Non-Categorical Constraints in Perception*

Emily Bender

Final version, April 8, 2000
To appear in *Penn Working Papers in Linguistics*
http://www.ling.upenn.edu/papers/pwpl.html

1 Introduction

Sociolinguistic inquiry, as pioneered by Labov (1966), has established that variation in linguistic form is subject to non-categorical but still systematic grammatical constraints. The existence of these constraints raises a question for generative linguistic theories: What is the relationship of such constraints to competence grammar?

Three possibilities have been proposed in the literature: On the Variable Rules model (Labov 1969, Cedergren and Sankoff 1974, Guy 1997), non-categorical constraints are another kind of arbitrary convention to be coded directly in the grammar. The second proposal is that non-categorical constraints follow from partial (Anttila 1997) or probabilistic (Boersma 1997) rankings of constraints in an OT grammar. On the OT proposal, the probability of a candidate occurring is derived from the number of complete rankings that select that candidate. Since OT constraints can be context-sensitive, different environments of a variable will interact with the constraints differently, deriving the pattern of probabilistically favoring or disfavoring environments found in the production data. The third proposal is that non-categorical constraints follow from functional considerations (Kiparsky 1972, Kiparsky 1988).

While there are many differences between these accounts, perhaps the most interesting is this: The Variable Rules model, but not the others, requires that speakers have direct (if tacit) knowledge of non-categorical constraints on variation. This paper reports on an experiment designed to investigate whether speakers do have such knowledge.[1]

*I would like to thank Penny Eckert, Tom Wasow, John Rickford and Stacy Fambro for their help in the design of this experiment, Jennifer Iljas and Mark Thomas for helping me to recruit students and providing a place to run the experiment, and Kristofer Jennings for his help with the statistical analysis. All errors remain my own.

[1]Space limitations prohibit a full discussion of the methodology and implications of this experiment. For all the details, please see Bender forthcoming.

2 The Variable: The AAVE Copula

The variable considered here is copula absence vs. presence in African American Vernacular English (AAVE). In addition to the full (1a) and contracted (1b) forms that it shares with other varieties of English, AAVE also allows sentences such as (1c) with no overt form of the copula at all.

(1) a. She is my piano teacher.
 b. She's my piano teacher.
 c. She my piano teacher.

This study contrasts the contracted and zero forms, although Labov (1969) argues that AAVE copula contraction is subject to similar non-categorical constraints.

AAVE copula absence/presence is subject to a number of non-categorical grammatical constraints. Here, I will focus on the effect of the following grammatical environment, i.e., the part of speech of the predicate. Production studies of AAVE across the U.S. have turned up remarkably effects of the following grammatical environment on copula absence. The ordering of environments, from least to most favoring, is as in (2). There is some disagreement across studies as to the ranking of following locatives and adjectives.

(2) __ NP < __ Loc < __ Adj < __ V+ing < __ gon[2]

Table 1, adapted from Rickford 1998, summarizes these studies.[3]

Studies			Following Environment				
Form	Place	Source	NP	Loc	Adj	V+ing	gon
is	NYC Thunderbirds	Labov 1969	.2	.36	.48	.66	.88
is	NYC Jets	Labov 1969	.32	.52	.36	.74	.93
is	NYC Cobras	Baugh 1979	.14	.31	.72	.59	.78
is+are	Detroit WC	Wolfram 1969	37%	44%	47%	50%	79%
is	LA	Baugh 1979	.32	.29	.56	.66	.69
are	LA	Baugh 1979	.25	.69	.35	.62	.64
is+are	Texas kids	Bailey & Maynor 1987	.12	.19	.25	.41	.89
is+are	Texas adults	Bailey & Maynor 1987	.09	.15	.14	.73	.68
is+are	EPA	Rickford et al. 1991	.29	.42	.47	.66	.77

Table 1: Copula absence in AAVE in different communities, by following environment. Adapted from Rickford 1998:190.

[2] *Gon* is a form of *gonna*.

[3] In this table, the values for the Wolfram study are percentages and all others are Varbrul weights. On Varbrul, see Guy 1988.

As pointed out by Mimi Lipson (p.c.) and others, the effects found in the production studies might actually be driven by semantic rather than syntactic properties of the predicate. However, semantic constraints would still be grammatical constraints. Further, it seems unlikely that a reanalysis of the production data in terms of semantic categories would turn up only categorical constraints, given minimal sets such as in (1). Although the syntactic effect shown in Table 1 may only be a reflection of a semantic effect, the fact remains that it is robust. Syntactic or semantic, it must therefore either be a part of the grammar of AAVE or follow from something else in the grammar.

3 Hypothesis

If speakers have knowledge of non-categorical constraints on variation, it is almost certainly tacit and inaccessible to introspection. Fortunately, Labov's (1963) finding that sociolinguistic variation is socially meaningful provides a jumping off point for constructing an experiment. If sociolinguistic variation is socially meaningful, then the social value of variants might interact with the non-categorical constraints. In particular, I propose the following two-part hypothesis:

I Copula absence/presence in AAVE is associated with some social value.

II Copula absence/presence in AAVE is more strongly associated with its social value the more marked the environment is for each variant.

For example, if copula absence sounds confident, then copula absence before a noun should sound especially confident and copula absence before a verb somewhat less so. Note that Part II of this hypothesis entails that speakers have knowledge of non-categorical constraints.

4 Methodology

The experimental design was based on the matched-guise methodology of Lambert et al. (1975). The same talker is recorded in different 'guises' (here, saying different sentences). Listeners rate the talkers on a number of characteristics. Scores are compared across sentences.[4]

4.1 Stimuli

4.1.1 Test Sentences

The test sentences were designed to meet the following desiderata: They should contrast two maximally distinct following grammatical environments, be matched for semantic and

[4] I will be using 'talker' to refer to the people who made the recordings and 'listener' to refer to those who judged the recordings. In some cases, it will be important to highlight the competence of certain listeners in certain varieties. In those cases, the term 'speaker' will be used to refer to the listeners.

phonetic content as closely as possible, contain no other stereotyped variables (which might swamp any effect of copula absence/presence), and be relatively short to reduce the possibility of uncontrolled variables. Since the most favoring environment (before *gon*, a variant form of *gonna*) involves another AAVE variable (*gon* itself), the test sentences contrast the NP and V+*ing* environments. The test sentences are those in (3).[5] These were matched with an equal number of filler sentences, given in (4).

(3) a. Yeah I know her. She's teachin me piano at Music World.
b. Yeah I know her. She's my piano teacher at Music World.
c. Yeah I know her. She teachin me piano at Music World.
d. Yeah I know her. She my piano teacher at Music World.

(4) a. Yeah I know her. She useta teach me piano at Music World.
b. Yeah I know her. She useta be my piano teacher at Music World.
c. Yeah I know her. She taught me piano at Music World.
d. Yeah I know her. She was my piano teacher at Music World.

Eight bidialectal African-American women were recruited from the Stanford community to be the talkers. Four spoke the test sentences and four spoke the filler sentences. In order to avoid reading pronunciations and to keep intonation and other factors constant across talkers, the talkers were recorded in groups as much as possible. The first group (three participants) discussed how the sentences should sound and practiced saying them. Later participants, some of whom also came in groups, were played the tape of the first recording so that they could hear how the other talkers decided to render the sentences. This strategy was successful in producing natural sounding AAVE tokens and avoiding reading pronunciation. Note that all of the sentences were rendered with similar intonation and pronunciation. It is not the case that the copula presence sentences represent the standard variety.[6]

To create the stimuli actually used for the experiment, the best tokens were selected out of the repeated recordings. I pasted the *yeah I know her* from a fifth recording onto the beginning of each stimulus, followed by 0.15 seconds of silence, and added leading and trailing 0.1 seconds of silence. The test stimuli were 2.482–2.698 seconds long (mean 2.592 seconds) and the filler stimuli were 2.360–3.208 seconds long (mean 2.711 seconds).

4.1.2 Test Scales

The Lambert study on which this methodology is based involved 20 scales representing personality traits. Since the stimuli were much shorter in the present study, the number of scales was also reduced to avoid listeners having to rate the talkers on scales after they'd forgotten what the talker sounded like. The seven seven-point scales given in (5) were used, each presented with a contextualizing question such (6).

[5] Ideally, such an experiment would involve multiple similar test sets. However, in order to keep the listeners' task to a manageable length, only the test set in (3) was used.

[6] Cf. Labov's 1969 finding that copula presence/absence is variable even in sounding (ritual insult).

(5) comical – not comical
confident – not confident
well educated – not well educated
good job – not a good job
likeable – not likeable
polite – impolite
reliable – unreliable

(6) "How likeable does this person sound?"

4.2 Participants

35 participants were recruited from a introductory psychology course at a community college in a community in California with a large African-American population. They earned extra credit in their course for their participation and were also paid $5. The participants fell into the ethnically and linguistically defined groups given in Table 2.

I	AAVE speakers (all African American)	N = 11
II	African Americans who do not identify as AAVE speakers but are familiar with AAVE	N = 5
III	participants who are familiar with AAVE but are not African American	N = 6
IV	native speakers of English not familiar with AAVE	N = 6
V	non-native speakers of English not familiar with AAVE	N = 7

Table 2: Groups of listeners

Familiarity with AAVE was determined on the basis of self-report data: In the debriefing I explained to the participants what I was looking for and added that I only expected to find it among people who are familiar with AAVE. I then asked the participants if they were familiar with AAVE. All of the African-American participants said they were familiar with AAVE. Some then went on to say they "sometimes talk that way at home", while others distanced themselves from AAVE with statements like "my daughter is picking that up at school, but I'm trying to teach her to speak correctly." Also, some non-African-American participants indicated that they were familiar with AAVE from having gone to a predominantly African-American high school or from Hip Hop culture. Note that native-speaker status in any variety of English is not required for membership in Groups II and III. Groups II and III are distinguished on the basis that the African Americans' experience with AAVE is substantially different from the others'.

4.3 Running of the Experiment

The experiment was run on a computer with PsyScope (Cohen et al. 1993), which presented the stimuli in a near-random order, beginning with one practice stimulus. Participants heard a stimulus and then were presented with the seven scales in a random order. (The placement of the positive end of the scale was also determined randomly.) Each scale stayed on the screen until the participant selected a point on the scale or it timed out at 20 seconds.

The description of the purpose of the study that participants saw was:

> You are invited to participate in a research study on how consistently people judge other people on the basis of their voices.

That is, the subjects were not told to attend to the copula. They could base their judgments on anything about the stimulus. The other factors they may have attended to were, however, controlled for to the extent possible. Would they attend to copula absence/presence and would it have a systematic effect on their ratings of the talkers?

4.4 Copula Presence vs. Copula Absence

This subsection evaluates the first part of the hypothesis, repeated here:

I Copula absence/presence in AAVE is associated with some social value.

To evaluate this hypothesis, I averaged the ratings across talkers, within listeners, sentences and scales. This gives each listener's average rating on each scale of each sentence. The four average ratings define an order of the sentences on each scale by each listener. Since there were 35 listeners and 7 scales, this gives a total of 245 such observations. (7) gives a hypothetical observation. In (7), the sentences are represented by symbols, according to the following key: P indicates copula presence, A copula absence, N indicates an NP predicate and V a V+ing predicate.

(7)									
			AN	AV		PV		PN	
	impolite	•	•	•	•	•	•	•	polite

Abstracting away from N and V, there are six possible orders of the two P sentences and the two A sentences, where the positive end of the scale (polite, comical, etc.) is aligned to the right.[7]

(8)
1. **A A P P**
2. A P P A
3. A P A P
4. P A A P
5. P A P A
6. **P P A A**

If the P and A sentences were strictly differentiated as in orders 1 and 6, then the listener judged the scale to be relevant to the social value of copula absence/presence. Of the 245 observations, there were 9 observations with order 6 and 115 with order 1. Accordingly, we will concentrate on the cases where the P sentences were rated strictly higher on a scale than the A sentences.

[7] Plus a few more where two or more sentences were given exactly the same rating.

Note that even among the AAVE speakers (Group I), there was a negative evaluation of copula absence and/or a positive evaluation of copula presence. This might be the result of linguistic insecurity (Labov 1966). However, it seems at least equally likely that it was the result of the experimental situation: at school, with a computer, with a white researcher.

Table 3 gives the percentage of listeners in each group that selected each scale as relevant (with order 1). The scale 'educated' was relevant for the most speakers, across groups. The scale 'comical' was the least relevant.[8] Group II listeners were most likely to judge a scale to be relevant to the social value of copula absence/presence. Group V listeners were the least likely to do so.

Scale	Group				
	I n=11	II n=5	III n=6	IV n=6	V n=7
educated	64%	80%	100%	67%	43%
job	45%	100%	67%	67%	29%
reliable	55%	100%	67%	50%	14%
polite	27%	100%	50%	50%	29%
confident	27%	60%	17%	50%	14%
likeable	45%	40%	17%	67%	29%
comical	18%	40%	17%	17%	29%

Table 3: P* > A*, by group and scale

There was no single scale selected by all of the listeners in group I. However, as Table 4 shows, most of the listeners in this group did select at least one scale as relevant. Further, there is fairly good agreement as to which scales are relevant, with the less frequently selected scales only being selected by listeners who also select other scales.

To summarize the findings of this section, across groups I-IV, the listeners attended to copula absence/presence and it influenced their ratings. This allows us to test part II of the hypothesis.

4.5 Effect of Following Grammatical Environment

This subsection evaluates the second part of the hypothesis, repeated here:

> II Copula absence/presence in AAVE is more strongly associated with its social value the more marked the environment is for each variant.

Here, the null hypothesis is that the following grammatical environment has no effect, so the orderings of the V and N sentences should be equally distributed. The alternative hypothesis is that the following grammatical environment systematically affects the ratings, and one ordering of the V and N sentences occurs significantly more than chance. In particular, the ordering in (9) should come up significantly more than chance. In this order, the

[8]But note that 4 of 9 A > P cases were for 'comical'.

	Listener											
Scale	A	B	C	D	E	F	G	H	I	J	K	Total
job	X	X	X		X	X	/[a]	/	/			8
educated	X	X	X	X	X		X	X				7
reliable	X	X	X	X	X		X					6
likeable	X	X	X	X		X						5
polite	X		X	X								3
confident	X			X		X						3
comical	X	X										2
Total	7	5	5	5	3	3	3	2	1	0	0	34

[a]X indicates P strictly greater than A, / indicates a scale where the lower P sentence and the higher A sentence were rated the same.

Table 4: P* > A* by listener and scale, Group I

two unusual cases (PV and AN) are rated more towards the end of the scale, while the two more common cases (PN and AV) are rated more towards the middle.

(9) PV > PN > AV > AN

In what follows, I will consider the P and A sentences separately, as any given group may in fact only be attributing social value to either copula presence or copula absence.

The Exact Binomial Test can distinguish between the null and alternative hypotheses in a dataset like this one. The results of this test are given in Table 5. The first column of this table gives the group. The second column gives the total number of observations for the group. For example, there were 11 listeners in Group I and 7 scales, giving 77 observations for Group I. The third column gives the number of observations in which the listener rated both of the P sentences higher than both of the A sentences on the scale. Since this part of the hypothesis concerns the effect of the following grammatical environment on the social value of the variable, it was only tested within those cases where the listener found the scale relevant to that social value. (For Group I, there were 31 such cases.)

Group	Total N	P* > A*	AV > AN	p	PV > PN	p
I	77	31	[a]12/28	0.2888	*18/23*	*0.0053*
II	35	26	*23/25*	*0.0000*	16/25	0.1148
III	42	20	[a]6/17	0.1662	9/18	0.5927
IV	42	22	[a]4/14	0.0898	8/14	0.3953
V	49	16	5/9	0.5000	[b]3/10	0.1719
I&II	112	57	*35/53*	*0.0135*	*34/48*	*0.0028*
III-V	133	58	[a]15/40	0.0769	[b]20/42	0.4388

[a]AN > AV more frequent [b]PN > PV more frequent.

Table 5: Effect of following grammatical environment, all groups

The fourth column gives the proportion of observations that matched the predicted order of AV and AN. The denominator in the fourth column is less than the number in the third column, as any cases where AV averaged the same as AN were discarded.[9] The next column gives the p values produced by comparing this proportion to 1/2 (chance distribution) with the Exact Binomial Test.

The only significant effect of the following grammatical environment on copula absence was for Group II, and it goes in the direction predicted. For Groups I and III-V, the distribution of the two orders of AN and AV is not distinguishable from chance. The last two columns give the results for the effect of the following grammatical environment on copula presence. In this case, the only significant effect was for Group I, again in the predicted direction. Finally, in order to make sure that the lack of significant results for Groups III-V was not due to the small sample size, the last row of the table gives the results for the combined group III-V. Even for this larger group, no significant result emerges.

Why should there be a difference between Groups I and II? Recall that Group I listeners self-identified as AAVE speakers, while Group II listeners said they were familiar with AAVE but did not self-identify as AAVE speakers. It appears that for Group I, copula presence is the meaningful variant, while copula absence doesn't carry any particular social value. For Group II, on the other hand, it is copula absence that is marked and meaningful. Crucially, both groups' responses are sensitive to the same non-categorical constraint: it is the unusual cases that are rated more towards the end of the scale, whichever end is in play. Further, no such effect was found among the control groups.

5 Conclusion

The experiment shows an interaction between the social value of a variant and its grammatical environment, perhaps mediated by knowledge of frequencies. This should be confirmed with more speakers, more constraints, and more variables. However, the results of this experiment suggest that people have direct (if tacit) knowledge of this interaction. Even if the production data can be explained in functional terms, these perception data require the listeners to know that copula absence in AAVE is marked before NP/copula presence is marked before V+ing.

The next question is: Is this linguistic knowledge part of these speakers' grammars, or is it represented separately? I would like to argue that that depends on what the grammar is supposed to model. If the grammar is only supposed to generate 'all and only the sentences of the language', then this kind of information is clearly not needed. However, the systematic aspects of language are clearly not restricted to those that affect acceptability/grammaticality judgments, and perhaps a more inclusive notion of grammar is warranted.

References

Anttila, Arto. 1997. Deriving variation from grammar. In F. Hinskens, R. van Hout, and

[9] That is, they were considered to be cases where the instrument was not sensitive enough to tell which way the order went.

W. L. Wetzels (Eds.), *Variation, Change and Phonological Theory*, 35–68. Amsterdam: John Benjamins.

Bailey, Guy, and Natalie Maynor. 1987. Decreolization? *Language in Society* 16:449–473.

Baugh, John. 1979. *Linguistic Style Shifting in Black English*. PhD thesis, University of Pennsylvania.

Bender, Emily. forthcoming. *Syntactic Variation and Linguistic Competence: The Case of AAVE Copula Absence*. PhD thesis, Stanford University.

Boersma, Paul. 1997. How we learn variation, optionality, and probability. In *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam*, Vol. 21, 43–58.

Cedergren, Henrietta J., and David Sankoff. 1974. Variable rules: Performance as a statistical reflection of competence. *Language* 50(2):333–355.

Cohen, Jonathan D., Brian MacWhinney, Matthew Flatt, and Jefferson Provost. 1993. PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers* 25:257–271.

Guy, Gregory R. 1988. Advanced Varbrul analysis. In K. Ferrara, B. Brown, K. Walters, and J. Baugh (Eds.), *Linguistic Change and Contact*, 124–136. Austin, Texas: University of Texis Department of Linguistics.

Guy, Gregory R. 1997. Competence, performance, and the generative grammar of variation. In F. Hinskens, R. van Hout, and W. L. Wetzels (Eds.), *Variation, Change and Phonological Theory*, 125–143. Amsterdam: John Benjamins.

Kiparsky, Paul. 1972. Explanation in phonology. In S. Peters (Ed.), *Goals of Linguistic Theory*. Englewood Cliffs NJ: Prentice-Hall, Inc.

Kiparsky, Paul. 1988. Phonological change. In J. Newmeyer, Frederick (Ed.), *Linguistics: The Cambridge Survey*, Vol. Vol. 1: Linguistic Theory: Foundations, 363–415. Cambridge: Cambridge University Press.

Labov, William. 1963. The social motivation of a sound change. *Word* 19:273–309.

Labov, William. 1966. *The Social Stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.

Labov, William. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45(4).

Lambert, Wallace E., Howard Giles, and Omer Picard. 1975. Language attitudes in a French-American community. *Linguistics* 158:127–152.

Rickford, John R. 1998. The creole origins of African-American vernacular English: Evidence from copula absence. In S. S. Mufwene, J. R. Rickford, G. Bailey, and J. Baugh (Eds.), *African-American English*, 154–200. New York: Routledge.

Rickford, John R., Arnetha Ball, Renee Blake, Raina Jackson, and Nomi Martin. 1991. Rappin on the copula coffin: Theoretical and methodological issues in the analysis of copula variation in African-American Vernacular English. *Language Variation and Change* 3:103–132.

Wolfram, Walter A. 1969. *A Sociolinguistic Description of Detroit Negro Speech.* Center for Applied Linguistics.

Department of Linguistics
Stanford University
Stanford, CA 94305-2150
bender@csli.stanford.edu