Introduction

The study of intonation has always been bound up with the study of meaning. But the myriad researchers who have taken up intonation have, unsurprisingly, found a number of ways to characterize both its structure and function. Below, I take an "ecumenical" approach to intonational meaning and function in one corpus of conversational data in Mandarin, an approach grounded in the study of variation, (socio-)phonetics, and interaction.

Many scholars, like Dwight Bolinger (Bolinger, 1989) and Gussenhoven (Chen, Gussenhoven and Rietveld, 2004; Gussenhoven, 2004), have made strong arguments for cross-linguistic universals in the interpretation of intonation. Bolinger notes that final rises are cross-linguistically associated with uncertainty and incompletion, while final falls signal wholeness and certainty. Drawing on earlier work linking intonation to the use of pitch in animal communication (Ohala, 1983), Gussenhoven has articulated a relatively more elaborate theory that bases intonational meaning on a set of "biological codes": frequency (see Ohala, 1983), effort, and production. A rise in fundamental frequency over a given domain, for example, indexes incompletion or smallness, which motivates certain assumptions about the speaker's physical or emotional state, e.g. uncertainty, or conveys information about the message. In addition, however, many if not most of the world's languages have grammaticalized rising intonation as a marker of yes/no-questions. However, as Gussenhoven's own work (Chen et al., 2004) has noted, even the paralinguistic, "affective" meanings of prosody can be culturally specific.
Intonational phonology's attention, from its early days, to both linguistic and "paralinguistic" intonational contributions to meaning, and thus to such concerns as "affect" and the "vividness of specific nuances in specific contexts" (Ladd, 1996, p. 40), have striking parallels in the present-day style of stylistic variation, which increasingly attempts to describe stylistic meaning in terms of speaker stance and with attention to interactional contingencies (Eckert, 2005; Eckert, 2008).

Sociolinguistic work in intonation has paid keen attention to local interpretations of contours. Podesva (2006) looked at rising, level, and falling contours among gay professionals, and Cheng and Warren (2005) at the use of rises in service encounters and other interactions in Hong Kong English. For one of Podesva's speakers, rising declaratives are used to establish a close, nonthreatening relationship between a doctor and his patient. In the Hong Kong study, the authors focus on the power differential between users of HRT -- the "service providers" -- and their interlocutors. Queen (2001) has looked at rising intonation contours among Turkish-German bilinguals. In all cases, intonation is used to construct certain kinds of relationships -- intimate or distant -- between interlocutors, to index affects for interactional effect, or to creatively invoke personas or identities.

Turning to Chinese, Chao (1968) remarked early on that unstressed syllables at the ends of sentences in Mandarin showed pitch movements that seemed to be linked to sentence pragmatics. Phonological analyses of Mandarin intonation have been greatly influenced by the autosegmental-metrical approach, popularized in part by the development of English ToBI, a system for transcribing English intonation (Beckman & Elam, 1997). The autosegmental approach posits that tone and intonation can be described as a series of discrete pitch events in time, and that the pitch contour is calculated largely on the basis of interpolating between pitch targets specified by the phonology.
Peng, Chan, Huang, Lee, and Beckman (2006) have made a proposal for a Mandarin ToBI in the vein of these autosegmental analyses, while also trying to account for instrumental data provided by Shen (1990) which seems to argue against an autosegmental account (see Ladd, 1996 for a discussion). Their model consists of a number of components. Among these are initial tones, which control pitch range and register across some prosodic domain. For example, the tone %e-prom specifies extended pitch range during emphasis. Phrase-final boundary tones, meanwhile, specify the final pitch target of the utterance. The tone H% signals a relatively high F0 at the end of the utterance, while L% denotes a relatively low phrase-final F0.

In an autosegmental approach, both the lexicon and sentence-level prosody specify pitch targets, and the interaction between these in the determination of the final fundamental frequency contour is not given a priori. In experimental paradigms, this problem is investigated and/or controlled for with carefully constructed stimuli. The use of naturalistic data may impose some limitations because of the inability to control for these factors.

How do we address the confounding effect of lexical tone on the F0 contour of the utterance? For this, I return to Chao's unstressed, atonic intonation carriers, because pitch movements on these syllables can be attributed to sentence-level intonation (pace accounts, e.g. Duanmu, 2007, where lexically specified tones "float" onto syllables such as these from nearby syllables). In the following example, the final syllable a is not specified for lexical tone, and so the pitch with which it is realized is likely due in large part to the intonation of the utterance, and mostly the choice of H% or L% boundary tones.

**Example 1. Right or not?**

\[ shuo1-de \quad bu4 \quad dui4 \quad a \]

Callier 3
The phenomenon of the phrase-final lengthening is less well studied in quantitative sociolinguistics, but it has been documented in the phonetics literature for languages as diverse as Chickasaw (Gordon & Munro, 2007), French (Smith, 2002), and Finnish (Nakai et al., 2009). Gordon and Munro's study demonstrates that the degree of final lengthening in Chickasaw varies with the size of the prosodic domain, with smaller domains (syllables) incurring less lengthening than larger ones (intonational phrases). In Finnish, phrase-final lengthening is progressive, that is, it may start before the final syllable and become progressively more extreme up until the end of the phrase. Interestingly, lengthening preserves phonologically contrastive length in that language, so that phrase-final short vowels are shorter than phrase-final long vowels (Nakai et al., 2009). With regard to lengthening phenomena in general, Ohala (1978) notes that speakers are able to articulate falling tones in a much shorter time than rising tones, which suggests a phonetically grounded connection between F0 manipulation and this dimension of prosody.

Lengthening may also be tied to sentence pragmatics. Smith (2002) demonstrates that lengthening in French is more extreme on yes/no-questions, particularly those where syntactic cues to interrogativity are absent. In perhaps the only variationist sociolinguistic work to address final lengthening, Kiesling (2005) documents lengthening occurring alongside other stylistic features of an immigrant variety of Australian English: backing and lowering of a mid-central vowel and, in particular, the so-called "high-rising tune" addressed by Podesva (2006) and Cheng and Warren (2005) in the studies reviewed above. The association of lengthening with questions in French and
forms which establish, among the speakers Kiesling studies, an "authoritative connection" suggests lengthening plays a role in interpersonal closeness in interaction, as well as with the management of turn-taking and in the articulation of stance. I will explore this possibility further with regard to my own data, below.

**Data and methods**

The data that I examine below come from a corpus of Mandarin telephone conversations available through the Linguistic Data Consortium (Fung et al., 2005). Speakers mostly do not know each other before their conversation, and five minutes of each recording are made available and transcribed. Although many studies, including sociolinguistic ones (Grabe, 2004), have made good use of constructed stimuli and read sentences in the study of intonation, "naturalistic" interaction provides more free reign for contingent and situated meanings of intonation to emerge.

In light of the research reviewed above as well as the need to minimize interference in the F0 contour from lexical tone, I will focus on a single "intonation carrier" and its realization on several phonetic dimensions as a way of investigating utterance-final intonation and prosody in Mandarin. The sentence-final particle *ma* is a non-referential clitic that occurs with a wide variety of sentence types, including yes/no-questions, assertions, directives, and topicalized constituents. Boya Li's (2006) dissertation argues that in all of these environments, *ma* has the fairly bleached core meaning of "heightened commitment" to the illocution.

This study looks at 125 utterances ending in this particle selected at random from the corpus. By "utterance," I generally mean the smallest intonational group that is also associated with a major syntactic boundary -- usually independent clauses, but also left-dislocated constituents and
sometimes, tag questions. I recorded the start and end time of each utterance, taking the presence of formant information for at least the first two formants on the final vowel as criterial in the placing of the endpoint.

I coded for characteristics of the speaker and situation, discourse and utterance properties, and phonetic and phonological dependent variables. The first of the dependent variables was phrase-final boundary tone, which I coded auditorily as either H%, a final high, or L%, a final low. Where I encountered difficulty -- often due to extremely compressed pitch range at the end of the utterance -- I tried to confirm my percept through inspection of a spectrogram.

The next variable was final syllable (ma) duration. Although research reviewed above demonstrates that final lengthening can occur on domains of variable size, lengthening of the final syllable is criterial and should suffice as a measurement of lengthening in general. Therefore I have decided to operationalize final lengthening as the duration of the final syllable only, normalized relative to speech rate. Speech rate is calculated by dividing the number of syllables in the utterance over the duration of the utterance in milliseconds. The final syllable is not included in the calculation. Final lengthening is measured by multiplying the utterance speech rate by the absolute duration of the final syllable, a value that is then scaled by the average syllable duration across the entire corpus, to yield a normalized measurement interpretable in milliseconds.

This measurement assumes that final lengthening is interpreted relative to speech rate, which I believe to be justified but does not find explicit support in previous investigations of final lengthening. Smith (2002) does not take speech rate into account, and Nakai et al. (2009) considered its effects but were able to control for it in the experimental design. Articulatory rates in this corpus are highly variable, and without the control that these experimentalists were able to exert
over the speaking situation, I consider the normalized measurement the safer option when compared to using absolute duration.

I explored a number of potential explanatory variables. These included: speaker and interlocutor sex, age -- which varied in this sample from around 18 to around 40 -- and the perceived accent of the speaker, as rated by the builders of the corpus. For this variable, they classified speakers as having standard-sounding and non-standard-sounding accents, on the basis of unspecified criteria. I also coded for the presence of a response from the other speaker after the token. Another variable was the clause type of the containing utterance, which could be declarative, interrogative (all the questions in this corpus were yes/no-questions), or left-dislocated topicalized constituent (topicalized). The distinction between declaratives and interrogatives, while not apparent in word order, was reliably clear in context, especially in the presence of specific constructions associated with questions. Another variable, illocutionary force, included possible values of assertion, information request, and tag question. Again, these were disambiguated with reference to the discourse context, and while other values (for example, command) are conceivable, they were quite rare and were not included in this corpus. Also recorded are preceding tone, utterance length in syllables, utterance duration in milliseconds, and utterance pitch range. All measurements of fundamental frequency were originally made in hertz, then converted to semitones, a logarithmic scale borrowed from music that more closely corresponds to perceived pitch. Pitch differences reported in semitones should reflect the auditory equivalence of different frequency intervals accurately.

Using exploratory data analysis and various statistical methods, I probed the relationship between these independent variables and the three dependent variables. Below I report regression
results for each of them, after arriving at a satisfactory model by starting with a number of predictors and removing insignificant variables from the equation while monitoring goodness-of-fit and other factors. In addition, based on the principle that more phonetically extreme variants of meaningful forms are more likely to highlight the social meanings they carry (Podesva, 2006), I examine a few such limit cases in their discourse context. In this way, I hope to shed more light on the full range of their meaning potential.

Results

Boundary tone

The first dependent variable I looked at was the choice of boundary tone. Table 1 gives a breakdown of boundary tone by clause type and illocutionary force. In Figure 1, we have a graphical representation. There is a clear segregation of declaratives and questions by the boundary tone that they take. Declaratives overwhelmingly use final low tones and occasionally although quite rarely choose the other tone, who H%. Questions are not as categorical, with high frequencies of both H% and L%, but they usually have a final high. Topicalization appears to behave much like declaratives, usually but not always using final low or L% tones.

<table>
<thead>
<tr>
<th>Clause Type</th>
<th>Illocution</th>
<th>Boundary Tone</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Women</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H%</td>
</tr>
<tr>
<td>Declarative</td>
<td>Assertion</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Question</td>
<td>Assertion</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Tag question</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Information request</td>
<td>10</td>
</tr>
<tr>
<td>Topicalization</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1. Boundary tone by clause type, illocutionary force, and speaker gender.
Figure 1. Boundary tone by clause type.
So what explains the increased variability in tone choice for questions? When we break questions down by illocutionary force (see Figure 2), we see that interrogatives are used for a wide variety of pragmatic ends. Information requests, the most canonical type of question, used predominantly H% or final high tones. Tag questions are more variable but still use mostly H% tones. Meanwhile, questions that were clearly interpretable as assertions in the discourse context always occurred with final lows or L%. The pattern that seems to emerge from this representation is that the proportion of high to low boundary tones is related to the need for a response from the interlocutor.

![Boundary Tones in Questions by Illocution](image)

**Figure 2.** Boundary tones in questions by illocutionary force.
The general finding relating boundary tones to sentence type and illocutionary force is confirmed by an ANOVA using a combination of both factors as the sole explanatory variable (chi-squared = 32.186, 4 DF, p < 0.001). Table 2 reports the regression results for this model. Declaratives used as assertions pattern with questions used as assertions and topicalized constituents, each with a relatively higher likelihood of an L%. Tag questions and information requests are less likely to have L%.

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.727</td>
<td>0.384</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Question</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assertion</td>
<td>15.838</td>
<td>1398.721</td>
<td>0.011</td>
<td>0.991</td>
</tr>
<tr>
<td>Tag question</td>
<td>-1.352</td>
<td>0.5483</td>
<td>-2.467</td>
<td>0.013</td>
</tr>
<tr>
<td>Info Request</td>
<td>-2.672</td>
<td>0.5879</td>
<td>-4.544</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Topicalization</td>
<td>-0.628</td>
<td>0.6932</td>
<td>-0.907</td>
<td>0.364</td>
</tr>
</tbody>
</table>

**Null deviance: 161.31, 128 df; Residual deviance: 129.12, 124 df**

Table 2. Regression results for boundary tone. Response is the log-likelihood of an L%. Positive coefficients indicate increased likelihood of L% relative to base category (Declarative Assertion), and negative coefficients, decreased likelihood.

In sum, Mandarin intonation follows universal tendencies to use falling tunes with declaratives and/or assertives and rising tunes with questions, at least for the yes/no-questions in this corpus. The ratio of high to low tones seems to vary with the need for a reply from the addressee. "Exceptional" pairings of sentence type and illocutionary force, such as questions used as assertions, use the intonation appropriate to the pragmatics, which is consistent with previous work linking intonation to pragmatic inference.

**Final lengthening**

Callier 11
The next dependent variable I looked at is the duration of *ma*, the final syllable in the intonational phrase.

The following example, *ni yizhi kandao wei ma* "did you watch it all the way to the end?" shows significant final lengthening and is in fact one of the longest final syllables in the corpus.

The duration of *ma* here is much greater than that of the other syllables in the utterance -- I will explore this example in more depth below. I calculated the relative duration of *ma* based on speech rate and set the result, transformed to better satisfy regression assumptions, as a response for a regression analysis. Table 3 gives a summary breakdown of lengthening by clause type, illocutionary force, and gender of speaker.
Table 3. Lengthening of final syllable by sentence type, illocutionary force, and gender. Measurements in milliseconds (standard deviations in parentheses).

Starting with a number of predictors and culling away variables until all terms have an appreciable contribution to the model, I arrived at the following analysis (Table 4). Clause type was the main contributing factor to the duration of the final syllable, with questions longer than any other sentence type, and no appreciable difference between declaratives and topicalized constituents.

<table>
<thead>
<tr>
<th>Clause Type</th>
<th>Illocution</th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declarative</td>
<td>Assertion</td>
<td>251.48</td>
<td>(69.77)</td>
</tr>
<tr>
<td></td>
<td>Assertion</td>
<td>401.22</td>
<td>(168.76)</td>
</tr>
<tr>
<td>Question</td>
<td>Tag question</td>
<td>259.29</td>
<td>(50.58)</td>
</tr>
<tr>
<td></td>
<td>Information request</td>
<td>315.08</td>
<td>(93.15)</td>
</tr>
<tr>
<td>Topicalization</td>
<td></td>
<td>197.57</td>
<td>(43.69)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.72</td>
<td>0.021</td>
<td>--</td>
</tr>
<tr>
<td>Clause type</td>
<td>Question</td>
<td>0.078</td>
<td>0.0288</td>
</tr>
<tr>
<td></td>
<td>Topicalization</td>
<td>-0.055</td>
<td>0.044</td>
</tr>
</tbody>
</table>

Null deviance: 2.972, 122 df; Residual deviance: 2.683, 120 df

Table 4. Regression results for final syllable duration. Response is log of the square root of normalized duration, coefficients give change in response relative to base category (Declarative).

The finding for clause type mirrors what Smith (2002) found for French, that yes/no-questions tended to lengthen, in the absence of other cues to interrogativity. Other factors such as turn-finality, speaker sex, and so on are either not significant or do not add enough explanatory power to be included in the model.
But of course, this does not mean that clause type accounts for all or even very much of the variability in final syllable duration. Variants like this always occur in a particular discourse context, and here increased duration co-occurs with signals of increased enthusiasm and involvement. In the case of this example, the lengthening occurs after stretch of particularly enthusiastic discourse, and arguably contributes to the impression of enthusiasm:

Example 2. Did you watch it all the way to the end?

1  
aha yi bai duo ji le!  
“More than a hundred episodes!

2  
zhen shi yi ge hen chang de Hanguo pianr a…  
That really is a very long Korean TV series,

3  
aiya wo shuo xuyao yiding de naixing la…  
wow I mean it requires some patience to keep watching...

4  
ni yizhi kandao wei ma::  
did you watch it all the way to the end?”

Example 2 gives some indication of the different linguistic features, besides lengthening, that make the speaker's contribution sound enthusiastic. These include interjections like *aha* 'whoa' and *aiya* 'wow,' intensifiers like *zhen* 'really' and *ye* 'really,' and affect-marking sentence-final particles *le*, *a*, and *la*, in addition to an overall expanded pitch range. Furthermore, lines 1, 2, and 3 are not as lengthened, and sound rather clipped compared to line 4. Thus the final lengthening here may serve as a turn-yielding device, coming after a series of "rushed-through" endings. Another example has extreme lengthening in an affect-laden context:

Example 3.  Aren't I broke?

1  
B danshi ba, ai dao xianzai hai mei zhaodao ren pei wo qu chi fan  
But, oh I still haven't found anyone to come eat with me!

2  
A na ni qingke bu jiu you ren zhaodao le ma  
Then if you offer to pay won't you have found someone then?

3  
B (laugh) wo bu shi mei qian ma  
Aren't I broke?
In the exchange illustrated in Example 3, the first speaker laments that she can't find anybody to eat out with her. The other responds that she might have more success if she offered to pay, to which the first speaker responds with a rhetorical question: "but aren't I (is it not true that I am) broke?" The final syllable here is lengthened to a considerable degree and, unlike Example 2, with a dramatic final fall. Although the speaker is not "enthusiastic," as in the earlier example, her apparent affective commitment to the utterance is high.

At the same time, both speakers' laughter and the almost histrionic, breathy whininess of speaker B mark this utterance as exaggerated. This over-production may be a lamination enacted to
achieve an ironic effect and help the speaker achieve distance from her words and the approbation they may occasion (for example, being accused of cheapness). Regardless of this possibility, the prosodic design of this utterance indexes emotion and particularly exasperation on the part of the speaker. The possible role of the sharp pitch movement across the last syllable is one that should be investigated further. In both of these examples, the lengthened utterances have something of an "other-oriented" quality, either in yielding the floor or soliciting a sympathetic reaction. Considering previous findings relating lengthening to "connection," this may be an integral part of the meaning of lengthening.

**Pitch Range**

As outlined in the methods section, the distances between the F0 maximum and minimum in the final syllable of each utterance were recorded. Table 5 summarizes some of the findings. In general, utterance-final pitch range was very variable, even within sentence types. Given the high variances of the sample means, it is hard to make hypotheses about the relationship between pitch range and clause type, illocutionary force, and gender. But the other possible independent variables I coded for may account for some of this variability.

<table>
<thead>
<tr>
<th>Clause Type</th>
<th>Illocution</th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Declarative</strong></td>
<td>Assertion</td>
<td>-1.50</td>
<td>(4.92)</td>
</tr>
<tr>
<td></td>
<td>Assertion</td>
<td>-10.30</td>
<td>(7.45)</td>
</tr>
<tr>
<td><strong>Question</strong></td>
<td>Tag question</td>
<td>-3.24</td>
<td>(4.83)</td>
</tr>
<tr>
<td></td>
<td>Information request</td>
<td>0.22</td>
<td>(4.32)</td>
</tr>
<tr>
<td><strong>Topicalization</strong></td>
<td></td>
<td>-0.64</td>
<td>(0.91)</td>
</tr>
</tbody>
</table>

Table 5. F0 movement (in semitones) over final syllable by clause type, illocutionary force, and sex of speaker. Standard deviations in parentheses.
Table 6 and Table 7 report the analysis I arrived at using ANOVA. Because I was interested in the gross quantity of pitch movement, the response was the absolute value of the F0 change between extrema in the final syllable. A square root transform was also necessary to help satisfy the regression assumptions.

<table>
<thead>
<tr>
<th></th>
<th>Chi-square</th>
<th>DF</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utterance F0 range</td>
<td>26.106</td>
<td>1</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Duration of final syllable</td>
<td>5.907</td>
<td>1</td>
<td>0.015</td>
</tr>
<tr>
<td>Non-standard accent?</td>
<td>8.688</td>
<td>1</td>
<td>0.003</td>
</tr>
<tr>
<td>Utterance F0 range × Final syllable duration</td>
<td>6.468</td>
<td>1</td>
<td>0.011</td>
</tr>
</tbody>
</table>

Table 6. ANOVA, utterance-final pitch range.

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.1977</td>
<td>0.5607</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Utterance F0 range (semitones)</td>
<td>-0.04479</td>
<td>0.04787</td>
<td>-0.936</td>
<td>0.3513</td>
</tr>
<tr>
<td>Final syllable duration (ms)</td>
<td>-0.002287</td>
<td>0.001827</td>
<td>-1.252</td>
<td>0.2130</td>
</tr>
<tr>
<td>Utterance F0 range × Final syllable duration</td>
<td>0.0004095</td>
<td>0.0001610</td>
<td>2.543</td>
<td>0.0122</td>
</tr>
<tr>
<td>Accent</td>
<td>Standard</td>
<td>-0.5009</td>
<td>0.1699</td>
<td>2.948</td>
</tr>
</tbody>
</table>

Null deviance: 91.401, 123 df; Residual deviance: 63.571, 119 df

Table 7. Regression results, final syllable F0 range. Response is the square root of the absolute value of the difference between absolute F0 extrema in the final syllable.

There is an interaction between the pitch range of the utterance and the relative duration of the final syllable. This effect is difficult to interpret from the coefficients, but the data show that as the pitch range of the utterance increases and the duration of the final syllable does the same, there is a multiplicative affect resulting in greater increases in final pitch range than if either factor were to
increase independently of the other. This stands to reason, as increased duration allows more time to hit an extreme articulatory target, and we would expect expanded pitch range on the final syllable to cooccur with expanded pitch range over the entire utterance.

More surprisingly, there is about a quarter-semitone difference between speakers based on accent. Speakers marked as standard by the corpus builders had decreased final pitch range. This is a bit puzzling, and the difference is small compared to the difference occasioned by the other, phonetic factors, where variability was more on the order of 10 semitones. Nevertheless, if pitch range is also implicated in displays of speaker affect, speakers with higher pitch range may have been indexing more intense emotional states. Recalling that Labov (2006) early on sought to identify switches into the vernacular by looking for paralinguistic cues of expressiveness and ease of communication (such as laughter, etc., and of which pitch range may be one), the corpus auditors may have been caught in a reverse observer's paradox. Speakers who evaded the somewhat awkward restraints on topic and addressee imposed by the corpus builders may have found themselves more at ease to make such paralinguistic displays, as well as to switch into vernacular varieties, which in some cases may have influenced auditors to rate the speakers themselves as "Nonstandard." In any event, I am quite hesitant to assign too much importance to the effect of accent on pitch range.

One surprising absence from the results is speaker gender. During coding, my subjective impression was that most of the extreme examples of pitch change came from women. The scatter plot in Figure 5 would seem to support this suspicion.
The variability that women display on the vertical axis, pitch range, far surpasses that of the men in the corpus. Unfortunately, despite its name, the analysis of variance is not designed to pick up on such differences. However, an F-test confirms that the difference between the two samples is, indeed, different ($F = 4.2267$, num df $= 61$, denom df $= 58$, p-value $< 0.001$). Thus gender may actually affect pitch range at the end of the utterance, but more investigation is needed to see to what degree this finding is compatible with the results from the regression analysis.
What is the relationship between boundary tones and final pitch movement? Because I coded impressionistically for boundary tone, I did not feel comfortable including it in the regression for fear of a circular analysis. Nevertheless, if boundary tones themselves have a pragmatic or indexical function, it stands to reason that final pitch range, as an implementation of the tone, would be related to this function -- although this relationship might be more complex and tortured than some intonation researchers would admit.

At least as an illustrative example, Example 4, supports the existence of such a link, which later work will need to expand on. One speaker has just asked the other if she would ever travel to Tibet.

**Example 4. What is plateau sickness?**

1 B xizang a, qishi wo man xiang qu de, Tibet? Actually I kind of want to go,
2 danshi wo haipa you gaoyuan fanying but I'm afraid of getting altitude sickness ["plateau reaction"].
3 A gaoyuan fanying, gaoyuan fanying shi Altitude sickness? What is altitude sickness?

4 B e, (lipsmack) xizang neibian bu dou shi Uh, isn't it true that Tibet is on a plateau? [L.%] gaoyuan ma
5 ranhou ta nabian jingchang shi dishi bijiao And over there the altitude is often rather high gao
6 ranhou kongqi bijiao xibao ma And the air is rather thin.
7 hen duo ren qu dou shiying bu liao A lot of people when they go can't get used to it
8 suoyi jiao gaoyuan fanying So it's called altitude sickness.

A question arises during the speaker's response -- what is altitude sickness? In her answer, she opens up with a rhetorical question, establishing as common knowledge the proposition that Tibet is located on a plateau. And on that line, the pitch falls over 15 semitones -- more than one full octave -- across the final syllable.
The L% boundary tone with which this utterance is realized helps mark it as an assertion. What follows is a fairly elaborate explanation in response to the other speaker's question, also composed of a series of assertions. The extreme fall may bolster this dimension of the pragmatics, increasing the epistemic commitment of the speaker to the utterance and enhancing the presuppositional quality and the sharedness of the proposition articulated in the rhetorical question. Thus the "informational" (Gussenhoven, 2004) function of this prosodic quality also has interactional consequences.

**Conclusion**

Throughout this paper I have tried to draw links where possible between the phonology and phonetics of the prosody, its "pragmatic" function -- narrowly defined -- and its deployment in interaction. I have taken a fairly ecumenical approach to doing so, and for good reason. While stable relationships between prosody and sentence pragmatics emerge with relative ease through common statistical methods -- and connections with "stable speaker characteristics" such as gender can be drawn too, with enough digging -- this only represents one dimension of the meaning of these forms. Entirely corpus-based investigations are best at revealing "text defaults" (Agha, 2007), regularities of meaning which can be played with and turned around in numerous ways when entextualized during interaction, as in the possible stylization of whininess in Example 3.

This study has been limited by several factors, chief among them the sample size, which should prevent any reader from taking too seriously any of the quantitative conclusions. The problems of segmentation and pitch tracking in the presence of devoicing and nonmodal voice qualities were not ones I confronted with the same methodological rigor as the phonetic studies cited.
in the introduction. In fact, given the literature on the interactional significance of voice quality, especially in domain-final positions, this should be a priority for future expansions of this research.

Nevertheless, this corpus investigation has uncovered some interesting questions which merit further attention. Chief among these is the gender difference in final pitch range variability. Extreme falls, which we saw in Example 3 and Example 4, did not serve a unitary expressive function. In the former case, a sharp fall in pitch contributed to an impression of whininess, while in the latter it bolstered the assertive quality of the utterance, and if anything painted the speaker in a knowledgeable or haughty light.

One question that a sociophonetic, rather than a (socio-)phonological, approach to intonational meaning raises is what exactly the role of phonetics is in contributing to, or modulating, the meanings of indexical forms. On Ladd's "Linguist's Theory of Intonational Meaning," "the elements of intonation have meaning" (1996, p. 39), in other words, meaning attaches to discrete, presumably phonological, quanta of structure. If a high-low sequence signals a declarative, then most nuances in the function of the declarative, inasmuch as those nuances are signaled in the prosody, indicate differences in phonological structure. Certain phonetic (usually scalar) dimensions of the realization of phonological structure, such as pitch range, pitch register, intensity, and so on, are often informative, but are essentially modulations of the signal.

But the particular nature of these modulations, as I and others have shown, directly affects the interactional uptake of linguistic categories themselves. What an assertion means in a particular context, whether it is whiny or authoritative, whether or not a speaker actually commits to it, is not just modulated by the phonetics in a scalar manner, but crucially mediated by it in a transformative one. Lengthening in the durational dimension plays a completely
different role from "lengthening" in the domain of pitch movement, and while these may index different "biological codes," this does not address how their meaning attaches to, transforms, or cancels out the meaning specified by the linguistic structure (see Agha, 2007).

It is for this reason that I remain uneasy with an account of intonational meaning, or indeed of any kind of stylistic meaning, that maintains a clear, hierarchical division between the linguistic and the paralinguistic. If such distinctions are to be made -- and they may be necessary for any sort of structural analysis to proceed -- it should always be with a clear goal of elucidating language as an activity engaged in by real people, where the regularities of function evinced by linguistic analysis are appropriated and transformed in ways that we may not yet fully understand.
References


