

By Susan Benesch
Senior Fellow, World Policy Institute
benesch@american.edu
(202)885-1611
susanbenesch Skype

January 12, 2012.

Dangerous Speech: A Proposal to Prevent Group Violence

This paper presents draft guidelines for monitoring speech and evaluating its dangerousness, i.e. the capacity to catalyze violence by one group against another. Inflammatory speech is of special interest for atrocity prevention since it tends to rise dramatically before outbreaks of mass violence, suggesting that it may serve as a basis for efforts to prevent such violence, including genocide. However such efforts must not infringe upon freedom of expression, a fundamental right whose exercise can, itself, prevent violence.

There are at least three ways in which inflammatory speech¹ presents an opportunity for violence prevention, and therefore three distinct applications for the guidelines. First, such speech can serve as a key indicator for early warning, since it is often a precursor - if not also a prerequisite - for mass violence.² Second, it may be possible to limit violence by finding ways to limit such speech or its dangerousness. Third, speakers may be held accountable for speech that constitutes crime.

1. The Importance of Dangerous Speech

The guidelines are designed to identify a subset of hate speech, which I have termed Dangerous Speech. Hate speech is variously defined in law and in common parlance, but is generally understood to mean speech that denigrates people on the basis of their membership in a group, such as an ethnic or religious group. This category of speech is too broad for successful early warning of mass atrocities, for two related reasons. First, hate speech is common in many societies, unfortunately, including those at minimal risk of genocide. Second, some hate speech does not appreciably increase the risk of mass violence, although it may cause serious emotional and psychological damage.

In other words, speech can harm directly or indirectly, or both. It may directly offend, denigrate, humiliate or frighten the people it purports to describe – such as when a racist shouts at a person of color. Speech can also bring about harm indirectly – and with equal or even greater brutality – by motivating others to think and act against members of the group in question.

¹ 'Speech' includes any form of expression, including images such as drawings or photographs, films, etc.

² It has long been widely assumed that inflammatory speech can catalyze and magnify mass violence, but the inference was unsupported by data until just recently. In what is apparently the first quantitative evidence of a link between speech and mass violence, David Yanagizawa's 2009 statistical study of the effects of the virulent propaganda outlet Radio Television Libre des Milles Collines (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 exogenous reasons such as topography) receive the signal. See David Yanagizawa, Propaganda and Conflict: Theory and Evidence from the Rwandan Genocide, available at <http://people.su.se/~daya0852/Rwanda_jmp.pdf>

When an act of speech has a reasonable chance of catalyzing or amplifying violence by one group against another, given the circumstances in which it was made or disseminated, it is Dangerous Speech.³ To my knowledge there is no other framework or methodology designed to identify this type of speech, either in the literature or in use in the field.⁴

2. Five Defining Variables

The guidelines are based on five variables which affect the dangerousness of a particular speech act in the time and place in which it was made or disseminated: the speaker, the audience, the speech act itself, the social and historical context, and the mode of dissemination.

The most dangerous speech act, or ideal type of dangerous speech, would be one for which all five variables are maximized:

1. a powerful speaker with a high degree of influence over the audience,
2. the audience has grievances and fear that the speaker can cultivate
3. a speech act that is clearly understood as a call to violence
4. a social or historical context that is propitious for violence, for any of a variety of reasons, including longstanding competition between groups for resources, lack of efforts to solve grievances, or previous episodes of violence
5. a means of dissemination that is influential in itself, for example because it is the sole or primary source of news for the relevant audience

Each speech act should be evaluated in terms of the information available, regarding the full set of variables. As in the case of the OSAPG's own Analysis Framework for gauging the risk of genocide, the variables are not ranked. Nor are they weighted equally across cases: in many circumstances, one or more variables will 'weigh' more than others. For example an especially outrageous or frightening speech act may be more important (i.e. more dangerous) than other factors in a particular instance. In other cases, an audience may be especially vulnerable to incitement, or the overweening influence of a speaker may contribute most to the force of his or her speech.

³ In law, speech intended to provoke another to commit a crime is incitement, or instigation. Dangerous Speech can also be understood, therefore, as incitement to mass violence, which is reasonably likely to be successful. We do not limit "mass violence" to genocide, since inflammatory speech is a precursor to outbreaks of mass violence including, but not limited to, genocide.

⁴ For further information and related research, see Susan Benesch, Vile Crime or Inalienable Right; Defining Incitement to Genocide, *Va J. Int'l L.*, 48 *Va. J. Int'l. L.* 485 (2008), available at <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1121926> and Susan Benesch, The New Law of Incitement: A Critique and a Proposal, available at <http://www.ushmm.org/genocide/spv/pdf/benesch_susan.pdf> and www.voicesthanpoison.wordpress.com For an address on the topic, given as a side event to the UNGA in October 2010, see <<http://www.unmultimedia.org/tv/webcast/2010/10/special-event-dialogue-on-dangerous-speech-on-the-road-to-genocide.html>>

Analysis may be conducted with varying degrees of information: the more information available, the more fine-grained the analysis will be. Since speech can only be well understood (and its influence gauged) in the psycho-social, historical, and cultural context in which it was made or disseminated, the analysis is best carried out with knowledge of the relevant language, culture, and social conditions – or with assistance from advisors who have such knowledge.

3. Guidelines for Evaluative Monitoring of Inflammatory Speech

The persuasiveness or force of speech has long been studied, albeit usually not in the context of mass violence. The first three of our variables were identified in a different form by Aristotle, who observed, in his *Rhetoric*, that the means of persuasion reside in the character of the speaker, the emotional state of the audience, and in the argument or *logos*, itself. Also searching for the keys to persuasiveness in their pioneering studies of propaganda in the aftermath of World War II, Carl Hovland and other social psychologists at Yale University focused on the same three variables to construct the so-called Yale attitude change model.

To evaluate a speech act's capacity to move one group to attack another – the capacity not only to persuade but to inspire action – in the contemporary world, I have added two key variables to the first three: the historical and social context in which a speech act is made or disseminated, and the medium or means of dissemination. Social context is an essential variable, since mass violence is unlikely in certain contexts, irrespective of other factors such as the intent and rhetorical skill of the speaker, and more likely in others. Finally, in the present day, the means of dissemination (limited in Aristotle's time to speech delivered in person) are increasingly varied, and can be influential in and of themselves. A message may be more compelling because it is delivered over a particular radio station, for example, or via SMS.

Any number of factors may affect the dangerousness of speech, but the most relevant factors are all related to the five variables that I have identified, based on fieldwork and research in literature including genocide studies, social psychology, and the philosophy of language. It is therefore possible to conduct a relatively streamlined analysis, taking each of the variables in turn. I have listed the variables together with exemplary factors presented as questions, since those would be most easily incorporated into a monitoring plan, and some explanatory notes.

1. THE SPEAKER

It should be noted that some cases present more than one speaker, such as when a radio talk-show host interviews a public figure.

- a. **Did the speaker have authority, power, or influence over the audience?** Influence or authority need not derive from a *de jure* political post; cultural and religious figures and entertainers often have more influence over an audience than political officials. Some speakers control resources needed by an audience, or can deploy force. Any of these factors can render their speech more dangerous.
- b. **Was the speaker charismatic or popular?** A speaker may be seen as popular only by a subset of the audience, but those listeners may also be most likely to react to the speech.

2. THE AUDIENCE

- a. **Who was the audience most likely to react to the speech at issue?** In many cases, the audience may be large or somewhat indeterminate, e.g. “the public” or all listeners of a radio station. The analysis should focus on the audience that is most likely to react with violence in response to the speech.
- b. **Was the speech directed primarily at members of the group it purported to describe, i.e. victims, or at members of the speaker’s own group, or both?** If the latter is true, the speech is more likely to be dangerous. The primary audience is often indicated by the language or venue in which the speech was delivered.
- c. **Did the audience have the means or capacity to commit violence against the group targeted in the speech?** If an audience is unable to commit mass violence, incitement cannot succeed, and is not Dangerous Speech.
- d. **Was the audience suffering economic insecurity, e.g. lacking in food, shelter, employment, especially in comparison with its recent past?**⁵
- e. **Is the audience characterized by excessive respect for authority?** This would make an audience more vulnerable to incitement.
- f. **Was the audience fearful?** Fear might be objectively reasonable or not; its impact may be equally large, and equally well exploited by a compelling speaker.

3. THE SPEECH ACT

Speech may be made in any number of forms, and disseminated by myriad means: a shouted command, a song broadcast at a rally, a newspaper editorial, a tweet, poster, web page, SMS blast, leaflet, film or photograph.

- a. **Was the speech understood by the audience as a call to violence?** Inflammatory speech is often expressed in elliptical, indirect language, which can be variously interpreted. For this analysis, the only relevant meaning is the way in which the speech was understood by the audience most likely to react, at the time when it was made or disseminated.
- b. **Did the speech describe the victims-to-be as other than human, e.g. as vermin, pests, insects or animals?** This is a rhetorical hallmark⁶ of incitement to genocide, and to violence, since it dehumanizes the victim or victims to be.

⁵ This factor, like the next two, is emphasized by the genocide scholar Ervin Staub, who has carried out extensive research into the causes of mass violence. See, e.g. *Overcoming Evil: Genocide, Violent Conflict, and Terrorism*, p. 208 *et seq.*

⁶ These hallmarks are discussed at greater length in *Vile Crime or Inalienable Right*, *supra* note 3.

- c. **Did the speech assert that the audience faced serious danger from the victim group?** Another hallmark of incitement, this technique is known as “accusation in a mirror.” Just as self-defense is an ironclad defense to murder, collective self-defense gives a psychological justification for group violence, even if the claim of self-defense is spurious.
- d. **Did the speech contain phrases, words, or coded language that has taken on a special loaded meaning, in the understanding of the speaker and audience?** Such coded language is typical of Dangerous Speech. It bonds the speaker and audience more tightly together. Familiar examples of this are the phrase “go to work,” used as code for killing during the Rwandan genocide, or the word “inyenzi” (Kinyarwanda for “cockroach”), used to refer to Tutsi or even to non-Tutsi who sympathized with Tutsi.

4. SOCIO-HISTORICAL CONTEXT

- a. **Were there underlying or previous conflicts between relevant groups?**
- b. **Were there recent outbreaks of violence following other examples of hate speech?** This would put both speaker and audience on notice that such speech can indeed lead to violence, thereby increasing the dangerousness of the speech.
- c. **Were other risk factors for mass violence present,** such as weak democratic structures and rule of law, and structural inequalities and discrimination against a group or groups? These risk factors have been described in greater detail elsewhere, e.g. in the OSAPG’s Analysis Framework.

5. MODE OF TRANSMISSION

- a. **Was the speech transmitted in a way that would reinforce its capacity to persuade, e.g. via a media outlet with particular influence or without competitors?** Other modes of transmission can be compelling in and of themselves, such as new media platforms that make members of an audience feel that they are part of a select and privileged group. Music can also increase the force or influence of a message.
- b. **Was the audience exposed to, or did it have access to, alternate views or sources of information?** Where there are no alternative sources of information, the impact of speech is much greater.
- c. **Was the speech frequently repeated, in similar form or content?** Repetition magnifies the impact of a message.