Coding for Gender and Sexuality
Penelope Eckert
Stanford University

Our gender order guarantees that by and large, the binary distinction male-female will be amply indexed by names, clothing, hair style, demeanor, voice pitch, making assignment of research participants into male and female categories unproblematic in most cases. And in the very rare cases of uncertainty, a sociolinguistic interview is likely to disambiguate an interviewee's sex class for us. The possibility that some of these people are transgender will presumably come out in the quantitative wash – unless we're actually interested in the speech of transgender people. But in that case, we aren't likely to use large corpora. On the other hand, sexual diversity isn't likely to come out in the wash, so people seem to recognize that ignoring it in large corpora can cause some disturbances in our findings.

By now most people recognize that indexing sexuality is not the same as indexing gender – that is, that gay men and lesbians do not simply emulate the speech of the other category in the gender binary. However, gender and sexuality are inseparable. The gender order is very much a sexual order, and the sexual order is very much a gender order. And both the gender and sexual orientations of transgender people are anything but simple (e.g. Zimman forthcoming). The fact is that both gender and sexuality are central and intertwined categories in the dynamics of variation, and the dynamics of both are ignored in the construction of large corpora. Gender dynamics are ignored because dominant ideology tells us we can ignore them, but also because we're not sure how to do otherwise. And besides, binary sex-class gets us pretty far. Sexuality (or I should say sexuality-class), like gender simplistically viewed as a binary, is ignored because we're afraid of offending someone. But the larger question is whether it makes sense to do otherwise – whether there's any way we can code large corpora for something subtler about gender than sex-class, and whether we can code for sexuality in any meaningful way. At the moment I believe not. If we're interested in getting under either binary, we aren't likely to do so on the basis of large corpora but through a variety of targeted ethnographic or, in some cases targeted survey, studies. But following what I believe to be the goals of this workshop, I will discuss what it is that we may have to ignore, and raise some issues that might eventually take us somewhere that can illuminate at least some aspects of gender and sexuality.

What would it mean to code for gender beyond the male-female binary? It has become clear over the years that gender does not have a consistent effect on variation across other categories. Gender often interacts with class (Eckert 1990, Labov 1991) with a crossover pattern in which the speech of upper middle class women is more standard than that of their male peers, and working class women's speech is more vernacular than that of their male peers. Assuming that class is a central factor in variation, this could reflect two significant things: that women are indexing certain class-related qualities or stances more than men, or that women overall use variation more to make social distinctions than men (Eckert 2000). Of course, the two amount to essentially the same thing, but they might be approached in different ways. One might look for those stances or qualities, or explore more carefully the nature of the individual's engagement in linguistic markets (e.g. Sankoff and Laberge 1976), but this is not something that can be meaningfully done across corpora. Or one might look for something else related to the greater distinctive use of variation.

I believe that the fact that women's patterns of variation constitute the envelope within which men's variability is contained is a fundamental fact in variation. Like all other relations between categories and practice, the next question is which women and which men are making this difference? Lesley Milroy (1998) once proposed that women are more innovative than men because they are less constrained by peer norms which, in her data, were norms against leveling. Certainly, men are more constrained by the gender order than women to the extent that their gender expressions are more closely monitored – and men rather than women are the primary monitors of gender expression. (This fits in with Tony Kroch's argument that the curvilinear pattern is due to working and lower middle class men avoiding sounding like women. While I don't agree with it across the board, it no doubt is an important factor.) Bill Labov, meanwhile, has speculated that the leaders in sound change are lower middle class women who are networked and gregarious. If we put this all together, we have a picture of gender differences having to do with differences in expressivity (Eckert 1989).

Similar considerations arise in the case of sexuality. The study of language and sexuality has by and large focused on a heterosexual – homosexual binary. (Sexual preference, of course, doesn't cover sexuality. However, I will ignore here the obvious and very interesting relation between variation and the expression of sexual desire and satisfaction, and stick to the kinds of identity categories that most sociolinguists like to consider.) The focus on the sexual binary has been pretty narrowly on men, and attempts to do similar binary studies of women have yielded the kind of unsatisfactory results that one would expect in a social category that has no popular speech stereotype. But although the studies of lesbian speech have begun with studying the speech of lesbians, the study of gay male speech has progressed quite far in the absence of any large-scale comparisons of the speech of gay and straight men. Rather, since gay men are in fact quite heterogeneous in their speech, the focus has been on speech that is *perceived* as gay. Smythe et al (2003) view gay-straight speech as indistinguishable from gender – with gay and straight speech lying along a gender continuum from female to male. But again, what they're actually studying is gay- and straight- sounding speech, correlating phonetic features with experimental judgments of speakers' sexual-class membership. There is no question that flamboyant gay speech uses many features that are associated with female speech, and in this study they've chosen to listen for variables that had more-or-less been found to correlate with gender.

The very existence of an enregistered gay male speech style as a prominent and long-standing stereotype and resource for humor sets it aside from other speech varieties, and sets it aside as well from the everyday speech of gay men. Hence, as Erez Levon has nicely put it (2007 p. 536) "… an empirical methodology should focus more on the semiotic practices people use to enact this social positioning … rather than focus on those positions themselves". The enregistered gay male style should be seen in terms of its historical origins in camp, which is drag in its most ironic sense. And in this context, one might wonder why we should distinguish this from the camp speech of, for example, Bette Midler. Putting gay speech on a male-female continuum makes sense only within this context, recognizing that some form of camp is an integral part of gay culture, and men who prefer same-sex partners may participate in this culture to varying degrees and depending on the context. Rob Podesva's work (e.g. Podesva 2007) on the speech of professional gay men shows that stereotypic gay speech emerges in situations that call for flamboyant displays of gay personae such as "gay diva" and "gay partier", and that this stylistic shift can involve the production of advanced variants of current sound changes (Podesva 2011). Elements of this flamboyant style may well bleed into some gay men's more conservative speech, but once

again the question is whether this is better predicted by something like sexual politics than by sexual preference.

This brings us back to the issue of Labov's leaders in sound change – the gregarious women of the lower middle class. I believe that Labov has his finger on the pulse of variation in this designation, and the question is what exactly is the connection between variation and gregariousness? We could be looking at something more like general social perspicacity. Alan Yu's research exploring the relation between language processing and measures of autism and empathy is highly suggestive in this regard. He posits a connection between cognitive and social dispositions, relating the potential for gregariousness to a potential for processing language "top down" – that is, attending to context at the expense of linguistic detail. This would make top town processors more likely to be sensitive to, and more likely to use, social patterns of variation. The apparent fact that women are more likely than men to score low on the autism scale may fall into a picture much like the likelihood of men being taller than women – that is, it may play a role in gender patterns. However, it will also play a role on both sides of the binary, and we might consider that Yu is looking at a small bottom-line and suggestive factor in social patterns of variation. At any rate, we aren't going to subject our interviewees to a battery of psychological tests. And we might also remember that social practice can have a feedback effect on biological processes (consider, for example, that research with species ranging from rhesus monkeys (Rose *et al*. 1972) to fish (Fox *et al*. 1997) has documented changes in hormone levels as a result of changes in social position.)

This leads me to the conclusion that if we want to get beyond the gender and sexuality binaries, we might begin by thinking about the nature and function of linguistic expressivity and flamboyance, and the relation of these to norms of gender, sexuality, class, ethnicity, and other categorization schemes we may use. What, then, are flamboyance, expressiveness and gregariousness? These terms are, in fact, relatively meaningless in the abstract and while I'm tempted to get excited about thinking about them, now is not the time. The real question is when we say that women are licensed to be expressive, what are they licensed to be expressive about? What other norms are there about female and male expressiveness, and which women and men conform more or less to them and why? If we think about affect, women are licensed to be expressive of positive emotion, men of negative. It may be, for example, in the exaggerated expression of positive affect, that female caregivers pass on certain sound changes to small children. Of course, I believe that affect is fundamental to the social meaning of variation, and that it is certainly in affective displays that small children begin to understand that variation is meaningful. And certainly affect colors all the social categories that are sufficiently important to constrain our language use.

While some people may BE more gregarious or expressive than others, we might consider that gregariousness or expressiveness manifest themselves in speech in more ways than one and that some might be useful as markers of expressiveness. It appears, for example, that within speakers the use of advanced variants correlates with pitch range (Drager, Eckert and Moon 2010, Podesva 2011) – no doubt a result of both pitch range and use of advanced segmental variables constituting greater expressiveness. In this case, one might consider that gender and sexuality – at least in terms of the way we've thought of variables with respect to them – are covering a factor that is actually somewhat orthogonal. In this case, we might find that coding for something like normalized pitch range – or allowing users of the corpus to do this for themselves – might be worth trying.

Beginning with the Martha's Vineyard Study (Labov 1963) and continuing with the more recent studies in the "third wave", it has been clear that on top of some fundamental values acquired in childhood, variation by and large indexes social categories only indirectly (Ochs 1991, Eckert 2008). Correlations with broader social categories can involve some claims about membership in those categories per se, but stylistic studies indicate clearly that they are more likely to result from cumulative expressions of stances and qualities that emerge as particularly salient in the lives of people who inhabit those categories. Given this expressive and social-constructive function of variation, it makes sense to delve beneath the categories that ideologically connect to those stances and qualities. And eventually, if we do our work well enough, we may dispense with some of those categories altogether. In the meantime, I'd say corpora should continue with m/f and pretend that sexuality doesn't exist.

Of course, asking interviewees whether they are gay, straight or bisexual would eventually get us a sufficiently large corpus that we could actually discover whether there are any correlations with willingness to check "gay" on a questionnaire. But my understanding is that people don't want to do this for fear of offending people or that they're afraid their IRBs are afraid. It's my firm belief that the people who would be offended would be the straight ones, but that's neither here nor there. Certainly looking to code multiple corpora in the same way calls for a uniform questionnaire at the end of the interview. While such questionnaires may be pretty blunt instruments, at least they constrain the responses to some extent in the same fashion. And presumably even the most conservative IRB would see the usefulness of asking interviewees to check a bunch of boxes, always with the option to not check any.

BOBERG, CHARLES. 2006. Sex, gender, and the phonetics of Canadian English. NWAV 35, Ohio State University, Columbus, Ohio

DRAGER, KATIE, ECKERT, PENELOPE and MOON, KYUWON. 2008. Style and prosodic variation. Paper presented at NWAV 37. Philadelphia: University of Pennsylvania.

ECKERT, PENELOPE. 1990. The whole woman: Sex and gender differences in variation. Language Variation and Change, 1.245-67.

—. 2000. Linguistic variation as social practice. Oxford: Blackwell.

—. 2008. Variation and the indexical field. Journal of sociolinguistics, 12.453-76.

—. 2011. Where does the social stop? Language Variation - European Perspectives III. Selected papers from the 5th International Conference on Language Variation in Europe (ICLaVE 5), Copenhagen, June 2009, ed. by Frans Gregersen, Jeffrey K. Parrott and Pia Quist, 13-29. Amsterdam: John Benjamins.

FOX, H. E.; WHITE, S. A.; KAO, M. H.; and FENALD, R. D. 1997. Stress and dominance in a social fish. *Journal of Neuroscience*, 17:6463–6469.

LABOV, WILLIAM. 1991. The intersection of sex and social class in the course of linguistic change. Language Variation and Change, 2.205-51.

LEVON, EREZ. 2007. Sexuality in context: Variation and the sociolinguistic perception of identity. Language in society, 36.533-54.

MILROY, LESLEY. 1998. Women as innovators and norm-creators: The sociolinguistics of dialect leveling in a northern English city. Engendering communication" Proceedings of the Fifth Berkeley Women and Language Conference, ed. by Suzanne Wertheim, Ashlee C. Bailey and Monica Corston-Oliver, 361-76. Berkeley: BWLG.

OCHS, E. 1991. "Indexing gender". Rethinking Context, ed. by A. Duranti and C. Goodwin. Cambridge: Cambridge University Press.

PODESVA, ROBERT J. 2007. Phonation type as a stylistic variable: The use of falsetto in constructing a persona. Journal of sociolinguistics, 11.478-504.

—. 2011. The California vowel shift and gay identity. American speech, 86.32-51.

ROSE, ROBERT; GORDON, THOMAS; and BERNSTEIN, IRWIN. 1972. Plasma testosterone levels in the male rhesus monkeys: influences of sexual and social stimuli. *Science*, 178:634–645.SANKOFF, DAVID and LABERGE, SUZANNE. 1978. The linguistic market and the statistical explanation of variability. Linguistic Variation: Models and Methods, ed. by David Sankoff, 239-50. New York: Academic Press.

SMYTH, RON, JACOBS, GREG and ROGERS, HENRY. 2003. Male voices and perceived sexual orientation: An experimental and theoretical approach. Language in society, 32.329-50.

ZIMMAN, LAL. forthcoming. The discursive construction of sex: Remaking and reclaiming the gendered body in talk about genitals among transsexual men. Queer excursions: Retheorizing binaries in language, gender, and sexuality, ed. by Lal Zimman, Joshua Raclaw and Jenny Davis. Oxford: Oxford University Press.