On the unobservability of structure:
Perception of phonology vs. phonetics in experimental data

Betsy Sneller, University of Pennsylvania

The ability of speakers to identify structural variables, such a phonological rule, has been the topic of some debate. Labov (1993) argues that linguistic structure is unobservable and that it is instead only the phonetic output that is subject to social evaluation. More recent work (Campbell-Kibler 2007, Dinkin 2015) carries this argument further with the assertion that listeners attach social meaning to the variant itself (such as the use of "like" across the different variables of quotatives and discourse markers), regardless of the structural composition of the variable. Philadelphia English provides a unique opportunity to test the observability of phonology, as it is currently undergoing an allophonic restructuring in the TRAP-MAN allophonic split (Labov et al. 2016), with younger speakers abandoning the local Philadelphia system (henceforth: PHL) in favor of the more geographically widespread Nasal system (henceforth: NAS). The difference between PHL and NAS is mainly found in four conditioning factors, outlined in Table 1. Crucially, both PHL and NAS produce phonetically tense and phonetically lax allophones; the main difference is in the structural conditioning factors that govern whether a word is produced as tense or lax.

	Conditioning Factor			
	Lexical mbg (e.g. "mad")	Fricative (e.g. "class")	nVn (e.g. "manage")	Engma (e.g. "hang")
PHL	**Tense**	**Tense**	**Lax**	**Lax**
NAS	**Lax**	**Lax**	**Tense**	**Tense**

Table 1: Conditioning factors that differ between PHL and NAS

In this paper, we test the observability of structure through two perception experiments. 68 participants from Philadelphia completed both experiments. The first experiment is a controlled matched guise task in which participants heard a story, wherein words containing /æ/ were manipulated to conform to PHL in the Philly condition and NAS in the Nasal condition. In both conditions, participants were exposed to the same number of tense and lax tokens; the only difference was in the phonological conditioning factors governing the tense-lax split. Participants were then asked to rate the speaker along a number of social attributes.

Figure 1: Attribute ratings by story condition and age of participant
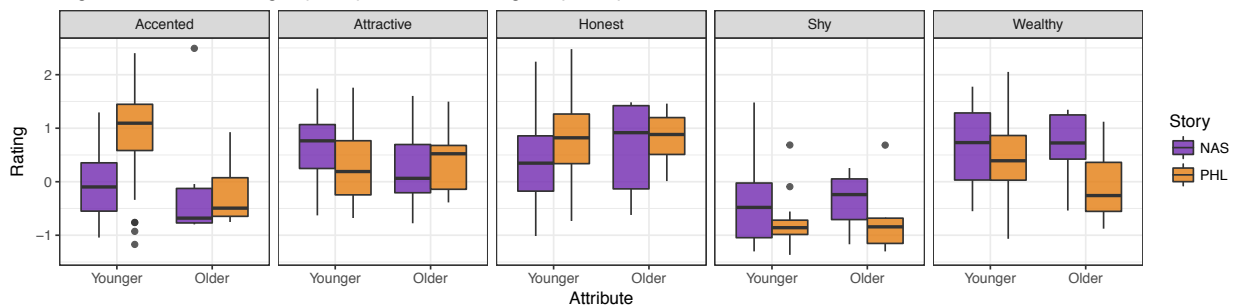
Results from Experiment 1 are shown in Figure 1. A fixed-effects regression model with Story Condition and Age of Participant as main effects was run separately over each attribute to determine whether the evaluation of /æ/ system changed across age group. We find a main effect

for four social attributes, with PHL rated as more "Accented" (p=.003) and "Honest" (p=.03), and NAS rated as more "Shy" (p=.03) and more "Wealthy" (p=.04). We find an interaction effect for "Attractive", with younger participants rating NAS as more attractive (p=.04). These results suggest that listeners do in fact attend to structural factors like allophonic system in their evaluation of speech, contra Labov (1993).

The second experiment was designed to test whether participants evaluated tokens differently based on their phonological conditioning environment. Experiment 2 was a modified magnitude estimation task, in which participants were asked to rate individual words for how "well pronounced" they sounded relative to a reference word. Each participant was played an equal number of lax and tense words from each conditioning factor, for a total of 48 target words per participant and 48 fillers. Responses were z-scored by participant.

Figure 2: 'well–pronounced' rating, by conditioning environment and phonetic production
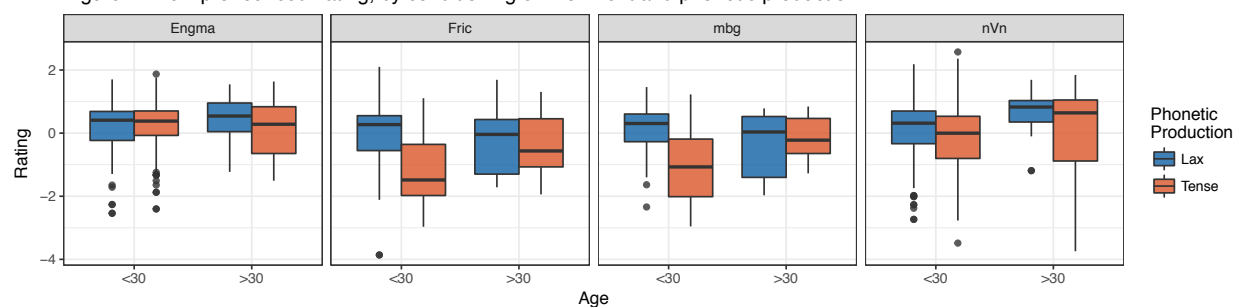

Figure 2 depicts the results of the results of the Magnitude Estimation task, broken down by conditioning factor, tenseness value, and age of responder. A fixed-effects regression model with was run separately over each conditioning factor to determine whether the evaluation of tense and lax production has changed between the older and younger participants. We find interaction effects such that older participants downgrade tense nVn (as in "manage", p=0.3) and tense engma (as in "hang", p=.04), while younger participants downgrade tense mbg (as in "mad", p=.01) and tense fricatives (as in "class", p=.01). In other words, older speakers downgrade NAS tense tokens but not NAS lax tokens. Likewise, younger speakers downgrade PHL tense tokens but not PHL lax tokens. These results suggest that what may look like phonetic evaluation when considering age groups separately is actually more complicated: participants do attend to underlying allophonic system, only downgrading the tense tokens of the allophonic system that is less widespread within their age group.

Taking the results of both experiments into account, we conclude that abstract linguistic structure is subject to speaker evaluation. Furthermore, these results highlight that while phonetic versus phonological effects may be confounded in synchronically stable data, these effects may be distinguished using data from a period of phonological change.

Selected References

Cambell-Kibler, K. 2007. Accent, (ING), and the Social Logic of Listener Perceptions. *American Speech* 82:1.

Dinkin, A. 2015. "The Like Conspiracy: Avoiding Accountability" Paper presented at PLC 39, University of Pennsylvania, March 2015.

Labov, W. 1993. "The Unobservability of Structure and its Linguistic Consequences." Paper presented at NWAV 22, University of Ottawa, October 1993.

Labov, W, A. Henderson, S. Fisher, D. Gylfadottir, and B. Sneller. 2016. Competing Systems in Philadelphia Phonology. *Language Variation and Change* 28:3.