

Variation and the Listener: The Contextual Meanings of (ING)

Kathryn Campbell-Kibler

1 Theorizing Listener Perceptions

This research addresses one of the basic questions about socially meaningful variation: how do listeners translate variation into social content? This question is, of course, important for understanding social structures and how they relate to language. But it is also important for understanding the use of linguistic variation and its spread.

By posing this question, I am positioning the process of listening as a fundamental of sociolinguistic behavior, equal in importance to speaking. Listening has not received as much attention from variationists as speaking, although this has been primarily a practical constraint rather than a theoretical position. Language attitudes, the primary area for the study of listeners, has focused on differences such as languages and language varieties. There have been multiple attempts to narrow the focus to sociolinguistic variables (Labov 1966, Rickford 1985), a trend which has strengthened in recent years (Fridland et al. 2004, Plichta and Preston 2005, Labov et al. 2005). Listeners merit this attention for several reasons: speakers can only learn about variation through listening; as speakers build performances, they are aware of their audience, and adjust their performances accordingly (Bell 2001); the reactions of listeners are part of the ongoing social work of using and defining forms.

Taking the listener as a basic sociolinguistic actor, I ask: how do tokens of a single variable contribute to listener perceptions of the speaker using it? Further, how does this effect shift for different speakers or in different utterances? I have chosen the English variable (ING) as a case study, since it is both widespread and well-studied. (ING) has been correlated in variation studies primarily with socioeconomic class and situational formality, typically measured through the manipulation of speech activities (e.g. Labov 1966, Trudgill 1974). (ING) has also been linked to age, race, gender and regional variation, particularly in the American South (see Hazen, 2005 for an overview).

There has not been much attention devoted to the question of how linguistic resources are used to build social perceptions. One possible way of understanding the relationship between social meaning and linguistic variation is as an additive process. Suppose that each variable (or, strictly speaking, each variant of the variable) indexes (connects to) a single meaning, or a small

handful of closely related meanings. Once a listener hears a linguistic performance, he or she takes note of the variants used and “translates” each to its social meaning. After decoding the variants into social content, the listener then combines the meanings obtained and forms a social picture of the speaker and the situation. This is a perfectly plausible model for processing variation and as perhaps the simplest process, it is a logical starting point in searching for a model.

My study demonstrates that this model is inadequate. Hearing variation is not an additive process. Rather than contributing a static piece of meaning to a social arithmetic procedure, the manipulation of a single variable such as (ING) shifts the relations of meanings which are connected in a flexible network. The effects of (ING) are mediated by contextual factors, including linguistic and extralinguistic cues from the speaker, message content, and individual listener reactions. I will present evidence of “peak points” where (ING) and one or more evaluative responses correlate with peaks in another response, indicating more extreme perceptions than each factor individually might suggest. The next section describes the methods of the study. After that, I will present the results, first demonstrating the general structure of responses, then discussing the role of (ING).

2 Methods

The data presented here come from a Matched Guise study on listener perceptions of (ING) in spontaneous speech. The Matched Guise Technique (MGT), developed by Lambert and colleagues in the 1960s (Anisfeld et al. 1962, Lambert et al. 1965, Lambert 1967), collects listener reactions to recordings controlled for content and speaker. By comparing reactions to linguistic performances differing only in specific ways, it is possible (ideally) to attribute differences in reactions to different attitudes toward the qualities under study. The MGT has been used to compare evaluations of different languages (Lambert et al. 1960, 1965, Sridhara 1984), language varieties (Callan and Gallois 1982, Giles et al. 1992, Purnell et al. 1999), paralinguistic cues (Smith et al. 1975, Apple et al. 1979, Ray and Zahn 1999), and sociolinguistic variables (Fridland et al. 2004, Plichta and Preston 2005). This study draws on several adaptations to the MGT, including the use of digitally manipulated stimuli (Fridland et al. 2004, Plichta and Preston 2005, Labov et al. 2005), the use of speech from interviews rather than reading passages, the inclusion of multiple examples from each speaker, the use of both open-ended interviews and a survey (Williams et al. 1976, Giles et al. 1990) and the investigation of (ING)’s

effects on interactions between responses as well as main effects.

The eight speakers in this study were university students in California and North Carolina. In all but one case, they had grown up within the state (the exception is the California speaker Elizabeth, originally from Seattle). Table 1 gives the names (pseudonyms) of the eight speakers.

	Women	Men
North Carolina	Bonnie	Robert
	Tricia	Ivan
California	Elizabeth	Sam
	Valerie	Jason

Table 1: Speakers, by region and sex.

To make the recordings used for the study, I met with each speaker twice. At the first meeting, the overall structure of the study was explained, although I did not tell them what linguistic features I would be looking at. The interview, approximately one hour in duration, was informal and dealt with both work topics and recreation or family topics. After transcribing the interview, I met again with the speaker, and explained in more detail the study's purpose, introducing (ING). We went through the transcript, and they produced alternate tokens using *-in* and *-ing* for each token of (ING) from the original interview. I showed them a given point in the transcript were they had said, for example, "I'm planning on going to grad school." and played the excerpt from the original interview, asking them to capture the speed and intonation of the original as much as possible. They then repeated the phrase twice, once with each of the two different variants.

Having a full hour of the original interview allowed flexibility in finding appropriate material for use in the study. I selected four short excerpts for each speaker, two on work- or school-related topics and two about family or recreation. Each recording contained two to six tokens of (ING). Excerpts were for the most part taken intact from the interview, although in a handful of cases I edited slightly for flow or to include more (ING) tokens. This typically consisted of removing particularly long pauses or intervening prompts from the interviewer, creating a single turn out of two adjacent turns at talk.

Using the software package Praat, I created minimal pairs by splicing in an alternate *-in* or *-ing* in place of the original tokens. When necessary, I used Praat's facilities for manipulating length, intensity and pitch to alter the inserted tokens to match the original token and each other and to maximize the

naturalness of the resulting manipulated speech. This process produced two guises for each of the 32 recordings, one with only tokens of *-in*, and one with only tokens of *-ing*.

Like the speakers, the listeners were university students in California or North Carolina, though not necessarily lifelong residents. I collected data in two phases: open-ended group interviews and a web-based survey. In conducting the group interviews I had two goals. The first was to determine what words and phrases were most commonly used to describe the speakers, both generally and for the different recordings. In the first part of each session, participants heard two recordings from each speaker and answered general questions about them:

- What can you tell me about Jason?
- Does he sound competent or good at what he does?
- Is he someone you would be likely to be friends with?
- Who do you think he's talking to? What is the context of the conversation?
- Where you think he is from?

I also wanted to solicit participants' conscious opinions regarding (ING) and its effect on the speech of these speakers. After going through the recordings once, I explained that the goal of the study was to understand the effect of (ING) on these excerpts. In the second half of the session, I played the same recordings, back to back in their minimal pairs, asking the participants to comment explicitly on the differences. A total of 55 people were interviewed, in settings ranging from one to six.

The group interview data on conscious perceptions formed the basis for selecting terms used in the experimental phase. I developed a survey (Appendix A) based on descriptions likely to be influenced by (ING) and those likely to be salient in distinguishing the speakers from one another. The survey began with a set of ratings, with which listeners indicated where speakers fell on six-point scales for qualities such as *casual/formal* or *educated/not educated*. It also included a series of checkbox descriptions such as *redneck* and *engineer*, allowing listeners to select those which they felt accurately described the speaker. 124 people finished the survey, each listening to one recording from each of the eight speakers.

The two types of questions on the survey required different forms of statistical analysis, since the ratings yielded numerical variables while the checkbox descriptions were binary responses. The connections between checkbox

descriptions and independent variables (speaker, recording, (ING) and listener demographics) were tested using the Chi Square test, as were relationships between multiple checkbox descriptions. The influence of the independent variables on the ratings and the connections between checkbox descriptions and ratings were examined using analysis of variance (ANOVA). I used a hierarchical version of ANOVA, allowing me to check connections between checkbox description and ratings while controlling for the influence of the independent variables.

3 Results

The results of the survey demonstrate that the additive model described earlier is inadequate. The social evaluations (ratings and checkbox descriptions) are impacted by the manipulation of (ING). They are also connected to each other, as different evaluations favor and disfavor the selection of others. While very few responses appear to be affected by (ING) across all the speakers, effects of (ING) more commonly differ for different speakers or recordings and shift interactions between multiple responses. These patterns show that the social contribution of (ING) is not a consistent piece of social information, but some kind of flexible force acting on a complex social network.

As the survey instrument in Appendix A shows, the response data incorporated many descriptions on many different topics. In order to demonstrate the point about the structure of the responses, I will limit myself to one area of meaning: the nexus of responses involving class, education, intelligence and formality. Because this terrain is complex, I will not be filling in the details connecting these responses to others. For example, region and accent, particularly with respect to the U.S. South, was a major theme in the study. All of the meanings I will be describing are also informed by this domain, particularly in that the Southerners were much more likely to be described as *working-class* and less likely to be described as *wealthy*. The relationship of these issues to region does not bear on the theoretical point I will be making but it will be useful to bear in mind in order to understand the social logic of the results.

Figure 1 shows the relationships between the responses in this domain. This graph is intended as a general picture, roughly indicating the strength of the favoring and disfavoring relationships. The graph merely indicates two levels of strength, rather than including statistical evaluations, as the connections between different responses were calculated differently. The ratings (in circles) were measured against each other using correlations and against checkbox descriptions (in hexagons) using ANOVA. Checkbox descriptions'

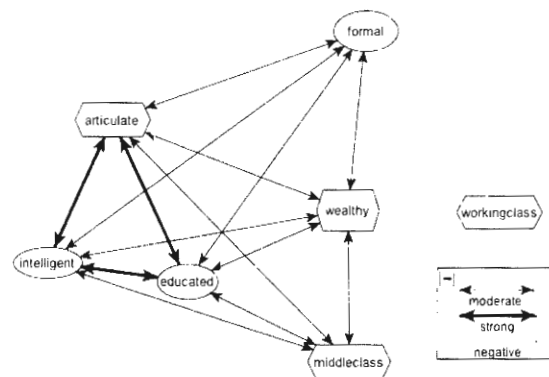


Figure 1: Meaning network

connections to each other, as mentioned above, were investigated using the Chi Square test. The darker lines between *articulate*, *educated* and *intelligent* indicate very strong connections, while the regular black lines show moderate (but statistically significant) positive associations. The lightest lines show the disfavoring relationships between *working class background* and *formal*, *wealthy*, *articulate*, and *educated*. Variationists have long been aware that social structures are connected to each other, tending to regard this as a hindrance to research rather than a piece of the overall puzzle. However, this network forms the social context within which linguistic resources operate.

Two of the responses in this domain (and two others outside of it) were significantly influenced by (ING) across all the data. Survey listeners were significantly more likely to describe speakers as *articulate* when they used *-ing*, as Table 2 shows¹ and more likely to describe them as *redneck* and *polite* when they used *-in*. Listeners also rated the *-ing* guises significantly more *educated* than the *-in* guises, as Table 3 shows.

These results may indicate that these qualities reflect general meanings of (ING). However, the many other effects without clear connections to these suggest that they are not the full picture. It's possible that these qualities (particularly *educated*, the most robust result) will appear as perceptual correlates of (ING) in other populations or with studies carried out in different ways, but we cannot assume that they will. The respondents were university students,

¹In this and other tables, numbers in italics indicate the significantly greater value at $p < 0.05$. The numbers in bold indicate the significantly greater value at $p < 0.01$.

	% listeners selecting checkbox		
	<i>-in</i>	<i>-ing</i>	<i>p</i>
<i>articulate</i>	21	27	0.037
<i>redneck</i>	12	8	0.033
<i>polite</i>	21	16	0.034

Table 2: *Articulate*, *redneck* and *polite* selections, by (ING).

	<i>-in</i>	<i>-ing</i>
<i>educated</i>	3.81	3.98

Table 3: *Educated* ratings, by (ING) ($p=0.007$).

asked to evaluate speakers as part of a linguistic study. These contextual factors may well have created an environment where education and articulateness were particularly highlighted and thus more available to be manipulated by (ING) across a variety of different recordings.

The importance of the networked structure of the social evaluations is best seen in the interactions of (ING) and two or more survey responses. As Figure 1 showed, evaluations on the *casual/formal* dimension were unevenly distributed with respect to class evaluations. Speakers described as from a *wealthy background* were rated as more *formal* than those who were not, while those marked as from a *working class background* were rated as more *casual* than those who were not. Table 4 shows that this latter pattern is primarily driven by responses to the recordings containing *-in*.

	<i>-in</i>	<i>-ing</i>
<i>not working-class</i>	2.84	2.79
<i>working-class</i>	2.29	2.71

Table 4: *Casual/formal* ratings, by (ING) and *working-class* ($p=0.025$).

We see a similar interaction involved in the description of the speaker as *in his/her 30's* and *intelligent*, shown in Table 5. In this interaction, however, the distribution shows a pattern I have called “peak points”: when speakers were heard using *-ing* and described as *in [their] 30s*, they received a peak in *intelligent* ratings, higher than all three other conditions. Indeed, when

	<i>-in</i>	<i>-ing</i>
not <i>in his/her 30s</i>	3.75	3.73
<i>in his/her 30s</i>	3.88	4.24

Table 5: *Intelligent* ratings by (ING) and *in his/her 30s* ($p = 0.006$).

either of these findings was missing, the other made no difference in perceived intelligence.

This effect is even more pronounced with the inclusion of the description *working class*, shown in Table 6. Speakers using *-ing* and described as *in*

		<i>-in</i>	<i>-ing</i>
not <i>in his/her 30s</i>	not <i>working-class</i>	3.86	3.73
	<i>working-class</i>	3.22	3.74
<i>in his/her 30s</i>	not <i>working-class</i>	4.00	4.45
	<i>working-class</i>	3.50	3.65

Table 6: *Intelligent* ratings by (ING), *in his/her 30s* and *working-class* ($p = 0.023$).

[*their*] *30s* and not *working-class* have an even greater peak in *intelligent* ratings. Conversely, *-in* users described as *working-class* and not as *in [their] 30s* have a noticeably lower mean rating of intelligence than any other category.

This finding likely relates to the connection between age and career path, given that half of the recordings have to do with work topics or experiences. From the perspective of university students, it is possible that the categories in question represent a particular life stage, that of a successful professional, an image which marks those in it as highly intelligent while those clearly outside it are seen as less so. (ING) represents one of many linguistic and extralinguistic ingredients which aid listeners in identifying a speaker as belonging to this category.

4 Conclusions

The additive process discussed earlier is a plausible model for the construction of a social image out of many linguistic cues but the complexity of this data shows it to be insufficient. New models are needed, particularly ones which account for the relationships among social structures. The fact that a single

variable changes the interactions of the social structures suggests that listeners are not evaluating each linguistic contribution individually prior (either temporally or logically) to combining them mentally. Instead, it seems likely that listeners are perceiving linguistic variation in groups of features, rather than individual tokens. It remains a priority to, as much as possible, untangle these connections in order to understand whether a given linguistic resource is associated with all of its social correlates or only some directly. But while doing so we must remember that listeners and speakers are aware of the connections between social concepts as well as the connections between the social and the linguistic. As a result, asking which social correlates are the “true” meanings may be missing the point.

Appendix A: Survey Instrument

This is Ivan:



Press the play button to hear the recording. You can play it as many times as you like. After listening to him, tell me as much as you can about Ivan, based on what you hear.

He sounds:

- | | | | | | | | |
|----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|---------------------|
| Not At All Masculine | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Very Masculine |
| Talking Very Slow | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Talking Very Fast |
| Very Shy | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Very Outgoing |
| Very Accented | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Not At All Accented |
| Educated | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Not Educated |
| Intelligent | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Not Intelligent |
| Casual | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Formal |

How old does Ivan sound (check all that apply, must choose at least one)?

- A Teenager College Age Under 30
 In His 30's Over 40

From what you heard, does Ivan sound like he might be (check all that apply):

- Lazy Hardworking Laidback
 Compassionate Knowledgeable Condescending
 Confident Articulate Religious

- Lonely Annoying Family-Oriented
 Funny Reliable Gay
 Hip/Trendy
 A Stoner A Valley Girl A Metrosexual
 A Jock A Redneck A Nerd
 A Farmer A Student A Banker
 An Artist An Engineer In A Fraternity

Other: _____

How well does he know the person he's talking to?

Best Friend ○ ○ ○ ○ ○ ○ Stranger

Right now, does he sound like he might be

(check all that apply):

- Nostalgic Bored Complaining
 Joking Arguing Chatting
 Bragging

 Selling Something Applying for a Job Giving a Lecture
 Being Polite Trying to Impress Hiding Something

 Other: _____

Where does Ivan sound like he might be from

(check all that apply, must choose at least one)?

- The South New England The Midwest
 The West Coast The East Coast The Southwest
 The North Anywhere
 The City The Country
 The Suburbs
 A Wealthy Background A Middle-Class Background
 A Working-Class Background
 Other: _____

Any other thoughts about Ivan? _____

References

- Anisfeld, Moshe, Norman Bogo, and Wallace E. Lambert. 1962. Evaluational reactions to accented English speech. *Journal of Abnormal and Social Psychology* 65:223–231.
- Apple, William, Lynn A. Streeter, and Robert M. Krauss. 1979. Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology* 37:715–727.
- Bell, Allan. 2001. Back in style: Reworking audience design. In *Style and Sociolinguistic Variation*, ed. Penelope Eckert and John R. Rickford, 139–169. Cambridge: Cambridge University Press.
- Callan, Victor J., and Cynthia Gallois. 1982. Language attitudes of Italo-Australian and Greek-Australian bilinguals. *International Journal of Psychology* 17:345–358.
- Fridland, Valerie, Kathryn Bartlett, and Roger Kreuz. 2004. Do you hear what I hear? Experimental measurement of the perceptual salience of acoustically manipulated vowel variants by Southern speakers in Memphis, TN. *Language Variation and Change* 16:1–16.
- Giles, Howard, Nikolas Coupland, Karen Henwood, Jim Harriman, and Justine Coupland. 1990. The social meaning of RP: An intergenerational perspective. In *Studies in the Pronunciation of English: A Commemorative Volume in Honor of A. C. Gimson*, ed. S. Ramgaran, chapter 17, 191–211. New York: Routledge.
- Giles, Howard, Karen Henwood, Nikolas Coupland, Jim Harriman, and Justine Coupland. 1992. Language attitudes and cognitive mediation. *Human Communication Research* 18:500–527.
- Labov, William. 1966. *The Social Stratification of English in New York City*. Center for Applied Linguistics: Center for Applied Linguistics.
- Labov, William, Sharon Ash, Maciej Baranowski, Maya Ravindranath, Naomi Nagy, and Tracy Weldon. 2005. Listeners' sensitivity to the frequency of sociolinguistic variables. Paper presented at NWAV 34.
- Lambert, Wallace E. 1967. A social psychology of bilingualism. *Journal of Social Issues* 23:91–109.
- Lambert, Wallace E., Moshe Anisfeld, and Grace Yeni-Komshian. 1965. Evaluational reactions of Jewish and Arab adolescents to dialect and language variations. *Journal of Personality and Social Psychology* 2:84–90.
- Lambert, Wallace E., R. C. Hodgson, R. C. Gardner, and S. Fillenbaum. 1960. Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology* 60.
- Plichta, Bartek, and Dennis R. Preston. 2005. The /ay/s have it: The perception of /ay/ as a North-South stereotype in US English. *Acta Linguistica Hafniensia* 243–285. Theme issue on “Subjective processes in Language Variation and Change”, Edited by Tore Kristiansen, Nikolas Coupland and Peter Garrett.
- Purnell, Thomas, William Idsardi, and John Baugh. 1999. Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology* 18:10–30.

- Ray, George B., and Christopher J. Zahn. 1999. Language attitudes and speech behavior: New Zealand English and Standard American English. *Journal of Language and Social Psychology* 18:310–319.
- Rickford, John R. 1985. Standard and non-standard language attitudes in a creole continuum. In *Language of Inequality*, ed. Nessa Wolfson and Joan Maines. New York: Mouton de Gruyter.
- Smith, Bruce L., Bruce L. Brown, William J. Strong, and Alvin C. Rencher. 1975. Effects of speech rate on personality perception. *Language and Speech* 18:145–152.
- Sridhara, A. 1984. A study of language stereotype in children. *Journal of Psychological Researches* 28:45–51.
- Trudgill, Peter. 1974. *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.
- Williams, Frederick, Nancy Hewett, Robert Hopper, Leslie M. Miller, Rita C. Naremore, and Jack L. Whitehead. 1976. *Explorations of the Linguistic Attitudes of Teachers*. Rowley, Mass.: Newbury House.

Department of Linguistics
Stanford University
Stanford, CA 94040
kcat@csl.stanford.edu