The Pragmatics of Slurs

RENÉE JORGENSEN BOLINGER*
University of Southern California

Abstract

I argue that the offense generation pattern of slurring terms parallels that of im-
politeness behaviors, and is best explained by appeal to similar purely pragmatic
mechanisms. In choosing to use a slurring term *rather* than its neutral counter-
part, the speaker signals that she endorses the term (and its associations). Such an
endorsement warrants offense, and consequently slurs generate offense whenever
a speaker's use demonstrates a contrastive preference for the slurring term. Since
this explanation comes at low theoretical cost and imposes few constraints on an
account of the semantics of slurs, this suggests that we should not require semantic
accounts to provide an independent explanation of the offense profile.

When we use slurs, we communicate information about ourselves and our attitudes
towards the targets. Recognizing this obvious fact requires no great insight, but
taking it seriously yields a simple and remarkably powerful explanation of how
and why slurs generate offense. Recent discussion of slurs has centered on their
offense-generation pattern, characterized by phenomena that cluster into roughly
five properties:

- ◇ *Offensive Autonomy*—slurs are offensive even when the speaker does not intend
  the use to be derogatory.[1]
- ◇ *Embedding Failure*—the offensiveness of slurs projects out of various forms of
  embedding, including indirect reports, negations, and mentions.[2]
- ◇ *Perspective Dependence*—use of a slur is taken to indicate that the speaker holds
  derogatory attitudes.[3]
- ◇ *Offensive Variation*—not all slurs, even if co-referential, appear to be equally
  offensive.[4]
- ◇ *Insulation*—despite all of the above, slurring terms can occasionally occur inof-
  fensively, and this is true even of particularly potent terms. [5]

With some notable exceptions, the dominant strategy thus far has been to try to
construct a semantics thick enough to account for this offense profile, and in turn

to take doing so as an adequacy condition on semantic theories. Croom (2011) appears to take offensiveness to be an indicator of derogation, so while he is primarily interested in explaining the derogatory aspect of slurs, he evaluates the success of his account by its ability to generate the right offense profile. Hedger (2012) frames his project of extending Kaplan's semantics straightforwardly as an attempt to account for the offense profile of slurs. Hom (2008) motivates his externalist semantics primarily by needing to account for offensive autonomy and variation, features which Williamson (2009) and Whiting (2013) aim to capture semantically with conventional implicature, and Jeshion (2013) explains by appeal to a semantic rule of use. Kennedy (2002) offers a radically contextualist picture in order to account for variation, insulation, and the offensiveness of mere mentions. Potts (2007), Saka (2007), Richard (2008), Bosivert (2008), McCready (2010), and Gutzmann (2013) each give (to varying degrees) expressivist semantics to explain embedding failures and perspective dependence.

Though there is plenty of work to be done in the semantics—slurs' conditions of application, connection to stereotypes, and relation to neutral counterparts are all promising topics—I argue that it is a mistake to hold our semantics hostage to offensiveness. The characteristics of the offense profile can be accounted for by a highly general, purely pragmatic process compatible with most theories of the semantics of slurs. This paper has three complimentary but separable aims: first, to establish that slurs and other offensive speech acts pattern together, and so should receive parallel explanation. Second, to develop and offer the *Contrastive Choice Account* as the best explanation, and finally to illustrate the explanatory power and generality of such a pragmatic account. In developing this account, I make two key assumptions:

   (a) slurring terms are marked as the dispreferred option for achieving reference to the target group
   (b) they are associated with the derogation of their targets.

(a) is relatively innocuous, requiring only that competent users of the language be aware that slurring terms are not polite, a typical requirement for competency with respect to slurring terms. (b) intentionally leaves open the question of whether the association is achieved semantically or otherwise. Most robust theories of the semantics of slurs address themselves to this question, and it is a virtue of this account that it is neutral between the various ways of specifying the nature of the association.[6] In the process of developing this account, it will be necessary to mention a number of potent slurs. I have tried to keep this to a minimum, using the terms directly only where required to elicit clear judgments or give illuminating examples. The status of slurs mentioned in academic work is an interesting question, so in §5.4.2 I trace the implications of my account for such occurrences.

§1 sets up some necessary preliminaries by examining the nature of the offense the account aims to explain. §2 reviews the central data to be accommodated concerning the offense generation patterns of slurring terms, and §3 examines some parallel data for impolite and rude speech. In §4, I develop and apply the Contrastive

Choice Account to explain this data. §5 summarizes the account's ability to explain the target pattern of offense generation. Finally, §6 showcases the theory's predictive power and generality by extending it to explain the insulation pattern of slurs which occur within works of fiction and theater.

1. Preliminaries

1.1 Warranted, Rational, and Actual Offense

In saying that an utterance is offensive, there are three distinct claims we might be making: that some hearer *actually* took offense, that the utterance *warranted* offense (whether or not any was taken), or that, regardless of whether offense was in fact warranted, it was *rational* for a hearer to take offense. Each of these bears some unpacking.

Actual offense depends on a hearer's interpretation of the utterance, and is neither necessary nor sufficient for either warranted or rational offense. An utterance may warrant, but fail to actually generate offense merely because either there is no hearer, or the hearer fails to find the utterance offensive (perhaps because she shares the offensive attitude, fails to take it seriously, or misinterprets the utterance). Similarly, mistakes concerning the semantic content, speaker's illocutionary intent, or pragmatic mechanisms activated by the utterance may result in a hearer taking unwarranted offense.

Not every such mistake is unreasonable, though all result in a mismatch between the offense taken and what was warranted. For instance, hearers may naturally take offense when a speaker uses 'nigger' to refer to them. Discovering that the speaker was ignorant of the word's derogatory nature should make us think that their offense was in some sense *in fact* unwarranted, but *nevertheless* the hearers were not irrational, or responding inappropriately to what they heard, in taking offense. Since hearers may be subject to a variety of sources of uncertainty concerning the level and degree to which offense is warranted (e.g., they may be unable to discern a speaker's intentions, or might be unfamiliar with the term used) let us use 'expected value of w' for the offense the hearer has reason to believe is warranted, given her evidence. In what follows, I will use 'rational' or 'licensed' to indicate when hearers are epistemically justified in taking offense, and reserve 'warranted' for morally justified offense.

Offense may be disproportionate either by taking more offense than is warranted, or by failing to take offense when warranted. On the assumption that rational hearers ought to proportion their offense to the severity of the offensive action (adjusted for their confidence that the action occurred), we may say that a level of offense is rational when it is appropriate given the hearer's evidence. Though warranted offense is the primary focus of this paper, the notion of rational offense helps clarify some of the murkier intuitive data, and will equip us to evaluate the Contrastive Choice Account's predictions of actual offense patterns insofar as they are rational.

1.2 Grounds for Offense

One more set of distinctions will prove useful in the discussion to follow. There are three identifiable grounds on which an utterance may warrant offense: intention, inappropriateness, and its associations. Utterances may be (and often are) offensive on multiple grounds simultaneously—when intending to offend, speakers commonly select tabooed insults ('shithead') or slurs ('dyke') to communicate their ire—but the severity of the offense warranted varies with the grounds for offense.

Utterances that violate a taboo warrant offense in virtue of their inappropriateness. Most common expletives, vulgarities, and general pejoratives warrant offense of this type.[7] Inappropriateness on its own grounds only a relatively shallow level of offense (such as that caused by a child trying out the word 'fuck' for the first time at the dinner table),[8] and is nearly always compounded with *intent* or *associations*.[9] *Speaker intent* is the most general warrant for offense: any phrase may be used to offend if the intention to do so is clear. However, it is inadequate as a full explanation of the offensiveness of slurs, since it cannot explain why a slur offends *despite* a speaker's good intentions. For that, we must appeal to the terms' offensive associations.

Associational offense exhibits huge variation, dependent on the severity of the association. Terms, phrases, or symbols that are closely associated with abhorrent attitudes or practices warrant the sort of deep offense commonly exhibited by slurs. The symbols of a swastika or burning cross belong to this category, as do terms that have come to be associated with various forms of racism, sexism, or more generally with a threatening program of discrimination against members of some group targeted by the term.[10] Elements of this category are often the subject of hate crime legislation, and singled out for their power to provoke their targets. It is not necessary that such terms be associated with or backed by formal social institutions (though they often are), just so long as there is an adequately visible practice associated with the terms.

2. The Offense-Generation Profile of Slurs

Before attempting to explain it, we should briefly recount the data constitutive of slurs' odd embedding pattern—specifically *insulation* and the various cases of *embedding failure*.

2.1 Insulation

The least controversial inoffensive occurrences of slurring terms are cases where the terms are only mentioned. In direct quotation, or when some contextual constraint (e.g. a hearer's insistence to 'tell me exactly what he said') leaves the speaker with no alternative to mentioning an offensive term, the offensiveness of the term remains embedded in its original context. Our indignation, if it is aroused, is directed at the individual whose utterance is being quoted, rather than the current speaker.[11] Similarly, when a slurring term T is mentioned in a dictionary entry, perhaps in the form

> T is a derogatory term for R..., often associated with/implying F...

we do not feel tempted to censure the dictionary compiler for tokening the term.[12]

Even mentioning several slurring terms in rapid succession does not necessarily warrant deep offense.[13] Consider a hypothetical corporate memo, advising employees that they must abide by a strict anti-slurring policy:

> MEMO: The following terms are not to be used by any Corp. employee, nor is their use to be tolerated in any Corp. classroom or workspace: 'chink', 'dyke', 'honky', 'nigger', 'spic'...[etc.]

It is doubtful that anyone would protest that the slurs as they occur in the memo are as offensive as they would be if they were simply used. So to at least some extent, mentioning slurring terms successfully insulates their offense potential. Perhaps you think, as I do, that there is still something strange (or offensive) about listing each of the slurs explicitly rather than giving a blanket admonition to avoid slurring terms. If so, that suggests that something other than a simple use/mention distinction is at work in mitigating, though not entirely neutralizing, the offensive potential of these terms. The positive account offered later in the paper may be able to explain this residual discomfort.

2.2 Embedding Failure, Perspective Dependence, & Offensive Autonomy

Insulation is neither unique to slurs nor especially puzzling, on its own. Embedding failure is a more interesting property, showcasing instances where slurs license offense despite forms of logical embedding. These contexts often simultaneously exhibit offensive autonomy and perspective dependence.

2.2.1 Negations and Denials

Aside from cases of metalinguistic denial,[14] speakers cannot distance themselves from the offensiveness of a slur simply by denying a slurring predication (as speaker B does in 1) or embedding it under negation (as 2):

(1) *Speaker A*: Henrik is a kraut.
    Speaker B: No he isn't.
(2) Henrik is not a kraut.

Speaker B's unspecified denial in (1) is by default a denial of speaker A's implication that Henrik is German; the offensiveness of the slur 'kraut' remains, and this is so even if the speaker bears no ill-will toward Germans.[15] Far from being rendered docile, the occurrence of the slurring term in (2) suggests that the speaker embraces the practice of calling, classifying, and referring to Germans by using the slur 'kraut', and is only disputing Henrik's nationality.

2.2.2 Indirect Reports, Conditionals, & Modals

Ordinarily, when content is embedded under a conditional or indirect speech report, the speaker is not held directly responsible for the embedded attitudes. Slurs are

an exception: the offensiveness of slurs is not insulated by embedding in indirect reports, and in nearly all cases also fails to embed under conditional constructions:

(3) Bob said he'll fire all the cunts.
(4) If I were racist, I probably wouldn't like niggers.[16]

(5) A less enlightened man than myself, a cruder man than myself, a man less sensitized to the qualities and charms of women—not me, but a man like that—just might call her a cunt.

The speaker of (5) is rightly censured, despite the counterfactual construction.[17] The occurrence of the slur in (3) and (4) suggests that the speaker is racist or sexist, despite the conditional construction of (4), and even though (3) is an indirect report.

3. A Parallel Phenomenon: Rudeness

The motivation for giving a thick semantics for slurring terms is a desire to accommodate this curious embedding pattern, together with the other features comprising slurs' offense profile. It stands to reason then that we have at least *prima facie* justification to expect the right explanation to generalize to any other offense-generating terms with the same offense profile. While slurs are a particularly clear case, they are not the only offensive expressions to exhibit the five features of interest. Rude and impolite expressions generate a remarkably similar (though typically less severe) pattern of offense. Absent a good argument in favor of treating slurs as *sui generis* among offensive expressions, the account we use to explain the offensiveness of slurs should generalize to rude expressions (and viceversa).

3.1 The Five Features, Redux
Christopher Potts (2007) gives a careful discussion of rude terms (vulgarities, expletives, or general pejoratives), noting that the expletives and general pejoratives in constructions like

(6) Herman believes that Hella's damn dog is dead.
(7) Sue believes that that bastard Kresge should be fired. (#I think he's a good guy.)[18]

are *perspective dependent*: the attitudes conveyed by the expressions are assumed to be the speaker's. (6) "conveys that the speaker of the sentence holds Hella's dog in low regard,"[19] and the continuation in (7) is infelicitous precisely because the earlier perspective—according to which Kresge is a bastard—is assumed to be the speaker's.[20] Rude speech is perspective dependent even when the expression used is unmarked, and impolite/rude only due to contextual features. For example, though addressing someone by a formal name or title is appropriate when the addressee and speaker are not on familiar terms, when directed at a lover this form of address insults the hearer. The insult is immediate and is assumed to reflect the perspective of the speaker, since by using a form of address appropriate to distant and formal relationships rather than one appropriate to close, personal relationships,

the speaker gives the hearer warrant to believe that the speaker considers their relationship to more closely resemble the former than the latter.[21]

Rude expressions also exhibit *embedding failure* in conditionals and indirect speech reports. In (8), the speaker appears to endorse the appropriateness of the description of 'all this' as 'shit', even though he ascribes it to Bob, and if we're offended by that, the offense is not lessened by the embedding:

(8) Bob said he was done with all this shit.
(9) If he doesn't shut up, that motherfucker is going to get a ticket.

Similar comments apply to the general pejorative in (9), despite its conditional construction.

Unsurprisingly, rude expressions also possess *offensive autonomy*: they cannot be rendered inert just by a speaker's ignorance, assurances that she means no offense, or even actual absence of offensive intent. In ordinary contexts, an utterance of "you motherfucker" signals aggressive rudeness, a signal that cannot be canceled by simply appending "no offense" or "I don't mean to be rude." Even if the speaker is unaware that the expression is marked as rude, and employed it only to express generalized frustration, that fact does little or nothing to reduce the rational offense licensed by the utterance. Similarly, speaker intent exacerbates, but is not necessary for offensiveness; it is possible to be offensively rude/impolite without even being aware that you have done so.[22] Nevertheless, in specialized contexts the expressions can be successfully *insulated*. In familiar, informal interactions where all parties are aware that normal signaling relations are suspended and the speaker intends no rudeness, the very same expression ('you motherfucker') can be used to signal intimacy rather than rudeness.[23]

Finally, and perhaps most obviously, there is considerable *variation* in the level of offense licensed by different rude expressions. Formulas associated merely with disrespect (e.g. 'that [moron/nutcase/idiot]') are less offensive and tend to be less inflammatory than those associated with aggression ('you [mother-fucker/cocksucker/goddamn idiot]').[24] Given these parallels, we have strong reason to want our account of the offensiveness of slurs to parallel our story for rude speech. This provides not just a constraint on theory selection; it also suggests where to go looking for a promising account.

3.2 Accounting for Rudeness: Co-Occurrence Expectations

Though they use different terminology, theorists explaining impoliteness have primarily focused on accounting for *offensive autonomy, variation,* and *insulation*. Recent work on these three aspects of the offense profile of impolite behavior has converged on pragmatic explanations invoking the contrast between a speaker's chosen performance and her relevant alternatives. I'll briefly gloss the work of two such theorists to draw out the structural similarities of their accounts, then argue that they illuminate a deeper connective thread which explains not only the features of interest to impoliteness research, but can be generalized to explain all five aspects

of the offense profile of both impolite behavior and slurs, and furthermore allows us to predict when offense will occur, if rational.

Marina Terkourafi (2005a, 2005b, 2005c) offers a frame-based account according to which expressions are associated with (im)politeness in virtue of "the regular co-ocurrence of particular types of contexts and [these] particular linguistic expressions."[25] Co-occurrence expectations arise from regularities between non-linguistic features of a context type and displayed preference for a particular expression over a semantically equivalent alternative.[26] Statistical regularity of co-occurrence with polite contexts associates an expression with politeness, with the result that displaying preference for the term becomes a conventional*ized* signal of politeness. The same analysis holds, *mutis mutandis*, for rudeness. Default frames for a context are generally minimally polite, so any expression that deviates from that—either impolite or overly polite—is a marked alternative. Use of a marked expression signals that the default frame does not fit the actual context, and use of a conventionalized signal activates the frame for the expression's associated context-type.

Terkourafi glosses 'conventionalized' signals as sitting halfway between pragmatic, purely inferential conversational implicatures and semanticised conventional implicatures. Conventionalized signals work to compress inferences in a way structurally similar to Bach and Harnish's notion of *standardization*,[27] and unlike conventional implicatures, these signals can be canceled or blocked in special contexts (where one or more of the steps in the compressed inference are blocked). Unlike conversational implicatures, conventionalized signal content cannot be unilaterally canceled by the speaker in ordinary contexts. Rude behavior grounds offense when the behavior/expression is associated with rudeness in the context type, and this is generally well known in the relevant linguistic community.[28]

Jonathan Culpeper (2011) notes that while politeness is conventionalized by statistical frequency of co-occurrence, impoliteness is less direct. What counts as impolite in a context is less a matter of the conventionalized meaning of the expression used, and more about the speaker's choice to flout contextual politeness expectations. Two conditions must be satisfied for impolite behavior to warrant offense in a context: the selected expression must contrast negatively with the expected or default (polite) behavior for the context, and this fact must be generally well known among members of the linguistic community.[29] When these two conditions are met, the speaker's choice to flout expectations signals a lack of concern for the "face", or social standing, of the recipient, who is licensed to take offense to the signaled attitude.[30]

4. The Contrastive Choice Account

Terkourafi and Culpeper apply their accounts to explain why rude/impolite behavior is both somewhat stable (exhibits offensive autonomy) and interestingly context-dependent (exhibits offensive variation and insulation). Though they invoke different specific mechanisms, the structure is the same: the offensiveness of impolite/rude behavior results from content signaled by the speaker's decision to

perform *that* particular behavior, rather than a comparatively polite alternative. Viewed at this level of abstraction, it is easy to see how to generalize this structure to give a unified account of all five features of the offense profile of both impolite behaviors and slurs. The type of signaling invoked in the impoliteness account is one instance of a broader signaling phenomenon, present whenever parties to an interaction face a free choice between referentially equivalent expressions. Signaling on this framework is factive: a speaker signals some content ϕ when her use of an expression satisfies the conditions, regardless of whether she intended to communicate ϕ, and independent of whether hearer uptake occurs.

4.1 Signaling and Contrastive Choice

Speakers competent with a language have knowledge not only of lexical items and grammar, but also a set of co-occurrence expectations that encode the social norms and conventions concerning the use of various terms and ways of speaking.[31] At their most general, such expectations are of the form 'the behavior α characteristically signals ϕ,' where ϕ ranges over some associated information. The possible values for ϕ exhibit substantial variety, ranging from endorsement of certain attitudes to membership in a (more or less loosely-defined) group. The signal content of a choice is not determined directly by patterns of use; whether (and how strongly) α signals ϕ depends on the strength of the association between choosing α and endorsing ϕ. Strong correlations in patterns of use is a natural way for such an association to come about, but is not the only one.[32]

For signals based in *contrastive choice*, the relevant behavior is the free selection of a marked expression, and performance signals that the speaker endorses a cluster of attitudes associated with the term (or, more precisely, a high probability that the speaker shares some or all of the attitudes in this cluster).[33] More formally,

For some content ϕ, when it is common knowledge in the linguistic community that

(i) α is an expression for ψ associated with ϕ, and
(ii) β is an expression for ψ not associated with ϕ

then in situations where the choice of expression is not forced, and the speaker is aware of (i) and (ii), selecting α in contrast to β signals that the speaker endorses or shares ϕ.

The information content of signals based in contrastive choice is linked to how marked the term is: if α is a term that is used almost exclusively by speakers who embrace ϕ, and this fact is well-known, then a contrastive preference for α is a high-information signal, raising the probability of the speaker's endorsing ϕ nearly to 1.[34] The more well-known the association between α and ϕ is, the higher the information content of the signal, and thus the more strongly the contrastive choice signals the speaker's endorsement of ϕ. For example, in polite contexts, the expression 'old lady' is associated with rudeness toward and disrespect for the social standing of the referent. Consequently when in such a context a speaker chooses to refer to his mother using the expression 'my old lady' rather than 'my mom/mother', he signals that he endorses or holds such an attitude of disrespect

toward her. Since this is an offensive attitude, his contrastive preference for the rude expression warrants offense in the context.

This pattern extends beyond expressive or evaluative terms: for instance, since use of the term 'cisgendered' is limited almost exclusively to members or allies of the trans∗ community, a speaker's selection of the term strongly signals that the speaker is sympathetic to the community's project of undermining the assumed synonymy between 'normal' and 'non-trans∗' experience. This information is *not* part of the semantic meaning of 'cisgendered', nor can it be convincingly glossed as a conventional implicature of the term. Rather, the signaling relationship results from the particularly low probability of freely selecting the term while failing to hold such an attitude.

4.2 Contrastive Choice and Warranted Offense

So, how does this mechanism account for the behavior of slurs? We have assumed that slurs are known to be dysphemistic (negatively marked) by any speakers competent with the terms, and are associated with the derogation of their targets. It is also a characteristic that for any slurring term there is some available alternative expression for the same target class. Where ϕ is the derogation or oppression of a class of people, offense is warranted when a speaker endorses ϕ.

Use of a slur α rather than a neutral alternative β is a defeasible indicator that the speaker endorses ϕ, and can be undercut by speaker ignorance or a forced choice. A choice is forced if something about the situation of utterance makes alternative expressions unavailable or inappropriate. In such cases, utterances of α do not signal that the speaker has a preference for α over β (and so do not signal that the speaker endorses ϕ).[35] Ignorance is similarly undercutting: when a speaker is ignorant of the association between α and ϕ, her use of α does not signal that she endorses ϕ. If she is unaware of any alternatives to α, then whether her use warrants offense depends on whether, if she had known of an alternative β, she would have refrained from using α.

This model predicts that *warranted offense* is subject to one main source of variation: the nature of the attitudes in the cluster associated with the term. If a use of α is an endorsement of ϕ, then it warrants offense proportional to the severity of the attitudes in ϕ. Different slurring terms may be associated with clusters of negative attitudes of varying degrees (ranging from bare contempt to a willingness or desire to kill or inflict great suffering on the target), and may have mixed associations.[36] The offense warranted by a slur is therefore sensitive to the tenor of the attitudes thereby endorsed: the more ambivalent or tempered, the lower the degree of warranted offense.

My account has significant parallels to suggestions made by a number of others that the offensiveness of slurs depends in an important way on the speaker's word choice.[37] Where it goes beyond them is in (i) specifying how and why the choice warrants offense, and (ii) accommodating the sort of fine-grainedness exhibited by the offensiveness of actual slurs. Because it allows contextual co-occurrence to define signal content, the contrastive choice account is able to explain how the offensiveness of some slurs varies depending on the speaker's apparent group

membership. For example, uses of 'nigger' by white middle-class males are associated with stronger and more aggressively negative attitudes than uses of the same term by upper-class black males, particularly if they happen to be hip-hop artists. This fact neatly explains the asymmetry in offensiveness of the term as used by members of these groups.

4.3 Introducing Uncertainty: Rational Offense

Our intuitions about offensiveness are guided not only by what seems to warrant offense, but also by our judgments concerning when hearers would be licensed or reasonable in taking offense. Ignorant slurring is a perfect example: on the one hand, the use of a slur by someone who has no idea that the term is a slur doesn't seem to warrant offense. On the other hand, we are unwilling to declare that a hearer who takes offense to it is unreasonable or oversensitive. Distinguishing carefully between the conditions for warranted offense and those for rational offense helps make sense of these (apparently conflicting) intuitions.

We earlier said that the rationality of a hearer's offense depends on how proportionate it is, given her evidence. In some cases, the offense licensed will be greater than the offense that is in fact warranted. Very informally, a hearer's expected value for w should be set by the severity of the associated content ϕ, and her confidence that the use of α constituted an endorsement of ϕ. This latter element will be a degreed matter, subject to variation based on her relative uncertainty concerning (a) whether the choice was forced or free, (b) whether the speaker is aware of the association between α and ϕ, and (c) whether α is in fact associated with ϕ. Greater signal strength will tend to increase hearer confidence in (c), and in (b) unless she has independent reason to think that the speaker is ignorant.

This picture receives support from our actual practices in cases of speaker ignorance. The default assumption is that the speaker is aware of the signaling relations, and consequently use of a slur is taken as grounds for the inference that the speaker endorses the derogation, which licenses offense in the context of utterance. Once we become aware of the speaker's ignorance, we no longer consider the speaker an appropriate target of censure, but neither do we consider ourselves wrong to have been offended. This behavior is easily explained if the speaker's utterance did not in fact *warrant* offense (due to his ignorance), but our reaction was *rational*, given the default presupposition that the speaker *is* aware of the signal. There is one notable exception to the pattern: when we judge that the speaker's ignorance is inexcusable—when she ought to have known—the speaker will be held liable for the offense regardless of her (professed) ignorance. This is most commonly the case when the speaker involved is a public figure, and is probably motivated by our suspicion that the speaker is not in fact ignorant, but is merely pleading ignorance to escape censure.

4.4 Limited Languages

One might object that contrastive choice cannot explain the offensiveness of all slurs, since slurs may license offense even when there is no alternative expression. To motivate the worry, suppose there were a language with only one expression,

α, for a set of people ψ, and either everyone in the linguistic community holds ψs in contempt, or the expression α is compositionally derogatory (we may imagine it translates roughly to 'swine-that-should-be-slaughtered' or some other suitably threatening and contemptuous phrase). In such a scenario speakers have no alternative expression for ψs, so if α is an offensive slur its offensiveness cannot be explained by appeal to a contrastive choice.

In these cases, I am inclined to say that it is not the use of the slur *per se* that grounds offense, but rather one of the other contextual factors. The setup provides two excellent candidates: first, one may infer from the fact that the speaker is a member of the community in question that he holds ψs in contempt; this is adequate grounds for offense.[38] Second, it may be that the compositional content of the expression α amounts to an explicit endorsement of offensive attitudes—which warrant as much (or more) offense as using a slur, but without in fact being slurs.[39] To eliminate this second explanation, we would need to stipulate that α has become idiomatic, but doing so makes it much more difficult to hear the term as warranting deep offense.

5. Explanatory Adequacy

The Contrastive Choice picture has a great deal of explanatory power, providing a coherent account of each of the major aspects of the offensiveness of slurs.

5.1 Offensive Autonomy
The degree to which speaker intent is necessary for a use of α to constitute an endorsement of ϕ varies inversely with the strength the signaling relation between the two. When the signal is strong—like the relationship between an out-group use of 'cunt' and derogatory attitudes towards women—then barring ignorance and forced choice contexts, use of the term constitutes an endorsement of ϕ, even absent direct speaker intent. However, when the signal is weak—as is the case for 'chick' and disrespect for women—then whether a given use is an endorsement does depend on the speaker's intentions. Offense will be warranted whenever a speaker's use of α constitutes an endorsement of the derogation of the target group.

For the vast majority of slurs, the signaling relation between choice of the term and endorsing derogation is both strong and well known. Since these signaling relations emerge from population-wide correlations, whether the use of a slur on a particular occasion signals offensive attitudes or not is largely independent of the speaker's actual intent. Consequently, we should expect that the use of a slur will license offense even when the speaker does not intend derogation.

5.2 Perspective Dependence
When a speaker chooses to use a slur rather than an available alternative, she signals that *she herself* endorses the associated derogatory attitudes. This makes sense of our practice of holding speakers themselves responsible for the derogatory attitudes associated with the terms, taking the use to reflect the speaker's own attitudes.

5.3 Offensive Variation

The variation in offense potential between co-referential slurs (e.g. 'nigger'/'colored' or 'faggot'/'fairy') can be understood as a variation in associated attitudes and signal strength, and this can be generalized to account for variation between slurs with distinct referents. Each slur is associated with a cluster of attitudes varying in degree and severity. Slurs with more extreme or more overwhelmingly negative clusters should be expected to be more offensive than those with mixed or more moderate clusters. Additionally, slurs vary in signal strength: expressions that are taboo (e.g. 'nigger', 'faggot') are far less likely to be used innocently than those that are not ('colored', 'fairy'). As a result, a use of one of the tabooed terms signals an offensive attitude much more strongly than use of the latter. These two factors—signal strength and severity—combine to create offensive variation.

5.4 Embedding Failure

5.4.1 Indirect Reports

The explanation this account gives for the offensiveness of indirect reports and mere mentions is slightly more involved than the story for direct uses. Indirect reports leave the hearer unable to distinguish whether the current speaker is merely reporting another speaker's use, or chose the term himself. As it occurs in 3a, 'cunt' is neither clearly mentioned, nor clearly used. To convince yourself of this, note the difference in offensiveness between (3a) and (3b), in which the slur *is* clearly mentioned:

> (3a) Bob said he'll fire all the cunts.
> (3b) Bob said he'll fire all the women, calling them 'cunts'.

Confronted with the ambiguous (3a), hearers attempt to discern from context whether the slur occurs as a use or a mention. Aware that (3b) is available to speakers who wish to make it clear that the term occurs as a mention, hearers are justified (absent stronger contextual cues) in concluding that this speaker is using the term, and therefore signaling an offensive preference. Consequently, a hearer's expected value for w is high, so offense is rational.

5.4.2 Offensive Mentions

In some cases, a speaker may rightly be censured for directly mentioning the slur by uttering (3b) instead of employing an easily available circumlocution like (3c):

> 3c. Bob said he'll fire all the women, calling them the C-word.

There are a few features of interest in these sorts of cases: first, the offense generated by a mention is invariably less severe than that generated by a use of the same slurring term. Second, offensive mentions typically occur only in sensitive contexts —public venues, in the presence of children, or contexts where it is not clear whether the speaker intends to derogate. The range of sensitive contexts varies between slurs, and is largest for those with dedicated circumlocutions (such as 'the N-word').

Explaining this on the Contrastive Choice account is fairly straightforward. In mentioning a slurring term, a speaker chooses between two (or more) expressions to refer to the slur α: the quotation name ('α'), a description ('a slur/epithet') or, in some cases, a dedicated circumlocution ('the α-word'). The attitudes associated with *using* α are invariably worse than those associated with mentioning it. While uses are associated with hostile, aggressive and threatening behavior, contextually inappropriate mentions appear to be associated with tamer (though not benign) attitudes, ranging from simple insensitivity to perverse pleasure at saying discomfiting words, and disregard for the risk of encouraging derogating uses of the slur. The speaker's contrastive choice to use the quotation name rather than an available alternative in a sensitive context signals that the speaker holds some or all of these less severe but still offensive set of attitudes, and so warrants offense.

This mechanism is what accounts for the fact that offensiveness is far less severe, but not wholly absent, in memo-type cases (*contra* the prediction of a mere mention/use distinction), where we feel that the author should have issued a blanket prohibition on slurring, rather than explicitly mentioning each banned term. Cases like this have instructive implications for the use of slur-mentions in academic work. The audience of a paper or presentation is typically controlled, and background conditions make it unlikely that the author intends to derogate, so we can expect slur-mentions in such contexts to be generally insulated. However, as memo-type cases indicate, the mere fact that the slurs are *mentioned* rather than used does not fully neutralize the offense potential. Scholars are not free to engage in limitless slur-mentioning without warranting offense: insulation extends only so far as plausibly necessary for the legitimate aims of the presentation (or paper).

Of course, this explanation of offensive mentions only works once a circumlocution is available; on its own, it does nothing to explain why circumlocutions are introduced to the language in the first place. For that, we may appeal to fact that rational offense can at times exceed warranted offense. Conscientious speakers may seek to insulate themselves from rational offense by removing any ambiguity concerning whether their utterances constitute endorsements, distancing themselves by using various tricks: "a slur", "the really bad slur", "the N-word", "the one slur I can't bring myself to say", etc. Most of these are not stable, having no clear referent, but those that are ('the C-word' and 'the N-word' are paradigm cases) become terms in their own right, easily available and non-ambiguous ways of referring to their respective slurs. Once these terms have been successfully introduced, the standard story above can explain why using the quotation name *in contrast to* such an available alternative will license offense.

5.5 *Insulation*

Finally, the account gives clear conditions for successful insulation, accurately predicting when mentions and uses of slurs will not license rational offense. When a speaker's selection of a slur is forced, as it is in direct quotation or dictionary-type mentions, the occurrence is insulated. This much is easily explained by nearly every theory, but where the pragmatic story earns its keep is in permitting the

possibility of insulated *uses*, when they occur in forced-choice contexts. The interesting question then is whether there are any such uses to be explained.

6. Extensions to Actual Offense Patterns

Up to this point, the project has been mostly a normative one, developing an account of the conditions under which it is rational or warranted to take offense at the occurrence of a slurring term. But the account can do more tricks than that; it allows us to predict patterns of actual offense-generation, insofar as the offense is rational. The accuracy of these predictions is striking confirmation of the view.

6.1 Art Uses: An Exceptional Case
Asked about the acceptability of uses of 'nigger', Dr. David Covin (Chair of the Sacramento Area Black Caucus) answered that

> "whether the use of the word is acceptable or not depends on the context. If it were from a quote of historical records, or litera[r]y work, or a theatrical context, that's different. If it were simply plain speech, it wouldn't be acceptable."[40]

The successful insulation of offense in historical records can be explained as direct quotation, but literary and theatrical works present more of a puzzle. Touré has an excellent discussion of the permissible uses of slurs in performance, which gestures at both why an artist may genuinely *need* to use the slur, and how a performance can justify such a use. He writes:

> "those deep emotions that 'nigger' taps into are precisely why I defend the right for artists onstage to use 'nigger'. (By stage I mean movies, TV, theater, stand-up comedy, visual art, and music.) *The stage is a special space where normal human laws and customs apply differently.* [. . .] Many whites have used nigger onstage to this end: to put nigger in the mouth of racists and losers and thus remind the audience that racism is dumb and deplorable." [41] (Emphasis mine)

Slurring terms that occur onstage in Touré's sense—call these 'art uses'—have a particularly interesting pattern of actual offense-generation. If the artist or performer is successful, and the use is well-taken, the artist is not censured for using the slur. However, the term is not inert: the use is accepted precisely because it evokes strong emotional response. Since art uses are still emotionally active as slurring terms, we cannot explain their acceptability by supposing that in art uses, slurs are really only *mentioned* by the artist. The slurs have their normal offense-generation patterns *within* the fiction, but do not generate offense or censure beyond it: the actor, writer, producer etc. are held innocent when art uses are deemed acceptable.

6.1.1 Locating the Insulation
One natural thought is that in such cases, we hold responsible and censure the character, while ignoring the artist on the grounds that she is merely the mouthpiece. This too simplistic: while we certainly do (typically) react negatively to the character using the slur,[42] we do not ignore the artist. In cases where the use is not well-taken,

it is the artist whom we censure, not merely the character or artwork. Those who found Quentin Tarantino's use of 'nigger' excessive in *Pulp Fiction* or *Jackie Brown* directed their censure at Tarantino himself, rather than at the actors who spoke the slurs or at the characters portrayed.[43] Similarly when a comedian's use of a slur—even while portraying a racist character—is not well-received, audience anger is directed at the comedian.

It appears then that when an art use causes actual offense outside the fiction, censure is directed at the person who decided that the slur should be used, and only when he happens to be the same person does it attach to the speaker. In no case is umbrage for an unacceptable use limited to only the *character* who used the slur within the fiction. So it will not do to explain the insulation of these terms by appeal to displacement alone.

If instead we gloss art uses as a special case of forced-choice context, the contrastive choice account predicts a pattern for rational offense that is a remarkably close match for the observed pattern of actual offense. It predicts that art uses will be insulated when two conditions are met (and license offense otherwise): (i) the use of the slur is required for the purposes of the artwork, and (ii) these purposes are good enough to justify the slur. Histories, social critiques, and works that function to improve the social position of the group targeted by the slur(s) are the most likely to satisfy both conditions, and so should be expected to be successfully insulated more reliably than art uses which occur in comedies or as the punchline in a stand-up routine. Importantly, the forced choice justification only protects the speaker from offense if we judge that the slur *really is* required and justified by the purpose of the performance. If the use is excessive, or not adequately motivated, then the performer's decision is less properly characterized as a forced choice, at which point the contrastive preference to use a slurring term warrants offense once more.

6.1.2 Conditions of Acceptability

A cursory glance at the reception of recent movies featuring slurring terms confirms these predictions: controversy over the acceptability of an art use tends to center on whether the artistic purpose was adequately justifying, rather than on whether it can *ever* be acceptable to token the offensive terms.[44] Two films are of particular interest in this respect: *Django Unchained* and *42*. Both follow the struggles of an ultimately victorious African American protagonist, supported and guided by an older white male. Quentin Tarantino's *Django* adopts the comic tone of a spaghetti western, while Brian Hegeland's *42* is a tribute to Jackie Robinson and verges on hagiography.

Tarantino's 110 uses of the slur 'nigger' are not driven by an attempt to accurately depict the speech patterns of the era; they function solely to shape the emotional impact of the script.[45] Many felt that this, coupled with the sheer quantity of slurs and the jocular tone of the movie, signaled a flippancy concerning the slur that is disrespectful and not outweighed by the movie's violent depiction and narrative condemnation of slavery.[46]

42 is remarkable for how little discussion of slurs it generated—particularly since it was released at the height of the controversy over the slur-ridden *Django*.

Hegeland's uses can plausibly be said to result from a historical necessity: set in 1947, the film includes 46 occurrences of 'nigger', 30 of them in a scene which re-enacts Ben Chapman's heckling of Robinson. Though Alan Tudyk (playing Chapman) reports that the scene was difficult to film, it was widely agreed that softening the slurs would sugarcoat history and belittle Robinson's struggle. Apparently the MPAA and general public agreed: the movie received only a PG-13 rating, and had the most successful opening weekend of any baseball movie to date.[47]

The difference in acceptability of the uses between the two films tracks the two conditions predicted by the Contrastive Choice account. The use of the slurs is crucial to *42*'s celebration of Jackie Robinson's historic triumph over racism, whereas it is unclear both whether the slurs were necessary to achieve the goals of *Django Unchained*, and whether those goals were sufficient to justify the slurs. Of course, analysis of the reception of these films is complicated by the many respects in which the movies differ—tone, historical approach, narrative style—and Tarantino's personal history with using 'nigger'.[48]

6.2 Accounting for Variation in Offendedness
By tying the rationality of offense to a hearer's evidence, the Contrastive Choice account yields an elegant explanation of faultless interpersonal threshold variance for offense at the use of slurs. It predicts that, if rational, a hearer's level of offense should vary with her confidence that the use of a term signals the speaker's endorsement of offensive atttiudes. If a given hearer has comparatively less confidence than her peers in the signaling relation, her tolerance for slurs should be higher than is typical among her peers, since she does not perceive them as warranting offense as reliably as do her peers (her expected value for w is low, as compared to her peer group).

6.3 Reclamation, Amelioration, & Transference
This model also provides a nice explanation of cases when the level of offense that is rational to take at an utterance of a given expression changes over time. There are three ways that this occurs: by reclamation, by disuse, and by transference of offensiveness from one term to another.

6.3.1 Reclamation
Successful reclamation works by undermining the signal strength of the slurring term. Initially, a group of speakers who reject the derogation of the target use the slur defiantly. As this group grows, the likelihood that a user of the term holds derogatory attitudes falls, causing the information in the signal to degrade. Reclamation is successful when the signal has been so diluted as to carry no information: the term is at least as likely to be used positively as it is to be used by those who hold objectionable attitudes. At that point, it is possible to exhibit a contrastive preference for the term without by default licensing rational offense.

That's not to say that a reclaimed slurring-term can be used by just anyone, or in just any context, without warranting offense. Recall that earlier we said that speaker intent is relevant when signal strength is weak: in the mouths of homophobes, use

of the reclaimed term 'queer' can still warrant deep offense. This is because the intent to use the word to endorse homophobic attitudes is contextually clear, and a use of a slur warrants offense precisely when it constitutes an endorsement of such derogatory attitudes.

6.3.2 Amelioration & Transference

The contrastive choice mechanism also yields a natural account of 'amelioration', or the diminished offensiveness of archaic slurs. Even after we are informed that 'Boche' is a derogatory term, expressing contempt for Germans and implying that they are all cruel, it is hard to feel the offensiveness of term.[49] This is difficult to explain if the offense is semantically encoded in the term, since it is not clear that the semantic meaning of the term has changed at all. What has changed, however, is the frequency of use: 'Boche' hasn't been reclaimed; it has simply fallen into disuse. If the offensiveness of a slur arises from a signaling relation between the slur and derogatory attitudes, then we should expect that if the term falls into disuse the signal will degrade, and with it the offensiveness of the term will fade.[50]

Transference is a closely related case, occurring when the offensiveness of one slur transfers over to a co-referential slur, leaving the first expression muted or toothless. We can understand transference as a two-part phenomenon: fading of the original term, and an attendant shift in social perception that the associations of the first term actually belong to the second. This appears to be what occurred with the terms 'chink' and 'coolie' in the United States: as 'coolie' faded, 'chink' gained prominence as the preferred slur against Chinese individuals. Speakers subsequently assumed that 'chink' had always been the dominant slur, and are now surprised to learn that it was 'coolie' rather than 'chink' that featured prominently in the historic oppression of Chinese nationals in the Americas.[51]

7. Conclusion

Invoking the signaling relations involved in contrastive choice fully accounts for all of the major features of offense generation, and gets the right results in some less well-known cases. Utilizing the account does not require postulating any novel pragmatic mechanisms, since signaling relations are well-accepted and prolific, and allows the explanation of the offensiveness of slurs to parallel our best accounts of the offensiveness of impolite and rude behavior. Since this explanation is compatible with, and therefore available to most theories of the semantics of slurs, absent a compelling reason we should not evaluate theories based on their ability to give a purely semantic explanation of offensiveness. Instead, perhaps we should focus on accounting for or explaining away other central characteristics, including *how* the terms came to be associated with derogation in the first place, their conditions of application, relation to their neutral counterparts, and the connection between slurs and stereotypes of the target group.

Notes

[1] Hom 2008:426, Anderson & Lepore 2013a:33.

[2] Anderson & Lepore 2013a, Camp 2013, Croom 2011:7, Kennedy 2002:75, Potts 2007, Schroeder 2009:305.

[3] Potts 2007:169, Hedger 2012:74, Boisvert 2008, McCready 2010.

[4] Hom 2008, Anderson & Lepore 2013a, Jeshion 2013, Whiting 2013.

[5] Hornsby 2001:130, Williamson 2009:139.

[6] Some semantic theories maintain that slurs express negative stereotypes (Hom, Tirrell) or attitudes (Potts), conventionally implicate derogatory attitudes or contents (Williamson, Whiting), semantically encode contempt, perhaps via a rule of use (Jeshion), or represent the target as contemptible (Richard, and perhaps many inferentialist accounts). These conditions can also be met by minimalist or primarily pragmatic theories, such as the one advocated by Nunberg (2013).

[7] Expletives are words generally used in outbursts, such as 'damn!'; vulgarities involve tabooed bodily references ('shit', 'fuck', and 'ass' are all vulgarities), and general pejoratives are insults used to condemn a particular person (as opposed to a group) for behavior at a given time ('asshole', 'idiot', and 'prick' are prime examples). I refer interested readers to Beller (2013) for an analysis of general pejoratives, and Hay (2013) for a careful discussion of the systematic differences between slurs and general pejoratives. The distinction is occasionally clouded by terms like 'bitch', which have both a slurring and a general pejorative use. Nunberg (manuscript) examines some such 'crossover' slurs, explaining how a slur can drift into use as a general pejorative and visa-versa.

[8] Allan & Burridge 2006:89; Culpeper 2010:207.

[9] Anderson & Lepore (2013a; 2013b) should be happy to allow this, since on their account the institution of a taboo subsequently associates use of the tabooed term with flagrant disregard for the wishes of the targeted community, and this disregard does the heavy lifting in their explanation of slurs' offensiveness.

[10] Feinberg 1985:93.

[11] Quotative occurrences of the slurring terms 'nigger' and 'cunt' are an exception in some high-profile cases, where the author is typically subject to criticism for failing to use an available circumlocution such as 'the N-word' (or 'C-word'). This exception to the rule is instructive, and I will return to it in §5.

[12] Himma (2002) discusses a case where a dictionary was censured for their definition of the term 'nigger', but the details of the case suggest that it was the apparent endorsement of the term, rather than its simple occurrence, that caused the offense. The *Merriam-Webster Dictionary* in 1998 gave the following as the first definition:

> D1: 'Nigger'—A black person. *–usu. taken to be offensive.*

This definition was interpreted by many (including the NAACP) as endorsing the application of the slur to African-Americans. (Himma 2002; Kennedy 2002:84.) The problem was not that the noisome word appeared in the dictionary; it was that

> "...the definition labeled me and anyone else who happened to be black or have dark skin a nigger. [...] the N-word needs a more accurate first definition, *reflecting that it is a derogatory term* used to dehumanize or oppress a group or race of people."(Delphine Abraham, 'Changing Webster's Dictionary', *Essence* March 1998. Emphasis mine.)

The proposed solution—which Abraham suggested would resolve what she and the NAACP found offensive—is *not* to refrain from even mentioning the term. Instead, it is to expressly acknowledge and label it as derogatory, perhaps by updating the definition to

> D2: 'Nigger'—A derogatory term for a black person. *–usu. taken to be offensive.*

D2 instances a docile occurrence: the term is mentioned, but not even implicitly endorsed, and consequently the entry is inoffensive despite tokening the slurring term.

[13] Some theorists dispute this possibility: Anderson & Lepore (2013b) contend that the sentence "'nigger' is a derogatory term" is offensive, so they would likely reject my suggestion that dictionary cases are docile.

[14] Utterances of the form "Henrik is not a kraut, no one is" are meta-linguistic comments on the range or permissibility of the term 'kraut', rather than denials in the standard sense (see Horn 1985 for an extended discussion.)

[15] The difficulty with negating or denying the offensive content of slurring terms has been well and widely observed; for more detailed analysis, see McCready (2010), Anderson & Lepore (2013a), and Camp (2013).

[16] This example is inspired by McCready's (2010) "If I didn't like niggers, I'd probably be racist." Both are offensive, though this is more puzzling in (4) given that the slur occurs embedded in a conditional, and other speech acts (i.e. promises) are inert under similar embeddings.

[17] *The Wire,* Season 1, Episode 5. The line in (5) is spoken by McNulty in reference to his ex-wife, and Detective Kima Greggs is quick to criticize him for "calling the mother of his children a cunt."

[18] Potts 2007:171, 175. The first is Potts' example (11), the second his (19).

[19] Potts 2007:171.

[20] Kratzer (1999) points out that with enough priming, we can get pejoratives like the one in (9) to attach to a perspective other than the speaker's, but these contexts are quite difficult to evoke, and are at best unclear whose perspective is being expressed.

[21] Brown & Levinson 1978, Anchimbe 2011.

[22] Culpeper 2011:69.

[23] Culpeper 2011; 2010:3237.

[24] Culpeper 2011:135.

[25] Terkourafi 2005a. A 'frame' is a holistic representation of a context type which tracks co-occurrence of linguistic expressions, on the one hand, and on the other hand a variety of non-linguistic elements, including social contingencies, speaker's intent, and the setting and history of the interaction.

[26] Terkourafi 2005a:247.

[27] Bach & Harnish 1979; Bach 1995, 1998.

[28] Terkourafi 2005a:248.

[29] Culpeper 2011:23; Terkourafi 2005b:213; Watts 2003:5.

[30] Fraser 1990:233; Locher 2006:263; Culpeper 2010:3239.

[31] D. Lewis, *Convention: A Philosophical Study* 1969.

[32] To make this clear, consider a hypothetical community where everyone believes that α is associated with ϕ, but as a matter of fact α has been used exactly once, by a foreigner who did not endorse ϕ and was unaware of the community's belief. In that community an informed unforced choice of α would still signal ϕ: the widespread belief that the term is associated with ϕ is sufficient to make it the case that the term is in fact so associated. (My thanks to an anonymous referee for pressing me to be more explicit on this point. In §6.3.2 I show how this dependence on the beliefs of the community provides an elegant explanation of phenomena like transference and amelioration over time.)

[33] Skyrms, 2010. Signaling relations between an expression α and some associated content ϕ naturally emerge when the correlation between preferring α and endorsing ϕ is sufficiently strong and stable; as a result, signaled content rarely attaches to the default or unmarked term, since the signal will be swamped by noise.

[34] The information carried by a signal is measured by how much it alters the relevant probabilities; a high-information signal alters the probabilities quite a lot, while low-information signals have a much more limited effect.

[35] This does not imply that *nothing* offensive can occur in forced choice contexts; it is consistent with this account that hearers could hold the speaker innocent while finding the utterance derivatively offensive, insofar as it (necessarily) makes salient an original utterance that warranted offense.

[36] This association could result simply from observed correlations between using α and holding some attitude in ϕ. Signaling relations can emerge in this way (Skyrms 2010), and specifying the association this way relieves us from having to clearly delimit the precise attitudes associated with a given slur.

[37] This suggestion is made to varying degrees by Camp (2013), Copp (2009), Croom (2011), Finlay (2005) and Väyrynen (2013). While my theory in no way depends on the frameworks developed by

these authors, it is largely consistent with them; those who favor these views can interpret my theory as a way of giving details for *how* and *why* the choice pragmatically communicates attitudinal facts. A distinctive element is that it captures the explanatory power without committing to much machinery or incurring the theoretic costs of other accounts: it need not define a 'slurring perspective' (as Camp must); it need not deny that slurring assertions have truth values (like Hedger); it does not depend on invoking stereotypes (unlike Croom's account), though it does not rule them out; it does not require some determinate proposition to play the role of implicated content, and it provides a mechanism that ties the offensiveness of slurs to word choice in a way that is systematic but still able to accommodate fine-grained differences that depend on features of the speakers (in that way it moves beyond suggestions by Copp, Finlay and Väyrynen).

[38] Though this explains the offensiveness of α as used by a community which *universally* endorses derogatory attitudes towards ψs, it cannot be generalized to explain the offensiveness of slurs in communities with mixed attitudes. Either in such cases we depend on extra-linguistic information (tone, gestures, etc.) to distinguish the contemptuous speakers from the non-contemptuous speakers—in which case the term itself is not underwriting offense—or an alternative term emerges. When some significant minority of speakers reject the offensive attitudes, they tend to linguistically innovate, introducing new terms or expressions to achieve reference to ψs without associating themselves with the dominant attitudes. These new terms are 'politically correct', and function as overtly *polite* alternatives, strongly signaling the speaker's rejection of the dominant attitude. As use of these alternative terms becomes more widespread, the signal content shifts to negatively mark speakers who continue to privilege the original term. (My thanks to a referee who encouraged me to elaborate on this alternative.)

[39] Using a slurring term is of course not the only way to signal that one endorses the derogation or persecution of the target class. There could be a polite racist, who scrupulously avoids using slurs while nevertheless endorsing offensive attitudes. He may do so by explicitly saying "I approve of the oppression, persecution, and subjugation of βs simply because they are β." In such cases, deep offense will be warranted, though it will not have been generated by the use of a slur. In extreme cases, such utterances may be *more* offensive than if the speaker had simply used the slur.

This is as it should be: while the use of a slur is strongly associated with endorsing *some* of the attitudes in the cluster ϕ, there is too much noise for any given use to signal which particular attitudes the speaker endorses. Specifically, the use of a slur does not reliably signal that the speaker endorses the *most* extreme such attitudes. If the use of a slur licenses offense by signaling (with probability < 1) that the speaker endorses offensive attitudes, then the explicit endorsement of the worst such attitudes (effectively raising the probability to 1) should be expected to be more immediately offensive.

[40] Interviewed in 'Too Much Use of The 'N' Word! TV Usage Grows', *Sacramento Observer* Jan 24, 1996.

[41] Touré, 'Can Whites Say the N-Word?' in *Time*, 2011.

[42] There are instances, particularly when slurs come from the mouth of children characters, that our ire is directed at the society around the character, rather than at the speaking character herself. Scout serves this function in *To Kill a Mockingbird*, as does a nephew of Pee Wee Reese in *42*.

[43] See Spike Lee's criticism of Tarantino for *Jackie Brown*. Amy Archerd, http://www.variety.com/article/VR111779698/'Lee has Choice Words for Tarantino', *Variety* Dec 16 1997.

[44] There have of course been some attempts to ban the word 'nigga/er' from the music industry, contending that the term should simply never be used. The following poem by performance poet Dean Atta is a prime example of the sort of appeal that is made:

"Rappers, when you use the word 'nigger', remember
That's one of the last words Stephen Lawrence heard
So don't tell me it's a reclaimed word
I am nobody's nigger So please, let my ancestors rest in peace..."

It is important to note (1) the authors of the movement make no negative comments about the uses in critical works (and in this case, the appeal even features an art-use of the term), and (2) the justification offered is that it is impossible to reclaim the word, and hence unacceptable to embrace and use the word to refer. This discussion is then more properly about direct primary uses, rather than the acceptability

of indirect art uses. (Poem reprinted as quoted in Akbar, Arifa. 'How Should we use the n-word?', *The Independent*, March 6, 2013.)

[45] The film's indifference concerning accurately representing the speech of the period is evidenced by the use of 'fuck' and its cognates in the film; though the film is set in 1858, the term did not acquire its modern status as an all-purpose expletive until at least 1901 (G. Nunberg, *Ascent of the A-Word*, 2012).

[46] Damien Willis, 'Airwaves: The N-Word', *Las Cruces Sun-News (New Mexico)*, December 27, 2003.

[47] Daniel Miller, "'42' director: I didn't soft pedal racially charged language', *LA Times*, April 12 2013.

[48] Spike Lee's criticisms of Tarantino's work allude to this, as does Louis C.K.'s podcast discussion of Tarantino wanting a "nigger pass".

[49] A number of authors in the slurs literature (Camp 2013, Jeshion 2013, Hedger 2012) have cited this phenomena as the reason why, despite their toxicity, discussions of offensiveness must use active slurs as examples, rather than the tamer archaic slurs.

[50] This is because signal content is reinforced through use. As use frequency decreases, the signal is reinforced less often, with the result that over time the signal strength falls below the minimum threshold for retention and is 'forgotten'. Alexander, Skyrms, & Zabell (2012) offer a basic model of this phenomenon; Alexander (forthcoming) discusses a more nuanced account that incorporates past-discounting to more accurately model natural languages by allowing for semantic drift. Either model predicts that signal content will fade as the term falls into disuse.

[51] 'Coolie trade' in the mid 1800s referred to the practice of shipping (often abducting) Chinese nationals to the Americas, and indenturing them to years of hard labor (Farley, 1968). The term also featured prominently in racist US political rhetoric surrounding the implementation of the 1882 Chinese Exclusion Act, outlawing the immigration of Chinese laborers into the United States (Junn, 2007; Herbst, 1997). Despite all this, it is 'chink', rather than 'coolie', that carries these associations.

References

Abraham, Delphine, 1998. 'Changing Webster's Dictionary', *Essence* March.

Alexander, J. McKenzie, Brian Skyrms, and Sandy Zabell. 2012. 'Inventing New Signals', *Dynamic Games and Applications*, Vol. 2 No. 1 pp. 129–145.

Alexander, J. McKenzie, forthcoming. 'Learning to Signal in a Dynamic World', *British Journal for the Philosophy of Science*.

Allan, Keith and Kate Burridge, 2005. *Forbidden Words: Taboo and the Censoring of Language*, Cambridge: Cambridge University Press.

Anchimbe, Eric A. 2011.'On Not Calling People by Their Names: Pragmatic Undertones of Sociocultural Relationships in a Postcolony', *Journal of Pragmatics* 43:6, 1472–1483.

Anderson, Luvell and Ernie Lepore, 2013a. 'Slurring Words' *Noûs* 47, 25–48.

———. 2013b. 'What Did You Call Me? Slurs as Prohibited Words', *Analytic Philosophy* Vol. 54 No. 3 pp. 350–363.

Akbar, Arifa, 2013. 'How Should We Use the N-word?', *The Independent*, March 6.

Archerd, Amy, 1997. 'Lee has Choice Words for Tarantino', *Variety* Dec 16.

Bach, Kent, 1995. 'Standardization vs. Conventionalization', *Linguistics and Philosophy* 18:677–86.

———. 1998. 'Standardisation Revisited', in Asa Kasher (ed.) *Pragmatics: Critical Concepts*. London: Routledge, pp. 712–722.

Bach, Kent and Robert Harnish, 1979. *Linguistic Communication and Speech Acts*. Cambridge, MA: MIT Press.

Beller, Charley, 2013. 'Manufactured and Inherent Pejorativity', *Proceedings of SALT* 23:136–155.

Brown, Penelope and Stephen Levinson, 1987. *Politeness: Some Universals in Language Usage*, Cambridge: Cambridge University Press.

Boisvert, Daniel. 2008. 'Expressive-Assertivism', *Pacific Philosophical Quarterly* 89:169–203.

Camp, Elizabeth, 2013. 'Slurring Perspectives', *Analytic Philosophy* 54:330–349.

Copp, David, 2009. 'Realist-Expressivism and Conventional Implicature', In Russ Schafer-Landau (ed.) *Oxford Studies in Metaethics, Volume 4*. New York: Oxford University Press.

Covin, David, 1996. Quoted in 'Too Much Use of the 'N' Word! TV Usage Grows', *Sacremento Observer* Jan 24.

Croom, Adam, 2011. 'Slurs', *Language Sciences* 33:3, 343–358.

Culpeper, Jonathan, 2010. 'Conventionalized Impoliteness Formulae', in *Journal of Pragmatics* 42:3232–3245.

———, 2011. *Impoliteness: Using Language to Cause Offense*, Cambridge: Cambridge University Press.

Farley, M. Foster, 1968. 'The Chinese Coolie Trade 1845-1875', *Journal of Asian and African Studies*, 3:3/4 (July/Oct.) 257–270.

Feinberg, Joel, 1985. *Offense To Others*, Oxford: Oxford University Press.

Finlay, Stephen, 2005. 'Value and Implicature', *Philosopher's Imprint*, 5:1–20.

Fraser, Bruce, 1990. 'Perspectives on Politeness', *Journal of Pragmatics* 14 (2):219–236.

Gutzmann, Daniel, 2013. 'Expressives and Beyond: An Introduction to Varieties of Use-Conditional Meaning', in Daniel Gutzman and Hans–Martin Gärtner (eds.) *Beyond Expressives: Explorations in Use-Conditional Meaning*, Leiden: Brill.

Hay, Ryan, 2013. 'Hybrid Expressivism and the Analogy between Pejoratives and Moral Language', *European Journal of Philosophy*, 21-3: 450–474.

Himma, Kenneth E., 2002. 'On the Definition of Unconscionable Racial and Sexual Slurs', *Journal Of Social Philosophy*, 33-3: 512–522.

Hedger, Joseph A., 2012. 'The Semantics of Racial Slurs: Using Kaplan's Framework to Provide a Theory of the Meaning of Derogatory Epithets', *Linguistic and Philosophical Investigations*, 11: 74–84.

Herbst, Philip, 1997. *The Color of Words: An Encyclopaedic Dictionary of Ethnic Bias in the United States*, Maine: Intercultural Press.

Hom, Christopher, 2008. 'The Semantics of Racial Epithets', *Journal of Philosophy*, 105:416–440.

Horn, Laurence, 1985. 'Metalinguistic Negation and Pragmatic Ambiguity', *Language*, Vol. 61, No. 1. (Mar.) 121–174.

Hornsby, Jennifer, 2001. 'Meaning and Uselessness: How to Think about Derogatory Words', *Midwest Studies in Philosophy*, 25: 128–141.

Jeshion, Robin, 2013. 'Expressivism and the Offensiveness of Slurs', *Philosophical Perspectives*, 27:231–259.

Junn, Jane, 2007. 'From Coolie to Model Minority: US Immigration Policy and the Construction of Racial Identity', *Du Bois Review*, 4:2, 355–373.

Kennedy, Randall, 2002. *Nigger: The Strange Career of a Troublesome Word*, New York: Pantheon.

Kratzer, Angelica, 1999. 'Beyond Ouch and Oops: How Descriptive and Expressive Meaning Interact. A comment on David Kaplan's Paper', Cornell Conference on Theories of Context Dependency.

Lewis, David K., 1969. *Convention: A Philosophic Study*. Harvard University Press, Cambridge, MA.

Locher, Miriam A., 2006 'Polite Behavior within Relational Work: the Discursive Approach to Politeness', *Multilingua* 25: 249–267.

McCready, Eric, 2010. 'Varieties of Conventional Implicature', *Semantics and Pragmatics*, 3:1–57.

Miller, Daniel, 2013. ''42' director: I didn't soft pedal racially charged language', *LA Times*, April 12.

Nunberg, Geoffrey. 2012. *The Ascent of the A-Word: A History of Assholism*. New York: Public Affairs.

———. 2013. 'Slurs Aren't Special', (manuscript, available online).

Potts, Christopher, 2007. 'The Expressive Dimension' in *Theoretical Linguistics* 33-2:165–198.

Richard, Mark, 2008. *When Truth Gives Out*, Oxford: Oxford University Press.

Saka, Paul, 2007. *How to Think About Meaning*, Dordrecht: Springer.

Schroeder, Mark, 2009. 'Hybrid Expressivism: Virtues and Vices', *Ethics* 119:257–309.

Skyrms, Brian, 2010. *Signals, Evolution, Learning, and Information*, Oxford: Oxford University Press.

Terkourafi, Marina, 2003. 'Generalised and Particularised Implicatures of Politeness', in Peter Kühnlein, Hannes Rieser, Henk Zeevat (eds.), *Perspectives on Dialogue in the New Millennium*, Amsterdam John Benjamins, pp. 151–166.

———. 2005a. 'Beyond the Micro-level in Politeness Research', *Journal of Politeness Research: Language, Behavior, Culture*, 1 (2): 237–62.

———. 2005b. 'Pragmatic Correlates of Frequency Use: the Case for a Notion of 'Minimal Context'. In Sophia Marmaridou, Kiki Nikiforidou and Eleni Antonopoulou (eds.) *Reviewing Linguistic Thought: Converging Trends for the 21st Century*. Berlin: Mouton de Gruyter, pp. 209–233.

———. 2005c. 'An Argument for a Frame-based Approach to Politeness: Evidence from the Use of the Imperative in Cypriot Greek', In Robin Lakoff and Sachiko Ide (eds.) *Broadening the Horizon of Linguistic Politeness*, Amsterdam: John Benjamins Publishing, pp. 99–116.

Tirrell, Lynne, 1999. 'Derogatory Words: Racism, Sexism, and the Inferential Role Theory of Meaning', in Christina Hendricks and Kelly Oliver Oliver (eds.) *Language and Liberation: Feminism, Philosophy, and Language*, Albany: SUNY Press, pp. 41–79.

Touré, 2011 'Can Whites Say the N-Word?' *Time*, Oct 12. Accessed January 2013.

Väyrynen, Pekka, 2013. *The Lewd, The Rude and the Nasty: A Study of Thick Concepts in Ethics*. Oxford: Oxford University Press.

Watts, Richard J., 2003. *Politeness*. Cambridge: Cambridge University Press.

Whiting, Daniel, 2013. 'It's Not What You Said, It's the Way You Said It: Slurs and Conventional Implicatures', in *Analytic Philosophy* Vol. 54 No. 3, 364–377.

Williamson, Timothy, 2009. 'Reference, inference and the semantics of pejoratives', in Joseph Almog and Paolo Leonardi (eds.) *The Philosophy of David Kaplan*, Oxford: Oxford University Press.

Willis, Damien, 2012. 'Airwaves: The N-Word', *Las Cruces Sun-News (New Mexico)*, December 27.