The Semantics of Climate Change and Global Warming

Timothy M. Gann (tgann@ucmerced.edu)
Sierra Nevada Research Institute and Cognitive and Information Sciences, University of California, Merced
5200 North Lake Rd. Merced, CA 95343 USA

Teenie Matlock (tmatlock@ucmerced.edu)
Sierra Nevada Research Institute and Cognitive and Information Sciences, University of California, Merced
5200 North Lake Rd. Merced, CA 95343 USA

Abstract

Creating consensus and facing the challenges of climate change requires effective climate communication. However, consensus about issues relating to climate science is unlikely to happen when there isn't a clear public consensus about which name is more appropriate, "climate change," or "global warming," and what those terms mean. Previous research has shown that perceptions of these terms varies, depending on factors such as the audience's political affiliation. To investigate this further, we analyzed two corpora from partisan online news using a high dimensional semantic analysis. This study found that while there is substantial semantic overlap between the terms "climate change" and "global warming," there is less overlap in the conservative media corpus. The results also show that there was a larger proportion of conservative articles that preferred to use "global warming" exclusively, whereas progressive articles tended to use "global warming" to supplement "climate change."

Keywords: Climate change; global warming; semantics; climate communication; high dimensional text analysis

Introduction

Understanding and adapting to climate change is one of the most serious issues facing our society. To overcome these challenges, it is important to learn to effective communicate about them. The magnitude of the difficulty of achieving this is highlighted by the fact that in the public's mind there isn't necessarily even any certainty about what to *call* the problem. Public preferences on the topic vary a great deal, with roughly equal numbers preferring either "climate change" or "global warming," and a plurality stating no preference (Akerlof & Maibach, 2011). This ambivalence is often reflected in survey research and the media, where it has been assumed in the past that they are functionally equivalent, or they are presented together like "climate change / global warming" (Whitmarsh, 2009). However, "climate change" prevails in the scientific community, and is seeing more and more use in the popular media (Boykoff, 2007).

Recent research, however, suggests that "climate change" and "global warming" are not semantically equivalent; they each have different connotations. Climate change is considered more amorphous than global warming (Whitmarsh, 2009), and is a more neutral rendering of the concept. Global warming has more in common with terms used in other [European] languages which focus on combative concepts. Global warming is also supposed to be more strongly associated with concepts related to temperature, the greenhouse effect, and human causes. Climate change tends to be associated with a diverse set of climate impacts, with a focus on effects that have already happened (Whitmarsh, 2009). These differences have concrete effects on how people comprehend speech about climate science, and how they respond to questions about it. For example, there is a correlation between outdoor temperature and people's belief in global warming (Joireman, Barnes Truelove, & Duell, 2010), which helps drive the intuitive appeal of arguments against global warming based on unseasonably cold weather (an argument that has less bite against "climate change"). Even one's choice of gum, either a cooling mint, or a warm cinnamon flavor, can prime concepts about cooling, warming, and temperature change that influence people's acceptance of climate science (Lewandowski, Ciarocco, & Gately, 2012). Due to the influence labels can have on perception, there have even been calls to switch to other terms, such as "climate disruption," because they may activate more favorable associates in the audience (Jaskulsky & Besel, 2013).

These labels also elicit varied reactions based on the cultural and political orientation of their audience, who also have their own preferences for which term is most appropriate (Akerlof & Maibach, 2011; Schuldt, Konrath, & Schwarz, 2011). Villar and Krosnick (2010) even argue that at a population level, the choice of words doesn't have a large effect on people's perceptions of the seriousness of the problem, but that among different subpopulations it can make a big difference. This disparity between how climate change and global warming are viewed by the partisan extremes creates a situation where people may be using similar words, but words that have substantially different semantics, making successful communicate difficult if not impossible (Hoffman, 2011).

Political polarization of views about climate science, especially in the United States, with an early report by Trumbo (1996) noting that the issue was getting increasingly politicized, a problem that has been getting worse as time passes (McCright & Dunlap, 2011). Over the last few decades a strong conservative resistance has emerged over past debates about the environment and to climate science (Armitage, 2005; Oreskes & Conway, 2010). Whitmarsh (2009) points out that peoples' understanding of different environmental issues tend to be linked together. For the most committed skeptics of climate change, climate science and environmentalism are inextricably bound to leftist politics in general (Hoffman, 2011). This divide appears in surveys that show both that there is a greater concordance between the scientific consen-

sus and Democrats' beliefs about climate science (McCright & Dunlap, 2011). There is also a difference in this disparity that depends on partisanship. Conservatives being more open to the scientific consensus when asked about climate change rather than global warming. Additionally, climate change is viewed as more serious (Villar & Krosnick, 2010; Schuldt et al., 2011). This may be one reason why conservatives appear to be ignoring the oft-cited advice by Luntz (2002) that they should favor the use of climate change over global warming. Global warming may be more attractive to conservatives because of this greater resistance (Schuldt et al., 2011). Progressive samples, on the other hand, tend to view the terms more interchangeably, with some indications that they might even see the opposite effect of taking global warming more seriously (Villar & Krosnick, 2010). Feinberg and Willer (2013) found that progressives cast environmental issues in moral terms in a way that conservatives do not. They suggest that one approach to bridging the divide between partisans would be through reframing these issues in ways that mesh better with a conservative moral framework (a call echoed by Lakoff, 2010).

Much of this work was done using surveys, but conducting content analyses have also been an important element of studying the language used in climate communication. Approaches have been varied, depending on the questions being asked. A simple approach by Schuldt et al. (2011) counted up the instances of "climate change" and "global warming" within the publications of a couple dozen partisan think tanks, finding the above mentioned preference for "global warming" amongst conservatives. Some have tracked how media representations have changed over time, for example, Boykoff and Boykoff (2007); Boykoff (2007) note that the American journalistic norm of trying to strike a balance between both sides, no matter how unbalanced (climate scientists, for example, express a remarkable consensus; Anderegg, Prall, Harold, & Schneider, 2010), has declined. Several researchers have done substantial human coding of corpora, determining whether articles expressed a supportive or dismissive position on climate science, and what kind of frames they used (e.g. Hoffman, 2011; Trumbo, 1996). Using this approach, Hoffman (2011) found that supporters of climate change in newspaper opinion pieces used more moral frames, mirroring the experimental findings mentioned above by Feinberg and Willer (2013). One problem with this kind of approach is that human coding of corpora can become extremely labor intensive as the size of the corpus increases. This has lead to recent attempts to example climate communications using more automated approaches. Grundmann and Krishnamurthy (2010); Grundmann and Scott (2012) analyzed a large corpus of news articles about climate change by compiling a list of all of the co-locations of words within the text, identifying the words that appeared beside climate change or global warming, and only then coding those words for whether they fit into particular frames.

We're interested in going one step beyond looking at simple co-locations of words by recruiting a more sophisticated model of word relationships. The Hyperspace Analogue to Language model (HAL) attempts to transform local co-occurrences between words into richer representations of meaning by encoding the contextual history of a word into a vector within a high-dimensional semantic space. What HAL does is read through a body of text using a moving window, creating a matrix of how closely the words in the corpus appeared to each other. Each word then corresponds to a vector in the matrix, with this vector encoding the contextual history of that word. The assumption is that words that appear in similar contexts have a categorical relationship (Burgess, 1998; Burgess & Lund, 2000). From these vectors, lists of nearest semantic neighbors can be generated. HAL has been shown to produce measures of word similarity that correlate with word priming data and lexical decision tasks (Lund & Burgess, 1996). More germaine to the current topic, it has also been used to analyze corpora related to political speech. (Murphy, Burgess, Johnson, & Bowler, 2012) used HAL to analyze the content of California ballot initiatives to see if there were systematic differences between the words used in pro- and con- arguments, and along partisan lines.

Our project aims to do the same thing with a corpus of language about climate drawn from partisan news sources to develop new insights into how the phrases "climate change" and "global warming" are used. Schuldt et al. (2011) propose that ideological effects for comparisons between the use of these terms may reflect different patterns of word association, and we'd expect this to be reflected in a semantic analysis of partisan corpora. For this reason, partisan news sources are more likely to show differences in the meanings behind these terms. Our primary hypothesis is that this difference should be apparent in the comparison of the relationship between these two words in both a progressive-oriented and conservative-oriented corpora. Both closer relatedness and greater overlap in the semantic neighborhoods surrounding "global warming" and "climate change" should be seen in the progressive corpus than in the conservative corpus. We should also find a preference for global warming over climate change for conservatives, clarifying the divergent results found between Akerlof and Maibach (2011) and Schuldt et al. (2011).

Method

Corpus Construction

To make a comparison of ideologically driven views of global warming and climate change in online media, we assembled two corpora composed of articles drawn from the websites of publications that typically reflect polar opposites of the right-left political spectrum. Articles for the conservative corpus (5006 articles, 4280346 words) were selected from sources such as *The National Review*, *Fox News*, and *The Wall Street Journal*; the progressive corpus (5859 articles, 5249624 words) were selected from sources including *Think Progress*, *The Nation*, *Mother Jones*, and *The Huffin-*

gton Post. To be included in either corpora, articles had to include either "global warming" or "climate change" in the body of the text. The primary topic of the articles from which the text was drawn are not necessarily primarily about global warming or climate change, however each use of those terms contributes a useful context within which they appear and from which meaning can be drawn. Extra text, such as the comments attached to an article, were removed. All articles were published between January 1, 2008 and December 31, 2012. These data include a period of increased skepticism regarding climate change, due in part to the "Climategate scandal" in 2009, and the beginning of a shift within the United States in preferences between the use of global warming and climate change (Hoffman, 2011; Grundmann & Scott, 2012).

The corpora were cleaned to remove misspellings (many were proper nouns). Words with especially low frequencies (< 100) were also removed from the corpus, because low frequency words tend to have inflated measures of relatedness in subsequent modeling (Bullinaria & Levy, 2007). In both cases, omitted words were replaced with a dummy symbol so the distance between words remained constant. The terms "climate change" and "global warming" were converted into single tokens by removing the space between them.

HAL Modeling

Two HAL matrices were generated from the corpora. We used the implementation of HAL and the distance metrics provided in the *S-Space* package described in Jurgens and Stevens (2010). The window size used was seven words wide. While other analyses using HAL (e.g. Murphy et al., 2012) have used a larger window size of ten, seven was chosen because it was a compromise between evidence that matrices using smaller window sizes tend to perform better and the possibility that a larger window size might make up for the relative small size of the corpora. The dimensionality of the resulting vectors was reduced to the 2000 dimensions that contained the greatest amount of information as evaluated by an entropy metric included in the *S-Space* package. While Bullinaria and Levy (2007, 2012) demonstrate little to no penalty for including more dimensions, 2000 were used because the number offered a good compromise between matrix size and outcomes.

Data and Analysis

Vector cosine angle distances were used as the distance metric. *S-Space* has implemented cosine angle distance and euclidean distance metrics, and some research by Qian, Sural, Gu, and Pramanik (2004) has shown that they perform similarly when applied to nearest neighbor queries. Because our topics of interest are climate change and global warming, we computed the cosine between the vectors for those tokens and every other word in each corpus (after data cleaning, there were 8370 unique tokens in the corpora). These cosines were in turn standardized through a z-transformation to aid in interpretation by making the units more comparable (a similar z-transformation was done in Murphy et al., 2012).

Results

Climate change is the favored term by frequency in the progressive corpus, used almost twice as often as global warming (63.7% vs 36.3%). In the conservative corpus, they are used in roughly equal proportions (53.1% vs 46.9%) with a slight edge to climate change. These results, at first glance, conflict with the findings of Schuldt et al. (2011), who found in their partisan think tank corpus an advantage for global warming use among conservatives. However, it is interesting to note that progressives rarely use global warming in isolation: only 15% of the time, on a by-document basis. Conservatives on the other hand will use global warming in isolation in 31% of documents. Progressives generally favor using both, 50% of the time, and conservatives use both in only 27% of documents. A test of independence finds that $\chi^2(2) = 592.2$ ($p < 0.0001$) with Cramer's V$= 0.235$, indicating a moderate dependence between ideological grouping and their choice of terms.

The HAL analysis showed that global warming and climate change were each others' closest associates (conservative corpus: $z = 2.26$; progressive corpus: $z = 2.57$) in both corpora. This suggests that their meanings were indeed considered to be very close–at least closer than any other tokens in the corpus. This does not preclude the possibility that there are still significant points of divergence between the two tokens. To test the degree of overlap between them, we created a list of words that was constituted from a union of the 100 nearest neighbors to both climate change and global warming within each corpus. We used this set of words because it would allow us to look at a meaningful comparison between relatively close associates (we're most interested in meanings that are relatively close to at least one of the targets) without cherry picking from the data. This resulted in a 115 word list for the progressive corpus, and a 127 word list for the conservative corpus (there was only a 75 word overlap between the two lists).

One sign that these nearest neighbors do not overlap completely is that the union of these two neighborhoods yield a list of 115 words for the progressive corpus, and 127 words for the conservative corpus (further, comparing these lists across corpora only find 75 that overlap). This is a tentative first indication that these neighborhoods may have differing degrees of overlap. In Figure 1, in the left column we present a scatterplot of the distances between the nearest neighbors for both corpora. The correlation between the words' distance to climate change and global warming were $r_{(115)} = 0.76$ for the progressive corpus, and $r_{(127)} = 0.45$ for the conservative corpus. The progressive corpus showed a high degree of overlap between the two terms, whereas the relationship was weaker in the conservative corpus. These two correlation coefficients reliably differed, $z = 3.87, p < 0.01$. So while they both show a large degree of overlap, the progressive corpus shows a slightly stronger concordance within the nearest neighbors of climate change and global warming.

Another sign of the divergence between their neighbor-

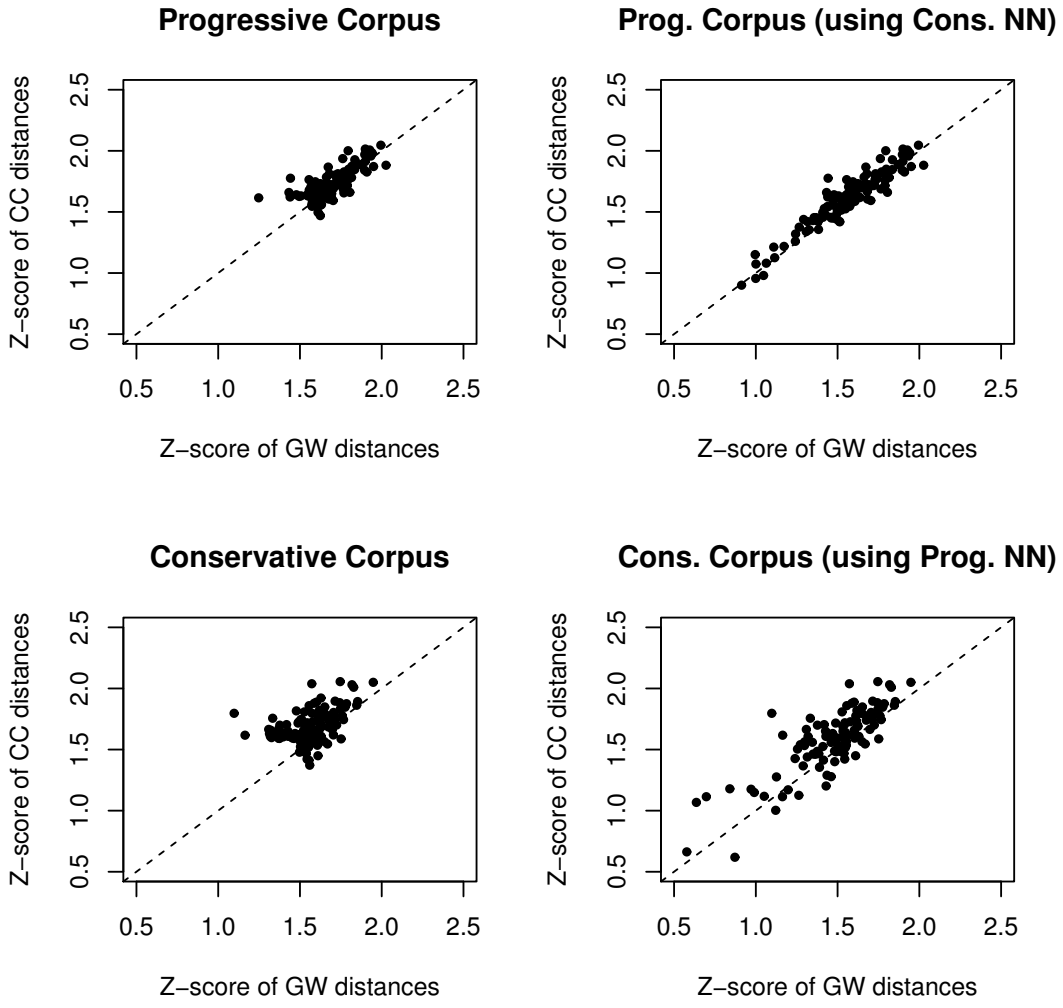

Figure 1: The left column plots the nearest neighbors to global warming and climate change. The right column shows the same relationship using the nearest neighbors for the other corpus, showing that many words that are highly related in one corpus are not as highly related in the other. In each plot, the dashed line signifies the line along which scores that are the same would follow.

hoods can be see in the two plots on the right in Figure 1. These show the relationship between the distances to global warming and climate change using the other corpus's wordlist. There was a very high relationship between the terms when looking at the conservative nearest neighbors in the progressive corpus ($r_{(127)} = 0.96$), and the progressive nearest neighbors in the conservative corpus ($r_{(115)} = 0.89$). While there are many words that overlap (75, as noted above), they show the addition of many words with far lower relatedness in each case, signifying the words that were close associates for one corpus, but not the other.

On a word-by-word basis, there are some interesting results, and some confirmations of findings seen in other literature. One of the larger comparisons between the corpora was that there was a difference of 1.33 standard deviations (z) between the average relatedness of the word "anthropogenic" to climate change / global warming, showing it to be much more closely related to those terms in the progressive corpus than in the conservative corpus. "Scam" was more closely related in the conservative corpus (diff = 1.13). Another interesting contrast was that "skeptical" was more highly related in the conservative corpus (diff = 0.63), while "skepticism" was more highly related in the progressive corpus (diff = 0.42). There was a closer relationship with "denier" in the progressive corpus (diff = 0.67). This makes sense because progressives will pair the target words with "denier," while conservatives will generally not label themselves as such. Similarly, they might express themselves as being skeptical, while progressives talk other people's skepticism.

Discussion

Our results confirm that while the meanings and nearest neighbors of "climate change" and "global warming" are closely associated, they do not perfectly overlap, and that this overlap is lowest within the conservative corpora. This fits with results from other sources that find few differences in progressive responses to questions about climate change / global warming, but larger number of differences in how conservatives will respond to those words (e.g. Villar & Krosnick, 2010). We didn't find the exact partisan differences in preference found in either Schuldt et al. (2011) or Akerlof and Maibach (2011). Progressives did tend to use climate change more often than global warming, but conservatives use them in roughly equal proportions. Contrary to Akerlof and Maibach (2011), progressives in our corpus used climate change more often. However, this may have arisen from the difference in samples. Akerlof and Maibach (2011) conducted a survey of general public preferences, not an analysis of media text. Journalists probably have a considered preference for particular terminology, and they don't have the option of saying "no preference" when choosing a term to include in an article.

However, we note that there were far more conservatives who *exclusively* used global warming. It is possible that our corpus was not quite partisan enough to pick up on this difference. Some of our sources for the conservative corpus, like Fox News' website for example, may not be sufficiently editorialized. This is in comparison with the corpus used by Schuldt et al. (2011) which used articles from think tank websites that are nakedly trying to present a particular ideological viewpoint. In future work on this corpus, we will expand the number of media sources included and develop a set of objective measures that get at ideological commitment. A clear measure of ideology would also make it easier to include a "mainstream" media corpus to act as another point of comparison. The possible segmentation of the corpus into subcorpora that look at the semantic neighborhoods surrounding these words when they are used together or in isolation would give us some additional insights into the meanings of these words if they are being deployed strategically in order to influence readers. Expansion of the corpus is a priority for these reasons, especially because our corpora were already relatively small by the standards of these kinds of models (Bullinaria & Levy, 2007, 2012). Similarly, the results of the models can be improved by using better distance estimates. In addition, we included traditional articles, opinion pieces, and blog entries in the corpus. The differing writing standards may have systematic qualitative differences that should be accounted for. Blogs and editorial pieces are much more likely to argue for a particular viewpoint and are more likely to contain a signal that corresponds to ideology differences in language use.

The individual word results showing wide cross-corpora difference between the relatedness of words like "scam" and "anthropogenic" can be seen in a couple ways. First, they act as a sanity check that there is some meaningful contrast between the corpora that tracks with attitudes presented in other studies, like there being more of a focus on human causation for climate change / global warming amongst progressives. Second, they offer an opportunity to support or elaborate on the findings of other content analyses done with exhaustive human coding. Our investigation was narrowly focused on the contrast between "climate change" and "global warming," but these tools can be applied toward investigating many other relevant concepts. For example, Nerlich (2010) sought examples of SCIENCE IS RELIGION metaphors. Using HAL, we can see which corpus those metaphors may be more prevalent in, and what other concepts they may be related to. In this instance, there is a stronger relationship with global warming (but notably not climate change) in the conservative corpus for the words "cult," and "religion." Finally, such results can act as guides for survey construction. Part of our ongoing research is to use these models to guide research with live participants in order to both validate the model and expand our knowledge of the ways in which the meanings of these climate terms.

Using high-dimensional semantic analysis to examine climate language in varied media sources is informative because it provides insights into how language use varies according to political ideology. The corpora used in this study were drawn from news articles posted on the internet, and it's possible that this limits the generalizability of the conclusions that can be drawn from it. Grabe, Kamhawi, and Yegiyan (2009), for example, found that people with lower education levels had a harder time remembering information from offline and online print media than they did for information they learned over television. Highly educated participants showed the inverse result. One goal moving forward is to integrate transcripts from broadcast media into our corpora in order to account for the differences in information presentation between the mediums.

Conclusions

Our work extends the established understanding of how conservatives and progressive media use the terms "climate change" and "global warming." Our findings indicate that there is a greater disparity in the meaning of the words within the conservative corpus, and that there is greater polarization within the conservative corpus in regards to which term is used. The use of a high dimensional vector space to aid in the examination of the semantics of key words within the corpora is a step forward within the domain of climate communication, and promises to be a productive tool for fostering future research.

References

Akerlof, K., & Maibach, E. W. (2011). A rose by any other name ...?: What members of the general public prefer to call climate change. *Climatic Change*, *106*(4), 699–710.

Anderegg, W. R. L., Prall, J. W., Harold, J., & Schneider, S. H. (2010). Expert credibility in climate change. *Pro-*

ceedings of the National Academy of Sciences, *107*(27), 12107–12109.

Armitage, K. C. (2005). State of Denial: The United States and the politics of global warming. *Globalizations*, *2*(3), 417–427.

Boykoff, M. T. (2007). Flogging a dead norm? Newspaper coverage of anthropogenic climate change in the United States and United Kingdom from 2003 to 2006. *Area*, *39*(4), 470–481.

Boykoff, M. T., & Boykoff, J. M. (2007). Climate change and journalistic norms: A case-study of US mass-media coverage. *Geoforum*, *38*(6), 1190–1204.

Bullinaria, J. A., & Levy, J. P. (2007). Extracting semantic representations from word co-occurrence statistics: a computational study. *Behavior Research Methods*, *39*(3), 510–26.

Bullinaria, J. A., & Levy, J. P. (2012). Extracting semantic representations from word co-occurrence statistics: stoplists, stemming, and SVD. *Behavior Research Methods*, *44*(3), 890–907.

Burgess, C. (1998). From simple associations to the building blocks of language: Modeling meaning in memory with the HAL model. *Behavior Research Methods, Instruments, & Computers*, *30*(2), 188–198.

Burgess, C., & Lund, K. (2000). The Dynamics of Meaning in Memory. In E. Dietrich & A. Markman (Eds.), *Cognitive dynamics: Conceptual and representational change in humans and machines.* Mahwaw, N. J.: Lawrence Erlbaum Associates.

Feinberg, M., & Willer, R. (2013). The moral roots of environmental attitudes. *Psychological Science*, *24*(1), 56–62.

Grabe, M. E., Kamhawi, R., & Yegiyan, N. (2009). Informing Citizens: How People with Different Levels of Education Process Television, Newspaper, and Web News. *Journal of Broadcasting & Electronic Media*, *53*(1), 90–111.

Grundmann, R., & Krishnamurthy, R. (2010). The Discourse of Climate Change : A Corpus-based Approach. *Critical Approaches to Discourse Analysis Across Disciplines*, *4*(2), 113–133.

Grundmann, R., & Scott, M. (2012). Disputed climate science in the media: Do countries matter? *Public Understanding of Science*, *0*(0), 1–16.

Hoffman, A. J. (2011). Talking Past Each Other? Cultural Framing of Skeptical and Convinced Logics in the Climate Change Debate. *Organization & Environment*, *24*(1), 3–33.

Jaskulsky, L., & Besel, R. (2013). Words That (Dont) Matter: An Exploratory Study of Four Climate Change Names in Environmental Discourse. *Applied Environmental Education & Communication*, *12*(1), 38–45.

Joireman, J., Barnes Truelove, H., & Duell, B. (2010). Effect of outdoor temperature, heat primes and anchoring on belief in global warming. *Journal of Environmental Psychology*, *30*(4), 358–367.

Jurgens, D., & Stevens, K. (2010). The S-Space package: an open source package for word space models. In *Proceedings of the acl 2010 system demonstrations* (pp. 30–35). Association for Computational Linguistics.

Lakoff, G. (2010). Why it Matters How We Frame the Environment. *Environmental Communication: A Journal of Nature and Culture*, *4*(1), 70–81.

Lewandowski, G. W., Ciarocco, N. J., & Gately, E. L. (2012). The Effect of Embodied Temperature on Perceptions of Global Warming. *Current Psychology*, *31*(3), 318–324.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, *28*(2), 203–208.

Luntz, F. (2002). *The environment: a cleaner, safer, healthier America.* Luntz Research, Alexandria.

McCright, A. M., & Dunlap, R. E. (2011). The politicization of climate change and polarization in the American public's views of global warming, 2001-2010. *Sociological Quarterly*, *52*(2), 155–194.

Murphy, C., Burgess, C., Johnson, M., & Bowler, S. (2012). Heresthetics in ballot proposition arguments: An investigation of California citizen initiative rhetoric. *Journal of Language and Politics*, *11*(1), 135–156.

Nerlich, B. (2010). 'Climategate': Paradoxical metaphors and political paralysis. *Environmental Values*, *19*(4), 419–442.

Oreskes, N., & Conway, E. M. (2010). *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*. Bloomsbury Publishing USA.

Qian, G., Sural, S., Gu, Y., & Pramanik, S. (2004, March). Similarity between Euclidean and cosine angle distance for nearest neighbor queries. *Proceedings of the 2004 ACM Symposium on Applied Computing*, 1232–1237.

Schuldt, J. P., Konrath, S. H., & Schwarz, N. (2011). "Global warming" or "climate change"?: Whether the planet is warming depends on question wording. *Public Opinion Quarterly*, *75*(1), 115–124.

Trumbo, C. (1996). Constructing climate change: claims and frames in US news coverage of an environmental issue. *Public Understanding of Science*, *5*, 269–283.

Villar, A., & Krosnick, J. A. (2010). Global warming vs. climate change, taxes vs. prices: Does word choice matter? *Climatic Change*, *105*(1-2), 1–12.

Whitmarsh, L. (2009). What's in a name? Commonalities and differences in public understanding of "climate change" and "global warming". *Public Understanding of Science*, *18*(4), 401–420.