This file is a record of ongoing thinking about how to experimentally investigate dogwhistle content.

A compiled PDF version can be found here.

Guidelines for this document

1. Let's cite the relevant papers using LaTeX and the BibTeX keys in our bibliography.
2. To compile this document using Pandoc and LaTeX open a command line shell and enter `make`. This requires that the Pandoc utility and LaTeX are installed.

Goals

First phase of SFB

- Ultimately, we aim to develop quantifiable criteria.
- Better understanding of discriminatory speech.

Second phase of SFB

- May also inform research on de-biasing AI language models. Current approaches seem simplistic and will let more subtle expressions slip through. Potentially an interesting connection to work in computational linguistics ⇒ idea for second stage of the SFB. (Inspired by Anne Lauscher talk at IMS.)

Examples of dog whistles

wonder-working power: Henderson/McCready 2019

big pharma: Henderson/McCready 2019

inner city: Henderson/McCready 2019, Stanley 2015

welfare: Henderson/McCready 2019, Stanley 2015

Definitions of dog whistle content

Merriam Webster

Definition according to Merriam Webster:

> An expression or statement that has a secondary meaning intended to be understood only by a particular group of people.

Stanley 2015 "How propaganda works", chapter 4

p.94: The third norma- tive ideal for public reason is that it is guided by equal respect for the perspective of everyone subject to the policy under de- bate. Following the recent political philosophy tradition, we shall call this the norm of reasonableness.

p.126: One moral of the previous chapter is that demagoguery in a liberal democracy takes the form of a contribution to public debate that is presented as embodying reasonableness yet in fact contributes a content that clearly erodes reasonableness. This form of propaganda is not merely a deceitful attempt to bypass theoretical rationality, on this view. It functions via an initial selection of a target within the population.

p.129: We should expect there to be linguistic means by use of which one can make an apparently reasonable claim, while simultaneously, merely by using the relevant vocabulary, wearing down the ideal of reasonableness.

p.138: One kind of linguistic propaganda involves repeated asso- ciation between words and social meanings. Repeated associ- ation is also the mechanism by which conventional meaning is formed; it is because people use "dogs" to refer to dogs, re- peatedly, that "dogs" comes to refer to dogs. My claim in this chapter is that when propagandists use repeated association between words and images, they are forming connections that serve as the basis of conventional meaning.

p.138: When the news media connects images of urban Blacks repeat- edly with mentions of the term "welfare," the term "wel- fare" comes to have the not-at-issue content that Blacks are lazy. At some point, the repeated associations are part of the mean- ing, the not-at-issue content. ... This does not mean that someone hearing the term "welfare" automatically comes to believe that Blacks are lazy. It does mean that they may have to shift to different vo- cabulary, or consciously resist the effects of the association, in conversation or otherwise, to deter the propagandistic effect.

p.139: The fact that the not-at-issue content cannot be can- celed is what makes it so effective.

p.140: According to the content model, one kind of paradigmatic propaganda in a liberal democracy would have a normal at-issue content that seems reasonable, and would also have a not-at-issue content that is not reasonable. Here is another way of thinking of the mechanism by which a contribution could lead to an erosion of empathy for a group. The contribution could express a perfectly ordinary at-issue content, but cause a decrease in empathy or respect

directly, as part of its not-at-issue function. The idea here is not, as on the content model of propaganda, that there is a not-at-issue content, acceptance of which decreases empathy for a group. It is rather that words have direct not-at-issue emo- tional effects. Let us call this the expressive model of propaganda. According to the expressive model, one kind of paradigmatic propaganda in a liberal democracy would have a normal at- issue content that seems reasonable, and would also have a not-at-issue effect that would decrease empathy for a group. Since decreasing empathy for a group runs counter to reason- ability, its not-at-issue effects would be unreasonable.

p.151f: not only politics but also ev- eryday discourse involve apparently innocent words that have the feature of slurs, namely, that whenever the words occur in a sentence, they convey the problematic content. The word "welfare," in the American context, is not on any list of prohib- ited words. Yet the word "welfare" always conveys a problem- atic social meaning, whenever it is used. A sentence like "John believes that Bill is on welfare" still communicates a problem- atic social meaning.

p.154: Mendelberg and her Princeton colleague Martin Gilens have both studied the effects of the use of the term "welfare" on political opinions. They have discovered that the use of the term "welfare" leads to a priming of white racial bias. In other words, the mere use of "welfare," and presumably also "food stamps," as well as some other expressions referencing social spending programs, primes racial bias against Blacks. A con- clusion from this research is that "any allusion to a racially tinged issue like welfare may racialize a campaign, even if it alludes to white recipients."30 Most interestingly for the topic of slurs, Mendelberg, via a compelling experiment with non- students in Michigan, shows that the racial-bias effects actu- ally decrease if a candidate's message is made explicitly racial in character.

p.156: On the picture I am sketching, certain words are imbued, by a mechanism of repeated association, with problematic im- ages or stereotypes. One can use these words to express ordi- nary contents, and explicitly deny complicity with the associ- ated problematic image or stereotype. ... Gingrich was allowed to act responsible just for the at-issue content of his utterance, and feign ignorance of the racial overtones of the expressions.

p.157f: Propaganda character- istically involves attaching problematic social meanings to seemingly innocuous words that are used to describe policy, in effect making the word "welfare" like the word "prostitution." The social meanings of these words are not-at-issue content. Because they are not-at-issue contents, they are "not negotia- ble, not directly challengeable, and are added [to the common ground]

3

even if the at-issue proposition is rejected." In short, even evaluating the proposal means that one must accept the social meaning. It is odd to challenge the social meaning; the social meaning associated with a word is accepted even if the claim made with the associated word is rejected.

p.165: When President Obama is described as being Muslim, the not-at-issue content, or social meaning, of the use of the term "Muslim" is of course related to terrorism, or some kind of anti-American sentiment. This is an attempt to challenge the president's sincerity,

Henderson and McCready (2019, p. 2)

language that sends one message to an outgroup while at the same time sending a second (often taboo, controversial, or inflammatory) message to an ingroup

But even people in the outgroup may recognize the message.

In broad strokes, we make the novel proposal that dogwhistles come in two types. The first concerns covert signals that the speaker has a certain persona. . . The second involves sending a message with an enriched meaning whose recovery is contingent on recognizing the speaker's covertly signalled persona

The use of dogwhistles is prompted by a desire to 'veil' a bit of content, but still to convey it in some manner. Deniability is essential. If a bit of content is conventional, it's not deniable any longer.

What do you have against...?'' diagnostic in (7) and (8); problem: does (8) give rise to DW content? Also:It's not cool to say Z'' – "I didn't say Z'' diagnostic on p.4.

Dogwhistles, by definition, are not needed when talking to an ingroup and can be disposed of, which wouldn't make sense if the subtext of dogwhistle were part of its conventional meaning for the in-group.

(i) dogwhistles are not part of conventional content, so speakers are able to avoid (complete) responsibility for what they convey, (ii) dogwhistles are semi-cooperative—that is, they are meant to convey part of their meaning to to only one segment of the audience while hiding it from the rest of the audience, and (iii) while deniable, dogwhistles are risky. Being detected using a dogwhistle by the wrong party should cost the speaker in some way.

knowledge about social personae can play a role in recovering intended meanings

4

a broader continuum of cases in which the rational use of language is utilized or manipulated by speakers for reasons of strategy, efficiency, or style

Dogwhistles share with conversational implicatures the property of being can- cellable (deniable), but differ from standard views of them in not following from anything but an extremely nonstandard construal of the Gricean maxims Grice 1975.

their interpretation arises from background assumptions about social meaning and how personae are linguistically expressed makes them quite different from ordinary implicatures. They are simultaneously conventional and socially dependent. In this sense, dogwhistles seem to occupy a genuinely new niche in the characterization of not-at-issue meaning.

There is a sense in which dogwhistles are an ubiquitous phenomenon: much communication involves underspecified meanings which can in part be resolved by learning more about the speaker and her intentions. Informa- tion about social categories often informs how such underspecification is resolved, but possibly in quite different ways in different contexts; this area is also a rich and complex one ripe for investigation.

Khoo (2017, footnote)

footnote 1: Code words are sometimes called "dog whistles" because (so the received view goes) they involve sending a message that can only be heard by audience members with suitably sensitive ears. Richard Morin is credited with introducing the term. He writes, "Subtle changes in question-wording sometimes produce remarkably different results . . . researchers call this the 'Dog Whistle Effect': Respondents hear something in the question that researchers do not." (Morin 1998).

p.33: Recently, a more subversive phenomenon has been noted and discussed in the social sciences— the use of code words to subvert norms of democratic deliberation.

p.34: The standard thought is that politicians use code words because they stand to gain from conveying certain messages implicitly, rather than stating them explic- itly.

p.34f: Both Mendelberg and Stanley propose varia- tions on the idea that code words encode as part of their meaning a hidden, or implicit, message, which the user of a code word communicates, along with some other explicit message. This idea keeps with the received view of dog whistles as devices of communicating secret messages to the suitably trained ear.

p.35: I think this kind of view is on the wrong track: what we identify as "code words" do not encode additional, implicit, meanings. Put oxymoronically, code words are not code for anything. Instead, someone using a code word exploits (intentionally or otherwise) their audience's stereotypical beliefs about what they are talking about, without explicitly communicating these beliefs. Thus, using a "code word" allows (or leads) the audience to draw additional inferences from the speech without it being clear that they are doing so—and this is what dis- tinguishes coded speech from speech where the relevant stereotypical beliefs are explicitly asserted. In a slogan: code words don't work by being vehicles of implicit communication; they work by triggering inferences which they are not used to communicate.

p.47: code words carry no "implicit" meaning at all. This theory offers the most straightforward account of why code words afford plausible deniability as discussed above.

p.48: though he does not communicate anything about race, since many hearers have preexisting beliefs about the subject matter of what he does say, Politician Z's speech has the result that hearers will infer some racial belief about the policy from what he says.

p.48: Notice here that the politician may or may not have conver- sationally implicated that the food stamp program will primarily benefit poor African Americans— whether he does will depend on whether he intended his hearers to come to infer (C) on the basis of his saying (A). ... It seems then that nothing implicit need be communicated by a code word for a use of it to create space for plausible deniability regarding the violation of the norm of racial equality

p.48: I should emphasize that the beliefs which drive these inferences need not be about the subject matter of the word—they could be about the word itself. Thus, in the case of "wonder-working power," a hearer who recognizes the expression might come to believe that the speaker uttering it is a Christian.

p.50: Explicit Statement: x is C. Existing Belief: If something is C, then it is R. Inferred: x is R.

This schema is too narrow. The existing belief can be not just about x but also about the speaker or the topic matter at hand.

Mendelberg (2001)

(from Khoo 2017) Tali Mendelberg proposes that using a code word allows the speaker and her audience to violate certain social norms while plausibly denying doing so.

This does not involve in/outgroup distinction, highlights plausible deniability.

Albertson (2015)

> multivocal appeals, meaning appeals that have distinct meanings to different audiences

Here the ingroup/outgroup distinction is avoided, "distinct" meanings is pretty vague, "different audiences" is also pretty vague (no requirement that the additional meaning is taboo, controversial, or inflammatory)

Wetts and Willer (2019)

> messages. . . in which race is not explicitly mentioned but instead is cued through coded language or accompanying visuals subtly connect racial prejudice to whites' views of policies and candidates, a process commonly referred to as "dog whistling" (Haney-López 2014; Mendelberg 2001; Valentino, Hutchings, and White 2002).

Dog whistles limited to racial prejudice? Controversial message is component of this definition, no in/outgroup, controversial message is not explicit but subtly coded.

Calfano and Djupe (2009)

> These cues, or what we term "the code," signal the in-group status of a GOP candidate to white evangelical voters. However, because the cues are so specific to evangelical culture, they are intended to pass unnoticed by other voters and therefore allow GOP candidates to avoid broadcasting very conservative issue positions that might alienate more moderate voters. Thus, the code is a highly sophisticated communication strategy that is designed to appeal to an in-group without rousing an out-group's suspicions.

In/outgroup difference is made, cues can be recognized by outgroup, signal belonging to ingroup without being explicit.

Our thinking

Dog whistle is an expression that is used by one party but not by another (lower bar, necessary condition, but perhaps not sufficient). The phenomenon itself may be probabilistic, rather than categorical. Perhaps this should be part of an improved definition of the term. It may also be that the term/message isn't neutral, but rather may also send a message for everybody to different degrees, and that message, when integrated with the person's ideology, results in different effects.

We're more sensitive to dog whistles that are outgroup!

Different aspect potentially contributing to a definition: Dog whistles are not necessarily expressions that carries secondary meaning but a strategy where a semantic/pragmatic gap is artificially created and to be filled by the listener using their social and contextual knowledge. For instance, when Gauland says "people don't want a Boateng as their neighbor", it's technically not clear what he means by "a Boateng". Does he mean a soccer player or a black person? But using social knowledge about the speaker, the audience can fill that missing meaning in. Compare to some soccer fan saying "people would love to have a Boateng as their neighbor", this would potentially be interpreted as meaning "a famous soccer player". In that way, dog whistles may be very similar to the rhetorical device where the speaker does not finish a controversial sentence because they know that the audience will be able to complete it in their minds without the speaker actually having to say the words: "If we let more migrants into the country, next new year's eve in Cologne will ..." Another closely related rhetorical device is where the offensive part of the statement is replaced by something that's effectively the opposite of the intended meaning and absurd given social knowledge about the speaker: "If we let more migrants in the country, next new year's eve in Cologne is going to be fun." What dog whistles offer in contrast to these rhetorical device is plausible deniability, because with dog whistles the (to-be-filled) gap meaning is not as obvious. Seen in this way dog whistles are the result of a natural evolution of ways to circumvent taboo speech. If true, this would explain why dog whistles are so difficult to characterize: The properties of the dog whistle expression are irrelevant. What matters is merely whether that expression leaves a meaning vacuum when used in a certain context. Corollary: "Clankriminalität" is a dog whistle in the presence of a meaning vacuum, but the same expression immediately looses its dog whistle property in the absence of such a vacuum, e.g. when the text is otherwise openly xenophobic. Question: How can we test for the presence of a meaning vacuum/gap?

Function(s) of dog whistles

One function of dog whistles is to communicate controversial ideas to ingroup members while having plausible deniability. Here's a related but different function: Anna Prysłopska made me aware of the wonderful world of alt-right mom blogs. Superficially, these blogs are about benign topics like childcare and cooking, but the true aim is to sneak ideological content into the readers' minds. (One author of such a blog, Mrs. Mid-west, admitted this openly.) Cooking recipes and parenting advice are just the honeypot. These blogs, according to Anna, are chock-full with dog whistle material. However, the purpose of dog whistles is not to hide controversial views from an outgroup but rather to accustom readers to these views in a subliminal way and to assimilate them ideologically. These dog whistles do not just work with pre-defined in/outgroups, but they are used to redefine these groups (= propaganda). The persuasion component is actually present in many of the examples that we considered, e.g., Gauland saying "einen Boateng". In sum, the core property, once again, seems to be plausible deniability. And the functions of dog whistles are: 1. Covert

communication with an ingroup. 2. Persuade outgroup members to become ingroup.

"Besonderheit" of dog whistles (Lukau)

Lukau wrote "dog whistles sind eine sprachwissenschaftliche Besonderheit". She's right and that's actually an interesting observation. Dog whistles are special in that their scope is very limited to political speech and perhaps even to some parts of the political spectrum (not sure). But why would that be? If dog whistles are such a clever communication strategy, they should be more common. They should even be taught in school. None of our considerations so far can explain why dog whistles are so special and limited. I think there's potentially an important aspect here that we (and perhaps others) have overlooked so far. Are dog whistles cognitively difficult? Are there social norms that dog whistles are violating?

Scope of dog whistles really limited to political speech?

Dog whistles are usually portrayed as a phenomenon occurring only in political speech. This is actually strange because if dog whistles are an effective communication strategy, they should also have uses in other areas. Raises the question: What are other areas where dog whistles may be used? And, if we don't find them in those areas, why are they not used there? One candidate may be advertising. In advertisements, companies are looking to convey potentially controversial ideas (e.g., some processed food is healthy when it's not) and where companies need plausible deniability because they might otherwise get sued.

Influence of discourse dynamics (in lack of a better term) / polarization of society

One of Lukau's examples is Jill Stein's use of the term "big pharma".

- "By the same token, being 'tested' and 'reviewed' by agencies tied to big pharma and the chemical industry, who sell unsafe vaccines to make a buck, is also problematic."

The claim is that "big pharma" signals anti-vaxx views. Personally, I didn't get that inference. "Big pharma" is certainly meant to be negative, but anti-vaxx? The fact that pharmaceutical companies are profit-oriented does not exclude the possibility that vaccines are effective. I think the US tendency to polarize may play a role here. Implicit criticism of the pharmaceutical industry is perhaps interpreted as a complete rejection of the pharmaceutical industry and its products. In this light, I wonder whether the peculiar discourse dynamics in the U.S. and the tendency to polarize may be a particularly fertile ground for dog whistles. Is this already discussed somewhere? Could be useful in motivating the investigation of dog whistles in German (where society and politics are not as polarized as in the U.S.).

Attempt at a taxonomy

Factors that distinguish types of dog whistles (incomplete list): - Whether the dog whistles content is central or ancillary. In the first case, the utterance doesn't make complete sense without the implied dog whistles meaning ("ein Boateng"). In the second case, the utterance is still discourse coherent without it ("wonder working power"). - Whether the goal is covert communication or persuasion. - Whether the dog whistle expression is conventionalized ("let's go Brandon", chiffre, requires that parties negotiate code beforehand) or ad-hoc ("einen Boateng", no prior negotiation of code necessary, addressee can infer meaning just using cultural common ground).

Related phenomena

Virtue signaling

E.g. gender neutral language: Virtue signaling and dog whistles are related in that they are both ways to convey social meaning. But they are different in that the social meaning is covert with dog whistles whereas virtue signaling is quite overt. One defining property of dog whistles seems to be that they provide plausible deniability, which makes them useful for subliminal persuasion. Virtue signaling like the use of gender neutral language is also a tool for persuasion, but it's not subliminal and offers no plausible deniability. The user of a dog whistle is working around a social taboo but the virtue signaler doesn't need to do that.

Taboos, sanctioned language, and degrees of explicitness:

Seen on Twitter in response to a Tweet by Elon Musk asking for more free speech on Twitter:

> Do you really think there should be NO limits on speech? Like, would *all* of these be acceptable on your platform? You wouldn't draw the line ANYWHERE? 1. Jews suck. 2. Jews deserve to die. 3. Jews should be murdered. 4. Jews in Cleveland should be murdered. 5. We should murder Jews in Cleveland. 6. We should murder Jews in Cleveland next Thusday at 3:00. 7. We should meet next Thursday at 3:00 at the Fountain of Eternal Life in Cleveland and go around murdering Jews. Bring a gun.

Code words (Chiffren) vs. steganography

Lukau distinguishes code words from dog whistles. Main difference: a code word requires that the parties agree on it in advance. Dog whistles seem to be more dynamic. Actually, as soon as a dog whistle is codified, it's no longer effective because it no linger provides plausible deniability. Code words are also overtly encrypted. When I write "88" on my t-Shirt, it's clear that there is a hidden meaning ("Heil Hitler"), but a dog whistle is more like steganography in that it's not even clear that a hidden message is present.

Tasks used in previous empirical research on dog whistles

Albertson (2015): Religious dog whistle "wonder-working power"; three conditions (dog whistle, obvious religious content+dog whistle, neither); participants viewed videos w speeches by candidate; response tasks: attitude toward candidate and whether they would vote for candidate (5-point Likert scale), emotion ratings regarding candidate (interested, upset, proud, moral...) on sliding scale, participants' religious denomination and beliefs about the Bible; outgroup was unfamiliar with religious connotation of "wonder-working power", they gave less favorable ratings with obvious religious appeal but not with WWP; for ingroup, both WWP and the obvious religious appeal lead to more favorable ratings of the candidate, the former resulted in even more favorable ratings.

White (2007): Conditions in experiment were explicitly racial ("poor, Black and Hispanic families"), implicitly racial ("poverty, ...layoffs, ...healthcare...affordable housing"), and non-racial; participants read articles (fictitious debate between two Ohio congressmen——one Democrat and one Republican——on the proposed war in Iraq. All of the experimental conditions presented two-sided messages about the war: the Republican expressed a consistent statement in support of using force to remove Saddam Hussein from power, and the Democrat voiced opposition. What changed across the conditions were the terms in which the Democrat articulated his argument against the war); answered an extensive battery of post-test questions that included questions about their support for the war in Iraq, issue importance ratings, and racial and political attitudes; results: COMPLICATED DESIGN, NEED TO SPEND MORE TIME READING THIS

Calfano and Djupe (2009): Participants were asked a series of questions concerning their demographics, political interests, political knowledge, and political attitudes; they were exposed to a picture and brief biographical sketch of a male candidate that included basic personal information (including age, family status, career background); candidates varied by race (white/black) and statement (land, worth, WWP, none); participants were asked to identify the candidate's party (all candidates were male) and whether they would vote for him; results: white evangelicals identify candidates using the code as Republicans and indicate their likely support for such candidates. Roman Catholics and mainline Protestants are generally not affected by the code, although mainline Protestants do link candidates using the code to the GOP. Instead, support for such candidates among mainline Protestants is driven largely by party affiliation and other concerns, not by the code itself. This is almost precisely the way the code is intended to work—to gain the recognition and support of the in-group evangelicals and pass unnoticed by others, or at least not engender opposition from other likely supporters

Wetts and Willer (2019): Participants assigned to the implicit racial appeal condition reported significantly less support for welfare than those assigned to

the no racial appeal condition (p = .02). Participants in the explicit racial appeal condition were in between, not differing significantly from either the control condition or the implicit racial appeal condition; participants assigned to the implicit racial appeal condition reported significantly lower levels of the view that welfare recipients are deserving of aid (p = .02). Again, participants in the explicit racial appeal did not report significantly higher levels of this view than participants in either the control or implicit racial appeal conditions. After reading an anti-welfare message where racial cues were subtly implied, liberals high in racial resentment expressed less support for welfare programs than those who read an excerpt about global warming, while liberals low in racial resentment voiced approximately equivalent support across condition (Figure 1a). In contrast, moderates and conservatives high in racial resentment showed little effect of the implicit racial appeal condition on welfare support relative to the control condition. THIS STUDY DOES NOT ASSUME THAT DOGWHISTLES ARE A MESSAGE THAT ONLY SOME PERCEIVE.

References from Wetts & Willer 2019: Welfare (Mendelberg 1997, 2001; Peffley, Hurwitz, and Sniderman 1997), food stamps (White 2007), government spending (Valentino et al. 2002), crime (Domke 2001; Gilliam and Iyengar 2000; Hurwitz and Peffley 2005; Mendelberg 1997, 2001), and health care (Tesler 2012). Other research, however, has called into question whether implicit racial appeals are more effective than explicit appeals, including two studies employing experimental designs and large, nationally representative samples (Huber and Lapinski 2006; Valentino, Neuner, and Vandenbroek 2018). These studies suggest that racial appeals may no longer need to be implicit to engage white prejudice as explicit references to racial stereotypes may have become more socially acceptable (Valentino et al. 2018). Still other work suggests that implicit appeals may be more effective than explicit or race-neutral appeals in harnessing white prejudice but only among particular demographic groups (Hutchings, Walton, and Benjamin 2010).

Starting point for an empirical investigation of German dog whistle expressions

- No clear definition of dog whistle content/expressions (in any language).
- No empirical investigation exists for German.
- A few expressions have been discussed: e.g., Clankriminalität.

Study to build test scales (from Beltrama & Staum-Casasanto 2017)

- We first conducted a preliminary study to construct the evaluation scales to be used to measure social meaning in the experiment.

- Sixty subjects, who self-declared to be native speakers of American English and between 18 and 35 years old, were recruited and paid $0.50 for participating.
- First, each subject saw a written sentence containing either an instance of lexical totally or speaker-oriented totally. The subjects were asked to provide four adjectives to describe the imagined speaker of the sentence.
- Based on the most frequently recurring adjectives in the responses, a total of eight evaluation dimensions were selected as particularly salient in connection with the use of the intensifier.
- Four of these dimensions, based on the participants' responses, appear to be positively associated with the use of the intensifier. Following the literature on language attitudes (Lambert et al. 1960), we call them 'Solidarity' attributes.
- The other four dimensions express identity categories that feature a negative association with totally. We label them 'Status' attributes.
- The resulting eight scales are summarized as follows:
  - Solidarity: Friendliness, Coolness, Outgoingness, Excitability
  - Status: Articulateness, Maturity, Intelligence, Seriousness

Concrete proposal for first mini-pilot with "Volk":

- Hypothesis: "Volk" is a dogwhistle: It's literal meaning is merely 'people', but it has a conservative connotation, evoking admiration for Nazi Germany; speakers who use it thereby signal alignment with conservative/nazi values.
- Expectation: Participants overall will be more likely to attribute right-wing stances to speakers using "Volk" than to speakers using the alternative formulations (no expectation that only participants with right-wing stances will attribute right-wing stances).
- Target materials:
  - "Kinder sind die Zukunft unseres Volkes." / "Kinder sind die Zukunft unserer Gesellschaft." (We need to specify that speaker is German politician.)
  - "Die Aufnahme in das deutsche Staatsvolk ist an strenge Bedingungen geknüpft" / "Der Erwerb der deutschen Staatsbürgerschaft ist an strenge Bedingungen geknüpft."
- Each participant sees one target with "Volk" and one replacement.
- 2 Controls (everybody sees these two, used to remove data from participants not attending to the task?):
  - "Alle Ausländer sind gewalttätig." Clearly racist, serves as baseline and sanity check.
  - "In Deutschland leben circa 22 Millionen Menschen mit Migrationshintergrund." Clearly harmless / unbiased.
- 6 Fillers (to distract from the 4 items concerning immigration, everybody sees these):
  - "Klimaschutz muss sozial gerecht sein."

- "Jedes Kind hat ein Recht auf gute Kitas und Schulen, egal wo es lebt."
- "Die Digitalisierung ist eine der zentralen Infrastrukturen unseres Lebens."
- "Klimaschutz darf nicht zu Nachteilen im internationalen Wettbewerb führen."
- "Gesundheitsschutz und Pflege brauchen einen größeren Stellenwert."
- "Die Kulturlandschaft soll nach der Pandemie zu neuer Lebendigkeit finden."

Proposal for a pilot experiment to identify German dog whistles

Research question

Which expressions are such that some but not all speakers of German perceive a message that is not part of the conventional content? In this pilot project, we focus broadly on the topic of migration/foreigners. Thus, the research question more specifically is: Which expressions are such that some speakers of German perceive that an expression conveys that the speaker has a racist attitude towards migrants / foreigners / refugees / asylum seekers?

Participants

- Recruit native speakers of German living in Germany on Prolific.
- How many?
- Questionnaire to identify relevant aspects of their socio-political profile, present at the end of the experiment.
    - Age
    - Gender: männlich, weiblich, divers, möchte ich nicht antworten
    - Höchster Bildungsabschluss: kein Schulabschluss, Hauptschule, Realschule, Gymnasium, Bachelor, Master, Promotion
    - Wenn am Sonntag Bundestagswahl wäre, welche Partei würden Sie wählen?
    - Relevante Wahl-O-Mat Fragen (stimme zu, neutral, stimme nicht zu):
        1. Das Recht anerkannter Flüchtlinge auf Familiennachzug soll abgeschafft werden.
           Stimme zu: AfD
           Neutral: FDP
           Stimme nicht zu: CDU/CSU, SPD, Links, Grün
        2. Das Tragen eines Kopftuchs soll Beamtinnen im Dienst generell erlaubt sein.
           Stimme zu: Links, Grün
           Neutral: FDP
           Stimme nicht zu: CDU/CSU, SPD, AfD,

3. In Deutschland soll es generell möglich sein, neben der deutschen
eine zweite Staatsbürgerschaft zu haben.
Stimme zu: SPD, Links, Grün
Neutral: FDP
Stimme nicht zu: CDU/CSU, AfD
4. Islamische Verbände sollen als Religionsgemeinschaften staatlich
anerkannt werden können.
Stimme zu: SPD, Links, Grün, FDP
Neutral: CDU/CSU
Stimme nicht zu: AfD
5. Asyl soll weiterhin nur politisch Verfolgten gewährt werden.
Stimme zu: CDU/CSU, SPD, AFD, FDP
Stimme nicht zu: Links, Grün

Wahl-O-Mat questions 1-5 allow us to identify people who are critical of foreigners
and migration (quantification: +1 (like AfD), 0 neutral, -1 (like Green/Left).

Further Wahl-O-Mat questions (purpose?): 1. Unternehmen sollen selbst entschei-
den, ob sie ihren Beschäftigten das Arbeiten im Homeoffice erlauben. 2. Der
kontrollierte Verkauf von Cannabis soll generell erlaubt sein. 3. Der Staat soll
weiterhin für Religionsgemeinschaften die Kirchensteuer einziehen. 4. Spenden
von Unternehmen an Parteien sollen weiterhin erlaubt sein. 5. Der für das Jahr
2038 geplante Ausstieg aus der Kohleverstromung soll vorgezogen werden.

Materials

- Materials are of the form [Sprecher] sagte: "X", where X is a short passage
with a potential dog whistle (PDW), or a minimally different version
without the PDW.

- Materials include PDWs that are conventionalized ("Volk", "Let's go
Brandon", "inner city") as well as ones that are not conventionalized but
need context to be PDWs ("Hilfe vor Ort"); TAXONOMY

Potential target stimuli (original with PDW / minimal variant without PWD):
- Volk: "Kinder sind die Zukunft unseres Volkes" / "Kinder sind die Zukunft
unserer Gesellschaft". - Hilfe vor Ort: "Hilfe vor Ort hat für die AfD höchste
Priorität." ⇒ "Hilfe vor Ort hat für uns höchste Priorität." / "Menschen zu helfen
hat für uns höchste Priorität." Needs context that makes clear that this is about
refugees: "Bei der Flüchtlingsfrage hat/ist . . . " "Eine neue Migrationspolitik,
die [. . .] für mehr Hilfe vor Ort in den Krisenregionen dieser Welt sorgt." –>
Diese Migrationspolitik sorgt für mehr Hilfe vor Ort in Krisenregionen." –> Diese
Migrationspolitik sorgt für mehr lokale Unterstützung in Krisenregionen."

- Kampf für Deutschland: Auch 2022 geht unser Kampf für Deutschland
weiter. –> Auch 2022 geht unser Kampf für Deutschland weiter. –> Auch
2022 setzen wir uns weiter für Deutschland ein.

- Identitätswahrende Migrationspolitik: "Vorbild einer identitätswahrenden Migrationspolitik können für Deutschland nicht klassische Einwanderungsländer von der Größe Kanadas oder Australiens sein, sondern eher Länder wie Japan, die eine ihrer Landesstruktur entsprechende Begrenzung und Steuerung der Migration verfolgen." ⇒ "Vorbild einer identitätswahrenden Migrationspolitik sind Länder wie Japan." / "Vorbild einer funktionierenden Migrationspolitik sind Länder wie Japan."

- Clankriminalität: "Wenn die Linke die Razzien gegen die Clankriminalität in Berlin-Neukölln beenden will . . . " ⇒ "Wer will Razzien gegen die Clankriminalität beenden?" / "Wer will Razzien gegen organisierte Kriminalität beenden?"

- Deutsches Staatsvolk: "die Aufnahme in das deutsche Staatsvolk, die definitiven Charakter hat, an strenge Bedingungen zu knüpfen". Source: https://www.afd.de/staatsvolk/ ⇒ "Die Aufnahme in das deutsche Staatsvolk ist an strenge Bedingungen geknüpft." / "Der Erwerb der deutschen Staatsbürgerschaft ist an strenge Bedingungen geknüpft."

Controls: - Type 1: Stimuli where everybody would agree that the expression is racist; this will allow us to exclude participants not attending to the task, but it also provides a point of comparison for PDWs (e.g., "Ausländer sind gewalttätig") - Type 2: Stimuli where everybody would agree that the expression is, e.g., green (exclude participants not giving high enough "green" ratings), inflammatory though not DW: e.g., "Steuererhöhung".

Fillers: to distract participants from the target stimuli.

Each participant reads: - Target stimuli with PDWs - Target stimuli without PDWs - Control stimuli (for exclusion) - Filler stimuli (for distraction)

Procedure

- On each trial, a participant reads a stimulus and rates the following:
  - Speakers' age, gender, education
  - Wenn am Sonntag Wahl wäre, wen würde Sprecher wählen?
  - Which attributes (on a slider?)? **establish through pretest?**
    * Rassistisch
    * Umweltbewusst
    * Gebildet
  - What would the speaker answer on the following questions
    * Das Recht anerkannter Flüchtlinge auf Familiennachzug soll abgeschafft werden.
    * Das Tragen eines Kopftuchs soll Beamtinnen im Dienst generell erlaubt sein.
    * In Deutschland soll es generell möglich sein, neben der deutschen eine zweite Staatsbürgerschaft zu haben.

    ∗ Islamische Verbände sollen als Religionsgemeinschaften staatlich anerkannt werden können.
    ∗ Asyl soll weiterhin nur politisch Verfolgten gewährt werden.

Analysis

Not yet started

Taxonomy (develop in phase 1)

- Conventionalized dog whistle: independent of context or speaker, "inner city (crime)"
- Speaker-dependent dog whistle: "Clankriminalität"? Used to activate stereotypes about certain ethnicities but Wikipedia (relatively neutral source) also has an entry for that term. Quote from Wikipedia: "Ein paar Wissenschaftler wiesen auf eine 'stigmatisierende' Verwendung des Begriffs 'Clankriminalität' hin und erklärten, dass staatliche Maßnahmen gegen Clankriminalität vorrangig politisch motiviert seien und sich negativ auf präventive Maßnahmen und auf die Integration von dem noch nicht kriminell gewordenen Teil in den Familien auswirken würden.[14]" So it's supposedly only certain uses that are stigmatizing, not the term itself?
- Content-dependent dog whistle: "Hilfe vor Ort" harmless in the context of the 2021 floodings in Germany, loaded when talking about migration.

Modeling

A formalization of reference systems that could be a suitable basis for modeling work:

1. A reference system P is a probability model that assigns a probability to each possible meaning m from a set M of meanings given a sign s from a set of signs S:
$$P(m|s)$$

2. Meanings can be discrete in which case P is a probability mass function, or continuous (e.g. vectors in some semantic space) in which case P is a probability density function.

3. Given a reference system P, we can quantify its confidence about the meaning of sign s as the entropy (discrete case):
$$H(M|s) = - \sum_{m \in M} p(m|s) \log p(m|s)$$

4. Given two reference systems P and Q, we can quantify their agreement on the meaning of s as the Bhattacharyya coefficient (discrete case):
$$BC(P,Q) = \sum_{m \in M} \sqrt{P(m|s)Q(m|s)}$$

5. Two reference systems P and Q can be combined (discrete case):

$$(P \circ Q)(m|s) = \frac{P(m|s)Q(m|s)}{\sum_{m \in M} P(m|s)Q(m|s)}$$

One interesting consequence of these definitions: We may have the intuition that the combination of two reference that disagree strongly results in a reference system that has low confidence about the meaning of a word, but that's not necessarily the case. For instance, for a sign s, reference system A may assign high probability to meanings a, b, and c, whereas reference system B may assign high probability to c, d, and e. The two reference system then disagree rather strongly, but their combination is really confident that the meaning must be c. Whether people combine references like that is of course an empirical question.

Specific examples of dog whistles

Clankriminalität

- Potentially stigmatizing certain ethnic groups.
- Contrary to dog whistle definition often not covert. The fact that it is used to refer to certain ethnicities and subcultures is rarely hidden.
- Definition by the Bund deutscher Kriminalbeamter: "Definition: Clankriminalität ist die Begehung von Straftaten durch Angehörige ethnisch abgeschotteter Subkulturen. Sie ist bestimmt von verwandtschaftlichen Beziehungen, einer gemeinsamen ethnischen Herkunft und einem hohen Maß an Abschottung der Täter, wodurch die Tatbegehung gefördert oder die Aufklärung der Tat erschwert wird. Dies geht einher mit einer eigenen Werteordnung und der grundsätzlichen Ablehnung der deutschen Rechtsordnung. Dabei kann Clankriminalität einen oder mehrere der folgenden Indikatoren aufweisen: 1. Eine starke Ausrichtung auf die zumeist patriarchalisch-hierarchisch geprägte Familienstruktur. 2. Eine mangelnde Integrationsbereitschaft mit Aspekten einer räumlichen Konzentration. 3. Das Provozieren von Eskalationen auch bei nichtigen Anlässen oder geringfügigen Rechtsverstößen. 4. Die Ausnutzung gruppenimmanenter Mobilisierungs- und Bedrohungspotenziale." See also section 3 in this document which adds more details (Heiratsregeln, Ehrverständnis, Paralleljustiz). This document assumes that this is a clearly distinguished phenomenon that requires dedicated responses, and perhaps that's also the justification for using the term (but that's not made explicit). Note that the existence of a clear and transparent definition (in the context of police work) is incompatible with plausible deniability. Other uses of "Clankriminalität" may still have a payload that requires plausible deniability.
- "Clan" is also used in the context of the Sicilian Mafia, see e.g. Wikipedia.
- "Clan" is also used in the context of Corsican organized crime, e.g. Süddeutsche Zeitung (2010, i.e. before Arab organized crime became a hot

topic in Germany). A Google search for "Korsische Clans" produces 1.2M hits.

Sources:

- Bund Deutscher Kriminalbeamter, "Clankriminalität bekämpfen: Strategische Ausrichtung – nachhaltige Erfolge", https://www.bdk.de/der-bdk/wer-wir-sind/positionen/2019-04-29-bdk-positionspapier-clankriminalitaet.pdf
- Clankriminalität (https://www.presseportal.de/pm/130241/5040654)
- https://www.afd.de/beatrix-von-storch-linkspartei-auch-unterstuetzer-der-organisierten-clan-kriminalitaet/
- AfD Wahlprogramm

"Sex is real"

- Implied meaning: Only people assigned female at birth (based on sex markers) are women.

"Das muss man auch einmal klar sagen." (Gysi, Bundestagsrede)

Similarly: "Jede Wahrheit braucht einen Mutigen, der sie ausspricht." (Campaign of the Bild-Zeitung, allegedly an Einstein quote but not substantiated, many times caricatured).

Implied meaning: - They speak the truth. - Others don't speak the truth. - The truth is suppressed. - Speaking this truth breaks a taboo is an inconvenient truth. - The one speaking those truths is virtuous and courageous. - Bild attributed their slogan to Einstein, Gandhi, ML King, which adds more layers of meaning.

But does any of that require plausible deniability? Do we expect sensitivity to meaning to differ between in- and outgroup?

"Kerneuropa" (Gysi, Bundestagsrede, 2015-06-18)

- Potentially alluding to the idea that France and Germany are trying to dominate the rest of the EU.

"Hilfe vor Ort"

- Dog-whistly when used in the context of immigration from the middle east.
- Not dog-whistly when used in the context of 2021 floodings in Rhineland-Palatinate and North Rhine-Westphalia.
- Hilfe vor Ort: "Hilfe vor Ort hat für die AfD höchste Priorität" (AfD Wahlprogramm)

"Volk"

- "Kinder sind die Zukunft unseres Volkes." (AdD Wahlprogramm)
- "Wir sind das Volk", Georg Büchner, /Dantons Tod/, 1. Akt, 2. Szene (1835)
- "Wir sind das Volk", DDR-Protestbewegung, 1989/190.
- "Wir sind ein Volk", Wendezeit, 1990-
- "Wir sind das Volk", PEGIDA-Bewegung, 2014-, Suggestion, dass PEGIDA die Mehrheit der Gesellschaft repräsentiert, aber auch Abgrenzung von Migranten.
- "Das Volk ist jeder, der in diesem Land lebt." Merkel speaking out against individual groups appropriating the concept of the people for themselves. Also Merkel's response to PEGIDA and possibly her rejection of the Dog Whistle use of the term. NZZ Article about Merkel's statement, its context, and the history of the concept and related concepts (e.g., "Staatsvolk").
- "Wirr ist das Volk!", Titanic, Die Partei, 2016.
- Zu den Begriffen „deutsches Volk", „Deutsche" und „deutsche Volkszugehörigkeit" im Grundgesetz

"Brauchtum"

- AfD Wahlprogramm

"Massenzuwanderung"

- "Die desaströsen Folgen der unregulierten Massenzuwanderung seit 2015 sind unübersehbar und verschärfen sich weiter: überproportionale Zuwanderkriminalität, Terroranschläge und islamischer Separatismus belegen dies ebenso wie dreistellige Milliardenkosten, Wohnraummangel und die hohe Arbeitslosigkeit unter den Zuwanderern." (AfD Wahlprogramm)

"Deutsches Staatsvolk"

- "die Aufnahme in das deutsche Staatsvolk, die definitiven Charakter hat, an strenge Bedingungen zu knüpfen" (https://www.afd.de/staatsvolk)

"Identitätswahrende Migrationspolitik"

- "Vorbild einer identitätswahrenden Migrationspolitik können für Deutschland nicht klassische Einwanderungsländer von der Größe Kanadas oder Australiens sein, sondern eher Länder wie Japan, die eine ihrer Landesstruktur entsprechende Begrenzung und Steuerung der Migration verfolgen." (AfD Wahlprogramm)

"Kampf (für Deutschland)"

- "Auch 2022 geht unser Kampf für Deutschland weiter." (https://afd.de)

Saxony's prime minister Michael Kretschmer demands "vernünftiges Verhältniss zu Russland" (reasonable relationship with Russia).

- https://www.zeit.de/news/2022-03/02/kretschmer-fordert-vernuenftiges-verhaeltnis-zu-russland?utm_referrer=https%3A%2F%2Fwww.google.com%2F
- Pragmatic Analysis: The literal content is self-evidently true because nobody would argue for an unreasonable relationship to any country. So there is a content gap / ambiguity / violation of maxim of quantity. Intended meaning not clear.
- Some might fill this gap with something like "we should not confront Russia over what they are doing in Ukraine". But this can't be said explicitly in the first week of war, it's a taboo.
- Further, Kretschmer can claim plausible deniability and pretend that he meant something more benign than what his accusers are saying.
- Summary:
  1. Presence of a taboo. ✓
  2. Violation of maxim of quantity. ✓
  3. Meaning can be enriched with a controversial meaning that's violating the taboo, i.e. literal meaning is not incompatible with pragmatically (not logically) implied meaning. ✓
  4. Plausible deniability. ✓
- Observation: It's not dog-whistly in the original sense because his outgroup immediately attacked him for this statement. This is perhaps more an instance of political persuasion than covert communication.

Sources of PDWs

- Here's someone's (failed IMO) attempt to elicit left-wing dog whistles on the "Ask the Donald" subreddit: https://www.reddit.com/r/AskThe_Donald/comments/b4qc5d/what_are_th
- https://www.afd.de/wahlprogramm-asyl-einwanderung/
- FDP
- https://www.tagesschau.de/ausland/amerika/usa-brandon-trump-republikaner-schlachtruf-101.html
- Bundestagscorpus und Reden der AfD
- https://die-linke.de
- Mindestsicherung
- Polizeistaat
- Ausdrücke in Anführungszeichen bei der AfD / Annotation von dog whistles?
- Gender-gerechte Sprache
- Klimadebatte
- Covid-Skeptizismus, Debatten zur Impfpflicht
- Stuttgart 21 Debatte?
- The Rhetoric Tricks, Traps, and Tactics of White Nationalism

References

Preliminary results of pilots

SUMMARY after looking at "Hilfe vor Ort", "Kampf für Deutschland", "deutsches Staatsvolk", and "Volk"

- "Hilfe vor Ort": less conservative participants rate DW speaker as less progressive, more racist, less helpful than more conservative participants (no difference between the two groups on control)
- "Hilfe vor Ort" exhibits trends:
  - DW speaker is judged younger and more friendly than control speaker (in both groups)
  - less conservative give lower intelligence and lower honesty rating than more conservative
- "Kampf für Deutschland": mostly no difference between more and less conservative participants!
  - both groups take DW speaker to be older, more racist, less helpful, less intelligent, and less friendly than control speaker
  - more conservative gave lower honesty rating to DW than control (no diff for less conservative)
  - more conservative gave higher christian rating to DW than less conservative
- "deutsches Staatsvolk": less conservative participants give lower progressive, lower helpful, and lower friendliness rating than more conservative participants (no differences on the control)
- "deutsches Staatsvolk" exhibits trends:
  - less conservative participants give higher racism, lower honesty, lower intelligence, and lower christian ratings than more conservative participants
- "Zukunft unseres Volkes": less conservative participants give lower honesty, lower helpful rating than more conservative
  - trends: less conservative gave lower intelligence and lower friendly rating than more conservative,
  - both groups gave low racism ratings, didn't differ in how progressive the speakers are
- Overall lessons:
  - "Hilfe vor Ort" is the clearest case of a DW: less and more conservative participants rate speakers who use this expression differently (especially relevant: progressive, racism)
  - "Kampf für Deutschland" is not a DW; it is obviously racist
  - "Volk" is not a conventionalized DW: "Zukunft unseres Volkes" didn't sound racist to either group, a trend towards a DW was observed for "deutsches Staatsvolk" (progressive; trend: racism)

NEXT: - embed the four sentences under clause-embedding predicates and

manipulate speaker identity, how projective are the contents? - test for plausible deniability with the "S but not racist" diagnostic

References

Albertson, B. L. (2015). Dog-whistle politics: Multivocal communication and religious appeals. *Political Behavior*, *37*(1), 3–26. https://doi.org/10.1007/s11109-013-9265-x

Calfano, B. R., & Djupe, P. A. (2009). God talk: Religious cues and electoral support. *Political Research Quarterly*, *62*(2), 329–339. https://doi.org/10.1177/1065912908319605

Henderson, R., & McCready, E. (2019). Secondary content: The semantics and pragmatics of side issues. Brill. https://doi.org/https://doi.org/10.1163/9789004393127_010

Khoo, J. (2017). Code words in political discourse. *Philosophical Topics*, *45*(2), 33–64. https://doi.org/10.5840/philtopics201745213

Mendelberg, T. (2001). *The race card: Campaign strategy, implicit messages, and the norm of equality.* Princton Universtiy Press. https://press.princeton.edu/books/paperback/9780691070711/the-race-card

Wetts, R., & Willer, R. (2019). Who is called by the dog whistle? Experimental evidence that racial resentment and political ideology condition responses to racially encoded messages. *Socius: Sociological Research for a Dynamic World*, *5*, 237802311986626. https://doi.org/10.1177/2378023119866268

White, I. K. (2007). When race matters and when it doesn't: Racial group differences in response to racial cues. *American Political Science Review*, *101*(2), 339–354. https://doi.org/10.1017/s0003055407070177