

Using Prosodic Completion Tasks to explore the Phonetics and Phonology of Intonation

Patricia Vermillion

School of Linguistics and Applied Language Studies
Victoria University of Wellington,
New Zealand
Patricia.Vermillion@vuw.ac.nz

Abstract

The aim of this paper is to present a methodological technique first developed in Vermillion (2001) and remodelled in Vermillion (anticipated 2005), which is used to investigate the meaning of intonation. Using this methodology, two contrasted sets of meanings are illustrated. First, we compare how 30 New Zealand English-speaking University students use their intonation to indicate a continuation request to their discourse partner, with how these same speakers cue that they want the conversation to stop. This pair was contrasted by the majority of speakers through modification of their final boundary tone (%); for 18 of the 30 speakers the contrast was between L% to show that the conversation should cease and H% to show that it should continue. The other 12 speakers used a higher H* pitch accent preceding the L% to indicate continuation. The second meaning contrast was selected to further explore how a H*LL% sequence may be used to convey two opposing meanings. The data showed that 77% speakers realised H*LL% when conveying both authority and submission. A comparison of the heights of H* and L% in these contours reveals that the height of the L% is significantly higher when conveying submission than authority. The example Prosodic Completion Tasks illustrate how PCTs may be a useful initial methodology in exploring the Phonetics and Phonology of intonation.

1. Introduction

If a speaker wants 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, if they want to emphasise how many sandwiches they have just eaten, they can raise the pitch of their H* pitch accent, for example, on the number *two*.

Experimental Phonologists have employed numerous methodologies in an attempt to determine the meaningful contrasts of such tonal manipulations. One type of experiment measures a listener's perceived meaning of a specific intonation. For example, Ladd and Morton (1997) examined whether variation in a H* pitch accent was perceived categorically, specifically examining how the listeners interpreted the degree of emphasis on different levels of H*. Alternatively, production experiments analyze the production of intonation in speech. Analysing production data is similar to using perception experiments inasmuch as the researcher interprets the meaning of the intonation use rather than eliciting the speaker's intended meaning. For instance, Britain and Newman (1992) used speech from a database to examine how High Rising Terminal

contours (HRT) are used in New Zealand English. They assessed the use of HRTs within the context of surrounding utterances, illustrating the potential likelihood (or unlikelihood) of certain meanings to apply to the specific situation.

The aim of this paper is to present a methodological technique first developed in Vermillion (2001) and remodelled in Vermillion (anticipated 2005), which examines the meaning of intonation. Different from the approaches discussed in the previous paragraph, this methodology attempts to compare the speaker's intended meaning. A description of this methodology is outlined below.

2. PCT Methodology

Similar to a Discourse Completion Task (see Brown, 2001), a Prosodic Completion Task (PCT) elicits a linguistic variable which illustrates how a speaker conveys meaning. Specifically, in a Discourse Completion Task, the speakers are given a scene or the background information, such as what the previous speaker had said. The scene also specifies the setting, and the speakers' relationship with one another. A specific feeling or emotion, which the speaker is then supposed to convey through his/her response, is also

prompted. The aim of a Discourse Completion Task is to elicit lexical and/or syntactic forms for conveying the prompted meaning. Therefore the lexical responses are provided spontaneously by the speaker rather than read aloud with a prepared script. The aim of a PCT, however, is dissimilar from the Discourse Completion Task; a PCT attempts to elicit the speaker's prosodic technique for conveying meaning. For this purpose, every PCT includes a lexical response, which therefore limits the speaker's own contribution of meaning to intonation.

PCTs include one *scene*, two *conditions*, and two *test items*.

| |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Scene: Your partner tells you "I want to go out to dinner this weekend. Would you like to join me?"</p> <p>Condition 1: You are interested in this idea and signal to him/her to continue: Test item 1: <i>Yeah</i></p> <p>Condition 2: You are not interested in this idea and signal to him/her to stop talking! Test item 2: <i>Yeah</i></p> |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Figure 1: Prosodic Completion Task (PCT) comparing conversational continuation and cessation cues.

| |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Scene: Your partner tells you "I want to go out to dinner this weekend. Would you like to join me?"</p> <p>Condition 1: You say, with authority: Test item 1: <i>No</i></p> <p>Condition 2: You say, submissively: Test item 2: <i>No</i></p> |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Figure 2: Prosodic Completion Task (PCT) comparing Authority and Submission.

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). Specifically, the participant responds to their "discourse partner's" utterances, or scene. Individual interpretations of the previous speaker's utterance may vary. Attempting to control such variation, the emotive use of intonation as used by the participant's discourse partner is identical for all participants; the scenes are recordings of a single speaker who was instructed to read aloud the sentences from a provided script "as naturally as possible" and "without emphasis".

Second, the *conditions* provide the meanings which the participant is to convey. Ideally, each condition should highlight just one meaning, so that it is clear which meaning the speaker should attempt to convey. This is the case for the example in Figure 2. Note however, that the example in Figure 1 illustrates a

contrast not only between a conversational *cessation* and *continuation* request, but also between *interest* and *lack of interest*. Although these two meanings are often coupled in speech, participants may focus on one meaning, such as interest, rather than the entire condition containing two meanings. Such combinations of meanings are discussed in more detail below. PCT contrasted conditions may be separated by other non-related conditions within the recording session, possibly inhibiting the conscious manipulation of such tunes. It should be noted, however, that the contrasted conditions were prompted in succession in the PCTs reported here. Thus the data may be illustrating the conscious manipulation of tunes employed to contrast meaning. Nevertheless, such data are useful in demonstrating what native speakers believe to exemplify such meaningful contrasts. For comparability, the scene remains the same for the two contrasted conditions.

Finally, utterances are supplied as a response for each condition, which serve as the *test items*. It is important to note that the test items which are used to assess the two contrasting conditions are lexically identical; an identical response inhibits the lexical expression of these types of meaning which, therefore, compels the speaker to convey the meanings prosodically.

3. PCT Experiment

In order to illustrate how this methodology may help our understanding of Intonational Phonetics and Phonology, we present here example data from a larger PCT experiment (Vermillion, anticipated 2005). For the purpose of brevity, two PCTs have been selected for discussion. These are illustrated in Figures 1 and 2 above. The first meaning contrast being examined here – continuation or cessation of the conversation – may be conveyed by rising phrase-final intonation. Hirst and Di Cristo (1998:27) claimed that rises often denote unfinished utterances. In addition to the speaker reserving his or her turn-at-talk, rising pitch may also signal conversational cessation or continuation requests to the speaker's discourse partner. I examine whether speakers use H% to convey such meanings in New Zealand English, and compare how participants use their intonation to indicate a continuation request to their discourse partner with how these same speakers cue that they want the conversation to discontinue (see Figure 1). The second meaning contrast investigated here – authority vs. submission - was selected to further explore how a H*LL% sequence may be used to convey two opposing meanings. Specifically, 77% (23) of the speakers realized H*LL% when conveying both *authority* and *submission* (see Figure 2).

3.1. Participants

Social variation between the participants, which has been found to be a factor influencing linguistic variation, was assessed using a questionnaire. The data

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 age range of 18-29 years (average 21.2 years). All listeners were either European descendent New Zealanders (non-Maori) or mostly non-Maori with some Maori ethnicity. Although all speakers participated on a voluntary basis, they were provided with a music/book voucher worth NZ\$5.00 for their time.

3.2. Procedure

The recording session started with an introduction, 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 additional contextual variation which the individual scenes for each PCT did not control. The introduction introduced the participant's "discourse partner" (i.e. the recorded speaker) and their relationship (the speaker's intimate partner).

Each scene is a recording of one native New Zealand English speaker (the participant's "discourse partner") who was instructed to read aloud the sentences from a provided script "as naturally as possible" and "without emphasis". The speech was recorded using identical equipment to that used for the introduction, which was then downsampled onto a signal processing system, namely Praat (Boersma and Weenink, 1996) with a sampling rate of 16kHz. The Total Rescaling method (Ladd & Morton, 1997) was applied by moving all of the pitch points in each utterance 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). Each resulting scene includes two pitch levels (one "male" and one "female") derived from one source utterance. The participants later confirmed both pitch levels as adequate ("believable") representations of both a male and a female speaker. Each male and female participant was played aloud a "female" and "male" scene, respectively. Thus the results illustrate how speakers respond to discourse partners of the opposite sex.

After each scene was played to each participant, they read aloud each test item from a provided 125 x 75mm 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. Each experiment was conducted individually in a recording room in the Language Learning Centre in Victoria University of Wellington.

3.3. Auditory and Acoustic Analyses

The recorded speech material was downsampled to a sampling rate of 16 kHz. With the help of the signal processing system *Praat*, the utterances were transcribed using Tone and Break Indices, or ToBI

(Beckman & Ayers, 1994; Pierrehumbert, 1980), which label "H"igh and "L"ow targets on rhythmically prominent syllables as "*" as well as targets at intonational phrase boundary locations ("%") and intermediate phrase boundaries ("."). The pitch values (ERB) of opposing targets occurring on pitch accents and intonational phrase boundaries were also noted in order to make further comparisons.

4. Results and Discussion

The results from the two PCTs indicated that the high pitch final target (H%) was used to contrast conversational continuation and cessation cues by the majority of participants; 18 of the 30 participants realised L% when conveying a lack of interest and a conversation cessation request (H*LL and L*LL), and all of these participants changed their phrase-final target to H% when indicating continuation.

The other 12 speakers used H*LL% to convey both meanings, suggesting that there must be another method of conveying such contrasts. Two Analyses of Variance (repeated measures design, Huynh-Feldt corrected) were performed on data from the 12 speakers who used L% for both meanings. The heights of the two opposing tonal targets were compared; the dependent variables for each analysis were (1) the height of L%, and (2) the height of H* (in ERB). The independent variables for both Analyses were Meaning (a within-subject factor with 2 levels: continuation and cessation) and Speaker Sex (a between-subject factor with 2 levels). Speaker Sex proved to be a significant factor in each Analysis performed on the pitch values within this paper. Specifically, female speakers realised significantly higher pitch targets than their male counterparts. Due to the omnipresent effects of Speaker Sex which reflect the physiological differences between female and male speakers, Speaker Sex as a main effect will not be discussed further in this paper.

The results illustrate that the height of H* was a significant factor ($F_{1,10} = 8.60$, $p < 0.05$); the speakers produced higher H* values when conveying continuation (6.228 ERB, s.e. 0.252) than a conversational cessation request (5.470 ERB, s.e. 0.157). Figure 3 illustrates the use of H* to convey the two contrasted meanings.

It should be noted that higher H* may not have been contrasting continuation. As illustrated in Figure 1, the conditions elicited two types of meanings - interest in the conversation and a continuation request. Higher H* may have been used to convey the Effort Code (Gussenhoven, 2004), emphasising the lexeme *yeah* and subsequently conveying interest in the conversation. Although such meanings may infer continuation cues to the speaker's discourse partner, this use of H* may not correctly illustrate how speakers use their intonation to contrast conversation cessation and continuation cues. Determining which meanings are being conveyed by a specific tonal manipulation may be resolved with supplementary studies, expressly comparing meanings

which are often coupled in speech, such as *interest* and *lack of interest*, with *continuation* and *cessation cues*.

When comparing the data from the 12 speakers who used H*LL% for both meanings, the height of L% was not found to be a significant factor, nor was there any significant interaction between the variables. It should be noted, however, that L% was found to be nearly significant ($p = 0.053$). Although not significant, continuation was signalled by higher L% values and lower L% values showed that the conversation should cease. The results may have been affected by the small data set (12 speakers), and statistical significance could be found if more H*LL% data were available for comparison.

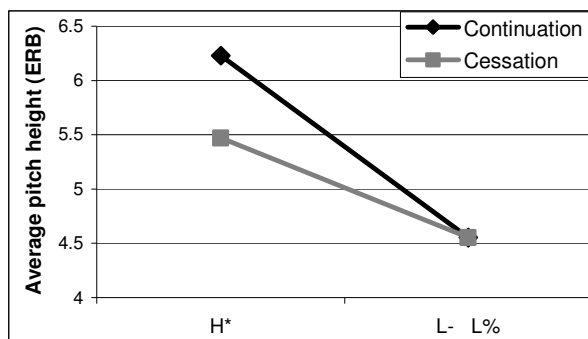


Figure 3: Illustration of “yeah” contrasting conversational continuation and cessation cues on H*.

The second meaning contrast investigated here was selected to further explore how similar tonal contours, including a phrase-final L%, may be used to convey two opposing meanings. Specifically, 77% (23) of the speakers realised H*LL% when conveying both Authority and Submission. Two Analyses of Variance (repeated measures design, Huynh-Feldt corrected) were performed on data from the 23 speakers who used L% for both meanings. The dependent variable for the two separate analyses was (1) the height of L%, and (2) the height of H* (in ERB). The independent variables for both analyses were Meaning (a within-subject factor with 2 levels: authority and submission) and Speaker Sex (a between-subject factor with 2 levels). A comparison of the heights of L% and the preceding peak in these contours reveals that the height of the L%, not H*, was significantly different when conveying Authority and Submission ($F_{1,21} = 11.874$, $p < 0.01$); the speakers realised significantly higher L% values when conveying submission (4.7307 ERB) than authority (4.2902 ERB). Once again, there were no significant interactions between the variables. Figure 4 illustrates the use of L% to convey both authority and submission.

This result appears to correspond with the meanings associated with the Frequency Code (Ohala, 1984: 340), namely that high pitch indicates smallness and subsequently submissiveness whereas low pitch implies largeness and dominance. Interestingly, the speakers may be preferring L% height rather raising other tonal

targets such as H* to signal the Frequency Code and contrast these two conditions. Such tonal preferences, however, would need to be explored further in a separate study.

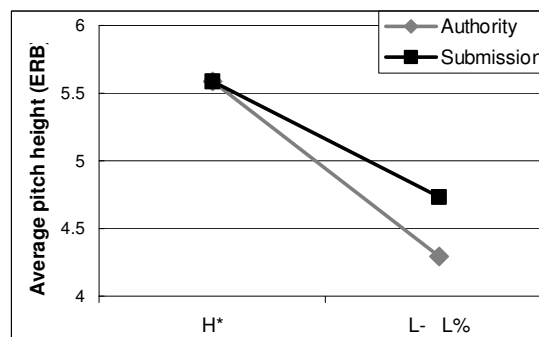


Figure 4: Illustration of “no” contrasting authority and submission on L-L%.

The difference in the height of L% may also indicate discourse completeness and finality. One meaning claimed for phrase-final falls is that they signal that the utterance is a statement rather than a question, the latter being indicated by a rise (or at least by 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 they do not want to discuss the topic further, and hence will 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.

The results may also 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”. Supplementary experiments may determine whether such interpretational differences had an effect.

5. Conclusions

Intonologists quickly discover that it is difficult to separate the many-layered meanings of intonation. PCTs do not evade such issues. The aim of the Prosodic Completion Task is to serve as an initial methodology in exploring the Phonetics and Phonology of intonation. As shown in the example PCTs, this methodology does not unveil all of the meaningful tonal contrasts used by the speakers of a language variety. The results illustrate possible contrasts, and exploring

such contrasts may prove useful for establishing a phonological system of intonation for this language variety. As illustrated above, the choice of either L% or H% may be determined by the speaker's intended meaning. In addition, the height of tonal targets such as H* and L% may be used to contrast specific meanings in New Zealand English. Such tonal contrasts may not be identified using an intonational labelling system created specifically for other varieties of English. The PCTs discussed here may be illustrating a phonological distinction between heights of H* and L%, for instance, in this variety of English. Nevertheless, supplementary experiments would need to be employed to confirm such suppositions.

PCTs may also prove beneficial for socio-phonetic/phonological comparisons between varieties of a language(s) or social categories. For instance, Vermillion (2001) used PCTs to illustrate how male and females use different peak alignments to contrast submission and authority, and pitch register to convey support for the conversation to continue. Prosodic Completion Tasks may also prove useful in demonstrating how varieties of one language use intonation differently to convey matched meanings.

6. Acknowledgements

I would like to thank Paul Warren for his assistance at various stages in the preparation of this paper. I would also like to acknowledge the financial assistance of Victoria University of Wellington.

7. References

- Beckman, M., and G. Ayers (1994). *Guidelines for ToBI labeling* (Vol. 2). Columbus: Ohio State University Linguistics Dept.
- Boersma, P., and D. Weenink (1996). *PRAAT: A system for doing phonetics by computer*. Amsterdam: University of Amsterdam.
- Britain, D., and J. Newman (1992). *High rising terminals in New Zealand English*. *Journal of the International Phonetic Association*, 22 (1/2): 1-11.
- Brown, J.D. (2001) *Pragmatic tests: Different purposes, different tests*. In K.R. Rose & G. Kasper (eds), *Pragmatics in language teaching*, pp. 301-325. Cambridge: CUP.
- Cauldwell, R. (1999). *Judgements of attitudinal meanings in isolation and in context*. *Phonetics Teaching & Learning Conference 1999*. University College London:UK.
- Cruttenden, A. (1997). *Intonation*. London: CUP.
- Crystal, D. (1969). *Prosodic systems and intonation in English*. London: CUP.
- Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge: CUP.
- Hirst, D., and A. Di Cristo (1998). *A survey of intonation systems*. In D Hirst & A. Di Cristo (eds), *Intonation Systems: A Survey of Twenty Languages*, pp. 1-44. Cambridge: CUP.
- Ladd, D.R., and R. Morton (1997). *The perception of intonational emphasis: continuous or categorical?* *Journal of Phonetics*, 25: 313-342.
- Ohala, J. (1984). *An ethological perspective on common cross-language utilization of F0 in voice*. *Phonetica*, 41: 1-16.
- Pierrehumbert, J. (1980). *The Phonology and Phonetics of English Intonation*. PhD Thesis: MIT, Cambridge, USA.
- Vermillion, P (2000). *Analysing the function of intonation in the discourse of men and women*. NWAV 29, Michigan State University, USA.
- Vermillion, P. (2001). *The Perception and production of intonational meaning by British men and women*, unpublished MPhil thesis, University of London, United Kingdom.
- Vermillion, P. (2003). *The Ups and Downs of Kiwis: An experiment investigating tonal cues which are used to identify NZE intonation*. *Wellington Working Papers*, 15: 17-30.
- Vermillion, P. (anticipated 2005). *The meaning of New Zealand English intonation*, unpublished PhD thesis, Victoria University Wellington, New Zealand.