Intonational variation and social meaning: Categorical and phonetic aspects

Robert J. Podesva
Intonational variation and social meaning: Categorical and phonetic aspects
1 Introduction

The field of sociolinguistics has in recent years witnessed an increasing interest in the social meaning of variation, examining the ways in which speakers exercise linguistic choices to construct identities and personae (Eckert 2000, Zhang 2005). While studies adopting this approach, known as the third wave of variation, have typically investigated the meaning of particular categories of variant (e.g., -in vs. -ing), the phonetic details differentiating tokens belonging to the same category have been disregarded. I argue here that social meaning is encoded not only in the choice of variant, but also in the phonetic properties characterizing variants. Tapping into the full range of social meanings conveyed by phonological variation requires looking at both categorical and phonetic aspects of variation, a point I illustrate here by considering variation in declarative intonation. In addition to arguing for the social significance of phonetic variation, this paper seeks to bring intonational variation into the third wave, as studies concerned with the social meaning of variation have thus far concentrated on segmental variation.

In order for linguistic forms to carry social meanings, they must be sufficiently salient when they are uttered. I draw a distinction between two kinds of salience: categorical salience and phonetic salience. Categorical salience, as summarized in (1), can be determined by the frequency with which a given category of variant occurs. Under this view, infrequent variants stand out more than frequent variants. Thus if the word ain’t is uncommon in my speech, it stands out when I use it. Ideologies linked to a variant may also endow it with categorical salience. For example, the alveolar form of the variable (ing) constitutes a stereotype in the Labovian sense (Labov 1972) because people explicitly comment on the feature, with phrases like ‘dropping your Gs.’

* Thanks to the following people for helpful comments on various aspects of this work: Kathryn Campbell-Kibler, Penelope Eckert, Zsuzsanna Fagyal, John Rickford, Mary Rose, Julie Sweetland, Malcah Yaeger-Dror, and Arnold Zwicky. While their feedback has surely improved this work, any errors that may remain are of course my own.
(1) Categorical Salience
   a. The frequency with which a variant occurs is inversely correlated with the degree of salience (infrequent variants are relatively salient).
   b. Ideologies associated with a variant contribute to the degree of salience (stereotypical variants are relatively salient).

Another way for a variant to stand out is through phonetic salience, as defined in (2). This type of salience is determined by a variant's phonetic character. In the pitch domain, for example, a fundamental frequency (f0) of 400 Hz stands out more than an f0 level of 200 Hz.

(2) Phonetic Salience
   The value along any acoustic dimension characterizing a variant is directly correlated with the degree of salience (variants exhibiting acoustic extremes are relatively salient).

These two types of salience will figure prominently in the discussion that follows. In the following section I provide an overview of the study, summarizing the variable, speakers, and data under analysis. I begin discussing the results in section 3 by examining each speaker's use of rising, falling, and level contours across situations. In section 4, I move from categorical to phonetic variation, focusing on cross-situational differences in the phonetic character of falling contours. Crucially, I show that meaningful phonetic variation surfaces in the absence of categorical variation patterns. I conclude by discussing some differences in the kinds of meanings encoded in categorical versus phonetic variation.

2 The Study

Taking advantage of the fact that the vast majority of intonational variation studies have targeted statement intonation (e.g., Guy et al. 1986, McLemore 1991, Lowry 2002, Warren 2005), this study examines patterns of nuclear accent variation in declarative sentences with canonical focus. By nuclear accent, I mean the pitch accent occurring on the last prominent word of the sentence. With the phrase declarative sentences, I refer to utterances with statement word order (i.e., without subject-auxiliary inversion) and with an explicit subject and predicate. Finally, canonical focus refers to primary sentential prominence on the final content word of a sentence.

This work targets intraspeaker variation across situations, concentrating on variation patterns within speakers. The question under investigation is not
so much what variants mean generally across speech communities, but rather what they mean to three particular speakers in their individual and often distinct lives. In other words, how do the three speakers under investigation employ these variants for social purposes? Two speakers, Heath and Jack, are medical school students in their mid-twenties, and the third, Regan, is a buyer in his early thirties. I have known all three very well for at least two years, in the case of Regan, and up to eight years, in the case of Jack. The speakers were asked to record themselves in a variety of naturally occurring situations. For each of the three speakers, I examined speech from three speaking situations: a one-on-one social situation, a social situation in a group, and a one-on-one professional situation. These three situations were chosen to facilitate obtaining a range of speaking styles from each speaker. I was not concerned with getting comparable situations across speakers, since I do not draw cross-speaker comparisons. The comparisons made here are within individual speakers across situations, the details of which will be introduced as they become relevant.

Using acoustic representations (primarily f0 contours), each token was categorized as having a falling, rising, or level contour. Yaeger-Dror and Hall-Lew (2003) have stated that as the situational variation for a given speaker increases, so too will the range of prosodic possibilities. Accordingly, I also conducted an acoustic analysis of the same contours, measuring the maximum and minimum f0, f0 range, duration, and f0 slope for each contour. Examining details such as these may bring to light more subtle, yet perhaps equally meaningful, aspects of variability in intonational behavior. A number of potential independent variables were also considered for each token, such as the situation, the previous speaker, the following speaker, the following syntactic environment, and whether one or more function words followed the nuclear accented word. In the interest of space and relevance, I report here on only the influence of speaking situation on the realization of intonational contour.

1 Although there are multiple ToBI transcriptions within each category of contour, I make only the rough distinctions identified here, as some intonation patterns arise only at higher levels of generalization. Grabe (2002) shows for varieties of British English that as the number of morphosyntactic cues for interrogativity declines (moving from wh-questions to yes/no questions to declarative questions), the likelihood of using a rising contour increases, a pattern that emerges only when different transcriptions for the rising contour are pooled. In point of fact, the patterns to be discussed here arise whether transcriptions are pooled or not. For the sake of simplicity, only pooled results are discussed.
3 Categorical Patterns

In this section I discuss variation patterns in choice of contour for each speaker and discuss the social meanings of these patterns. Figure 1 summarizes the distributions for the falling, level, and rising contours across situations for each speaker. Beginning the discussion with Regan, the chart shows that the falling contour is the most common variant (as it is for the other two speakers as well) and that there are no differences in the frequency with which Regan uses particular variants across situations.

Heath, on the other hand, uses the rising contour relatively more frequently in his professional situation ($df=4$, $\chi^2=11.59$, $p<0.025$). In this situation, Heath is meeting with a patient with Parkinson's disease. As part of a routine examination in the degenerative nerve clinic, Heath asks his patient questions about his medical history and also briefly examines the patient’s motor skills and short-term memory.

Previous literature on the meaning of rising intonation provides a useful starting point for determining how and why Heath uses the contour with his patient. In her foundational work on language and gender, Lakoff (1973)
suggested that rising terminals could index uncertainty. McConnell-Ginet (1983) questions whether the rising contour necessarily bears such pejorative connotations, proposing that it could be employed to politely challenge, to catch someone’s attention, or to sustain talk. McLemore (1991) provides empirical support for the connective capacity of rising terminals, identifying its use in a sorority to encourage group involvement. Although these studies offer many divergent meanings of rising contours, all possibilities share a core ‘non-threatening’ meaning.

I argue that Heath draws on the non-threatening meaning of rising contours to put his patient at ease in the nerve degeneration clinic. The discourse contexts in which rising contours appear support this interpretation, as Heath uses the variant most often when offering explanations. This pattern contrasts sharply with Guy et al.’s (1986) finding for Australian English, which indicates that explanations and facts did not favor rising contours. Here explanations are the predominant type of text for rises, with representative examples appearing in (3). In (3a) we can rule out the ‘uncertain’ meaning of rises because Heath, a successful medical school student, knows what MRI stands for. In fact, all of these explanations are statements of fact, and Heath could have just as easily used a falling contour. In these examples, the meaning of the rising contour approximates the question: ‘Can you understand me?’

(3) Explanations with the rising contour (Heath)
   a. Ahh, MRI stands for magnetic resonance imaging.
   b. What causes you tremors and stuff like that is the decreased dopamine in your brain, the decreased amount of that chemical.
   c. Um, bu-, so what we do is give you the Levidopa and Carbidopa.
   d. It checks your reflexes.

Heath also uses rises to give warnings and to ask for permission, as in (4). The utterance in (4b) is simultaneously and implicitly a request for permission to proceed and a warning about what the next stage of the exam entails. While the precise meanings of the rising contours in (3) and (4) are open to debate, a general ‘non-threatening’ meaning common to all the possibilities is available. Heath’s rising intonational behavior combines with nonlinguistic stylistic practices like smiles and encouragement, together forming Heath’s pleasant bedside manner.

(4) Requests for permission/warnings with the rising contour (Heath)
   a. Alright, I’m gonna ask you some questions.
   b. Um, but I’m just gonna do a quick exam.
Figure 1 reveals that Jack, like Heath, also distributes the three contours differently across situations. For Jack, the main point of interest lies in the frequency of level intonation in his professional situation, in which he uses a strikingly high percentage of the variant \((df=4, \chi^2=11.98, p<0.025)\).

In order to understand the meaning of Jack’s level contour, let us examine his professional situation in greater depth. Jack is meeting with one of his medical school professors, with whom he is practicing interpreting EKGs. Although the purpose of the meeting strictly speaking is to develop skills, the meeting bears a stronger resemblance to a testing situation than a teaching situation, since the professor asks Jack a barrage of questions. Jack does not know many of their answers. Although Jack is very serious about medical school, he is dedicated to striking a healthy balance between studying and leisure activities. This ideological commitment enables Jack to pursue passions outside of medicine, like figure skating, gourmet cooking, and raising poison dart frogs, though sometimes these activities come at the expense of studying.

I suggest that Jack uses the level contour in his testing situation as an attractive alternative to using either of the others. The rising and falling contours have strong pragmatic meanings, both highly incompatible with Jack’s experience in this situation. He avoids the falling contour because he is too uncertain about the accuracy of his answers. As shown in Figure 1, the increase in the level contour comes at the expense of the falling contour. At the same time, Jack avoids using more rising contours so as not to sound too uncertain, particularly to his professor.

The discourse contexts in which Jack’s level variants occur support this interpretation. As shown in (5), Jack on two occasions uses the level variant to admit to not knowing the answer to a question. In nearly all other utterances pronounced with a flat contour, representative examples of which appear in (6), Jack is answering a question. In essence, Jack does not commit to a dynamic contour. One of these examples offers some additional evidence that Jack is suppressing pitch movement. The syntactic form of (6c) should necessitate that focus appear on the topicalized word, two, yet Jack displaces the sentential focus, putting it instead in sentence-final position. By doing little prosodically, when a pitch excursion is expected, Jack exposes his uncertainty.

(5) Admissions of uncertainty with the level variant (Jack)
   a. I don’t know.
   b. I made it up.
(6) Answers to questions (Jack)
   a. Normally the standardization is two blocks.
   b. So it’s not LBH.
   c. Two is the one to look for.
   d. I think that was the same patient you saw a little earlier.
   e. We can still interpret the prechordials.

To conclude this section, none of the speakers use the most frequent variant—the falling contour—more frequently across situations. This pattern is predicted, as frequent variants cannot be categorically salient and are therefore ill suited for indexing social meaning. Two of the speakers, however, exhibit stylistic patterns for the less frequent variants. Heath uses the rising contour, and Jack the level contour more often in their respective professional situations, taking advantage of the social meanings they can index. Because the rising and level contours appear less frequently in general, they are categorically salient and can thus be used stylistically for social purposes.

4 Phonetic Patterns

I now move away from categorical variation to consider phonetic variation in contour shape. Rather than focusing on the distributional properties of variants, I examine the phonetic properties differentiating tokens within a class of variant. Identifying patterns of cross-situational style shifting at the phonetic level, I situate particular phonetic forms in discourse and infer the social meaning of phonetic axes of variation.

I begin by considering the phonetic properties of the contours that appeared with different frequencies across situations. For Heath, even though the rising contour appeared more frequently in his professional setting than the other situations, the phonetic properties of the rising contours were indistinct across situations. Likewise for Jack, even though the level contour was more common in his professional situation, the phonetic quality of the level contours was not significantly different across situations. Since these variants already stand out, due to their categorical salience, the production of acoustic extremes is unnecessary. Put another way, categorical salience is sufficient for indexing social meaning. I am not trying to claim that the phonetic character of the contour could not contribute to the contour’s social meaning, only that it need not.

The question remains: Where might we expect stylistic phonetic variation? Since common variants cannot be categorically salient, because we expect to hear them, the only way for them to stand out is phonetically. The remainder of this section will therefore focus on the phonetic properties of
the most common variant, the falling contour.

Table 1 displays the acoustic properties of Jack's falls by situation, indicating that two of the acoustic dimensions — maximum f0 and f0 range — have significantly higher values in Jack's one-on-one situation, a dinner out with his boyfriend, than in his professional situation, an EKG reading session. However, as these values do not differ significantly from those in the group social situation, no clear stylistic pattern emerges, other than perhaps a weak tendency for Jack's social situations to be characterized by higher peak f0 levels and wider f0 ranges.

<table>
<thead>
<tr>
<th>Acoustic Dimension</th>
<th>Social in a Group</th>
<th>Social One-on-One</th>
<th>Professional One-on-One</th>
</tr>
</thead>
<tbody>
<tr>
<td>f0 max (Hz)*</td>
<td>165</td>
<td>172</td>
<td>137</td>
</tr>
<tr>
<td>f0 min (Hz)</td>
<td>115</td>
<td>120</td>
<td>106</td>
</tr>
<tr>
<td>f0 range (Hz)*</td>
<td>50</td>
<td>52</td>
<td>31</td>
</tr>
<tr>
<td>f0 slope (-Hz/ms)</td>
<td>0.414</td>
<td>0.289</td>
<td>0.192</td>
</tr>
</tbody>
</table>

* significantly higher for social one-on-one than for professional one-on-one.

Table 1: Acoustic properties of Jack's falling contours across situations

Although Jack did not utilize the phonetic properties of his falls for stylistic purposes, both Heath and Regan exhibit robust cross-situational patterns. I begin by discussing Regan's falls. The mean values for four of the acoustic dimensions characterizing Regan's falling contours across situations appear in Table 2. For the maximum f0, the pitch is highest in the group social setting, significantly lower for the one-on-one setting, and lowest in the professional setting. The same is true for the minimum f0 level. With regard to the f0 range, the two social settings exhibited significantly higher means than the professional setting. A similar pattern emerges for the f0 slope, which indicates how quickly the f0 falls; the higher the number, the more sharply the contour falls. For Regan, the f0 falls more sharply in the two social situations than in the professional situation. Table 2 points to the fact that Regan uses the most extreme values for these four acoustic dimensions in his group social situation.

Having established that the acoustic properties of Regan's falls are most extreme in his group social situation, let us now examine what these acoustically extreme falls mean. One way of tapping into the meaning of acoustically extreme falls is to inspect the conversational contexts in which the outliers, or most extreme versions of the falls, were uttered.2 If a particular axis

---

2 For the sake of consistency, I am using the statistical definition of an outlier, referring to values lying two standard deviations or more from the mean.
of phonetic variation indexes a particular social meaning, then outliers on that axis can be understood as the strongest indicators of the meaning. For example, if release bursts index a meaning of ‘articulate,’ relatively long and intense release bursts index articulateness more strongly than weaker bursts. In other words, I argue that the amount of phonetic substance indexes the strength of the meaning.

<table>
<thead>
<tr>
<th>Acoustic Dimension</th>
<th>Social in a Group</th>
<th>Social One-on-One</th>
<th>Professional One-on-One</th>
</tr>
</thead>
<tbody>
<tr>
<td>f0 max (Hz)*</td>
<td>224</td>
<td>191</td>
<td>150</td>
</tr>
<tr>
<td>f0 min (Hz)*</td>
<td>163</td>
<td>136</td>
<td>112</td>
</tr>
<tr>
<td>f0 range (Hz)**</td>
<td>60</td>
<td>56</td>
<td>38</td>
</tr>
<tr>
<td>f0 slope (-Hz/ms)**</td>
<td>0.431</td>
<td>0.413</td>
<td>0.237</td>
</tr>
</tbody>
</table>

* three-way significant difference between social in a group (highest), social one-on-one (next highest), and professional one-on-one (lowest)
** two-way significant difference between social in a group and social one-on-one (higher) and professional one-on-one (lower)

Table 2: Acoustic properties of Regan’s falling contours across situations

Before contextualizing Regan’s outliers in conversation, it would be useful to describe the social situation in which the majority of Regan’s outliers occur. Regan is out to dinner with six friends as part of his weekly ‘Boys’ Night Out’ ritual, when most of the core members of Regan’s friendship group, all gay men, get together for dinner and catch up on the events of the previous week. This particular week, Regan has returned from a wedding in Mexico and spends a good deal of the meal telling others about his trip. Regan is the main attraction because he has been away, but also because he is usually closer to the center of attention. At least among this group of friends, Regan takes on the ‘partier’ persona. He is regarded as perhaps the Wittiest member of the group, usually smiling, and always ready with a smart, funny, and often off-color joke. At the time of this recording, he went out to bars and clubs multiple times weekly and was not shy about alcohol consumption when he did. Regan frequently mentions drinking and partying, allowing him to establish and perpetuate his reputation as a partier.

Outlying falls, exemplified in (7), typically occur on utterances that enable Regan to construct the partier persona. In (7a), Regan uses outlying f0 values to express excitement about a tequila bar, which in turn reinforces his reputation as a partier, as drinking is the kind of activity a partier would endorse. The meaning paired with extreme f0 levels could be something as

---

3 Words appearing in italicized text were uttered with outlying f0 values on at least two acoustic dimensions.
general as ‘animated,’ which is consistent with previous experimental results showing that animated emotions like happiness and anger are marked by higher f0 levels than less expressive emotional states like sadness and neutrality (e.g., Pell 2001). This meaning, though vague, has the potential of being used as a component of the partier persona, but also to express surprise or to offer an opinion, as in (7b) and (7c), respectively.

(7) Acoustically extreme falling contours (Regan)
   a. They had a tequila bar. It’s like all different kinds of tequila. It was so classy. You know?
   b. That [train] service is pretty nice, that non-stop one...I was pretty surprised. Like, I could commute that way.
   c. It’ll be fun.

As is the case for Regan, the acoustic properties of Heath’s falls also differ from one situation to another, as shown in Table 3. The four measures are most extreme in the group social situation, with a significantly higher value for each phonetic axis of variation. One could make the argument that higher pitch properties are expected in social settings, because individuals are expected to be more animated in social contexts. The data in Table 4.1, however, reveal that Heath’s f0 values in the one-on-one social context more closely resemble those characterizing his professional setting (no acoustic dimension differentiated the social one-on-one setting from the professional setting). It appears, then, that sounding animated is not a property of social situations, but rather a property of how Heath behaves in some social situations versus others.

Before discussing the contexts in which Heath uses outlying falls, let us consider the details of the situation. In this recording, Heath is attending a barbecue with four close friends. The barbecue attendees are a rather close-knit group of people sharing many things in common. All in their mid-twenties, they all hail from middle class backgrounds and suffer through the same classes and rotations as they make their way through their second year of medical school. In spite of their shared characteristics, their interests lie in their differences, and it is the group-internal differences between Heath and his friends that determine the personae they take on. Heath distinguishes himself as a flamboyant diva, a persona markedly different from the caring doctor persona he adopts in his professional setting. The website divamind.com defines a diva as “anyone who can employ a tiara and blowtorch with equal effectiveness,” a definition which Heath embodies in some situations. Heath has an attitude, uses profanity frequently, and does not shy from expressing his often unwarranted opinions. He is also image conscious,
wearing eyeliner and blush when going out and forbidding everyone from touching his hair.

<table>
<thead>
<tr>
<th>Acoustic Dimension</th>
<th>Social in a Group</th>
<th>Social One-on-One</th>
<th>Professional One-on-One</th>
</tr>
</thead>
<tbody>
<tr>
<td>f0 max (Hz)*</td>
<td>190</td>
<td>119</td>
<td>128</td>
</tr>
<tr>
<td>f0 min (Hz)*</td>
<td>121</td>
<td>95</td>
<td>93</td>
</tr>
<tr>
<td>f0 range (Hz)*</td>
<td>69</td>
<td>24</td>
<td>35</td>
</tr>
<tr>
<td>f0 slope (Hz/ms)*</td>
<td>0.408</td>
<td>0.153</td>
<td>0.156</td>
</tr>
</tbody>
</table>

* significantly higher for social in a group than other two situations

Table 3: Acoustic properties of Heath’s falling contours across situations

Outlying f0 values, due to their associations with expressiveness, allow Heath to perform his ‘diva’ persona, as illustrated by the examples in (8). In the utterances in (8a) through (8c), Heath uses extreme falls when talking about grooming and appearances, both of which are major concerns for a diva. For example, when Heath discovers that his friends use the same shampoo everyday to wash their hair, he produces the sentence in (8b). This turn, which features an outlying fall on shampoos, expresses surprise and indignation while hinting at an interest in proper hair cleansing habits. What is more, it does so with an attitude. While these utterances may assist in Heath’s production of a diva persona, they may also serve a more general affective purpose. In (8d), for example, an acoustically extreme fall on the word park is used to offer an opinion in an animated way. Heath’s use of acoustically extreme falls to mean ‘animated’ resemble Regan’s prosodic moves in (7b) and (7c), made for the same purpose.

(8) Acoustically extreme falling contours (Heath)
   a. It’ll survive with salon quality hair gel in it.
   b. You wanna rotate your shampoos. What’s the matter with you?
   c. Ooh, ooh, a little vent thing. Oh cool, I like it. I’m so excited by your little vent thing.
   d. There’s a cute little park.

To conclude this section on the phonetic properties of falling contours, Regan and Heath both produce falls with higher f0 levels, wider f0 ranges, and steeper f0 slopes in their group social situations. Although it could be argued that such acoustic levels are to be expected for all speakers in group

---

4 In (8c), Heath expresses his excitement about a vent feature in his friend Eliza’s dress which allows for the circulation of air.
settings in order to facilitate being heard, extreme falls are not a necessary part of group interactions, as Jack’s falls are acoustically less distinct in group settings than one-on-one settings. Phonetically salient falls may aid Regan and Heath in constructing and maintaining their individual and unique personae. These and other personae may be built by drawing on a more general meaning of acoustically extreme falls, such as ‘animated,’ since partiers and divas share this characteristic. My analytical practice of pairing extreme falls with such a vague meaning allows the same linguistic form to be used in accomplishing rather different kinds of social work, such as performing the diva persona versus the partier persona. In Podesva (2006), I demonstrate how Heath and Regan build these different personae by exhibiting distinctive variation patterns at the level of the segment and in phonation type, and this comes in conjunction with non-linguistic stylistic practices.

5 Conclusion

This paper has illustrated that social meaning can attach to two kinds of linguistic form: (1) a particular category of variant (e.g., rises for Heath when examining a patient, levels for Jack when discussing EKGs with a professor) or (2) the phonetic properties of a particular token of a variant (e.g., acoustically extreme f0 excursions for Heath and Jack when performing their respective personae). The meanings conveyed by these two kinds of form may be qualitatively different. As should be evident from the data discussed in this paper, intonational variation carries a wide array of social meanings. In a recent study on rising intonation in New Zealand English, Warren (2005: 225) notes that “intonation conveys a range of linguistic and paralinguistic meanings. It is not simply used to distinguish questions and statements, but also to convey attitude and emotional involvement.” Gussenhoven (2002) explicitly divides this range of meaning into two classes: informational (pragmatic) meanings, which refer to the message being encoded, and (2) affective meanings, which refer to the speaker of the message. It may be that the phonetic details of intonation are used for more affective than pragmatic purposes. However, this distinction between affective and pragmatic meaning may be of no great consequence to the variationist. Whether affective or pragmatic, the meaning is social, and an approach to variation committed to social meaning needs to consider both phonetic and categorical aspects of variation.

In conclusion, it has been shown that in order to tap into a variable’s full range of social meanings, both categorical (phonological) and scalar (phonetic) characteristics of the variable should be examined, as different kinds of linguistic substance can index different meanings. In addition, accessing
the social meaning of variables requires an understanding of how speakers position themselves in particular contexts, both across situations and within them in conversational moments. In the future, I will examine additional speakers to determine whether the results reported here can be generalized. While the focus here has been on producing intonation, just as important is an understanding of the social perception of intonation. Whatever shape the perceptual theory takes, it will need to incorporate the two notions of salience identified here. Whereas merely using a categorically salient rising contour is sufficient for indexing social meanings, doing other kinds of social work with falling variants requires phonetic creativity. I hope that as the body of work on the social meaning of intonational variation grows, the interest in phonetic creativity and its social meaning will grow as well.

References


Department of Linguistics
Georgetown University
Washington, DC 20057
rjp39@georgetown.edu