

Dogwhistles as Inferences in Interaction

Ellen Breitholtz and Robin Cooper

Centre for Linguistic Theory and Studies in Probability (CLASP)

Department of Philosophy, Linguistics and Theory of Science

University of Gothenburg

ellen.breitholtz@ling.gu.se robin.cooper@ling.gu.se

Abstract

In this paper we propose an approach to dog-whistle communication which combines ideas from Khoo (2017) where inferences associated with particular expressions give rise to dog-whistle messages, and ideas from Henderson and McCready (2018) who use game theory to account for dogwhistle communication. Our view focuses on the dialogical nature of dog-whistle communication where participants can draw different conclusions based on communicative events. By modelling content in terms of intensional types in TTR we avoid the potential problem of Khoo's inferential account where distinct expressions that have equivalent content are analysed as generating the same dogwhistle inferences.

1 Introduction

In recent years the term “dogwhistle” (also spelled as two words: “dog whistle”) has become increasingly common in media discourse, not least on the internet. The term originally refers to a whistle producing a high pitched sound perceived by dogs but not humans. Metaphorically it has come to mean an expression that conveys one message to one group and a different message to another group. Such dogwhistles have been used in politics to for example express loyalty with a particular group while not alienating the public at large. The phenomenon (and its effects on voter sympathies) has been discussed since at least the eighties in political opinion research and related fields (Lamis, 1984). Much of the focus of this research is on the invocation of racial prejudice, famously part of the so called “Southern Strategy” to gain Republican voters in the previously Democratic southern states of the USA. Many of the examples in the literature are thus related to race. One example is the expression “inner city” as a dogwhistle for “African American”. In a series of experiments

where subjects were asked how favourably they viewed a welfare program proposed by a fictitious politician depending on the language used, White (2007) showed that the term “inner city families” triggered more negative responses in subjects who also had a negative attitude towards African Americans. One example where the term “inner city” is used in political discourse is the quote below (taken from Khoo, 2017) from former president Trump's 2011 book “Time to get tough: Making America Great Again”:

- (1) If we keep on this path, if we reelect Barack Obama, the America we leave to our kids and grandkids won't look like the America we were blessed to grow up in. The American Dream will be in hock. The shining city on the hill will start to look like an inner-city wreck (Trump, 2011)

However, despite being interesting from a linguistics perspective, dogwhistles have not been studied extensively in linguistics or philosophy of language, with a few exceptions, such as Stanley (2015) who analyses dog whistles as conventional implicatures, Khoo (2017) who thinks of them in terms of inferences and Henderson and McCready (2018), who use techniques from game theory to account for dogwhistle meaning.

In this paper we will argue that the nature of dogwhistle communication is essentially dialogical, and that to account for dogwhistle meaning we must consider interactive events in which dialogue partners can draw different conclusions based on communicative events. This leads us to a theory based on inference, similar to that of Khoo. However, as identified by Khoo himself and emphasised by Henderson & McCready, a potential problematic aspect of this approach is that expressions that have a similar meaning are analysed as generating the same dogwhistle inferences, which appears not

always to be the case. By modelling meaning in terms of intensional types in TTR, we avoid this problem. In addition we think there are examples where expressions with similar meanings to other expressions commonly associated with dogwhistle communication can also develop dogwhistle meaning. Combining an inferential approach with an intensional account of meaning can provide an analysis of dogwhistles that allows for this without having all coextensional expressions generating the same dogwhistle messages.

An important ingredient in a complete analysis of dogwhistles introduced by Henderson & McCready is the game theoretic aspect of using dogwhistles. This approach is particularly useful in the context of a theory based on communicative events (Breitholtz and Cooper, 2019). An important part of this analysis is accounting for social meaning in terms of sociolinguistic personae. We suggest developing this account by linking inferences and personae via sets of topoi which are associated on one hand to particular personae, on the other to particular inferences.

Henderson & McCready also argue against Stanley (2015) who analyses dogwhistles in terms of conventional implicatures, as conventional content cannot be denied by the speaker and deniability is a key aspect of dogwhistle communication. We sympathise with this argument. However, we believe that a dogwhistle can develop according to a life cycle, where the dogwhistle content becomes increasingly conventionalised. We argue that this process is similar to how non-controversial words can develop into slurs as associated but initially deniable inferences become part of the conventional meaning and thus not deniable.

In order to account for this we need both an analysis incorporating communicative events and a notion of gradation using probability estimation. We will look at some problems introduced for these theories by a number of authentic examples of dogwhistle communication, and suggest a way of combining the inferential and game theoretic approach.

2 Reasoning in interaction

Different people will draw different inferences from the same utterance. This may be because they interpret the utterance in a different way. But it may also be the case that even though two different people interpret a given utterance in the same way, they still draw different conclusions. The reason

for this may be because they are aware of different inferences that can be drawn from the content of the utterance. Suppose *A* says:

- (2) The lights are on and there is music coming from the house

B might reasonably come to the conclusion:

- (3) Somebody is in the house

C, however, who knows that the people who live in this house often go out and leave the lights on and music playing would not draw this conclusion. If, in addition, *C* saw everybody in the house leave ten minutes ago and has been watching to see if anybody has come back, *C* might draw the conclusion:

- (4) The people in the house left the lights and the music on when they left

This is an example in which the knowledge that *B* and *C* have leads them to draw different (in this case incompatible) conclusions. There are other cases where different conclusions are drawn because of differing opinions or prejudices. Policeman *A* says:

- (5) The guy driving the Mercedes is black

Policeman *B*, who is of the opinion that black people do not in general drive expensive cars, comes to the conclusion:

- (6) The guy driving the Mercedes might have stolen it

Policeman *C*, however, who is of the opinion that black people do sometimes drive expensive cars, comes to the conclusion:

- (7) There is not enough evidence to stop and search the driver of the Mercedes

Often we are led to draw different inferences because of the mental state that we are currently in. Child *A* says:

- (8) I just heard a strange noise at the back door

Parent *B*, who has just been watching a horror movie, draws the conclusion:

- (9) Someone or something is trying to get into the house

Parent *C*, who has just completed a meditation session, concludes:

- (10) The sound was made by the wind at the back door

Both parents, in different circumstances, could have come to the conclusion of the other parent.

It seems that the reasoning that we perform under the time pressure of interaction is often of the kinds illustrated here (and there are probably many more kinds of such reasoning). This is not logical, or even necessary, inference but rather snap judgements made on the spur of the moment. In some cases they may easily be abandoned when we are presented with evidence that falsifies them. In other cases they may represent deeply entrenched prejudices which are hard to get rid of and we may be barely conscious of the reasoning that led to the conclusion. It is for these reasons that we treat such inference in terms of the theory of *topoi* developed in Breitholtz (2020). This theory has its origins in rhetoric where arguments are often designed to tap into rules of thumb warranting the argument. For example, if a politician were to say “I love freedom – therefore I will lower taxes” they count on the audience accepting this argument based on the notion that lower taxes are associated with a higher degree of freedom, or a more general principle of reasoning – that having more money increases your freedom (and if you pay less tax you will have more money). These rules of thumb are often referred to as *topoi*, and the theory of *topoi* goes back at least to Aristotle’s “Topics” and “Rhetoric”. Aristotle sees *topoi* as important tools to be employed in inventing coherent and persuasive dialectic and rhetorical argumentation. *Topoi* can also be seen as a basis of sense making, in the sense that infants pick up basic pre-linguistic *topoi* which function as scaffolding for extending the set of *topoi* the child has access to. In modern times the theory of *topoi* has been picked up by Ducrot (1988) and Anscombe (1995) who see *topoi* as an essential part of semantics. One *topos* may be used to warrant many different types of argument, and different *topoi* may be instantiated as the same argument. This means that a theory of *topoi* allows for different dialogue participants drawing different conclusions from a single utterance. This is a central and pervasive feature of dialogic interaction, which often goes unnoticed but sometimes is made visible to us for example through processes of repair (Clark, 1994; McRoy and Hirst, 1995). The ability to associate

and draw inferences is an important part of what makes dialogic communication efficient and also interesting. Imagine, for example, a bizarre world in which our dialogue partners were incapable of drawing conclusions from what we say. *A* says to *B*:

- (11) Put the book on the bookshelf

B, however, remains motionless because *B* does not reason that putting the book on the bookshelf involves picking up the book, that picking up the book involves stretching your hand towards the book, . . . We would not be able to function in a world where a dialogue partner was incapable of drawing their own conclusions. The fact that different dialogue partners draw different conclusions can be a source of interest and can also give pause for reflection and consideration on whether we also want to accept their conclusion:

- (12) *A*: I’ve been seeing a lot of Kim recently
B: He’s after your money

3 Exploiting expected inferential behaviour in communication

We see dogwhistles as parasitic on this basic fact of dialogue interaction. Dogwhistle events are those where a dialogue participant, *A*, exploits the fact that different agents interpreting *A*’s utterance will draw different conclusions. This is essentially the theory expressed by Khoo (2017). Let us see how it might work in our terms.

Let *A*, *B* and *C* be agents and φ_0 , φ_1 and φ_2 be message contents. In terms of a treatment in terms of TTR we can think of the φ_i as being types. To make the example concrete we can think of the φ_i as representing the following message contents:

- (13) φ_0 We need to solve the problem with inner cities
 φ_1 We need to solve the problem with social support for poverty in inner cities
 φ_2 We need to solve the problem with lazy out of work black populations in inner cities

Suppose that *A* wishes to communicate φ_1 to *B* and φ_2 to *C*. Suppose furthermore that *A* estimates that there is a high likelihood that *B* has, and would currently be disposed to apply, an inferential resource, a *topos* in the sense of Breitholtz (2020), which we represent informally here as:

(14) $\varphi_0 \rightarrow \varphi_1$

That is, given an utterance, u , whose content is φ_0 , B is highly likely to draw the conclusion φ_1 . Cooper (in prep) analyzes utterance situations as signs, that is records including fields labelled by ‘s-event’ (for speech event) and ‘cont’ (for content). Thus to say that φ_0 is the content of u is to say that the value in the ‘cont’-field of u is φ_0 , in notation: $u.\text{cont} = \varphi_0$.

Suppose that in addition A has a similar view of C , namely that there is a high likelihood that C has, and would currently be disposed to apply, a topos informally represented as:

(15) $\varphi_0 \rightarrow \varphi_2$

That is, given an utterance, u , such that $u.\text{cont} = \varphi_0$, C is highly likely to draw the conclusion φ_2 .

If A talks to B and C , it seems that any of the following four scenarios would represent success with respect to the goals of communicating φ_1 to B and φ_2 to C :

- (16) a. A addresses u_0 to B
 B judges that $u_0.\text{cont} = \varphi_0$
 B therefore receives the message φ_0
and draws the conclusion φ_1
- b. A addresses u_1 to B
 B judges that $u_1.\text{cont} = \varphi_1$
 B therefore receives the message φ_1
- c. A addresses u_0 to C
 C judges that $u_0.\text{cont} = \varphi_0$
 C therefore receives the message φ_0
and draws the conclusion φ_2
- d. A addresses u_2 to C
 C judges that $u_2.\text{cont} = \varphi_2$
 C therefore receives the message φ_2

Following a gametheoretical approach to dogwhistles (Henderson and McCready, 2018), we say that A gets a reward of 1 for such successful scenarios. We will take rewards to be in the interval $[-1, 1]$. Negative rewards are costs. If A ’s goal is to communicate φ_1 , but not φ_2 , to B and φ_2 , but not φ_1 , to C , then A ’s rewards/costs might be given by a table such as the following:

(17)

| Message | To | Reward |
|-------------|-----|--------|
| φ_1 | B | 1 |
| | C | -.5 |
| φ_2 | B | -1 |
| | C | 1 |

From what we have seen so far it seems that if A wishes to maximize the reward obtained when talking individually to B or C it would be equally good to choose any of the successful scenarios in (16). Actually, however, things are not this simple. The topoi $\varphi_0 \rightarrow \varphi_1$ and $\varphi_0 \rightarrow \varphi_2$ are hidden and, as we pointed out above, A has to estimate the probability that they are available to B and C and whether they will actually make the inference in the given situation. Another probability involved is the probability that B and C will understand the utterances u_i in the way that A intends, that is, that the appropriate message contents will be received. If we assume that the probabilities of understanding are higher than the probabilities of drawing the intended inference, then the best way for A to maximize the reward is to choose the respective utterances u_1 or u_2 which directly indicate the message to be communicated (φ_1 or φ_2) rather than run the risk that B or C will not draw the desired inference.

There may be other factors that indicate that this is not the best strategy, however. For example, while an agent may be able and disposed to draw an inference, they may not wish to publically admit to drawing the inference. Thus while C may well privately draw the inference “We need to solve the problem with lazy out of work black populations in inner cities”, they may well be unwilling to be forced to acknowledge an utterance which explicitly has this as its content. (“I’m not a racist.”) The advantage of communicating through inference is that there need not be any explicit commitment on the part of either dialogue participant to the inferred content. (“Know what I mean?”) This phenomenon seems related to euphemistic uses of language: “Is there somewhere I can wash my hands?” uttered in the hope that the interlocutor will infer the real need and point to a bathroom (*sic!*) rather than a kitchen sink.

4 Dogwhistles in multiparty dialogue

Another motivation for using an utterance which is less explicit that the message you wish to convey is multi-party dialogue where you wish different parts of your audience to draw different conclusions, that is, the dogwhistle scenario. Consider the following scenario:

- (18) A addresses u_0 to B and C
 B and C judge that $u_0.\text{cont} = \varphi_0$

B therefore receives the message φ_0 and draws the conclusion φ_1
 C therefore receives the message φ_0 and draws the conclusion φ_2

We can compute the reward that A receives by summing the rewards that would be received for each of the other dialogue participants and dividing by the number of the other dialogue participants, that is, in this case:

$$(19) \frac{1+1}{2} = 1$$

The scenario (18) is exactly as desired if A is wishing to dogwhistle. Note that what makes this scenario a dogwhistle is A 's desire to achieve the scenario (18). It is not clear that we would want to call this a dogwhistle (or that A would get the reward) if A had not reckoned with a positive likelihood that making the utterance u_0 would have this effect. One might want to claim that there is no such thing as an unintentional dogwhistle. Alternatively, one might say that whenever such a scenario occurs this should be called a dogwhistle (and also possibly in addition that A is responsible even it was unintentional – A should have known that this would be the effect).

5 Dogwhistles and (hyper)intensionality

At this point we should take up a potential problem for this inferential approach to dogwhistles which [Khoo \(2017\)](#) raises and which [Henderson and McCready \(2018\)](#) pursue, calling it a reason to reject any approach to dogwhistles like Khoo's or ours which is based on inference. The potential problem is that the required inferences may be drawn only when the message is expressed in a certain way. Saying something different, though equivalent, may not trigger the inference. For example, there may be inferences associated with *inner city* which are not associated with *metropolitan area*. We do not see this as a problem within the kind of approach we are advocating using TTR types. Recall that the messages φ_i which we have been discussing would be construed as types in TTR. Types in TTR are intensional (or what would be called hyperintensional in traditional possible world semantics). What this means is the you can have two distinct types which have exactly the same witnesses. As discussed in [Cooper \(in prep\)](#) and elsewhere, this has important consequences for linguistic semantics. For example, any situation of the type “the glass is half full” will also be of the

type “the glass is half empty” and *vice versa*, yet you can be glad that the glass is half full and sorry that it is half empty. Other examples are discussed in [Cooper \(in prep\)](#). Given that natural language semantics seems to be sensitive to the difference between equivalent types it should not come as a surprise to us that this sensitivity is also showing up in dogwhistle scenarios.

However, the situation with dogwhistles and equivalent or near-equivalent expressions is not as clear-cut as we might have thought. One can question to what extent there is true equivalence in a logical sense between such expressions. One can also question whether it is always the case that there is no carry-over of inferences between expressions of even similar content. An example of this is the dialogue below, taken from an episode of the TV-show Last Week Tonight with John Oliver. The episode was dedicated to black hair and ways in which it is often the target of discrimination. In the story a young woman, Destiny Thompkins, describes her experience of being approached by her manager at the Banana Republic shop where she had been working for a month:

(20) D: He was like, yeah so like the district manager came in and she pointed out something about your hair. And I'm like, okey so what's wrong with my hair? He said it's a little too urban and unkempt for our look and our image, we were just wondering like, if you could just take them out?

(Last Week Tonight, May 10th 2021, accessed August 20th 2021)



In the example above “urban” seems very likely to be a code word for “black” or “African American”, as is also pointed out by John Oliver in his commentary of the interview. In this case then, it seems like an expression with a very similar extension to “inner city” also invokes a similar dogwhistle message.

The fact that TTR types are intensional does not prevent us from relating similar types in terms of

dogwhistle inferences.

6 When dogwhistles go wrong

Let us now consider what happens in a dogwhistle scenario where the different addressees have knowledge or beliefs about each other's access to topoi. In a dogwhistle scenario it is not only important that the intended recipients receive the messages directed to them, but also that they do not receive the messages not intended for them. This is reflected for our example by the negative rewards in the reward table (17). Even if they are not disposed to draw a certain conclusion themselves, they may well realize that somebody else in the audience may draw this conclusion and therefore suspect that the speaker is trying to communicate different messages to different members of the audience. That is, they may realize that they are part of a dogwhistle scenario. If the recipient gets more than one message we can say that the reward for that recipient is the sum of the rewards for each message divided by the number of messages. Thus for B in this new scenario A would get the reward:

$$(21) \frac{1-1}{2} = 0$$

If C gets only the intended message, the overall reward for A , summing over the reward for the dialogue participants, would be:

$$(22) \frac{0+1}{2} = .5$$

This is, of course, a very simplistic way of calculating the reward which does not take into account other factors such as the degree of importance attached to B realizing what message is being passed to C which could have consequences that could range from mildly embarrassing to career or even life-threatening.

7 Conclusions and future work

In this paper we propose an approach to dogwhistle communication which follows the work of Khoo (2017) in that it draws on inferences that language users may draw from the use of particular expressions based on their being associated with topoi warranting those inferences. An important part of our view of dogwhistles is that they exploit standard principles of inference in communication. This means that we would like our model to be able to capture the *life cycle* of a dogwhistle where an expression evolves from not conveying a dogwhistle message, via conveying a dogwhistle message

to a subgroup of the audience to eventually being conventionally associated with that message. We believe that this process is similar to those in play in the context of semantic change in general.

We think of dogwhistles as dialogue events. We propose that the strategy involved in deciding to attempt to create a dogwhistle event can be modelled using techniques from game theory, a proposal originally made by Henderson and McCready (2018). However, their approach focuses on the relevance of personae in calculating the social meaning of an expression. While we agree that personae are relevant for inferring dogwhistle messages we also believe that being aware of potential topoi associated with the expression itself are also important. In future work we would like to also take persona into consideration in order to be able to account for the interplay between topoi and personae in dogwhistle communication.

Acknowledgments

This work was supported by the Wallenberg AI, Autonomous Systems and Software Program – Humanities and Society (WASP-HS) funded by the Marianne and Marcus Wallenberg Foundation through the project Gothenburg Initiative for Politically Emergent Systems (GRIPES). It was also supported by the Centre of Linguistic Theory and Studies in Probability (CLASP) (VR: 2014-39). We want to thank our colleagues in the GRIPES project, Elina Lindgren, Gregor Rattenegger, Björn Rönnerstrand and Asad Sayeed, for valuable discussions.

References

- Jean-Claude Anscombre. 1995. La théorie des topoi: Sémantique ou rhétorique? *Hermès*, 15.
- Ellen Breitholtz. 2020. *Enthymemes and topoi in dialogue: The use of common sense reasoning in conversation*, volume 41 of *Current Research in the Semantics/Pragmatics Interface*. Brill.
- Ellen Breitholtz and Robin Cooper. 2019. Integrating personae in a TTR account of interaction. Presented at *Integrating Approaches to Social Meaning*.
- Herbert H Clark. 1994. Managing problems in speaking. *Speech communication*, 15(3-4):243–250.
- Robin Cooper. in prep. *From perception to communication: An analysis of meaning and action using a theory of types with records (TTR)*. Draft of book chapters available from <https://sites.google.com/site/typetheorywithrecords/drafts>.

- Oswald Ducrot. 1988. Topoi et formes topique. *Bulletin d'études de la linguistique française*, 22:1–14.
- R. Henderson and E. McCready. 2018. How dogwhistles work. In *New Frontiers in Artificial Intelligence*, pages 231–240, Cham. Springer International Publishing.
- Justin Khoo. 2017. Code Words in Political Discourse. *Philosophical Topics*, 45(2):33–64.
- Alexander P. Lamis. 1984. *The Two-Party South*. Oxford University Press.
- Susan W McRoy and Graeme Hirst. 1995. The repair of speech act misunderstandings by abductive inference. *Computational linguistics*, 21(4):435–478.
- Jason Stanley. 2015. *How propaganda works*. Princeton University Press.
- LastWeek Tonight. 2021. [Hair: Last week tonight with john oliver.](https://www.youtube.com/watch?v=Uf1c0tEGfrU) <https://www.youtube.com/watch?v=Uf1c0tEGfrU>.
- Donald J Trump. 2011. *Time to get tough: Make America great again!* Simon and Schuster.
- Ismail K White. 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.