Amelioration vs. Perversion

Teresa Marques

Abstract

Words change meaning, usually in unpredictable ways. But some words' meanings are revised intentionally. Revisionary projects are normally put forward in the service of some purpose - some serve specific goals of inquiry, and others serve ethical, political or social aims. Revisionist projects can ameliorate meanings, but they can also pervert. In this paper, I want to: (i) draw attention to the dangers of meaning perversions; (ii) argue that the self-declared goodness of a revisionist project doesn't suffice to put forward a viable new meaning. I'll briefly consider whether power or authority play a role in viable revisions; (iii) show that the self-declared goodness of a revisionist project doesn't suffice to avoid meaning perversions. The road to Hell, or to horrors on Earth, is paved with good intentions. (iv) I will then point to problems for ameliorative conceptual engineering projects. Finally and more importantly, (v), I want to demarcate what meaning perversions are, and offer a working hypothesis about how they contribute to "destroy facts of shared reality".

1   Introduction

Theories about the meaning of words for social roles and categories raise interconnected questions. One is metaphysical: What is the nature of the social things we talk about? Another is conceptual: How do we think about the social world, and how does our thinking relate to the nature of social reality? And how should it relate? The latter question concerns so-called ameliorative projects. These projects answer other additional questions, for instance, what is the point of having the concept in question? What concept (if any) would better serve our social or political goals? And who are *we*? These questions are illustrated in recent debates about race, for instance (Appiah (2006), Glasgow (2003), Andreasen (2000, 2005), Haslanger (2003, 2006), Machery et al. (2009), Diaz-Leon (2015)).

In this paper, I argue that we should add one more question to this list, namely: *How can we assess the legitimacy of ameliorative purposes and projects?* A secondary question is: what theory of our concepts better accommodates the phenomena of meaning revision? I will try to answer these questions by drawing from deeply problematic historical cases, which I'll call *meaning perversions*. I do not have an answer to the legitimacy question, but will try to point to some possible answers.

2   Lessons from the Past and the Present

In LTI, *The Language of the Third Reich*, Klemperer offers a chilling description of the corrupting power of language. Klemperer witnesses how the use of language under the Third Reich corrupted and perverted a "huge number of concepts and feelings", for instance 'heroic' and 'fanatic':

> *Heroic*, *Heroism*: "young people in all their innocence, and despite a sincere effort to eliminate the errors in their neglected education, cling to Nazi thought processes... as soon as this concept (*heroism*) was touched upon, everything became blurred in the fog of Nazism"...

> *Fanatical*: "...and then replaced it with 'fanatical'": people become convinced that "no one can be a hero without fanaticism".

> *Words that were previously common property*: "[Nazi language] commandeers for the party that which was previously common property and steps words, groups of words, and sentence structures, in poison." (Klemperer, 2000, 14)

In an article from 2017, "The Autocrat's Language", Masha Gessen warns of the damage to public discourse and political reality caused by what Donald Trump says. She does so by drawing a parallel between Trump's discourse and the discourse of autocrats in the former Soviet Union and presently in Russia.

> "Freedom," on the other hand, was, as you know, slavery. That's Orwell's 1984. And it is also the USSR, a country that had "laws," a "constitution," and even "elections," also known as the "free expression of citizen will." The elections, which were mandatory, involved showing up at the so-called polling place, receiving a pre-filled ballot—each office had one name matched to it—and depositing it in the ballot box, out in the open... Calling this ritual either an "election" or the "free expression of citizen will" had a dual effect: it eviscerated the words "election," "free," "expression," "citizen," and "will," and it also left the thing itself undescribed. When something cannot be described, it does not become a fact of shared reality. Gessen (2017)

I call these uses of "freedom", "free expression of citizen will", "election", "heroic", "fanatical", etc. *meaning perversions*. Meaning perversions are attempts to hijack the language of justice, politics, social roles, moral or epistemic virtues, in a way that makes people worse off. In other words, meaning perversions are revisionary or engineering projects that produce harm.

3   Meaning revisions in the wild

3.1   Viable meaning revisions?

In her recent book, Bicchieri (2017) shows how social norms change in the real world – in the wild, as it were. Word meanings also change. Glasgow (2018) nicely illustrates a variety of changes in word meaning and reference:

> 'Sick' used to just mean ill; now it (like 'bad') also means good. Apparently we can now use 'literally' to mean figuratively. And then there are terms whose old meaning

or referent is simply lost and taken over by a new one—outright replacement. Famously, 'Madagascar' had been the name for the part of the Somali peninsula around Mogadishu, not that big island off of Africa's eastern shore. But apparently because 'jazira' is Arabic for both 'island' and 'peninsula,' Marco Polo got the location wrong. When map-makers subsequently doubled down on his mistake, the referent of that word underwent its change, and we now know Madagascar to be the island, not the peninsular region (Burgess (2014)). This instance of referential replacement is hardly unique. 'Awful' once meant inspiring awe. It now instead means really bad. Even our hallowed 'bachelor' meant young knight, and then a certain rank of university achievement, before it referred to a marital status. 'Fantastic' (from imaginary to wonderful), 'silly' (from worthy to foolish), 'senile' (from the senescent generally to those with dementia specifically), 'tool,' and on and on: language is littered with conceptual change.

These examples illustrate the malleability of words' meanings. Political history of the last century offers various examples that illustrate how political power has tried to take advantage of that malleability. As was the case in the USSR, or now in Russia under Putin or in the US under Trump, people with power or influence often try to *pervert* word meanings in order to advance their own agenda and manipulate the public.

The end aim of meaning revisions can be, and often is, self-interested. But the proposed end aim may be *prima facie* good. When authors discuss *conceptual engineering*, *conceptual ethics*, *conceptual amelioration*, often they are concerned with the concepts our words should express, given the specific aims of a theory. For instance, Tarski (1943) considered that 'truth' as used in natural languages is an incoherent notion, and that it gives rise to paradox. He argued that 'true' and 'truth' should be defined in a way that would allow it to play its foundational role in a semantic theory, while insulating the theory from the contradictions and paradoxes that the natural language use gives rise to. We could say that, for the purposes of the theory, a revision of the meaning of 'true' is viable when the new meaning fulfills the aims set by the theory.

We can think that similar theoretical goals can be pursued in social and political domains. A theorist may have the aim of offering a political theory of gender identities and relations, for instance. A concept that will play a role in that theory will be the concept of *woman*, and some of the theoretical decisions that will need to be weighed concern biology, social norms and roles, social hierarchies, and self-identity. Depending on the theoretical choices, a theorist may or may not count trans-women as women (Haslanger (2006), Saul (2006), Bettcher (2013, 2014), Díaz León (2016)). Also here, we could say that, for the purposes of the theory, a revision of meaning of 'woman' is viable when the new meaning fulfills the aims set by the theory.[1]

There are differences between theoretical domains and how the meaning revisions they promote transpires into language use in the wild. Whereas the notion of truth in Tarski's theory is not meant to replace the use of 'true' and 'truth' in informal natural language use, meaning revisions in social ontology about sex and gender also have social justice aims *in the real world*. In this sense, redefining 'marriage' purports to have impacts in actual legislations, and thereby change the extension of 'marriage' in the real world. Likewise, whether or not 'woman' refers to trans-women has an impact on who gets to use which public toilets, for instance. In this sense, even if conceptual engineering offers a viable new meaning *in a theory*,

---

[1]I'm grateful to Pablo Rychter for discussion about what a viable new meaning may be.

given its aims, it's not automatic that there is a viable new meaning out there in the world.

What do viable meaning revisions "in the wild" require? One possibility is to say a meaning revision of a word S is successful just in case the 'meaning engineer', i.e. proponent of the revision, has the power or the influence to change the use of the word in a way that meets the social or political aims that are set. Recent research by Galinsky et al. (2013) suggests that the successful appropriation of a derogatory term occurs when the perceived *power* or *social status* of the target group increases. But it would seem that power or influence are at best necessary for meaning revisions, but not sufficient. The power of the Politburo was not sufficient to change the meaning of 'free elections' or 'citizen will'. Forced attempts of meaning revisions clearly can fail. Moreover, they share a negative feature: the resulting language is *destitute* and its enforcement results in an impoverished experience.

3.2   The road to Hell is paved with good intentions

The end aim of meaning revisions can be self-interested (as is, allegedly, Trump's case), but often meaning manipulations have social or political justice aims, or at least, aims that their proponents *believe* will bring about a better world. Revisionists may *believe* in their purpose and methods. Yet, as we know, there is a gap between what may actually be the best course action to achieve a certain end, and our epistemic capacity in assessing that course of action as a means towards the desired end. Moreover, the fact that we may desire certain ends does not make those ends desirable all things considered, nor things that are desirable independently of the means we set to achieve them, or independently of the costs that trying to achieve them may bring.

In recent work, Maynard and Benesch (2016) discuss the conditions under which so-called *Dangerous Speech* can occur. They characterize Dangerous Speech as speech acts whose force is capable of encouraging approval of violence by the audience.

After WWII, many people asked themselves *how was the Holocaust possible?* Sadly, the Holocaust is not the only case of mass violence in the history of the 20th century. The field of genocide studies has covered atrocities in the former Soviet Union, China, Cambodja, Rwanda, the former Yugoslavia, etc. In spite of socio-cultural differences, a common denominator in mass violence appears to be the use of inflammatory propaganda and discourse.

In recent work on dangerous speech and ideology, Maynard and Benesch (2016) define *dangerous speech* as speech acts whose force is capable of encouraging approval of violence by the audience. They argue that dangerous speech is a product both of the *context* and the *content* of discourse, and these feed into and overlap each other. Inflammatory speech uttered in a context where it cannot be disseminated, because audience strongly disapproves of it or the speaker has no influence, does not amount to dangerous speech. I.e., in contexts where there is *no uptake*, inflammatory speech does not amount to dangerous speech. And an influential or authoritative speaker addressing a volatile audience through powerful means of dissemination is not dangerous speech if the content is not inflammatory or hateful. I.e., unless the content persuades audience that violent actions are permissible, because they are represented *as morally justified by the circumstances*, speech is not dangerous (this does not mean that the whole audience will engage in violent actions, just that it will at least appear permissible and justified).

The specific features of *contexts* of dangerous speech are the *speaker*, the *audience* in its the socio-historical environment, and the availability of *means of dissemination* (Maynard and Benesch, 2016, 77). The speaker may be someone with authority, or charismatic, with enough power to influence the audience. For instance, before the 2016 presidential election in the US, Donald Trump was already an influential personality, because of his presence in popular reality TV shows, in political commentary in FOX News, involved in the "birther" movement, active on Twitter, etc.

The audience is the second important component. Although audiences are not homogenous, certain members are more susceptible to inflammatory speech. Dangerous speech often occurs in social and historical contexts that increase the likelihood that the audience accepts that violent actions against certain people is permissible, as punishment for presumed past crimes, or to prevent presumed future threats. Among the possible relevant aspects of socio-historical contexts include longstanding grievance, resentment, and memory of (real or imagined) historic injustice, a weak or dysfunctional justice system, competition for resources, and land disputes. Influential speakers may manipulate and exacerbate the resentment against members of another group to capitalize on populist rhetoric and take advantage of those grievances for political gains.

The final element in the characterization of the context of dangerous speech is the medium of dissemination, and the *correlation* between dangerous speech and mass violence could be established thanks to the study of the use of mass media. If a community relies on one source of news, the message spread by that source is more influential. In Rwanda, the station RTLM was the main source of the inflammatory messages. Yanagizawa (2014)'s statistical study of the effects of the virulent propaganda RTLM in Rwanda indicates that killings were 65-77% *higher* in Rwandan villages that received the RTLM signal, compared with those that did not (for reasons like topography) receive the signal.

Maynard and Benesch (2016) characterize six features of the *content* of dangerous speech. They call the first *dehumanization*, a notion that Tirrell (2012), Jeshion (2016), and Snyder (2017) also rely on. These are forms of discourse that can do direct harm by the offense, denigration, or derogation of members of a target group as less human. The second feature of the content of dangerous speech is *guilt attribution*: members of a group are said to be guilty (as members of the group) of past crimes, e.g., rape or murder, of stealing, responsible for current difficulties, destruction of economy, occupation or oppression, etc., and this is seen as a moral justification for resentment and retribution. The third feature is *threat construction*: an in-group accuses an out-group of being a threat – of planning violent attacks against the group. The fourth feature is the *destruction of alternatives* and the representation of the proposed course of action as a historical necessity to achieve a certain end. For instance, Figes (2002) describes how citizens of the Soviet Union thought that the violence of the Stalinist era was the only possible and necessary path to Communism, reporting someone that said:

> I had my doubts about the Five Year Plan... but I justified it by the conviction that we were building something great... a new society that could not have been built by voluntary means. (Figes, 2002, 111)

Maynard and Benesch (2016) say that the destruction of alternatives may be achieved through other means, for instance by describing alternatives as impractical or inefficacious.

The fifth feature of the content of dangerous speech is what they call *Virtuetalk*, which is the feature I'm particularly interested in connection to amelioration and meaning perversions. Through VirtueTalk, the audience is motivated by deep and unreflected feelings that something feels "good" or "bad", in particular inducing positive moral self-appraisal, a "satisfactory mental image of themselves... often shaped by notions of ideal group-identities, that produces considerable self-esteem" (Maynard and Benesch, 2016, 84). In a blog post written two weeks after the 2016 presidential election, writer Kelly J Baker wrote:

> In my north Florida hometown, white people (no matter their class orientation) perfected nice racism. Now, there were some white folks, who were belligerent and unapologetic racists. But most white folks hid their racism behind civility until provoked. (They also didn't acknowledge that white supremacy was a structure that organized our lives, but rather imagined that racism just appeared in particular racist words and actions.) These white people seemed very nice and decent until they felt they had to respond to (or were provoked) by the existence of people of color. Racism existing under smiles and small talk. (Baker (2016))

The self-representation of white southern Americans as "nice decent folks" contrasts with their condonation of violence against black Americans and minorities, and it is captured in the alternating lines of the second verse of the song *Strange Fruit*. The song was written by Abel Meeropol after seeing a widely circulated photograph of the 1930 lynching and hanging of Thomas Shipp and Abram Smith in Marion, Indiana, two black men accused of murder and rape. The second verse of *Strange Fruit* reads:

> *Pastoral scene of the gallant south*
> *The bulging eyes and the twisted mouth*
> *Scent of magnolias, sweet and fresh*
> *Then the sudden smell of burning flesh*

The verse contrasts two strongly opposed images. On the one hand, it presents the pastoral scene, the sweet and fresh scent of magnolias, symbols of the "gallant south": a self-representation compatible with the revisionist historical myth of the Lost Cause of the American secession war, which presents the motives of the Confederate States as honorable and brave. On the other hand, it presents the stark reality of the human costs of the South's white privilege, i.e., of the South's entire economic dependence on cheap or free labour, and the racial subordination that underlies this economic structure.

The final feature of the content of dangerous speech is *Future-bias*, i.e., the promise of future goods. Future-bias is presumed to outweigh the short term difficulties the audience may have to endure, or the moral costs of the violence against others. As they say,

> But the anticipated benefits can also be extravagant and utopian—promises that a positive transformation of society will be brought about through a temporary violent transition, or that national unity and prosperity for a long-mistreated people can be obtained. In light of the expectation that Soviet violence would protect the revolution and usher in Communist utopia, Lenin assured his followers that in the future "the cruelty of our lives, imposed by circumstance, will be understood and pardoned. Everything will be understood, everything." Soviet ideology and the justification of massive violence and cruelty in the name of a promised future society of abundance convinced millions in the Stalinist era. The novelist Boris Pasternak wrote in a letter in 1935: "The fact is, the longer I live the more firmly I believe in what is being done, despite everything. Much of it strikes one as being

savage [yet] the people have never before looked so far ahead, and with such a sense of self-esteem, and with such fine motives, and for such vital and clear-headed reasons."(Maynard and Benesch, 2016, 85-86)

It suffices for present concerns to highlight the fact that *Virtue Talk* and *Future-bias* can be misguided: either in the means that agents see as justified to achieve their ends, or in desiring to achieve certain ends as good.

4   Consequences for Ameliorative Projects

4.1   Haslanger on Externalism, Descriptive Analysis, and Amelioration

Suppose we are interested in understanding the role of the language of justice and politics, e.g., in understanding the social and political role of words like 'democracy', 'freedom', 'fair elections', 'citizen', 'the people', etc. Sally Haslanger has persuasively made the case that we can mean different things by "conceptual analysis", and as such, our reflection on the role of the language of politics can take different approaches.

In a 2006 article, Sally Haslanger distinguishes between *manifest*, *operative*, and *target* concepts. A manifest concept is the concept one thinks guides one's categorizing, whereas the operative concept is the concept that corresponds to actual categorization patterns. The *target* concept is the concept that, *all things considered*, we should employ *given our interests, facts, etc.* (Haslanger, 2006, 99). It is the concept that, in the end of the *ameliorative* project, we should be using.

Haslanger (2006) embraces an externalist account of the meaning of social kind terms. A *descriptive analysis* of social kind terms is one that tracks the social kinds, properties, or entities. An *ameliorative* approach to the language of politics itself will ask which concepts we *should* use, given our aims. Although Haslanger uses "ameliorative" to talk about conceptual engineering approaches, she says, "whether or not an analysis is an improvement on existing meanings depends on the purposes of the inquiry." (Haslanger, 2012, fn 1, p. 367) I take it that it would be too simplistic an answer to say that an analysis is an improvement on existing meanings if the engineers of the conceptual shift aim at positive social or political purposes. *Virtue Talk* and *Future Bias* show that the (perceived) goodness of a purpose does not guarantee the goodness of the outcome. The purposes of the outcome don't suffice to make a revisionist project legitimate.

I take it that there is some tension between *the concept we should be using, all things considered* and *the concept we should be using, given our aims, interests, facts.* The intersection of what we should do, all things considered, and what we should do, given our aims, etc., leaves open the possibility that the answer to what the question of what ameliorative projects are legitimate may remain epistemically unavailable. It would be desirable to have some constraints or guidelines on how to assess a projects' legitimacy to guide our assessment of *all things considered* questions.

4.2 The Viability of Ameliorative Projects

Earlier, I distinguished two ways of understanding the viability of an ameliorative project. We can understand a revisionist project as successful insofar as the meaning of a word in the theory contributes to serve the aims of the theory. But we can also understand a revisionist project not just by its theoretical aims, but by the adoption of the revised word meaning at large.

Haslanger wants to say that there is room for contextual variation in "our evolving practices". Is a social matrix that partly determines a word meaning *restricted* to a historical time and region, and the practices therein, or does it transcend the borders of a socio-historical period narrowly conceived? Does it depend on how a word is used within the borders of a limited practice, or is it constrained by its genealogy?

Understood narrowly, a social practice where 'free expression of citizens will' and 'election' are used can pick the ritual practice in the USSR. But if the narrow delimitation were all there was to the meaning of "free elections", the impoverished experience phenomena should not occur. But Gessen's testimony reveals a mismatch between the *presumed target concept* of 'free election' and the *resilient operative concept*. A lesson here is the extent of the "substantial work" needed for a viable new meaning revision.

This lesson *may* be compatible with a semantic externalist view. What the crucial words in the language of justice and politics mean is determined externally by their long history across a variety of countries and cultures. The intention to hijack those words to give legitimacy to the undermining of moral or social values does not guarantee a successful shift in the words' meanings. It also does not guarantee that the attempt will be morally or politically *legitimate*. This answer preserves *externalism* at the cost of sacrificing the weight of *local social practices of meaning revision*.

There are non-externalist options. Glasgow (2018) offers an internalist, dispositional account of meaning: the meaning of a word is fixed by language users having certain dispositions to use the term in certain ways: "Consequently, meanings change—concepts shift—when the relevant dispositions change". Yet, dispositions to resist a meaning revision show the limits of a word's meaning. Negotiation dispositions are not just concept indicators but determine which concepts and referents are assigned to which words. *Dispositions to refuse to negotiate*, rather than actual refusals to negotiate, set meaning. If actual use dictated meaning, then it would be impossible to use old words in new ways without changing the subject. Glasgow's dispositional internalism offers a natural explanation of why the meaning of 'free elections' didn't really change in the USSR, although actual use did change. People's actual use of the words was impacted by what the party/politicians dictated. Since the extensive (in area and time) use of 'free election' to refer to the ritual of casting pre-filled ballots did not actually change the meaning of 'election' or of 'freedom', we should say that the USSR made a perverted use of those words that did not actually succeed in changing meaning. Glasgow's theory can explain *resistance* to meaning revisions.

Ameliorative projects, understood as redefinitions made for the purpose of advancing ethics and justice still face a problem: if meanings are determined by use dispositions, and these can be so resilient to change, what effective ameliorative projects can we engage in? Moreover, since unsuccessful meaning revisions have as consequences the 'loss of a shared experience', this alone suffices for ameliorative projects to be intrinsically problematic, given their possible

consequences.

4.3   The Legitimacy of Ameliorative Projects

4.3.1   Harmful perlocutionary effects and constitutive norm erosions

The main question this paper addresses is *wow can we assess the legitimacy of ameliorative purposes and projects?* I want to claim that meaning revisions can be ameliorations or perversions. Ameliorations are improvements on existing meanings. Meaning perversions, in contrast and by analogy, would be corruptions of existing meanings.

This is still fairly imprecise, but it can be understood in two ways. First, it can be understood in terms of the harmful consequences of meaning revisions, i.e. harmful *perlocutionary effects* of meaning revision. These could include the "impoverished experience", "destitute language", a lost "sense of shared reality". The political dangers of these perlocutionary effects should not be neglected or minimized. These effects are autocrats' intended perlocutionary effects, since they diminish a population's capacity to resist the autocrat's control of social reality.

Second, meaning perversions can be understood *constitutively*. Here, a defining constitutive feature of meaning perversions is *undermining norm enforcement*. A meaning perversion is a wolf in sheepskin:

A speaker S perverts the meaning of a word w (of the language of justice, politics, morality, or epistemic virtues) just in case S's use of w is presented as an enforcement or application of norms or values (of justice, politics, morality, or epistemic virtues) but erodes those very same norms or values.

The constitutive meaning of meaning perversions and their perlocutionary effects are related. One sense in which we can pinpoint the nature of meaning perversions is that they have the harmful perlocutionary effects they have precisely because they undermine fundamental norms or values. For instance, when speech is used in a way that creates exclusion, or derogates some people, in an us vs. them way, it undermines norms of equality.

This provides a way of assessing the legitimacy of meaning revisions that need not wait to assess the actual harm caused by those revisions, but that can lay conditions for morally legitimate ameliorations.

People may contribute to pervert meanings unwittingly, either by aiming at purposes that are problematic in themselves, or through means that have a high potential to cause harm. Being alert to the risks of undermining norms can contribute to identify what types of *all things considered* deliberations should be made when we engage with ameliorative projects.

4.3.2   Motivational set and norm erosion

I now advance a hypothesis about how norms can be eroded. Normative and conative aspects of words' meaning (connotations) are not merely propositions that are taken as true as updates to the *common ground*. Traditionally, *common ground* is understood as the set of propositions that are accepted as true by an audience. But I believe there is a better understanding of conversational contexts.

We can think of conversational common ground and conversational score following Langton (2012):

> I want to propose, in an exploratory spirit, the idea that the phenomenon of accommodation might extend beyond belief—beyond conversational score, and common ground, as originally conceived—to include accommodation of other attitudes, including desire and hatred. My remarks here will inevitably be programmatic. But to convey the general idea: just as a hearer's belief can spring into being, after the speaker presupposes that belief, so too a hearer's desire can spring into being, after the speaker presupposes the hearer's desire; and so too a hearer's hatred can spring into being, after the speaker presupposes that hatred. Stalnaker's common ground can perhaps be extended to include not just common beliefs, and other belief-like attitudes, but common desires, and common feelings, as well. Speakers invite hearers not only to join in a shared belief world, but also a shared desire world, and a shared hate world.(Langton, 2012, 140)

I suggest that we adapt Bernard Williams (1979)'s term 'motivational set' for the set of dispositional attitudes, plans, intentions, emotions, etc., that explain the reasons agents have for acting, as that contributes to their *motivated reasoning*. The notion of a motivational set can be generalized to be part of the *common* motivational set (which is part of common ground), i.e., those motivational attitudes that are common. And it can be generalized to be part of the 'motivational score', i.e. those attitudes that are part of the conversational score. This distinction is useful. It is good to distinguish common ground and score when we consider the motivational set. People often accept to follow norms that they don't agree with or that they disavow, and regard widely shared and permissible attitudes that they do not actually have.

What goes on with meaning perversions, like Putin's 'the dictatorship of the rule of law', the Nazi 'fanatic hero', Medvedev's 'managed democracy', Soviet elections as 'free expressions of citizen's will', the Catalan 'unilateral declaration of independence' (DUI) that as it turns out is merely 'symbolic', or their slogan 'we vote to be free!' which obliterates the democratic representation of at least half the voting people in Catalonia, etc., is that they take advantage of the different ways we communicate, and the different types of contents we can convey. There are, on the one hand, legitimate referents of these words, and there is, on the other hand, the normative or evaluative connotation that these phrases express. A sincere literal use of 'democracy' made by a competent speaker denotes a social institution that meets certain minimal constraints, and moreover expresses that institutions that meet those constraints are *desirable as good*.

Take 'free elections'. Free elections are good things, they are the basic component of democracy, a recognition of the sovereignty of the citizens of a country through their representation in the institutions of their countries' State. It has a positive connotation. All that positive aspect of the meaning of 'free election' is taken for granted, at least for a good amount of time, to manipulate people. By describing the ritual practice of the so-called elections in the Soviet Union (where people were forced to vote, although there was only one pre-filled ballot, and which as a result were neither free nor a real exercise of citizenship) the Soviet regime was perverting the meaning of 'free elections'.

The normative and conative aspect of the meaning a word, which presumably is part of the shared motivational conversational set, pragmatically contradicts with the actual application

of a phrase to something that does not meet the minimal constraints for being a free election. By doing this, the Soviet regime was eroding the positive connotation of 'free election', and normalizing the new practice. People lived under a pragmatic contradiction between the official normative ideal they were presumably respecting, and the normative reality they were forced to inhabit. This pragmatic contradiction left them, as Masha Gessen said, with an impoverished experience, one where resistance to the power of the State was more difficult. There's conflict between different norms, plans, or evaluative dispositions when they can't be jointly implemented, followed, or satisfied. Accepting one, when another is already in place, *corrodes* preexisting norms.

Similarly, the campaign for an independence referendum in Catalonia that failed to meet the requirements of the Venice Code of Good Practice on Referendums, and which was organized after a change in the Catalan constitution that violated the 2/3 majority requirement for constitutional revisions, and defied Court orders against its realization, claimed to defend the referendum under the slogan 'we vote to be free!'. The slogan conversationally implicates that, unless people vote, they are not free. Since people desire to be free, this conditional presents the act of not participating in an illegal, undemocratic, and unconstitutional act as a defense of democratic liberties. This is a blatant lie, since Spain is one of the 19 countries recognized as full democracies in the Democracy Index, Catalonia has enjoyed self-government for approximately 40 years, is one of the richest regions in Spain, and enjoys greater levels of autonomy than almost any other European region. However, it is much harder to engage and refute content that is not explicitly asserted, especially when doing so seems to conflict with things people that people strongly desire. Questioning people who engage in an illegal action, convinced that they are exercising their fundamental civic rights, without appearing to disrespect or violate their rights is extremely difficult.

The positive normative and conative connotations of 'freedom' and 'vote' are part of the shared motivational conversational set. But it pragmatically contradicts with the use of these words in actions that undermine citizen's actual freedoms and rights to political representation. This has brought about a social division that arises from the pragmatic contradiction between the official normative democratic ideal, and the way the terms have been used to create a new political reality.

4.3.3 Contrast between meaning perversions and code words

Meaning perversions contrast with *code words* (or dogwhistles). Mendelberg (2001) argued that using a code word allows the speaker and her audience to *violate certain social norms*. For example, it is a socially shared norm that *being racist is bad*. We can convene in treating this as a Norm of Racial Equality. But with a code word a speaker can flaunt the norm while *plausibly denying* doing so: e.g. 'inner city', 'welfare'.

(1)   Trump in 2011: [Explicit] If we reelect Obama, America will start to *look like an inner-city wreck.*

(2)   Trump in 2011: [Implicit] If we reelect Obama, America will come to be *dominated by poor, lazy, and criminal black people*

Code words are words that *may carry a positive or negative connotation, which is not explicit.* One way of representing this is by saying that the evaluative connotation of code words is not part of the shared conversational record. There are *overt code words*, which are designed with intent to allow two plausible interpretations – (i) one that is private and is aimed at a desired target audience that interprets the implicit racist content; (ii) a plausible non-racist content for a broader audience. There are also *covert code* words, which are designed with intent to get an audience that would disapprove or reject an explicitly racist speech, because a Norm of Racial Equality is officially in the conversational score, but who feels ok with apparently reasonable claims that are implicitly discriminatory.

(3)   Paul Ryan: We need to reform our *welfare system* to put incentives in place *for people to work hard.*

Jason Stanley (2015) claims that uses of code words undermine norms of reasonable democratic discourse. The speaker explicitly *asserts its nonracial at-issue content*, and also *communicates the not-at-issue content* that poor blacks are lazy. Since (3)'s nonracial at-issue content is "reasonable", the contribution is felt to be reasonable. Its not-at-issue content is expected to be accepted without discussion, and as such typically will be common ground after audience accepts (3) (unless it is explicitly blocked). Accepting this not-at-issue content tends to decrease empathy for poor black people ("they are lazy, so they don't deserve government assistance"), and produce belief that the perspective of a poor black person is not worthy of respect.

During the third presidential debate with Hillary Clinton in 2016, Trump made a statement that deployed both code words and meaning perversions. I'll signal code words with a *c* and meaning perversions with a *p*.

> But when I started this campaign, I started it very strongly. It's called *Make America Great Again$_c$*. We're going to make America great. We have a depleted military. It has to be helped. It has to be fixed. *We have the greatest people on Earth$_c$* in our military. *We don't take care of our veterans. We take care of illegal immigrants*, people that come into our country illegally better than we take care of our vets. That can't happen. Our policemen and women are disrespected. *We need law and order$_p$*, but we need *justice$_p$* too. Our *inner cities$_c$* are a disaster. *You get shot walking to the store.* They have no education. They have no jobs.

Here, Trump is mixing two different mechanisms, and engaging in Dangerous Speech 101.Trump is the *powerful speaker* that explores unresolved tensions or aggrieved entitlement and resentment. His *audience is susceptible* in the socio-historical context, mostly out of a sense of *aggrieved entitlement*, and *resentment*. Trump has at his disposal various *media*: TV, newspapers, radio, Twitter, campaign rallies.

He is attributing *guilt*: "they come into our country illegally, we take care of them and not our own","they have no education, they have no jobs" (i.e., they are dumb and lazy and that's their fault). He is constructing immigrants and poor black people as a *threat*: "you get shot walking to the store". He is *destroying alternatives* – policies that promote taking care of illegal immigrants are bad because we are taking care of undeserving people, while we don't take care of "our military veterans", or policemen and women, we need justice".[2] Trump is also

---

[2]Trump has repeated this message at the end of May 2018 in a rally in Nashville: "Chuck and Nancy don't want

engaging in *Virtuetalk*: "we have the greatest people", "law and order." Finally, his campaign slogan is essentially a display of *future-bias*: Make America Great Again.

Code words and meaning perversions differ. *Code words* induce the acceptance into the motivational common ground of evaluative dispositions, plans, and norms, that *conflict* with pre-accepted dispositions/plans/norms that are part of a shared conversational score. But the code word is used to talk about the *proper referent* of the word.

In contrast, the normative and conative connotation of a word that is *perverted* is part of the conversational record. Hence, speech that uses those words is easily accepted in a context. However, that positive connotation is pragmatically in conflict with the use of the word to talk about an *improper referent*. This corrodes e.g. the positive connotation of democratic elections. This *pragmatic conflict* produced by meaning perversions leaves people, as Masha Gessen said, with an impoverished experience, one where there's no coherent shared practical reality, and where power abuses are difficult to resist.

BS strategies of speakers who use code words or meaning perversions also differ. Code words allow *plausible deniability*: "I wasn't saying anything about race!". Meaning perversions allow *undermining norm enforcement*: "How can you be against our freedom? How can you oppose democracy?"

4.3.4   Consequences for Ameliorative Projects

Ameliorative projects, understood as redefinitions made for the purpose of advancing ethics and justice face important problems:

1. If meanings are externally determined by their socio-historical contexts broadly understood, our revisionary intentions as theorists or activists may fail to produce new meanings.

2. If meanings are determined by use dispositions, and these can be so resilient to change, what effective ameliorative projects can we engage in?

3. Through this paper, I hope to contribute to identify guidelines and constraints for legitimate ameliorative projects.

4. I also try to delineate what meaning perversions are, in term of their harmful perlocutionary effects, and their constitutive properties, and their differences from close linguistic phenomena like uses of code words.

References

Andreasen, R. O. (2000). Race: Biological reality or social construct? *Philosophy of Science 67*(3), 666.

Andreasen, R. O. (2005). The meaning of 'race'. *Journal of Philosophy 102*(2), 94–106.

Appiah, K. A. (2006). How to decide if races exist. *Proceedings of the Aristotelian Society 106*(3), 363–380.

---

the wall; they're more interested in taking care of criminals than in taking care of *you*."

Baker, K. J. (2016). Nice, decent folks. *Cold Takes 11-17-2016*, n/a. http://www.kellyjbaker.com/nice-decent-folks/#.

Bettcher, T. M. (2013). "trans women and the meaning of 'woman'. In A. Soble, N. Power, and R. Halwani (Eds.), *Philosophy of Sex: Contemporary Readings, Sixth Edition*, pp. 233–250. Rowan & Littlefield.

Bettcher, T. M. (2014). Selves have sex: What the phenomenology of trans sexuality can teach us about sexual orientation. *Journal of Homosexuality 61*(5), 605–620.

Bicchieri, C. (2017). *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. Oxford University Press USA.

Burgess, J. P. (2014). Madagascar revisited. *Analysis 74*(2), 195–201.

Diaz-Leon, E. (2015). What is social construction? *European Journal of Philosophy 23*(4), 1137–1152.

Díaz León, E. (2016). Woman as a politically significant term: A solution to the puzzle. *Hypatia 31*(2), 245–258.

Figes, O. (2002). *The Whisperers: Private Life in Stalin's Russia*. London: Penguin Books.

Galinsky, A., C. Wang, J. Whitson, E. Anicich, K. Hugenberg, and G. Bodenhausen (2013). Reappropriation of stigmatizing labels: The reciprocal relationship between power and self-labeling. *Psychological Science 24*, 2020–2029.

Gessen, M. (2017). The autocrat's language. *The New York Review of Books May 13, 2017*.

Glasgow, J. (2003). On the new biology of race. *Journal of Philosophy 100*(9), 456–474.

Glasgow, J. (2018). Conceptual revolutions. In T. Marques and Åsa Wikforss (Eds.), *Shifting Concepts: The Philosophy and Psychology of Conceptual Variability*. Oxford: Oxford University Press. Forthcoming.

Haslanger, S. (2003). Social construction: The "debunking" project. In F. Schmitt (Ed.), *Socializing Metaphysics*, pp. 301–325. Oxford: Rowman and Littlefield.

Haslanger, S. (2006). What good are our intuitions? philosophical analysis and social kinds. *Aristotelian Society Supplementary Volume 80*(1), 89–118.

Haslanger, S. (2012). *Resisting Reality: Social Construction and Social Critique*. Oxford University Press.

Jeshion, R. (2016). Slur creation, bigotry formation: the power of expressivism. *Phenomenology and Mind* (11), 130–139.

Klemperer, V. (1957/2000). *The Language of the Third Reich – LTI Lingua Tertii Imperii, a Philologist's Notebook*. London: Continuum.

Langton, R. (2012). Beyond belief: Pragmatics in hate speech and pornography. In M. K. McGowan and I. Maitra (Eds.), *Speech and Harm: Controversies Over Free Speech*, pp. 72. Oxford University Press.

Machery, E., R. Mallon, S. Nichols, and S. Stich (2009). Against arguments from reference. *Philosophy and Phenomenological Research 79*(2), 332–356.

Maynard, J. and S. Benesch (2016). Dangerous speech and dangerous ideology: An integrated model for monitoring and prevention. *Genocide Studies and Prevention: An International Journal 9*(3), 70–95.

Mendelberg, T. (2001). *The Race Card: Campaign Strategy. Implicit Messages, and the Norm of Equality.* New Jersey: Princeton University Press.

Saul, J. (2006). Gender and race. *Aristotelian Society Supplementary Volume 80*(1), 119–143.

Snyder, T. (2017). *On Tyranny: 20 lessons for the 20th century.* New York: Duggan Books.

Stanley, J. (2015). *How Propaganda Works.* New Jersey: Princeton University Press.

Tarski, A. (1943). The semantic conception of truth: And the foundations of semantics. *Philosophy and Phenomenological Research 4*(3), 341–376.

Tirrell, L. (2012). Genocidal language games. In I. Maitra and M. K. McGowan (Eds.), *Speech and Harm: Controversies Over Free Speech*, pp. 174–221. Oxford University Press.

Williams, B. (1979). Internal and external reasons. In R. Harrison (Ed.), *Rational Action*, pp. 101–113. Cambridge University Press.

Yanagizawa, D. (2014). Propaganda and conflict: Theory and evidence from the Rwandan genocide. *The Quarterly Journal of Economics*.