant, for this variety, of the shape suggested in ToBI. The results in Chapter 6 suggested that the level pitch movement from H\* to the boundary target would be grouped with the other H% stimuli. In addition Chapter 5 showed that NZE speakers may distinguish between Authority, which may be conveyed with a wider pitch movement from H\* to L%, and Submission, which may be signalled with a narrow pitch movement between these two tonal targets. The distinction between a pitch movement which does not fall to the F0 minimum and a movement which does fall to this low value would be best categorised separately. The labels H\*L-L% and H\*H-L% may adequately describe these two categories.

Finally, the results have also illustrated that the height of specific tonal targets may more adequately describe semantic contrasts than other tonal descriptors (e.g. pitch range or F0 shape) that have been claimed to indicate meaning. This is an important claim as it differs from previous assertions made on the meaning of intonation, which have presumed that such meaning is conveyed by pitch excursion size (Gussenhoven 2004) and the shape of the F0 movement (O'Connor and Arnold 1961). What is interesting to note is that the Effort Code claims that a wide pitch width may convey both authority and the speaker's enthusiasm. However, it appears that several meanings ascribed to pitch width would be more accurately described in terms of pitch target height, with higher H\* targets indicating a speaker's enthusiasm to continue the conversation (Chapter 7) and lower L% conveying authority (Chapter 5). For instance, if a NZE speaker wants to indicate that they have finished speaking while still showing enthusiasm, they may raise the height of the H\* without adjusting the level of L% (as shown in Chapter 7) whereas if

the same speaker wants to indicate finality, they may lower the level of the L% without adjusting the height of the pitch accent (as shown in Chapter 5).

It was suggested at the beginning of this thesis (Chapter 2) that NZE speakers may raise the pitch heights of both their high and low values within an intonation phrase rather than widening their pitch range by raising the high values and lowering the low values. It was claimed that such a manipulation in pitch would permit a NZE speaker to retain their “natural” narrow pitch range within the phrase (Chapter 3) whilst using higher high targets to convey the meanings ascribed to pitch range, such as enthusiasm. It may be the case that NZE speakers adjust an individual tonal target rather than raising the high targets and lowering the low targets in order to retain their Kiwi-ness via their relatively narrow phrase-internal pitch range whilst still conveying the appropriate meaning.

### **8.3. Further considerations**

It should be noted here that – as in any experimental research – the limitations of the procedure limit the interpretation of the results. In this thesis, an experimental approach to the study of NZE intonational phonology was adopted in an attempt to determine the phonological status of certain intonational features in this language variety. The detailed and controlled analyses illustrate a number of semantic contrasts in NZE. However, the restrictive experimental procedures also limit the interpretation of these results. For instance, the results presented in Chapter 6 suggested the existence of a categorical boundary which is determined by IP-Final target height. However it could not be determined whether this category can be moderated by the height of the preceding pitch accent. Further research would be required in order to confirm such suggestions.

As a result of the experimental procedures adopted, many additional questions have been raised during the course of this research. Several of these questions are listed here:

1. (Chapter 3) Are there additional boundary pitch levels other than the two assessed high and low pitch levels which are indicative of NZE?

2. (Chapter 4) How does the content of the reading material and speech style (read aloud speech vs. dialogue) affect the pitch range produced by male and female NZE speakers?
3. (Chapter 4) How are utterance types, such as questions and statements, produced in a reading text in comparison to a natural dialogue between two speakers?
4. (Chapter 5) How are other variables (e.g. overall pitch level, peak alignment, etc) used to convey meaning in this language variety?
5. (Chapter 6) How do other factors (e.g. height of pitch accent and tonal alignment) affect the categorical boundary in NZE?
6. (Chapter 7) How do other factors (e.g. height of pitch accent and tonal alignment) affect the interpretation of conversational cues in NZE?

These questions are not an exhaustive list of the questions raised by the experiments presented in this thesis. They are an indication of research questions that require to be answered in order to fully understand the phonology of NZE intonation. Future research into the meanings and forms of NZE intonation may find it productive to investigate such matters further.

## References

---

- Addington, David W. 1968. The relationship of selected vocal characteristics to personality perception, *Speech Monographs* 35, 492-503.
- Agresti, Alan 2002. *Categorical Data Analysis*, 2<sup>nd</sup> edition. Hoboken: John Wiley and Sons, Inc.
- Ainsworth, Helen 1993. *Rhythm in New Zealand English*. Unpublished BA (Honours) Dissertation, Wellington: Victoria University.
- Ainsworth, Helen 1994. The emergence of the High Rising Terminal contour in the speech of New Zealand children, *Te Reo* 37, 3-20.
- Ainsworth, Helen 2003. How she says it and how he says it - differences in the intonation of dairy farming women and men in South Taranaki, *Wellington Working Papers in Linguistics* 15, 1-15.
- Ainsworth, Helen 2004. *Regional variation in New Zealand English: The Taranaki Sing-song Accent*. PhD Thesis: University of Wellington, New Zealand.
- Allan, Keith 1984. The component functions of the high rise terminal contour in Australian declarative sentences, *Australian Journal of Linguistics* 4, 19-32.
- Allan, Scott 1990. The rise of New Zealand intonation. In Allen Bell and Janet Holmes (eds), *New Zealand Ways of Speaking English*. Clevedon: Multilingual Matters, 115-128.
- Aronovich, Charles D. 1976. The voice of personality: Stereotyped judgements and their relation to voice quality and sex of speaker, *Journal of Social Psychology* 99, 207-220.
- Austin, William M. 1965. Some social aspects of paralanguage, *Canadian Journal of Linguistics* 2, 31-39.
- Barrett, Rusty 1997. The Homo-genius Speech Community. In Anna Livia and Kira Hall (eds), *Queerly Phrased: Language, Gender & Sexuality*. Oxford: Oxford University Press, 181-201.
- Bauer, Laurie 1986. Notes on New Zealand English Phonetics and Phonology, *English World-Wide* 7, 225-258.
- Bauer, Laurie 2000. The dialectal origins of New Zealand English. In Allan Bell and Koenraad Kuiper (eds), *Varieties of English around the World: New Zealand English*. Amsterdam/Philadelphia: John Benjamins, 40-52.
- Beckman, Mary and Gayle Ayers 1994. *Guidelines for ToBI labeling* (Vol 2). Columbus: Ohio State University Linguistics Dept.
- Beckman, Mary and Janet Pierrehumbert 1986. Intonational Structure in Japanese and English, *Phonology Yearbook* 3, 255-309.

- Blum-Kulka, Shoshana. 1983. Interpreting and Performing Speech Acts in a Second Language a Cross-Cultural Study of Hebrew and English. In N. W. E. Judd (ed), *Sociolinguistics and Language Acquisition*. Rowley: Newbury House Publishers, 36-55.
- Boersma, Paul and David Weenink 1996. *PRAAT: A system for doing phonetics by computer*. Amsterdam: University of Amsterdam ([www.praat.org](http://www.praat.org)).
- Bolinger, Dwight 1972. *Intonation*. USA: Penguin Books.
- Bolinger, Dwight 1989. *Intonation and its uses*. Stanford: Stanford University Press.
- Brazil, David 1975. *Discourse intonation. Discourse analysis monographs no. 1*. Birmingham: The University of Birmingham, English Language Research.
- Brend, Ruth 1975. Male-Female Intonation Patterns in American English. In Barrie Thorne & Nancy Henley (eds), *Language and Sex: Difference and Dominance*. MA: Newbury House, 84-87.
- Britain, David 1992. Linguistic change in intonation: the use of HR terminals in New Zealand English, *Language Change and Variation* 4, 77-104.
- Britain, David and John Newman 1992. High rising terminals in New Zealand English, *Journal of the International Phonetic Association* 22, 1-11.
- Brown, Gillian, Karen Currie and Joanne Kenworthy 1980. *Questions of intonation*. Baltimore: University Park Press.
- Brown, James Dean 2001. Pragmatic tests: Different purposes, different tests. In K. Rose and G. Kasper (eds), *Pragmatics in Language Teaching*. Cambridge: Cambridge University Press, 301-325.
- Bruce, Gösta 1977. *Swedish Word Accents in Sentence Perspective*. Lund: Gleerup.
- Carleton, M. and John J. Ohala 1980. The Effect of Pitch of Voice on Perceived Personality Traits, *the Meeting of Kroeber Anthropological Society*, Berkeley.
- Cauldwell, Richard T. 1999. Judgements of attitudinal meanings in isolation and in context, *Phonetics Teaching & Learning Conference 1999*, University College London.
- Caspers, Johanneke 2000. Looking for melodic turn-holding configurations in Dutch, *Linguistics in the Netherlands 2000*, Amsterdam: John Benjamins, 45-55.
- Caspers, Johanneke 2001. Testing the perceptual relevance of syntactic completion and melodic configuration for turn-taking in Dutch, *Proceedings of the 7<sup>th</sup> European Conference on Speech Communication and Technology*, 1395-1398.
- Chafe, Wallace L. 1974. Language and consciousness, *Language* 50, 111-133.

- Chen, Aoji, Carlos Gussenhoven and Toni Rietveld 2002. Language-specific Uses of the Effort Code. In Bernard Bel and Isabelle Marlien (eds), *Proceedings of Speech Prosody 2002*, 215-218.
- Chen, Aoji, Carlos Gussenhoven and Toni Rietveld 2004. Language-Specificity in the Perception of Paralinguistic Intonational Meaning, *Language and Speech* 47, 311-349.
- Chen, Aoji, Toni Rietveld and Carlos Gussenhoven 2001. Language-Specific Effects of Pitch Range on the Perception of Universal Intonational Meaning, *Proceedings of Eurospeech 2001*, 1403-1406.
- Ching, Marvin 1982. The question intonation in assertions, *American Speech* 57, 95-107.
- Cohen, A and J. t Hart 1967. On the anatomy of intonation, *Lingua* 19, 177-192.
- Coulthard, Malcolm 1992. The significance of intonation in discourse. In M. Coulthard (ed), *Advances in Spoken Discourse Analysis*. London: Routledge, 35-49.
- Couper-Kuhlen, Elizabeth 1996. The prosody of repetition on quoting and mimicry. In E. Couper-Kuhlen and M. Selting (eds), *Prosody in Conversation*. Cambridge: Cambridge University Press, 366-405.
- Couper-Kuhlen, Elizabeth and Margret Selting 1996a. *Prosody in Conversation*. Cambridge: Cambridge University Press.
- Couper-Kuhlen, Elizabeth and Margret Selting 1996b. Towards and interactional perspective on prosody and a prosodic perspective on interaction. In E. Couper-Kuhlen and M. Selting (eds), *Prosody in Conversation*. Cambridge: Cambridge University Press, 11-56.
- Cruttenden, Alan 1986. *Intonation*. Cambridge: Cambridge University Press.
- Cruttenden, Alan 1995. Rises in English. In J. Windsor-Lewis (ed), *Studies in General and English Phonetics: Essays in Honour of Professor J.D. O'Connor*. London: Routledge, 155-173.
- Cruttenden, Alan 1997. *Intonation*, 2<sup>nd</sup> edition. Cambridge: Cambridge University Press.
- Cruttenden, Alan 2001. Mancunian intonation and intonational representation, *Phonetica* 58, 53-80.
- Crystal, David 1969. *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press.
- Crystal, David 1975. *The English tone of voice*. London: Arnold.
- Daly, Nicola and Paul Warren 2001. Pitching it differently in New Zealand English: speaker sex and intonation patterns, *Journal of Sociolinguistics* 5, 85-96.
- de Jong, Ken J. 1995. The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation, *Journal of the Acoustic Society of America* 97, 491-504.



- den Ouden, Hanny and Jacques Terken 2001. Measuring pitch range, *Proceedings of Eurospeech 2001* 1, 91-94.
- Eble, Connie 1975. How the speech of some is more equal than others. In B. Thorne & N. Henley (eds), *Language and Sex: Difference and Dominance*. MA: Newbury House, 246.
- Fernald, Anne and Thomas Simon 1984. Expanded intonation contours in mothers' speech to newborns, *Developmental Psychology* 20, 104-113.
- Fernald, Anne, T. Taeschner, J. Dunn, M. Papousek, B. de Boysson-Bardies and I. Fukui 1989. A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants, *Journal of Child Language* 16, 477-501.
- Firenzuoli, Valentina 1999. Ordine e istruzione-Espressione di incredulità e contrasto Descrizione di profili intonativi dal corpus di italiano parlato LABLITA. Atti delle X Giornate del Gruppo di Fonetica Sperimentale "Il parlante e la sua lingua", Istituto Universitario Orientale, Naples, Italy.
- Fletcher, Janet, Esther Grabe and Paul Warren 2004. Intonational variation in four dialects of English: the high rising tone. In Sun-Ah Jun (ed.) *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press, 390-409.
- Fletcher, Janet and Jonathan Harrington 2001. High Rising Terminals in fall-rise tunes in Australian English, *Phonetica* 58, 215-229.
- Gay, T. 1978. Physiological and acoustic correlates of perceived stress. *Language and Speech* 21, 347-353.
- Gibbon, D. (1976). *Perspectives of Intonation Analysis*. Bern: Lang
- Goldman-Eisler, Frieda 1967. Sequential Temporal Patterns and Cognitive Processes in Speech, *Language and Speech* 10, 122-132.
- Goldman-Eisler, Frieda 1968. *Psycholinguistics: Experiments in Spontaneous Speech*. London: Academic.
- Gordon, Matthew 1999. The intonational structure of Chickasaw, *Proceedings of the 14<sup>th</sup> International Congress of the Phonetic Sciences* 3, 1993-1996.
- Grabe, Esther 1998. *Comparative intonational phonology: English and German*. MPI Series in Psycholinguistics 7, Wageningen, Ponsen en Looien.
- Grabe, Esther, Greg Kochanski and John Coleman 2005. The intonation of native accent varieties in the British Isles - potential for miscommunication? In Katarzyna Dziubalska-Kolaczyk and Joanna Przedlacka (eds.), *English pronunciation models: a changing scene*. Linguistic Insights Series, Peter Lang: Bern.
- Grabe, Esther, Francis Nolan and Kimberley Farrar 1998. IViE - a Comparative transcription system for intonational variation in English. *Proceedings of the 5th Conference on Spoken Language Processing* 4, 1259.

- Grabe, Esther, Francis Nolan and Kimberley Farrar 2000. Pitch accent realization in four varieties of British English, *Journal of Phonetics* 28, 161-185.
- Grabe, Esther, Brechtje Post and Francis Nolan 2001. *The IViE Corpus*. Department of Linguistics, University of Cambridge. <http://www.phon.ox.ac.uk/~esther/ivyweb>.
- Graddol, David 1986. Discourse specific pitch behaviour. In C. Johns-Lewis (ed), *Intonation in Discourse*. London: Croom Helm, 221-237.
- Graddol, David and Joan Swann 1983. Speaking fundamental frequency: Some physical and social correlates, *Language and Speech* 26, 351-366.
- Greenberg, Joseph H. 1966. *Universals of language*. 2nd edition. Cambridge, MA: MIT Press.
- Grosz, Barbara and Julia Hirschberg 1992. Some intonational characteristics of discourse structure, *Proceedings of the International Congress on Spoken Language Processing* 19, 867-879.
- Gussenhoven, Carlos 1983. *On the Grammar and Semantics of Sentence Accents*. Dordrecht: Foris.
- Gussenhoven, Carlos 1999. Why question intonations rise and why they sometimes don't, *Second International Conference of the North-West Centre for Linguistics, 'Questions'*. University of Liverpool, 12-14 November.
- Gussenhoven, Carlos 2002. Intonation and interpretation: Phonetics and Phonology. In Bernard Bel and I. Marlin (eds) *Proceedings of the First International Conference on Speech Prosody 2002*, 47-57.
- Gussenhoven, Carlos 2004. *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.
- Gussenhoven, Carlos and Aojun Chen 2000. Universal and Language-Specific Effects in the Perception of Question Intonation, *Proceedings of the International Congress on Spoken Language Processing 2000* Beijing, 91-94.
- Gussenhoven, Carlos and Toni Rietveld 1988. Fundamental Frequency declination in Dutch: testing three hypotheses, *Journal of Phonetics* 16, 355-369.
- Gussenhoven, Carlos and Toni Rietveld 1994. Intonation contours and the prominence of F0 peaks, *Proceedings of the International Congress on Spoken Language Processing* Yokohama 1, 339-342.
- Gussenhoven, Carlos and Toni Rietveld 1998. On the speaker-dependence of the perceived prominence of F0 peaks, *Journal of Phonetics* 26, 371-380.
- Gussenhoven, Carlos and Toni Rietveld 2000. The Behavior of H\* and L\* Under Variations in Pitch Range in Dutch Rising Contours, *Language and Speech* 43, 183-203.
- Gussenhoven, Carlos, Toni Rietveld and Jacques Terken 1999. *ToDI - Transcription of Dutch intonation*, version 1.1. University of Nijmegen and University of Eindhoven (<http://lands.let.kun.nl/todi/todi/home.htm>)

- Guy, Gregory, Barbara Horvath, Julia Vonwiller, Elaine Daisley and Inge Rogers 1986. An intonation change in progress in Australian English, *Language in Society* 15, 23-52.
- Guy, Gregory and Julia Vonwiller 1984. The meaning of an intonation in Australian English, *Australian Journal of Linguistics* 4, 1-17.
- Guy, Gregory and Julia Vonwiller 1989. The high rising tone in Australian English. In P.Collins (ed) *Australian English: the Language of a new Society*. Brisbane: University of Queensland Press, 21-34.
- Haan, Judith and Vincent J. van Heuven 1999. Male vs. female pitch range in Dutch questions, *Proceedings of the 14<sup>th</sup> International Congress of Phonetic Sciences*, San Francisco, 1581-1584.
- Halliday, Michael A.K. 1967. *Intonation and grammar in British English*. The Hague: Mouton.
- t' Hart, Johan, René Collier and Antonie Cohen 1990. *A perceptual study of intonation: An experimental phonetic approach to speech melody*. Cambridge: Cambridge University Press.
- Hays, W.L. 1981. *Statistics*, 3<sup>rd</sup> ed. New York: Holt, Rinehart & Winston.
- Hermes, Dik J. and Joost C. van Gestel 1991. The frequency scale of speech intonation, *Journal of the Acoustical Society of America* 90, 97-102.
- Hirschberg, Julia and Janet Pierrehumbert 1986. The intonational structuring of discourse, *Proceedings of the Twenty-fourth Annual Meeting of the Association for Computational Linguistics*, 136-144.
- Hirschberg, Julia and Gregory Ward 1992. The influence of pitch range, duration, amplitude and spectral features on the interpretation of the rise-fall-rise intonation contour in English, *Journal of Phonetics* 20, 241-251.
- Hirst, Daniel and Albert di Cristo 1998. A survey of intonation systems. In D Hirst and A. Di Cristo (eds), *Intonation Systems: A Survey of Twenty Languages*. Cambridge: Cambridge University Press, 1-44.
- Holmes, Janet 1993. 'New Zealand women are good to talk to': an analysis of politeness strategies in interaction, *Journal of Pragmatics* 20, 91-116.
- Hudson, Amelia and Anthony Holbrook 1981. A study of the reading fundamental vocal frequency of young black adults, *Journal of Speech and Hearing Research* 24, 197-201.
- Johns-Lewis, Catherine 1986a. *Intonation in Discourse*. London: Croom Helm.
- Johns-Lewis, Catherine 1986b. Prosodic differentiation of discourse modes. In C. Johns-Lewis (ed.), *Intonation in discourse*. London: Croom Helm, 199-220.
- Jun, Sun-Ah 2005. *Prosodic Typology, The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press.
- Knowles, G., A. Wichman and P. Alderton 1996. *Working with Speech*. London: Longman.

- Kochanski, Greg, Esther Grabe and John Coleman 2004. The difference between a question and a statement: a survey of English dialects, *The Journal of the Acoustical Society of America* 115, 2398.
- Kochanski, Greg, Esther Grabe, John Coleman and Burton Rosner, accepted. Loudness predicts prominence; fundamental frequency lends little, *Journal of the Acoustical Society of America*.
- Kraayeveld, Hans 1997. *Idiosyncrasy in Prosody: Speaker and speaker group identification in Dutch using melodic and temporal information*. PhD Thesis, University of Nijmegen: the Netherlands.
- Ladd, D. Robert 1980. *The Structure of Intonational Meaning*. Indiana: Indiana University Linguistics Club.
- Ladd, D. Robert 1988. Declination "reset" and hierarchical organization of utterances, *Journal of the Acoustical Society of America* 84, 530-544.
- Ladd, D. Robert 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.
- Ladd, D. Robert and Rachel Morton 1997. The perception of intonational emphasis: continuous or categorical?, *Journal of Phonetics* 25, 313-342.
- Ladefoged, Peter 1976. *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago Press.
- Lakoff, Robin 1975. *Language and Women's Place*. New York: Harper & Row.
- Lehiste, Ilse 1982. Some phonetic characteristics of discourse, *Studia Linguistica* 36, 117-130.
- Lieberman, Mark and Janet Pierrehumbert 1984. Intonational in-variance under changes in pitch range and length. In M. Aronoff and R. Oehrle (eds), *Language Sound Structure*. Cambridge, MA: MIT Press, 157-233.
- Lieberman, Philip 1967. *Intonation, Perception and Language*. Cambridge: MIT Press.
- Local, John K. 1986. Patterns and problems in a study of Tyneside intonation. In C. Johns-Lewis (ed.), *Intonation in Discourse*. London: Croom Helm, 181-198.
- Local, John K., John Kelly and Bill Wells 1986. Towards a phonology of conversation: Turn-taking in Tyneside English, *Journal of Linguistics* 22, 411-437.
- Local, John K., Bill Wells and Mark Sebba 1985. Phonology for conversation: Phonetic aspects of turn-delimitation in London Jamaican, *Journal of Pragmatics* 9, 309-330.
- Meyerhoff, Miriam 1991. *Grounding and overcoming obstacles: the positive politeness motivations of High Rise Terminals*, Wellington, NZ: manuscript.
- O'Connor, J. D., and G.F. Arnold 1961. *Intonation of Colloquial English: A Practical Handbook*. London: Longman.
- Ohala, John J. 1983. Cross-Language Use of Pitch: An Ethological View, *Phonetica* 40, 1-18.

- Ohala, John J. 1984. An ethological perspective on common cross-language u-tilization of F0 in voice, *Phonetica* 41, 1-16.
- Ohala, John J. 1994. The frequency code underlies the sound-symbolic use of voice pitch. In L. Hinton, J. Nichols and J. Ohala (eds), *Sound Symbolisms*. Cambridge: Cambridge University Press, 325-347.
- Olshtain, Elite and Shoshana Blum-Kulka 1984. Cross-Linguistic Speech Act Studies: Theoretical and Empirical Issues. In L. MacMathuna and D. Singleton (eds), *Language across Cultures*. Dublin: IRAAL, 235-248.
- Pierrehumbert, Janet 1979. The perception of fundamental frequency declination, *Journal of the Acoustical Society of America* 66, 363-369.
- Pierrehumbert, Janet 1980. *The Phonology and Phonetics of English Intonation*. PhD Thesis: MIT, Cambridge, USA.
- Pierrehumbert, Janet 1998. Tonal elements and their alignment. In M. Horne, (ed) *Prosody: Theory and Experiment. Studies Presented to Gosta Bruce*. Kluwer: Dordrecht
- Pierrehumbert, Janet and Julia Hirschberg 1990. The meaning of intonation contours in the interpretation of discourse. In M. Pollack & P. Cohen (eds), *Plans and Intentions in Communication and Discourse*. Cambridge: MIT Press, 160-197.
- Pike, Kenneth L. 1945. *The Intonation of American English*. Michigan: University of Michigan Press.
- Post, Brechtje 2000. *Tonal and phrasal structures in French intonation*. The Hague, Holland: Academic Graphics.
- Rice, John A. 1995. *Mathematical Statistics and Data Analysis*, 2<sup>nd</sup> edition. Belmont, CA: Duxbury.
- Rietveld, Toni and Patricia Vermillion 2001. Cues for Perceived Pitch Register, *Proceedings of Eurospeech 2001*, 1399-1402.
- Rietveld, Toni and Patricia Vermillion 2003. Cues for Perceived Pitch Register, *Phonetica* 60, 261-272.
- Sacks, Harvey, Emmanuel Schegloff and Gail Jefferson 1974. A simplest systematics for the organization of turntaking in conversation, *Language* 50, 696-735.
- Scherer, Klaus R. 1979. Personality markers in speech. In K. Scherer and H. Giles (eds), *Social Markers in Speech*. Cambridge: Cambridge University Press, 147-209.
- Silverman, Kim, Mary Beckman, John F. Pitrelli, Mari Ostendorf, Colin Wightman, Patti Price, Janet Pierrehumbert and Julia Hirschberg 1992. ToBi: A Standard for Labeling English Prosody, in *Proceedings of the Second International Conference on Spoken Language Processing*, 286-270.
- Sluijter, Agaath M. and Jacques Terken 1993. Beyond sentence prosody: paragraph intonation in Dutch, *Phonetica* 50, 180-188.



- Swerts, Marc and Julia Hirschberg 1998. Prosody and Conversation: An introduction, *Language and Speech* 41, 229-234.
- Uldall, Elizabeth 1960. Attitudinal Meanings conveyed by intonation contours, *Language and Speech* 2, 223-234.
- Uldall, Elizabeth 1964. Dimensions of meaning in intonation. In D. B. F. D.N Abercrombie, P.A.D. MacCarthy, N.C. Scott and J.L.M Trim (eds), *In Honour of Daniel Jones: Papers Contributed on the Occasion of his Eightieth Birthday*. London: Longman, 271-279.
- Van Bezooijen, Renee 1984. *Characteristics and recognizability of vocal expressions of emotion*. Dordrecht: Foris.
- van Leeuwen, Theo 1999. *Speech, music, sound*. London: Macmillan.
- van den Berg, Rob, Carlos Gussenhoven and Toni Rietveld 1992. Downstep in Dutch: Implications for a model. In G. J. Docherty and D. R. Ladd (eds), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*. Cambridge: Cambridge University Press, 335-359.
- Vermillion, Patricia 2000. Analysing the function of intonation in the discourse of men and women, *NWAV* 29, Michigan State University, USA.
- Vermillion, Patricia 2001. *The Perception and Production of Intonational Meaning by British Men and Women*. MPhil Thesis: University of London, UK.
- Vermillion, Patricia 2003. The Ups and Downs of Kiwis: An experiment investigating tonal cues which are used to identify NZE intonation, *Wellington Working Papers in Linguistics* 15, 17-30.
- Warren, Paul 1998. Timing patterns in New Zealand English rhythm, *Te Reo* 41, 80-93.
- Warren, Paul 1999. Timing properties of New Zealand English rhythm, *Proceedings of the 14<sup>th</sup> International Congress of Phonetic Sciences*, San Francisco, 1843-1848.
- Warren, Paul 2005a. Patterns of late rising in New Zealand: intonational variation or intonational change?, *Language Variation and Change* 17, 209-230.
- Warren, Paul 2005b. Issues in the study of intonation in language varieties, *Language and Speech* 48, 345-358.
- Warren, Paul and David Britain 2000. Intonation and Prosody in New Zealand English. In Allan Bell and Koenraad Kuiper (eds), *Varieties of English around the World: New Zealand English*. Amsterdam/Philadelphia: John Benjamins, 146-172.
- Warren, Paul and Nicola Daly 2000. Sex as a factor in rises in New Zealand English. In J. Holmes (ed), *Gendered Speech in Social Context: Perspectives from Gown and Town*. Wellington: Victoria University Press, 99-115.
- Warren, Paul and Nicola Daly 2005. Characterizing New Zealand English Intonation: broad and narrow analysis. In A. Bell, R. Harlow & D. Starks

- (eds) *Languages of New Zealand*. Wellington: Victoria University Press, 217-237.
- Wells, John C. 1982. *Accents of English*. Cambridge: Cambridge University Press.
- Wells, Bill and Sue Peppé 1996. Ending up in Ulster: prosody and turn-taking in English dialects'. In E. Couper-Kuhlen and M. Selting (eds) *Prosody in Conversation*. Cambridge: Cambridge University Press, 101-130.
- Wichmann, Anne and Johanneke Caspers 2001. Melodic cues to turn-taking in English: evidence from perception. In J. van Kuppeveldt and R. Smith (eds), *Proceedings of the 2nd SIGdial Workshop on Discourse and Dialogue*, 180-185.
- Zraick, Richard and Julie M. Liss 2000. A comparison of equal-appearing interval scaling and direct magnitude estimation of nasal voice quality, *Journal of Speech, Language and Hearing Research* 43, 979-988.
- Zwartz, Joel and Paul Warren 2003. This is a statement? Lateness of rise as a factor in listener interpretation of HRTs, *Wellington Working Papers in Linguistics* 15, 51-62.



## Appendices

---

<b>A.</b>	<b>Chapter 3 experimental materials</b>	<b>179</b>
	1. Recording materials used for source utterance and instructions	179
	2. Consent and feedback forms	180
	3. Instructions	182
	4. Questionnaire and answersheet (Perception experiment)	183
	5. Questionnaire and answersheet (Attitude experiment)	187
<b>B.</b>	<b>Chapter 4 and 5 experimental materials</b>	<b>194</b>
	1. Consent forms	194
	2. Questionnaire	196
	3. Instructions (Chapter 4)	197
	4. Instructions (Chapter 5)	197
	5. Recording material used for scene	197
<b>C.</b>	<b>Chapter 6 experimental materials</b>	<b>198</b>
	1. Consent and feedback forms	198
	2. Recording material used for source utterance	200
	3. Instructions (Task 1)	200
	4. Instructions (Task 2)	200
	5. Questionnaire and answersheet	202
<b>D.</b>	<b>Chapter 7 experimental materials</b>	<b>208</b>
	1. Consent and feedback forms	208
	2. Recording material used for source utterance	210
	3. Instructions (Task 1)	210
	4. Instructions (Task 2)	211
	5. Questionnaire and answersheet	213

**A1. Recording material used for source utterance and instructions  
(different speakers)**

---

- a. Read the following sentences aloud *as naturally as possible*:

**Manny won in rummy.**

- b. Read the following sentences aloud *as naturally as possible*:

Rocca computer's vision is to empower people through great software - any time, any place and on any device. As one of the worldwide leaders in software development for business computing, Rocca strives to produce innovative products and services that meet our customers' evolving needs. In order to "be there" for our British and New Zealand customers, we have satellite offices in both countries. Our largest offices are in Wellington, New Zealand and Reading, England.

## A2. Consent and feedback forms

---

Memorandum to: Class members  
From: Patricia Vermillion  
PhD student  
School of Linguistics and Applied Language Studies  
Re: Experiment on New Zealand English intonation

During your *Introduction to Language Studies tutorial* during the week starting 4 October until the 12 October, I will be conducting an experiment on New Zealand English intonation. You will hear a number of speakers saying one sentence, which will be played over a loud speaker. If you would like to participate in the experiment, you will be asked to indicate whether the speaker is a native New Zealand speaker or not. We envisage the experiment to take approximately 10 minutes.

As this will complement your lesson, I hope that you will decide to take part in this 10-minute exercise. However, you should be aware of the following (Human Ethics Committee Guidelines, pg.21):

- Your participation is voluntary
- Your participation or non-participation will not be recorded or regarded by the department in either a positive or a negative way
- If you participate, you have a right to withdraw at any time and you do not have to state a reason
- Your consent to participate is implied by your filling out the questionnaire.

The current Human Ethics Committee Guidelines can be found at:  
<http://www.vuw.ac.nz/home/research/index.html#hec>

If you would like to take part in the experiment, but not include your questionnaire as part of the collected data, you may withhold your questionnaire after the experiment is completed. Responses from this experiment will serve as the starting point of my Ph.D. thesis, and consequently be a part of a written report and possibly a publication. The questionnaire will include questions such as whether your first language is New Zealand English or whether you are from the Wellington area. However, it will not be possible to identify who completed which questionnaire, as the questionnaires are anonymous. The collected data will be included as group scores only. No other person, besides my supervisor, Dr. Paul Warren, and me will see the completed questionnaires. The thesis will be submitted for marking to examiners and subsequently a copy will be lodged in the Victoria University of Wellington Library, and all datasheets will be destroyed at the completion of the 3-year project.

If you have any queries, please feel free to contact me via email ([Patricia.Vermillion@vuw.ac.nz](mailto:Patricia.Vermillion@vuw.ac.nz)) or by telephone (04-463-5632).

## **Feedback for Participants taking part in Experiment on New Zealand Intonation**

I am a PhD student in the School of Linguistics and Applied Language Studies at Victoria University of Wellington. As part of my degree, I am conducting a series of experiments which are investigation New Zealand intonation. As part of your tutorial today, you have just taken part in one of my experiments.

The aim of the experiment was to investigate which features of intonation New Zealand speakers use when distinguishing New Zealand English. The stimuli which have heard during the experiment was derived from one native New Zealand English speaker. The recorded utterance was *Manny won in rummy*. Using a computer program, this utterance was low pass filtered and manipulated in various ways in order to assess which tonal features you were using when distinguishing New Zealand English intonation. Specifically, we manipulated several tonal features, including height of H and L values, height and type of utterance onset and coda, peak alignment, pitch dynamism and rhythm. This information will be used to determine whether you use these cues as determiners for New Zealand English.

If you have any queries regarding this experiment, please feel free to contact me at [Patricia.Vermillion@vuw.ac.nz](mailto:Patricia.Vermillion@vuw.ac.nz). Thank you for your time and participation in this experiment.

Patricia Vermillion  
School of Linguistics and Applied Language Studies  
Victoria University of Wellington

### A3. Instructions

---

The aim of this experiment is to find out what makes New Zealand English sound like New Zealand English. We are doing this in an attempt to synthesise a realistic sounding New Zealand accent on the computer. In order to make an utterance created by a computer sound as if it were spoken by a Kiwi, however, we need to first find out what a Kiwi accent is.

In this experiment you will hear a number of utterances. Each utterance was read aloud by one of two British English speakers: one male and one female. The utterance that you will hear have been changed in some way. We want you to tell us whether the utterances that you hear sound like a Kiwi or not.

Look on page 2 of your questionnaire. (5secs). You will see two example scales on the page. With your pen or pencil, mark on the scale given how much the utterance sounds like a New Zealand English speaker. I should also tell you that the utterances have been filtered. In other words, you cannot distinguish *what* the speaker is saying. We did this so that the words would not influence your perceptions. However, you will still be able to hear important aspects of each utterance. For example, you will hear **[APPENDIX A1a]** or **[APPENDIX A1a]**. If you think the utterance sounds very much like a Kiwi, you may indicate this by marking toward the right side of the line how much more you think that it sounded like a Kiwi. If you think the utterance does not sound like a New Zealand English speaker, mark on the line to the left side of the scale. If it sounds nothing like a New Zealand English speaker, mark on the far left of the line. The space in the middle is for variations therein: sounds "somewhat" like a New Zealand English speaker. Please remember that your mark must *cross* the line, just like in the two examples.

The variety of British English that we used in this experiment is the standard variety. Here is an example of the type of British speech that the utterances was taken from **[APPENDIX A1b, British English]**.

And here is an example of the variety of New Zealand English that we are looking for. Our sample speaker is reading from the same script as our British speaker. She is Pakeha, from the Wellington area of New Zealand and who has a University level education. **[APPENDIX A1b, New Zealand English]**

Turn to the next page of your questionnaire. You will see a number of scales, just as in the examples. On these scales, mark on the line how much each utterance sounds like a New Zealand English speaker. There is a 4 second delay between each utterance which should be used to indicate your impressions on the given scales. At the end of every page, you will hear a short beep. This was done so that you know whether you are on the right number or not. If you find that you were on the wrong number, put an X over the page. Also, make sure that you indicate your opinions on all of the utterance, even if you are uncertain. We want to check your guesses too. Finally, in order to make the task a little easier for you, you will hear all of the female speakers first and then male speakers second.

If you have any questions, please ask them now. Otherwise, the experiment will begin in 10 seconds. Good luck.

#### A4. Questionnaire and Answersheet (Percetion experiment)

---

##### *Questionnaire*

1. How old are you?

☐

Up to 17 years old

☐

18-22

☐

23-29

☐

30-35

☐

36 or older

2. Are you

☐

Male

or

☐

Female

3. Is your first language New Zealand English?

☐

Yes

or

☐

No

4. Are you a full-time or part-time university student?

☐

Full-time

or

☐

Part-time

5. Do you have any known hearing impairments?

☐

Yes

or

☐

No

Please try to remember your answers for next week. If you believe you may not remember your answers, please write them down somewhere.

1. What is your favourite colour? \_\_\_\_\_

2. What is your favourite number? \_\_\_\_\_

3. What is your favourite name? \_\_\_\_\_

4. What is your favourite animal? \_\_\_\_\_

### Instructions

Please mark on the following lines how much each utterance sounds like a New Zealand English speaker. The far right of the line is if the second utterance sounds very much like a New Zealand English speaker, whereas the far left indicates that the utterance sounds nothing like a New Zealand English speaker. The space in the middle is for variations therein: sounds “somewhat” like a New Zealand English speaker. For example, if the utterance sounds quite a bit like a Kiwi, you may mark the following:

**EXAMPLE:**

**EXAMPLE:**

Does **not** sound like a NZE speaker

Somewhat like a NZE speaker

Sounds **very much** like a NZE speaker

However, if you think that the utterance sounds nothing like a New Zealander, then you mark the following on the line:

**EXAMPLE:**

**EXAMPLE:**

Does **not** sound like a NZE speaker

Sounds **very much** like a NZE speaker

Somewhat like a NZE speaker

**Please remember, make sure that your mark CROSSES the line**



FEMALE SPEAKERS

Does <b>not</b> sound like a NZE speaker			Sounds <b>very much</b> like a NZE speaker
Somewhat like a NZE speaker			

Does <b>not</b> sound like a NZE speaker			Sounds <b>very much</b> like a NZE speaker
Somewhat like a NZE speaker			

Does <b>not</b> sound like a NZE speaker			Sounds <b>very much</b> like a NZE speaker
Somewhat like a NZE speaker			

Does <b>not</b> sound like a NZE speaker			Sounds <b>very much</b> like a NZE speaker
Somewhat like a NZE speaker			

Does <b>not</b> sound like a NZE speaker			Sounds <b>very much</b> like a NZE speaker
Somewhat like a NZE speaker			

MALE SPEAKERS

Does <b>not</b> sound like a NZE speaker			Sounds <b>very much</b> like a NZE speaker
Somewhat like a NZE speaker			
Does <b>not</b> sound like a NZE speaker			Sounds <b>very much</b> like a NZE speaker
Somewhat like a NZE speaker			
Does <b>not</b> sound like a NZE speaker			Sounds <b>very much</b> like a NZE speaker
Somewhat like a NZE speaker			
Does <b>not</b> sound like a NZE speaker			Sounds <b>very much</b> like a NZE speaker
Somewhat like a NZE speaker			
Does <b>not</b> sound like a NZE speaker			Sounds <b>very much</b> like a NZE speaker
Somewhat like a NZE speaker			

**A5. Questionnaire and Answersheet (Attitude experiment)**

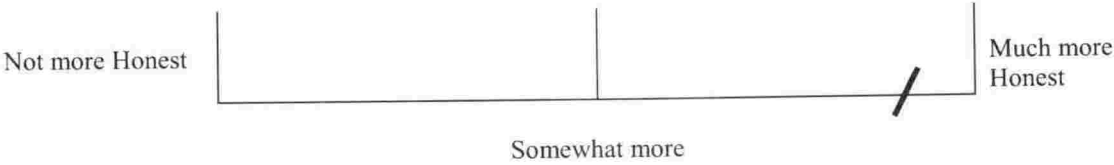
**Questionnaire**

TRY TO REMEMBER YOUR ANSWERS TO THE FOLLOWING FOUR QUESTIONS. IF NECESSARY, WRITE THEM DOWN. YOUR FOUR ANSWERS WILL BE USED FOR MATCHING UP THIS QUESTIONNAIRE WITH A SECOND ONE, WHICH WILL BE COMPLETED IN CLASS IN SEVERAL WEEKS TIME.

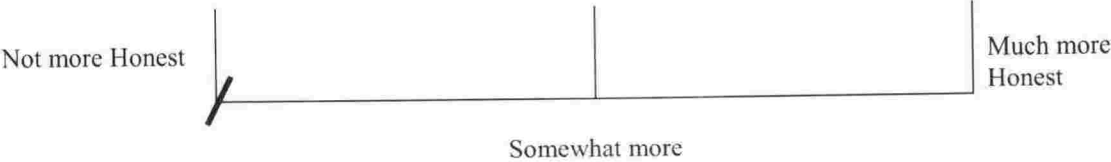
- 5. What is your favourite colour? \_\_\_\_\_
- 6. What is your favourite number? \_\_\_\_\_
- 7. What is your favourite name? \_\_\_\_\_
- 8. What is your favourite animal? \_\_\_\_\_

NEXT, TRY TO GIVE YOUR OPINION ABOUT NEW ZEALAND PEOPLE IN COMPARISON TO BRITISH PEOPLE.

If you believe Kiwi men are more honest than British men, mark on the scale how much more honest you believe them to be. For example:



If you *don't* think New Zealand men are more honest than British men (*they are equal with regards to honesty*), mark on the line:



**MAKE SURE YOUR MARK CROSSES THE LINE!!!**

**AND REMEMBER – BE HONEST!**

**KIWI MEN ARE \_\_\_\_\_ THAN BRITISH MEN:**

<i>Kiwi men are</i>		<i>than British men</i>
Not more LIVELY		Much more LIVELY
Somewhat more		
<i>Kiwi men are</i>		<i>than British men</i>
Not more DULL		Much more DULL
Somewhat more		
<i>Kiwi men are</i>		<i>than British men</i>
Not more ENCOURAGING		Much more ENCOURAGING
Somewhat more		
<i>Kiwi men are</i>		<i>than British men</i>
Not more GLOOMY		Much more GLOOMY
Somewhat more		
<i>Kiwi men are</i>		<i>than British men</i>
Not more RESENTFUL		Much more RESENTFUL
Somewhat more		
<i>Kiwi men are</i>		<i>than British men</i>
Not more APPRECIATIVE		Much more APPRECIATIVE
Somewhat more		
<i>Kiwi men are</i>		<i>than British men</i>
Not more APOLOGETIC		Much more APOLOGETIC
Somewhat more		

<i>Kiwi men are</i>		<i>than British men</i>
Not more STRONG WILLED		Much more STRONG WILLED
Somewhat more		
<i>Kiwi men are</i>		<i>than British men</i>
Not more CONFIDENT		Much more CONFIDENT
Somewhat more		
<i>Kiwi men are</i>		<i>than British men</i>
Not more SELF DOUBTING		Much more SELF DOUBTING
Somewhat more		
<i>Kiwi men are</i>		<i>than British men</i>
Not more DOMINANT		Much more DOMINANT
Somewhat more		
<i>Kiwi men are</i>		<i>than British men</i>
Not more SUBMISSIVE		Much more SUBMISSIVE
Somewhat more		
<i>Kiwi men are</i>		<i>than British men</i>
Not more THREATENING		Much more THREATENING
Somewhat more		
<i>Kiwi men are</i>		<i>than British men</i>
Not more WELCOMING		Much more WELCOMING
Somewhat more		

<i>Kiwi men are</i>		<i>than British men</i>
Not more INSISTENT		Much more INSISTENT
Somewhat more		
<i>Kiwi men are</i>		<i>than British men</i>
Not more COMPLIANT		Much more COMPLIANT
Somewhat more		

**KIWI WOMEN ARE \_\_\_\_\_ THAN BRITISH WOMEN:**

<i>Kiwi women are</i>		<i>than British women</i>
Not more LIVELY		Much more LIVELY
Somewhat more		
<i>Kiwi women are</i>		<i>than British women</i>
Not more DULL		Much more DULL
Somewhat more		
<i>Kiwi women are</i>		<i>than British women</i>
Not more ENCOURAGING		Much more ENCOURAGING
Somewhat more		
<i>Kiwi women are</i>		<i>than British women</i>
Not more GLOOMY		Much more GLOOMY
Somewhat more		
<i>Kiwi women are</i>		<i>than British women</i>
Not more RESENTFUL		Much more RESENTFUL
Somewhat more		
<i>Kiwi women are</i>		<i>than British women</i>
Not more APPRECIATIVE		Much more APPRECIATIVE
Somewhat more		
<i>Kiwi women are</i>		<i>than British women</i>
Not more APOLOGETIC		Much more APOLOGETIC
Somewhat more		



<i>Kiwi women are</i>		<i>than British women</i>
Not more STRONG WILLED		Much more STRONG WILLED
Somewhat more		
<i>Kiwi women are</i>		<i>than British women</i>
Not more CONFIDENT		Much more CONFIDENT
Somewhat more		
<i>Kiwi women are</i>		<i>than British women</i>
Not more SELF DOUBTING		Much more SELF DOUBTING
Somewhat more		
<i>Kiwi women are</i>		<i>than British women</i>
Not more DOMINANT		Much more DOMINANT
Somewhat more		
<i>Kiwi women are</i>		<i>than British women</i>
Not more SUBMISSIVE		Much more SUBMISSIVE
Somewhat more		
<i>Kiwi women are</i>		<i>than British women</i>
Not more THREATENING		Much more THREATENING
Somewhat more		
<i>Kiwi women are</i>		<i>than British women</i>
Not more WELCOMING		Much more WELCOMING
Somewhat more		

<i>Kiwi women are</i>		<i>than British women</i>
Not more INSISTENT		Much more INSISTENT
Somewhat more		
<i>Kiwi women are</i>		<i>than British women</i>
Not more COMPLIANT		Much more COMPLIANT
Somewhat more		

THANK YOU! THANK YOU! THANK YOU! THANK YOU! THANK YOU!

---

**B1. Consent forms**

---

**Consent Forms**

Thank you for agreeing to participate in this study. The recording sessions that you will take part in are all to be carried out in English, and consist of the following two parts:

- (i) You will be asked to read aloud a number of sentences from sheets that will be given to you. Each sentence includes a prompt, which should be used as a guide on how to read the sentence.
- (ii) You will be asked to read out a paragraph from sheets that will be given to you.

For each of these tasks, we ask you to speak clearly, but as naturally as possible. We are interested in normal conversational speech, rather than careful 'newsreader-style' speech, but it is important for our analysis that the recordings are clear.

A researcher will remain in the room with you while you make these recordings. The entire recording session should last about 10-15 minutes. There will be an opportunity for questions prior to the commencement of the recording, as well as at the end of the session.

Your recording will be analysed in the School of Linguistics and Applied Language Studies, VUW, and may contribute to research publications, which will be publicly available. Your identity will remain anonymous, and no personal data will be stored with the recordings. As recompense for your time, you will be offered a \$5 music or book voucher.

Patricia Vermillion  
School of Linguistics and Applied Language Studies  
Victoria University of Wellington  
Ph. 4635632  
Patricia.Vermillion@vuw.ac.nz

## CONSENT TO PARTICIPATION IN RESEARCH

NAME (please print clearly): \_\_\_\_\_

I understand the procedure of this research as set out in the INFORMATION SHEET. I have had an opportunity to ask questions and have had them answered to my satisfaction.

I understand that recordings of my speech, together with written transcriptions, will be analysed and serve as part of a PhD thesis, as well as submitted to conferences and journal publications. I also understand that my contribution will remain anonymous, and that no personal data that may be used to identify me uniquely will be stored with the recordings and transcriptions.

I understand that I may withdraw myself or any information I have provided from this project, without having to give reasons.

I understand that, if I so request, details of the purpose of the research will be discussed with me once I have completed my participation, and that I may ask for a summary of results from the study to be sent to me at a later date.

I agree to take part in this research.

Signed: \_\_\_\_\_

Date: \_\_\_\_\_

Please tick here if you wish to discuss with the experimenter the purpose of the research after you have completed your participation:

☐

Please tick here if you wish to receive a copy of a short summary of the study (and give below an address to which this summary can be sent):

☐

## B2. Questionnaire

---

### *Questionnaire*

1. How old are you? \_\_\_\_\_ years

2. Are you ☐ Male or ☐ Female

3. Is your first language New Zealand English?

☐ Yes or ☐ No

4. Are you a full-time or part-time university student?

☐ Full-time or ☐ Part-time

5. What is your ethnicity?

☐ Māori      ☐ mostly Māori  
with some Pākehā      ☐ mostly Pākehā  
with some Māori      ☐ Pākehā      ☐ Other  
Specify: \_\_\_\_\_

5. Do you have any known hearing impairments?

☐ Yes or ☐ No

### B3. Instructions (Chapter 4)

---

Please read silently index card 2 and become familiar with the words. Please make sure that you understand this card. If you have any questions, please feel free to stop and ask.

Index card 2 is a monologue. I want you to imagine that you are talking to your partner about your friend George. When you feel that you are ready to read the card aloud, as if you were speaking to your partner, please do so as naturally as possible.

### B4. Instructions (Chapter 5)

---

In this session, I want you to imagine that you are having a conversation with your partner. For example, when you hear "What did you think of the film?", imagine that your *partner* has just said this to you. In response, you should **read aloud**, as naturally as possible, what is on the given index cards – just as you would say in response to your partner.

Let's look at the example on the first index card. First you will hear "What did you think of the film?"

In response, I want you to read aloud the first sentence on the card according to the prompt. For example, the first prompt wants you to **emphasise** how much you enjoyed the film. *I* may say such a response to my partner like this: **It was helpful.**

If I am asked to say the same sentence **without** emphasising, I might say it like this:

**It was helpful.**

Let's practice. First you will hear: "What did you think of the film?"

Then you will read aloud from the card the two sentences according to the prompts. Go ahead, give it a go. Let's begin.

### B5. Recording material used for scene (Chapter 5)

---

Please read the following sentences aloud, as naturally as possible and without emphasis, as you would if you were speaking to your partner:

**Did you enjoy the film?**

**I would like to see the film again. Would you like to join me next week?"**

**I want to go out to dinner this weekend. Would you like to join me?"**

**I told my boss off at work today. Did I tell you?**

---

**C1. Consent Forms**

---

Memorandum to: Class members

From: Patricia Vermillion  
PhD student  
School of Linguistics and Applied Language Studies

Re: Experiment on New Zealand English intonation

During the week starting 23 September, I will be conducting an experiment on New Zealand English intonation. You will hear a number of speakers saying two different sentences, which will be played over a headset. For each sentence, you will be asked to make a judgement about what you hear. We envisage the experiment will take approximately 20 minutes. For your time, you will be provided with a \$10 book/music voucher.

You should be aware of the following (VUW Human Ethics Committee Guidelines, pg.21):

- Your participation is voluntary
- Your participation or non-participation will not be recorded or regarded by the school in either a positive or a negative way
- If you participate, you have a right to withdraw at any time and you do not have to state a reason
- Your consent to participate is implied by your filling out a questionnaire.

The current Human Ethics Committee Guidelines can be found at:  
<http://www.vuw.ac.nz/home/research/index.html#hec>

Responses from this experiment will be a part of written reports, such as my PhD thesis and possibly publications. The questionnaire will include questions such as whether your first language is New Zealand English. However, it will not be possible to identify who completed which questionnaire, as the questionnaires are anonymous. The PhD thesis will be submitted for marking to examiners and subsequently a copy will be lodged in the Victoria University of Wellington Library, and all datasheets will be destroyed at the completion of the 3-year project.

If you have any queries, please feel free to contact me via email ([Patricia.Vermillion@vuw.ac.nz](mailto:Patricia.Vermillion@vuw.ac.nz)) or by telephone (04-463-5621).

## **Feedback for Participants taking part in Experiment on New Zealand Intonation**

I am a PhD student in the School of Linguistics and Applied Language Studies at Victoria University of Wellington. As part of my degree, I am conducting a series of experiments which are investigation New Zealand intonation. Thank you for agreeing to participate in this study.

The aim of the experiment was to investigate what meanings are conveyed by the intonation of a phrase-final rise or fall. The stimuli which you have heard during the experiment were derived from one native New Zealand English speaker. The recorded utterance was *Manny won in Rummy*. Using a computer program, this utterance was manipulated in various ways in order to assess which meanings are conveyed by the different height of the phrase final point. This information will be used to determine which meanings are conveyed by this intonation for New Zealand English speakers.

If you have any queries regarding this experiment, please feel free to contact me at [Patricia.Vermillion@vuw.ac.nz](mailto:Patricia.Vermillion@vuw.ac.nz). Thank you for your time and participation in this experiment.

Patricia Vermillion  
School of Linguistics and Applied Language Studies  
Victoria University of Wellington

---



## C2. Recording material used for source utterance

---

Please read aloud the following sentence as naturally as possible:

**Manny won in rummy.**

## C3. Instructions (Task 1)

---

If you can hear this, please raise your hand now.

Thank you for agreeing to take part in this experiment. If you haven't already completed the questionnaire, please do so after the experiment has ended.

You will hear two sentences spoken aloud by a Kiwi man. For example, you will hear [insert sample sound here]. We want you to listen to the pairs of sentences and determine whether the first and second sentences are identical, which you would indicate by ticking the SAME box, or whether the final pitch point is different in the two sentences, which you would indicate by ticking the DIFFERENT box. Even if you are uncertain, make sure to mark ALL of the given scales --- In other words, do not skip any of the sets of sentences. Finally, to make sure you are keeping up with the experiment, you will hear a beep after seven pairs of sentences. If you find that you were not keeping up, put a large X over the previous page. If you have any questions, please raise your hand now. Otherwise, the experiment will begin in 10 seconds

## C4. Instructions (Task 2)

---

During this experiment, you will hear a speaker saying two sentences, which will be played over your headset. For each group of two sentences, you will be asked to make a comparison judgement about what you hear. Specifically, you will be asked to tell how much more lively, confident or insistent the second sentence sounds in comparison to the first sentence.

You should see two example scales in front of you. We want you to mark on the given scale how much more or less lively the second sentence sounds in comparison to the first sentence. For example, you will hear [insert sample]

The far right of the line should be marked if you believe the second sentence sounds much **more** lively than the first sentence, such as in example 1. The far left should be marked if you believe the second sentence sounds much **less** lively than the first sentence, as shown in Example 2. The centre point is to be marked if you think the second sentence conveys the **same** degree of liveliness as the first sentence. Everything in between these points is for variations therein of perceived liveliness. Please remember that your mark

**must cross** the line! If you have any questions, please raise your hand now. Otherwise, please turn to the next page where you will see more scales.

Even if you are uncertain, make sure to mark ALL of the given scales --- In other words, do not skip any of the sets of sentences. Finally, to make sure you are keeping up with the experiment, you will hear a beep at the beginning of each new page. If you find that you were not keeping up, put a large X over the previous page. If you have any questions, please raise your hand now. Otherwise, the experiment will begin in 10 seconds

For each group of two sentences, please mark on the given scales how much more or less lively the second sentence is in comparison to the first.

For each group of two sentences, please mark on the given scales how much more or less confident the second sentence is in comparison to the first.

For each group of two sentences, please mark on the given scales how much more or less insistent the second sentence is in comparison to the first.

C5. Questionnaire and Answersheet

---

Questionnaire

1. How old are you?

☐

Up to 17 years old

☐

18-22

☐

23-27

☐

28-35

☐

36 or older

2. Are you

☐

Male

or

☐

Female

3. Is your first language New Zealand English?

☐

Yes

or

☐

No

4. Are you a full-time or part-time university student?

☐

Full-time

or

☐

Part-time

5. Do you have any known hearing impairments?

☐

Yes

or

☐

No

Is the *second sound* the **SAME** or **DIFFERENT** from the *first sound*?

- |    |                               |                                    |
|----|-------------------------------|------------------------------------|
| 1. | Same <input type="checkbox"/> | Different <input type="checkbox"/> |
| 2. | Same <input type="checkbox"/> | Different <input type="checkbox"/> |
| 3. | Same <input type="checkbox"/> | Different <input type="checkbox"/> |
| 4. | Same <input type="checkbox"/> | Different <input type="checkbox"/> |
| 5. | Same <input type="checkbox"/> | Different <input type="checkbox"/> |
| 6. | Same <input type="checkbox"/> | Different <input type="checkbox"/> |
| 7. | Same <input type="checkbox"/> | Different <input type="checkbox"/> |

**BEEP** ----- **BEEP**

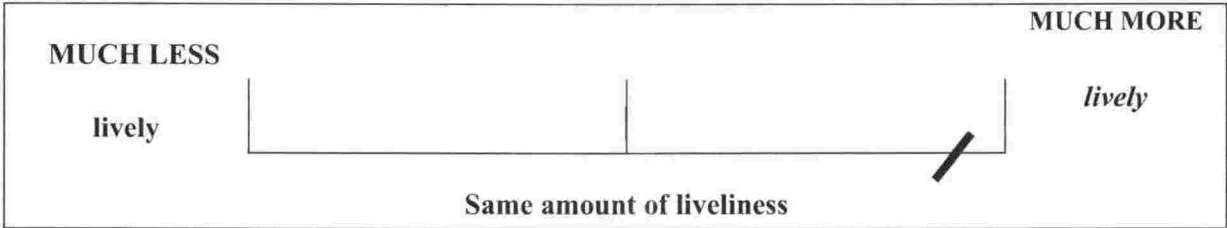
- |     |                               |                                    |
|-----|-------------------------------|------------------------------------|
| 8.  | Same <input type="checkbox"/> | Different <input type="checkbox"/> |
| 9.  | Same <input type="checkbox"/> | Different <input type="checkbox"/> |
| 10. | Same <input type="checkbox"/> | Different <input type="checkbox"/> |
| 11. | Same <input type="checkbox"/> | Different <input type="checkbox"/> |
| 12. | Same <input type="checkbox"/> | Different <input type="checkbox"/> |
| 13. | Same <input type="checkbox"/> | Different <input type="checkbox"/> |
| 14. | Same <input type="checkbox"/> | Different <input type="checkbox"/> |

**BEEP** ----- **BEEP**

- |     |                               |                                    |
|-----|-------------------------------|------------------------------------|
| 15. | Same <input type="checkbox"/> | Different <input type="checkbox"/> |
| 16. | Same <input type="checkbox"/> | Different <input type="checkbox"/> |
| 17. | Same <input type="checkbox"/> | Different <input type="checkbox"/> |
| 18. | Same <input type="checkbox"/> | Different <input type="checkbox"/> |
| 19. | Same <input type="checkbox"/> | Different <input type="checkbox"/> |
| 20. | Same <input type="checkbox"/> | Different <input type="checkbox"/> |

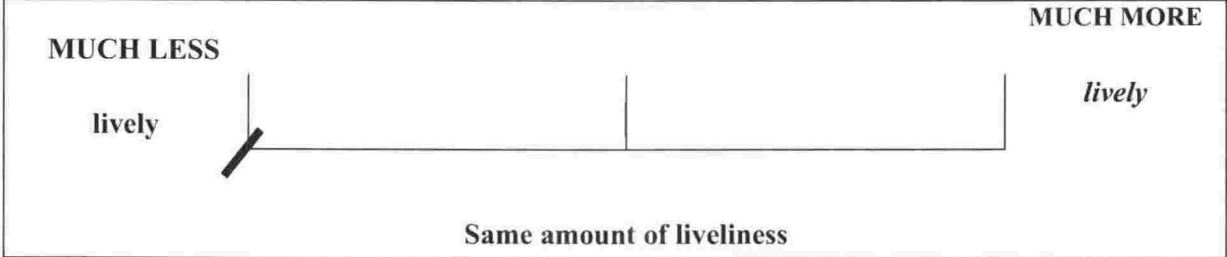
The far right of the line should be marked if you believe the second sentence sounds much **more** lively than the first sentence, such as in Example 1.

**EXAMPLE 1:**



The far left should be marked if you believe the second sentence sounds much **less** lively than the first sentence, as shown in Example 2.

**EXAMPLE 2:**



The centre point is to be marked if you think the second sentence conveys the **same** degree of liveliness as the first sentence. Everything in between these points is for variations therein of perceived liveliness.

Please remember that your mark **must cross** the line!

HOW MUCH MORE/LESS **CONFIDENT** DOES THE SECOND SENTENCE  
CONVEY IN COMPARISON TO THE FIRST SENTENCE?

<b>MUCH LESS confident</b>			<b>MUCH MORE confident</b>
<b>SAME</b> amount of confidence			
<b>MUCH LESS confident</b>			<b>MUCH MORE confident</b>
<b>SAME</b> amount of confidence			
<b>MUCH LESS confident</b>			<b>MUCH MORE confident</b>
<b>SAME</b> amount of confidence			
<b>MUCH LESS confident</b>			<b>MUCH MORE confident</b>
<b>SAME</b> amount of confidence			
<b>MUCH LESS confident</b>			<b>MUCH MORE confident</b>
<b>SAME</b> amount of confidence			

HOW MUCH MORE/LESS **INSISTENT** DOES THE SECOND SENTENCE  
CONVEY IN COMPARISON TO THE FIRST SENTENCE?

<b>MUCH LESS insistent</b>			<b>MUCH MORE insistent</b>
<b>SAME amount of insistence</b>			
<b>MUCH LESS insistent</b>			<b>MUCH MORE insistent</b>
<b>SAME amount of insistence</b>			
<b>MUCH LESS insistent</b>			<b>MUCH MORE insistent</b>
<b>SAME amount of insistence</b>			
<b>MUCH LESS insistent</b>			<b>MUCH MORE insistent</b>
<b>SAME amount of insistence</b>			
<b>MUCH LESS insistent</b>			<b>MUCH MORE insistent</b>
<b>SAME amount of insistence</b>			

HOW MUCH MORE/LESS **LIVELY** DOES THE SECOND SENTENCE  
CONVEY IN COMPARISON TO THE FIRST SENTENCE?

<b>MUCH LESS</b> lively			<b>MUCH MORE</b> lively
<b>SAME</b> amount of liveliness			
<b>MUCH LESS</b> lively			<b>MUCH MORE</b> lively
<b>SAME</b> amount of liveliness			
<b>MUCH LESS</b> lively			<b>MUCH MORE</b> lively
<b>SAME</b> amount of liveliness			
<b>MUCH LESS</b> lively			<b>MUCH MORE</b> lively
<b>SAME</b> amount of liveliness			
<b>MUCH LESS</b> lively			<b>MUCH MORE</b> lively
<b>SAME</b> amount of liveliness			



---

**D1. Consent Forms**

---

Memorandum to: Class members

From: Patricia Vermillion  
PhD student  
School of Linguistics and Applied Language Studies

Re: Experiment on New Zealand English intonation

I will be conducting an experiment on New Zealand English intonation. You will hear a speaker saying a sentence in a number of different ways. For each sentence, you will be asked to make a judgment about what you hear. We envisage the experiment will take approximately 15 minutes. For your time, you will be provided with a \$5 book/music voucher.

You should be aware of the following (VUW Human Ethics Committee Guidelines, pg.21):

- Your participation is voluntary
- Your participation or non-participation will not be recorded or regarded by the school in either a positive or a negative way
- If you participate, you have a right to withdraw at any time and you do not have to state a reason
- Your consent to participate is implied by your filling out a questionnaire.

The current Human Ethics Committee Guidelines can be found at:  
<http://www.vuw.ac.nz/home/research/index.html#hec>

Responses from this experiment will be a part of written reports, such as my PhD thesis and possibly publications. The questionnaire will include questions such as whether your first language is New Zealand English. However, it will not be possible to identify who completed which questionnaire, as the questionnaires are anonymous. The PhD thesis will be submitted for marking to examiners and subsequently a copy will be lodged in the Victoria University of Wellington Library, and all datasheets will be destroyed at the completion of the 3-year project.

If you have any queries, please feel free to contact me via email ([Patricia.Vermillion@vuw.ac.nz](mailto:Patricia.Vermillion@vuw.ac.nz)) or by telephone (04-463-5621).

## Feedback for Participants taking part in Experiment on New Zealand Intonation

I am a PhD student in the School of Linguistics and Applied Language Studies at Victoria University of Wellington. As part of my degree, I am conducting a series of experiments which are investigation New Zealand intonation. Thank you for agreeing to participate in this study.

The aim of the experiment was to investigate what meanings are conveyed by the intonation of a sentence. The stimuli which you have heard during the experiment were derived from one native New Zealand English speaker. The recorded utterance was *Manny went to the union*. Using a computer program, this utterance was manipulated in various ways in order to assess how conversational cues are interpreted by different pitch heights and movements. This information will be used to determine which meanings are interpreted by this intonation for New Zealand English speakers.

If you have any queries regarding this experiment, please feel free to contact me at [Patricia.Vermillion@vuw.ac.nz](mailto:Patricia.Vermillion@vuw.ac.nz). Thank you for your time and participation in this experiment.

Patricia Vermillion  
School of Linguistics and Applied Language Studies  
Victoria University of Wellington

## **D2. Recording material used for source utterance**

---

Please read aloud the following sentence, as naturally as possible and without emphasis.

**Manny went to the union.**

## **D3. Instructions (Task 1)**

---

You are going to hear one sentence said aloud by a native New Zealand English speaker in a number of ways. Imagine that he is talking to you. I would like you to tell me whether he is indicating for you to say something or whether he is not indicating for you to say something. If you believe the speaker would like you to say something after the sentence, please tick the "indicating: say something box". Alternatively if you believe the speaker is not asking you to add anything more to the conversation, tick the "Not indicating" box. You will have four seconds in between each sentence in order to give your answer. Also you will hear a short beep before every fifth sentence. After you hear the beep, you should be at the highlighted box. This was done to make sure that you are keeping pace with the recording. If you find that you are not at the same place at the recording, just place a big X over the previous set of five answers. If you have any questions, please raise your hand now. Otherwise, please turn the page now and the experiment will begin in 10 seconds.

#### D4. Instructions (Task 2)

---

Once again, you will hear the same speaker say the same sentence to you. This time what I would like you to do is to tell me whether he is going to say something more after the sentence or whether he is finished speaking. For example, the speaker may say:

*Manny went to the union.*

The speaker may be telling you that Manny had gone to the student union and have nothing more to say. He may also use the sentence at the end of monologue such as:

*Do you know that Manny can get free beer at the student union, he has a friend there that works in the bar who always gets free beer. You have to love free beer. So anyway. Manny went to the union.*

Either way the speaker is finished talking and has nothing more to say. Alternatively the speaker can say the same sentence but have more to say. For example:

*Manny went to the union. He gets free beer at the student union because he has a friend there who works in a bar. He always gives him free beer.*

You are going to hear the speaker say to you Manny went to the union in a number of different ways. All I want you to do is to tell me whether he is finished talking and has nothing more to say by ticking the "Finished turn at talk" box. Alternatively if you believe the speaker is going to say something more, tick the "More to say" box, just as in the example. Please make sure that you indicate an answer for every sentence. Even if you are uncertain, make your best guess. You will have four seconds in between each sentence in order to give your answer. Also you will hear a short beep before every fifth sentence. After you hear the beep, you should be at the highlighted box.

This was done to make sure that you are keeping pace with the recording. If you find that you are not at the same place at the recording, just place a big X over the previous set of five answers. If you have any questions, please raise your hand now. Otherwise, please turn the page now and the experiment will begin in 10 seconds.

**D5. Questionnaire and answersheet**

---

*Questionnaire*

1. How old are you?

☐

Up to 17 years old

☐

18-22

☐

23-29

☐

30-35

☐

36 or older

2. Are you

☐

Male

or

☐

Female

3. Is your first language New Zealand English?

☐

Yes

or

☐

No

4. Are you a full-time or part-time university student?

☐

Full-time

or

☐

Part-time

5. What is your ethnicity?

☐

Māori

☐

mostly Māori  
with some Pākehā

☐

mostly Pākehā  
with some Māori

☐

Pākehā

☐

Other

Specify: \_\_\_\_\_

5. Do you have any known hearing impairments?

☐

Yes

or

☐

No

1.	INDICATING: Say something <input type="checkbox"/>	NOT INDICATING <input type="checkbox"/>
2.	INDICATING: Say something <input type="checkbox"/>	NOT INDICATING <input type="checkbox"/>
3.	INDICATING: Say something <input type="checkbox"/>	NOT INDICATING <input type="checkbox"/>
4.	INDICATING: Say something <input type="checkbox"/>	NOT INDICATING <input type="checkbox"/>
5.	INDICATING: Say something <input type="checkbox"/>	NOT INDICATING <input type="checkbox"/>
6.	INDICATING: Say something <input type="checkbox"/>	NOT INDICATING <input type="checkbox"/>
7.	INDICATING: Say something <input type="checkbox"/>	NOT INDICATING <input type="checkbox"/>
8.	INDICATING: Say something <input type="checkbox"/>	NOT INDICATING <input type="checkbox"/>
9.	INDICATING: Say something <input type="checkbox"/>	NOT INDICATING <input type="checkbox"/>
10.	INDICATING: Say something <input type="checkbox"/>	NOT INDICATING <input type="checkbox"/>
11.	INDICATING: Say something <input type="checkbox"/>	NOT INDICATING <input type="checkbox"/>
12.	INDICATING: Say something <input type="checkbox"/>	NOT INDICATING <input type="checkbox"/>
13.	INDICATING: Say something <input type="checkbox"/>	NOT INDICATING <input type="checkbox"/>
14.	INDICATING: Say something <input type="checkbox"/>	NOT INDICATING <input type="checkbox"/>

---

