Of friends and foes: Communicating foreign policy alignment via statements of opinion at the United Nations General Assembly*

Zuhaib Mahmood
Michigan State University
mahmoo21@msu.edu

February 20, 2019
(Working paper): Please do not cite.

Abstract

The conventional wisdom in international relations holds that foreign policy speech is largely cheap talk, and by extension does not contain useful information about the underlying interests of states. Yet little empirical research to date has directly scrutinized this claim, largely due to a dearth of large-scale empirical data on speech and state interests. In this paper, I argue that leaders face a tradeoff between the costs of aligning with unfriendly peers and the incentives they have to pander to multiple audiences. I test this theory using original data scraped from the web on individual speeches delivered at the UN General Assembly, compared against voting patterns on UN General Assembly resolutions. The results show that on the most controversial issues, where camps of friends and foes are most clearly defined, countries communicate the alignment of their interests with other countries most clearly. Given the ubiquity of speech in international politics, relative to behaviors such as voting and conflict, this opens up new avenues for scholars and policymakers to optimize the informative value of foreign policy speech and gain a more robust understanding of the alignment between nations in international politics. Further, it has implications for both the study of diplomacy by showing that statements of opinion communicate useful information, and the study of international institutions by showing that providing a public forum for discussion is an important, informative function of such institutions.

*Parts of this paper were presented at the PoliInformatics workshop in Seattle, WA (August 2017), the Annual Meeting of the Midwest Political Science Association in Chicago, IL (April 2017), the Annual Meeting of the International Studies Association in San Francisco, CA (April 2018), and the Annual meeting of the American political science association in Boston, MA (August 2018).
In September of 2013, Hasan Rouhani gave what the *New York Times* called “the most widely awaited speech” at the United Nations General Assembly.\(^1\) This speech was widely received as signaling a potential transition in Iranian foreign policy—particularly toward the United States, its allies, and Israel—and more specifically a transition in the direction of moderation.\(^2\) This perception of moderation, notably *before* Rouhani had taken any tangible policy actions, resulted in a historic opening of diplomatic between the United States and Iran via a 15 minute phone call between the two countries\(^3\)—and led in part to the capacity for Rouhani to negotiate a historic international agreement on his country’s nuclear program in a way his more hardline predecessor could (or would) not.\(^4\)

The co-incidence between Rouhani’s shift in speech with an eventual foreign policy change points to a deficit in the study of diplomacy in international politics. Namely, it suggests that these types of speeches—otherwise innocuous statements of opinion, interests, and priorities—can communicate important information about the landscape of foreign policy interests among nations. In this case, the shift in Rouhani’s speech *aligned* him closer with the US and the rest of the international community, relative to his predecessor whose rhetoric tended to *differentiate* him from them. This seems at odds with a range of previous literature, which has suggested that statements of this nature by politicians at the United Nations likely do not contain useful information (*Voeten*, 2005; *Farrell and Gibbons*, 1989).

The question of how speech patterns could potentially map to future behavior is of critical importance. Especially on high-stakes issues such as nuclear proliferation, policymakers and diplomats would benefit from understanding how to interpret statements of opinion, which are otherwise prolific in international politics and often occur on a continuous basis. For scholars of international relations, understanding this type of information in a systematic way allows for a more expansive study of

---


\(^2\)"Obama tells UN that diplomatic path must be tested with Iran“. The Guardian. Julian Borger and Ed Pilkington. 24 September 2013.

\(^3\)"Obama holds historic phone call with Rouhani and hints at end to sanctions“. The Guardian. Dan Roberts and Julian Borger. 28 September 2013.

\(^4\)"Former President Ahmadinejad reacts to Iran nuclear deal“. Iran Front Page News Service. 16 July 2015.
diplomacy and international communication—particularly at international institutions such as the UN.

In this paper, I examine the conditions under which speeches at the UN General Assembly can communicate useful information about the alignment pattern of foreign policy interests. I argue that leaders face a tradeoff between the the cost of aligning with unfriendly peers and the incentive they have to pander to multiple audiences. An analysis of UN General Assembly speeches between 1984 and 2012 shows that on more controversial issues, where a history of disagreement separates between friends and foes more easily, the alignment of speech patterns map most clearly on to the alignment of behavioral patterns, suggesting that in these cases leaders restrict themselves to speaking mostly with friendly peers.

**Speech, foreign policy alignment, and the United Nations**

A rich literature exists in international relations on how states communicate. From the credibility of threats (Schelling, 1960; Waltz, 1979; Guisinger and Smith, 2002; Fearon, 1994, 1995; Sartori, 2005; Kydd, 2005; Press, 2005), to the alignment of policy interests (Morrow, 1994, 2000; Crescenzi, 2007), to the salience of a demand (Trager, 2011), there are a range of things that leaders need to communicate to other state leaders. This communication generally takes the form of either actions or speech, such as conflict and troop deployment in the first category, or verbal threats and verbal demands in the second category.

While the scholarly interest in these types of communication has grown over the past several decades, a specific form of communication has remained largely under-theorized: public expressions of opinion and sentiment. This is particularly important since these types of speeches are so prolific in politics writ large: from media interviews to the United Nations, state leaders communicate an almost continuous stream of opinions and sentiment over time to a host of audiences on a range of issues. While there has been a fair amount of work done on communication within states by political elites
(Monroe, Colaresi and Quinn, 2008; Quinn et al., 2010; Proksch and Slapin, 2012), including some theorizing on the data generating process thereof (Budge, Robertson and Hearl, 1987), there has not been an analogous effort in international politics to understand speeches between states. Not only is this type of speech prolific in international politics, its informative value is often assumed in many scholastic studies of international politics, either in case study narratives (Press, 2005; Thompson, 2009) or as a variable of interest which affects other processes Thyne (2009); Hayes and Guardino (2013).\footnote{Ramsay (2011); Bils and Spaniel (2017) are notable exceptions; they provide promising frameworks to understand how uncertainty over the alignment of interests can lead to credible communication, both in terms of policy coordination and preventing militarized conflict.}

This informative value of speech, however, has not been subject to systematic testing. More explicitly: what, if any, information is being communicated in these expressions of opinion and sentiment? To better understand this, we can look to the literature on domestic party politics. Proksch and Slapin (2012), for example, argue that public debates serve to “communicate...policy positions to their parties, other parties, and voters”, arguing that parties have incentives to “maintain [their] brand” by “preventing [members] from undertaking activities that contradict the party’s primary message” (p.521-522). In this way, the choice to prioritize some issues over others – for instance, emphasizing symbolic and social issues near an election (Quinn et al., 2010) – or the choice of one “brand” over another “brand” (in terms of the sentiment they express on an issue) can serve an informative purpose.\footnote{This is also similar to Trager’s (2011) theory of multi-dimensional diplomacy, wherein he argues that during bargaining, leaders can communicate issue salience by prioritizing some issues over others—for example, choosing to demand a more difficult concession over an easier one, suggesting that the leader is willing to risk negotiations falling apart and potential war. While his theory exists in the context of private negotiations and the communication of resolve, it is directly applicable to public deliberation and the communication of alignment, in that the choice to emphasize some issues over others reflects a prioritization that differentiates one leader/country from another.}

Yet international relations is markedly different than domestic politics. Most importantly, formal political parties do not exist the way they do in countries, and so the eponymous “brand” analogy made by Proksch and Slapin (2012) breaks down since these brands ultimately do not exist in a structured way.\footnote{Lake (2009) provides a compelling counterargument to this, where he argues states do in fact exist in “hierarchies”. An interesting theoretical link exists here in bridging the literature on party politics and international relations, with}
to deliberation at forums like the United Nations, where there are only “a shallow set of common values... [and] obvious incentives to misrepresent their positions, as the stakes are clear and the relevant actors few”. This would seem to render it difficult to map particular opinions and argument to some kind of identifiable position (p.537). Ultimately, without some kind of *supra-national* lexicon by which to identify a particular country’s branding, or without a mechanism linking statements to a clear incentive structure, the conventional wisdom among international relations scholars is generally that these statements ultimately have little informative value.

However, there are two areas for improvement. First, especially in public arenas like the United Nations, the number of “relevant actors” are in fact quite large. These public arenas broadcast opinions not only to other state leaders, but also to interested political elites *across the world*. Be they NGOs, domestic elites, or elites in other countries, the scope of relevant actors is large and heterogenous. Since a given speech is finite in both size and content, a leader must pick and choose between audiences – creating a strategic environment and incentive structure. Second, while there is no *supra-national* lexicon to objectively identify any given speech, there is an *endogenously defined* lexicon by which to compare speeches to one another. More explicitly, while the absolute content of a speech may not be meaningful, the relative content can be: an emergent lexicon exists as leaders speak on any given issue over time, providing a medium in which to compare any set of speeches.

The UN General Assembly is the most natural starting point for analyzing how statements of opinion and sentiment on foreign policy communicates foreign policy interests, and how this maps respect to messaging and branding.

---

8 For arguments defending the UN’s discursive and persuasive roles, see Johnstone (2003, 2005); Krebs and Jackson (2007).

9 While a more detailed, descriptive lexical analysis of the UN is beyond the scope of this paper, there is useful work on how the language of politics changes over time (Greene, Park and Colaresi, forthcoming). For a simple example, consider the Dynamic Topic Model which I describe later in the paper. One of the uncovered topics is consistently composed of words like “international”, “tribunal”, and “law”, which based on post-hoc domain knowledge likely maps on to an international law topic. Starting in the middle of the dataset, toward the early 2000s, the terms “rome” and “icc” begin to appear. This is almost certainly due to the fact that the Rome Statute establishing the ICC as an arbiter and player in international law was adopted (1998) and implemented by 2002. In this way, the emergent lexicon surrounding the issue of international law shifted slightly to reflect changes in the relevant players. To cite the ICC during an ongoing crisis, then, becomes associated with the politics which surround the institution–and the countries who invoke it during that crisis.
onto observable behavior. Since its inception in 1945, scholars of international politics have studied the UN from a variety of angles, including its role in legitimizing foreign policy (Claude, 1966; Voeten, 2005; Thompson, 2009; Chapman, 2011), the Security Council’s authorization of collective foreign policies (Fortna, 2008; Beardsley and Schmidt, 2012), and with respect to UN General Assembly voting Alker (1964); Alker and Russett (1965); Vincent (1971, 1972); Kim and Russett (1996); Voeten (2000); Bailey, Strezhnev and Voeten (2015).

The General Assembly in particular is useful here for several reasons. First, it includes all countries in the world. This membership scope provides a unique opportunity to observe statements of opinion in a context where all countries speak on similar issues, and in a similar environment with a commonly understood emergent lexicon at any given time. This makes cross-country comparison more robust. Second, the wide membership combined with the General Assembly’s high profile provides a large set of “relevant actors” (to use Voeten’s (2005) terminology) from which leaders must pick and choose an audience. This creates a strategic environment and incentive structure whereby leaders must prioritize some interests of some audiences over others. Third, the UN General Assembly provides a very useful empirical opportunity to map speech patterns to relevant behavior, issue by issue. In this case, as I discuss in more detail below, the UNGA provides an opportunity to map speech patterns onto foreign policy behavior across knowable issues by using voting patterns to measure foreign policy behavior. Finally, the General Assembly is useful from a theoretical perspective in this study, since it has little if any true legal power. In other words, the null hypothesis and conventional wisdom – namely, that speech patterns reflect a babbling process, conveying little or no systematic information about foreign policy interests – should be favored most strongly in a forum like the UN General Assembly, which might otherwise be seen as merely a “contemptible talk-shop” (Claude, 1966, p.372).
A scale of two interests: Alignment through relative content

When leaders speak, audiences listen. Though this axiom has mostly been used to understand how domestic audiences can make threats more credible (Fearon, 1994; Weeks, 2008; Levendusky and Horowitz, 2012; Tarar and Leventoğlu, 2013; Kertzer and Brutger, 2016), there is also work to suggest that it applies more broadly to international audiences as well–both overseas intra-state audiences (Hayes and Guardino, 2013) and other countries (Sartori, 2005; Ramsay, 2011). In a high-profile setting like the UN General Assembly, this means that all public actions–such as voting and speeches–are subject to this simple but important axiom.

What, if any information can these audiences glean from speech, and why would leaders care? A critical issue for onlookers to resolve from these speeches is that the underlying scale is, *a priori*, undefined. For instance, consider three countries, A, B, and C. This issue can be illustrated by a hypothetical situation where A states that he is a “7" on the issue at hand. Without additional information, the content of that speech is essentially indeterminate, since as Voeten (2005) points out there is only a shallow “common set of values” in which to interpret this statement. That is, a “7” could mean vastly different things depending on the context: compare Figure 1 and Figure 2, where the same statement maps A on to very different sides of the ideological scale (holding B and C constant in their own positions).

![Figure 1: Three countries speaking in similar ideological spaces](image1)

![Figure 2: Three countries speaking in very different ideological spaces](image2)

Drawing from the terminology used earlier, this problem occurs from reading the *absolute* content of speeches: a “7” has little to no meaning in an absolute sense without some clarifying set of values in which to interpret it. However, as Figure 1 and Figure 2 allude to, A’s speech does still contain *relative* content. That is, it remains possible to make robust statements about $\Delta AB$, $\Delta BC$, and
ΔAC. For a more concrete example, if A and C emphasize their “concern about the humanitarian situation” in some civil war i, and B tends to emphasize his concern about “maintaining institutional stability” in the same civil war i, while it is difficult to say whether A, B, and C are different in some objectively meaningful way, it is much easier to say that ΔAC is smaller than ΔAB—meaning that A and C are more aligned in their foreign policy speech. The informative nature of difference and similarity is not a novel ideal in the study of diplomacy: Jervis (1970, p.21) argues a similar idea, noting that “all secret messages and most public ones are signals...these signals may involve following an established routine, as in the use of diplomatic language, or may entail breaking that pattern” (emphasis added). This pattern is defined by the community at large, or by a sustained pattern over time, subject to change.

Interestingly, under this framework the information about alignment in speech depends only on content, and requires no additional statements about the truth or falsehood of that content. For instance, whether or not A is “truly” concerned about the humanitarian situation in i has no bearing on the fact that in his speech, A is communicating a closer alignment to C than he is to B. This is a critical takeaway from the theory presented here: there might be valuable information contained in speeches even if leaders are not faithfully representing their true interests. The emphasis on “true” here is important, since it is often a red herring that can obfuscate the process of gleaning information when leaders deliver public statements of opinion.

Also noteworthy is that the scales are defined at the bounds by the most polarized content: given some endogenously defined vocabulary, countries are identified in alignment relative to one another in a space defined by those countries who speak most differently than one another on that particular issue. This leads to interesting theoretical opportunities to study patterns of polarization more broadly in international politics, though this is beyond the scope of this paper.
Mapping speech to friends and foes

Political actors care about other leaders’ interests. For domestic actors, there is a lingering concern that leaders will forego the interests of the public in order for a leader to pursue his/her own private interests (Stasavage, 2004; Svolik, 2012; Colaresi, 2014), or more generally that a leader may forego the interests of one domestic group in favor of another. For international actors, a similar concern applies: decisions about foreign policy making are driven by the anticipation of what other leaders will do (Findley and Teo, 2006), and more specifically by the anticipation of whether these other leaders will implement favorable (or “preferred”) policies (Singer, 1963). Learning about the alignment of interests becomes a vital part of coordinating foreign policy, particularly when the stakes involved with pursuing a difficult or risky policy–where counter-action might be extremely costly (Gent, 2007)–are very high. As Bils and Spaniel (2017) argue, uncertainty over ideological positions can make cheap talk a useful component of pre-crisis bargaining and the ultimate decision to go to war. They show that the presence of many potential opponents with disparate policy preferences can render cheap talk effective, and more importantly that the variance of perceived alignment in latent ideological space can have critical implications—including, in some cases, whether war breaks out. Thus, there is wide agreement in the scholarly literature that the audiences who are listening to this relative content care about (and can potentially use) the information they glean.¹⁰

Connecting relative content to alignment, this creates a decision for leaders to make. Consider, for example, a simple setting where two interested onlookers—CA and CB—are watching A and B. Assume that A and B disagree on some issue i, and that their behavior reflects a decision to reflect the interests of CA or CB respectively, whose preferences are conflicting on issue i. There are two possibilities with regards to the speech patterns of A and B on issue i, given their conflicting behavior on that issue. The first possibility is that A can speak in a way that also differentiates him from the way B speaks on that issue. Under this scenario, A is using his speech to double down, or

¹⁰ Also see (Ramsay, 2011), who argues that pre-crisis signals about potential alignment of interests can affect peace by leading to bargaining in the first place.
reinforce his disagreement with B. The second possibility is that A can speak in a way that aligns him with the way B speaks on the issue. Under this scenario, A is using his speech to hedge against his disagreement, or pander to the interested onlooker C_B who sees A siding with a foe.

The key, however, is whether C_B is able to identify that A is in fact siding with a foe. Here, variation across the nature of issue i becomes relevant. On the most controversial issues, there is a consistent pattern of disagreement among nations: on these issues, the prevalence of disagreement creates the ability to separate between friends and foes. For example, on the issue of the Middle East, there is a consistent pattern of disagreement between countries on multiple possible questions. This creates separation: determining whether Libya and Iran are a friend or foe to the United States on this issue is relatively simple. Similarly, returning to Hassan Rouhani’s speech in 2013, his rhetoric on the nuclear issue represented a marked change in alignment patterns from his predecessor, calling international concerns “reasonable” and stating “unambiguously” that Iranians sought peaceful technology and embraced “international cooperation” and “transparency”.\textsuperscript{11} Compare this to rhetoric in 2012, where on the nuclear issue then-president Mahmoud Ahmadinejad focused on “intimidation by nuclear weapons... by the hegemonic powers”, the “continued threat by [Israel]”, and the “state of mistrust...[where] no one feels safe.”\textsuperscript{12} This shift was radical in its stark differentiation from his predecessor and its move toward reconciliatory language, communicating information on the new regime’s priorities and foreshadowing the eventual cooperation that ended up occurring.\textsuperscript{13} More importantly, the high-stakes nature of this issue meant that the combination of a radical shift in rhetoric and the intense polarization that had arisen around the issue created an environment where this shift in language communicated a substantial amount of information to the international community.

Thus, on these controversial issues, there is much more possible information to be gleaned from speeches, in that when A speaks then C_B is more readily able to determine whether A is speaking

\textsuperscript{11} Transcript accessed from The Times of Israel website. Accessed December 2018.
\textsuperscript{12} Full transcript available here.
\textsuperscript{13} Interestingly, the subsequent termination of the Iran agreement by the Trump administration in May of 2018 (source: BBC online) is also predictable from another shift in rhetoric, though this time originating from the United States.
similarly or differently than friends and foes on that issue. Contrast this with much less controversial issues, where the information is more muddled: when nations generally agree on an issue—such as world peace in more abstract terms—the differences in $\Delta AC$ or in $\Delta AB$ are less informative, since it remains unclear who is a friend or foe of whom (or if there are any foes at all). Further, in these cases politicians should have a less restrictive environment in which to speak freely—for instance, to pander to multiple audiences, or to explore the ideological space more freely. This should result in more randomness with regard to who countries are aligned with at any point in time, regardless of their eventual behavior. Thus, controversial issues should be more informative overall, in that speeches in these environments should map most closely with eventual behavior.

**Hypothesis:** Higher controversy issues will see the most reinforcing speech.

**Empirical strategy**

To measure speech patterns, the data comes directly from the United Nations General Assembly, where I have scraped the transcripts for all plenary sessions in PDF format, translated them to text, and identified each speech with its corresponding country. This corpus was collected independently of the one by Baturo, Dasandi and Mikhaylov (2017), whose corpus focuses only on speeches delivered by heads of state under *General Debate*.\(^{14}\) To measure behavior corresponding with those speeches, I rely on the raw text of UN General Assembly resolutions; I use a scraper built by Erik Voeten\(^{15}\) to extract the full text of resolutions starting in UN General Assembly session 48, corresponding to the year 1993, and I collect the remaining resolution text back to 1984 myself. Thus, the time frame for this paper is 1984-2012.\(^{16}\)

The goal of this paper is to measure whether or not similarities and differences in speech map onto similarities and differences in observed behavior, and the direction in which this mapping occurs

---

\(^{14}\)For example, their data from the year 2000 (session 55) includes the speech made by Madeleine Albright in the 10th plenary meeting of the 55th session, but leaves out President Clinton’s address during the 3rd plenary meeting, which occurs outside the confines of “General Debate” and under different agenda topics.

\(^{15}\)Email correspondence, 5 May 2017

\(^{16}\)Voting data is only mapped to resolutions through 1984; see previous email correspondence.
(inverse or direct). The dependent variable, therefore, is the similarity of voting on some resolution between any two countries \( A \) and \( B \). This requires some specific decisions, however. There are 4 possible ways a country can act on a particular resolution: a *yes* or *no* vote, an *abstention*, or an *absence*. To collapse this into a binary decision, I code *abstentions* as *no* votes, and I treat *absences* as missing data.\(^{17}\) Finally, since the purpose here is to predict similarity and differences in votes, I convert these to an [undirected] dyadic dataset covering all dyads, where \{1\} indicates disagreement in voting within the dyad and \{0\} indicates agreement in voting. Finally, given that the covariates of interest (controversy, speech alignment) vary at the level of topic-year, I collapse all observations to the topic-year unit of analysis, with the dependent variable of interest being the count of disagreements and agreements on all resolutions within a particular dyad.

The independent variable of interest is the similarity between any two countries’ speeches. There has been a growing literature in measuring similarities and differences from political text (Grimmer and Stewart, 2013), including with applications to measuring similarities and differences between groups such as countries, political parties, or party manifestos (Budge, Robertson and Hearl, 1987; Laver, Benoit and Garry, 2003; Slapin and Proksch, 2008; Monroe, Colaresi and Quinn, 2008). To measure differences between countries on speech, I draw upon previous work, drawing from the aforementioned literature and treating words as data, to estimate ideal point estimates drawing from the text of the UN speeches (Mahmood and Colaresi, 2016, 2017). I estimate ideal point measurements using a dynamic linear model in the form:

\[
\begin{align*}
y_{jt} & = \begin{pmatrix} z_t \ x_{jt} \ w_{jt} \end{pmatrix}_{(N \times 1)} + \begin{pmatrix} \epsilon_{jt} \ \end{pmatrix}_{(N \times 1)} \\
x_{jt} & = \begin{pmatrix} A \ x_{jt-1} \ \end{pmatrix}_{(p \times 1)} + \begin{pmatrix} e_{jt} \ \end{pmatrix}_{(p \times 1)}
\end{align*}
\]

where \( N \) represents each word spoken in the full vocabulary of all speeches and \( p \) represents the number of dimensions in which the ideal point estimates \( x \) are estimated per country over time. In

\(^{17}\)The raw voting data also includes a placeholder for countries that are not members of the international system at the time of voting. I treat these as missing as well.
this model, \( y_{jt} \) represents a noisy projection of latent ideal points \( x_{jt} \) in \( p \)-dimensional space, where the state equation represents the evolution over time of \( x_{jt} \), for country \( j \) at time \( t \). I use Euclidean distance as the distance metric between the estimated positions; in the main body of the paper, I present results from a 2-dimensional model.\(^{18}\)

**Mapping speeches to resolutions**

Mapping speeches to corresponding resolutions is a complicated task. For example, a resolution can be on different topics, from nuclear proliferation to the Middle East to a UN humanitarian mission. However, a single speech can encompass multiple topics. This makes it difficult to generalize from a broad, multi-topic speech to a single resolution. To rectify this problem, I separate each General Assembly speech into paragraphs. Next, I use Latent Dirichlet Allocation (LDA) (Blei, Ng and Jordan, 2003) to classify paragraphs and resolutions into distinct topics. LDAs classify documents – in this case, paragraphs and resolutions – into topics via a mixed-membership model, where a document can be composed of many topics at once. The model does so by estimating \( pr(topic|word) \) across all words, allowing for post-processing by the analyst to label topics (Quinn et al., 2010). For this project, I use a 20-topic model and classify resolutions and paragraphs into the topic with the highest probability.\(^{19}\) I follow Blei, Ng and Jordan (2003) and stem terms to root as well as remove stop-words such as “the” and “and”.\(^{20}\)

Figure 3 shows the count of resolutions across topics, separated by those during the Cold War period (recall, back to 1984) and those after. Figure 4 shows the total number of words spoken on each topic, aggregated across time. Finally, using the measure of controversy as described in the

---

\(^{18}\)The results remain substantively similar at 3 dimensions, as well as if instead of this model I use Proksch and Slapin’s (2012) WordFish scaling model.

\(^{19}\)For example, a 20-topic model might return a vector \([0.0001, 0.0001, 0.0001, 0.15, 0.0001, 0.0001, 0.0001, 0.20, 0.0001, 0.0001, 0.0001, 0.0001, 0.0001, 0.0001, 0.0001, 0.0001, 0.0001, 0.0001, 0.0001, 0.0001, 0.65, 0.0001, 0.0001]\), representing a single point on the 20-dimensional simplex. In this paper, I would classify this document into topic 17. I remove any resolutions which are less than 10% composed of any single topic.

\(^{20}\)I opt to use the LDA model as opposed to the Dynamic version, DTM, for two reasons. The first is that a comparison of LDA and DTM for the years 1993-2014 yielded extremely similar results; relatedly, then, the second reason is that the DTM is far more computationally intensive. I calculate LDA using the GENSIM package in Python 3.
next section, Figure 5 shows the average level of controversy—across all years—per topic, with bars representing the 25th and 75th percentiles of controversy.

![Figure 3: Number of resolutions, by topic](image)

![Figure 4: Word count, by topic](image)

**Measuring levels of controversy**

Recall that one of the predictions of the theory is that the direction of relationship between speech patterns and voting patterns depends on the strategic incentives present on a given issue. To operationalize these incentives, I use the proportion of dyads who disagreed, constructed for each
topic, to rank order topics every year by their respective levels of controversy (using voting patterns from the previous year). This provides a scale of least to most controversial issues.

To justify this, consider the role of controversy in the theory presented earlier. Controversy specifically pertains to the ability to identify two countries $A$ and $B$ as either tending to be friends or tending to be enemies. This is a difficult concept to measure in any objective way. While a thorough discussion of identifying objective enmity or friendship between countries is beyond the scope of this paper, I focus primarily on the UN as a way to most closely approximate friends and enemies in the context of the UN. To do this, I consider each resolution $R_i$ as representing a single possible question in the broader array of possible questions that any given issue $i$ could potentially produce. It is reasonable, then, that as a given issue $i$ sees a larger and larger tendency for countries to disagree on more and more questions, it can be considered more “controversial”. On these issues, countries are displaying a tendency to have more polarized views—such as the Palestine issue above—whereas on less controversial issues the sorting between friends and foes is much less clear. Most relevant to the research design, it also provides a rigorous way to measure a yearly, issue-by-issue level of controversy at the UN General Assembly over time, mapping those to corresponding speeches and resolutions in order to test the theory presented earlier.

21 For example, in the year 1995, the measure of controversy would be created by calculating the proportion of dyads who disagreed, based on all dyads in the dataset within 1994, for each topic.

---

Figure 5: Controversy level, by topic
**Research design**

The proposed hypothesis reflects an interaction between speech patterns and the level of disagreeableness on an issue. I model this relationship as follows:

$$Y_{ABi} = \alpha + \beta_1 X_{sp} + \beta_2 X_{disag} + \beta_3 X_{sp} X_{disag} + \beta_c X_c$$

Where $Y_{ABi}$ represents the count of disagreement across the total number of resolutions on a particular topic $i$ in a particular year, between countries $A$ and $B$. $\beta_c X_c$ represent additional coefficients and linear parameters included as control variables, which I discuss further below. Given that the goal is to model the probability of any two countries disagreeing, I use a binomial regression to model resolutions as a series of Bernoulli trials, using a logit link to transform $Y_{ABi}$ to continuous space.

To gain as accurate an estimate as possible for evaluating whether the theoretically predicted patterns actually show up in the data, I include several additional indicators for modeling the underlying relationships of countries to the observed outcome of voting. Table 1 provides an overview of these variables, categorized into a rough grouping of *Structural*, *Interdependent behavior*, and *Unobserved Variance*. The first category captures those dynamics which the literature has understood to affect the calculus of whether two leaders will pursue similar policies, such as material power differentials (Voeten, 2001), alliance ties, or joint-democracy (Oneal and Russet, 1997; Gartzke, 1998).

The second category of *interdependent behavior* is very important and worth elaborating upon. Given that the research question examines whether *speeches* at the UNGA are linked to *votes* at the UNGA, an important caveat is to emphasize that often, both activities are carried out in
blocs. It is highly common practice at the UNGA for blocs of countries—such as the Arab League, European Union, Non-Aligned Movement, etc.—to deliver foreign policy speeches as a group, which by definition results in an observed similarity in their speech patterns, $X_{sp}$ in Equation 2. These groups are often formed on the very fact that these countries share a common foreign policy context, and so they must be accounted for when trying to link foreign policy behavior and speech. Therefore, during the process of parsing and organizing speeches, I identified speeches which were given “on behalf of” an identifiable group, or those which are given by a country “in their capacity as” the chairperson or leader of one of these groups. I include a binary variable for whether two countries are members of the same group, across all groups for which I have data.

As a matter of theory, however, it is worthwhile to note that this phenomenon complements the broader argument of this paper: namely, that foreign policy speech is systematically related to underlying foreign policy interests, and that the political incentives involved in crafting and delivering a speech are much stronger than previous research has claimed. Indeed, the very fact that these communities take the time and resources to agree upon and craft joint statements suggests that these speeches are vehicles for communicating some kind of information relevant to that foreign policy community. Moreover, a surprisingly common phenomenon is to see countries who are not members of a group signing onto statements—such as former Soviet states or aspiring EU states signing on to statements delivered by the EU. Thus, while I do take these groups into account empirically, it is noteworthy that the presence of these groups is a direct, explicit result of the argument presented in this paper. Future research on the dynamics of foreign policy speech would do well to examine the politics of crafting these speeches, to see how these politics balance between sending reinforcing and compensatory (or clarifying) content with respect to the foreign policies carried out by member countries.

Finally, the third category of Unobserved variance includes variables like the year of the resolution, and whether one or both countries are either P3 (United States, UK, and France) or P2 (China, Russia), representing the 5 major powers at the UN Security Council. To understand the robustness of the theory, I also include fixed effects for the issue under discussion; these labels were generated.
post-hoc from the results of the Dynamic Topic Model discussed earlier. Importantly, while labels for these topics were carefully considered based on the lexical output from the DTM as well as domain knowledge of the United Nations, the labels of the topics themselves do not affect the fixed effects or model estimation (except for robustness checks when a particular issue was excluded).22

Table 1: Control variables, linking speech to UNGA voting patterns

<table>
<thead>
<tr>
<th>Structural</th>
<th>Interdependency</th>
<th>Unobserved variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alliance ties (COW)</td>
<td>Joint group members</td>
<td>Fixed effects (topic &amp; year)</td>
</tr>
<tr>
<td>Material power (CINC)</td>
<td>Proportion words delivered</td>
<td>UNGA ideal point difference</td>
</tr>
<tr>
<td>Democracy (Polity IV)</td>
<td>as member of group</td>
<td>(Bailey, et al 2015)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P3 (USA/UK/FRA)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P2 (RUS/CHN)</td>
</tr>
</tbody>
</table>

22 More explicitly, it would make no difference to the results presented in this paper whether the issues are labeled with their respective contextual labels, or whether they were simply labeled as issue 1, issue 2, issue 3, ... etc.
Results and evaluation

The results of the model are provided in Table 2. To present the interaction effects, I primarily rely on visualizations given the limited information conveyed by tables—particularly for Generalized Linear Models and interaction effects (Brambor, Clark and Golder, 2006). Figure 6 plots the change in conditional mean (with 95% confidence intervals) across levels of controversy, on the scale of the dependent variable, with all other variables held at their median values. The results are consistent with expectations that as the level of controversy increases, the slope of the relationship between speech difference and behavioral difference becomes more positive—indicating a greater incidence of reinforcing speech.23 One noteworthy part of this is the fact that this is true at the highest levels of controversy: above the 90th percentile. This is due to the fact that disagreement at the UN General Assembly is exceedingly rare: over 2/3 of resolutions are adopted unanimously, and on roll call votes disagreement is still relatively rare. Moreover, at lower percentiles of controversy—i.e. relatively low levels—the baseline probability of disagreement is much lower by definition, which further accentuates the slope differences at very low probabilities of disagreement.24 As a result, the variation of interest in issues occurs at the outlying cases; this raises important statistical questions about how to handle these outlying cases, but as a matter of appropriateness for testing the theory, the role of these extremes is quite important. It suggests that ultimately, the best information exists in these exceptional cases, which is useful for scholarly study in narrowing down which issues to focus on moving forward.

23I also run all analyses excluding the United States and Israel from the analysis, given their unique role with the topic of Israel at the UN over the past 70 years. I also run these models using both 2- and 3- dimensional ideal point models. Next, I also estimate the results using Proksch and Slapin’s (2012) WordFish model instead of the one presented above. Since their model does not include a dynamic component, I estimate a different model per year across all countries. Finally, I estimate all results using alternative measure for the theory: the UNGA ideal point differences. Since the theory states that reinforcing speech is a function of the identifiability of friends and foes, it stands to reason that we should expect highly reinforcing speech between countries with enduring and very high levels of disagreement at the UN General Assembly. This is best captured by ideal point differences. Thus, I also estimate a model varying speech patterns between low and high levels of UN ideal point differences (90th percentile difference). All results are substantively similar to those reported here.

24The same plot on the scale of the linear predictor, for example, shows a much more evenly distributed change in slopes from low to high controversy.
**Figure 6**: Change in conditional mean of Disagreement, over Speech differences, across levels of controversy (95% confidence intervals)
<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>−2.40***</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
</tr>
<tr>
<td>defense</td>
<td>−0.05***</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
</tr>
<tr>
<td>jointgroup</td>
<td>−0.52***</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
</tr>
<tr>
<td>CINCdiff</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
</tr>
<tr>
<td>UNGAipdiffCENT</td>
<td>0.73***</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
</tr>
<tr>
<td>Speech dist</td>
<td>−0.02***</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
</tr>
<tr>
<td>Controversy</td>
<td>10.46***</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
</tr>
<tr>
<td>jointdemL1</td>
<td>0.20***</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
</tr>
<tr>
<td>proptotalWCgroup</td>
<td>−0.76***</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
</tr>
<tr>
<td>oneP2</td>
<td>−0.05***</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
</tr>
<tr>
<td>jointP2</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
</tr>
<tr>
<td>oneP3</td>
<td>0.01***</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
</tr>
<tr>
<td>jointP3</td>
<td>0.48***</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
</tr>
<tr>
<td>Speech*Controversy</td>
<td>0.18***</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
</tr>
</tbody>
</table>

**Topic FE**  Yes
**Issue FE**   Yes
**Num. obs.**  2172726

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>4721220.70</td>
</tr>
<tr>
<td>BIC</td>
<td>4721862.87</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-2360559.35</td>
</tr>
</tbody>
</table>

* ***p < 0.001, **p < 0.01, *p < 0.05

**Table 2:** Interaction model, Speech difference and Controversy
Out-of-sample validation

Given the high number of observations and the strict assumptions of dyadic independence, there remain some questions about the validity of these results. Concerns about overfitting, for example, may lead us to wonder whether this model is only chasing noise; moreover, as discussed earlier, the rarity of disagreement and the skew of controversy levels—while theoretically and substantively justified—still leaves statistical questions open. To address this matter, I include an additional test: whether the addition of speech, interacted with the controversy levels, adds predictive power to classifying which countries will disagree on an issue and which will not. Evaluating the out-of-sample predictive power of the model can help alleviate concerns about overfitting (Colaresi and Mahmood, 2017), and can also provide an additional robustness check on whether useful information is being communicated in these speeches. Finally, it also serves as an additional check on the measurement itself, since a performance boost at minimum suggests that the measurement is providing some kind of useful information—even for those skeptical of its validity otherwise.

To perform this test, I split the data into two time periods: a training set during the time period 1984-2006, and a test set during the years 2006-2012. I present results from a 10-fold cross validation in the training set (performance in the held-out-samples), as well as results from predictions in the test set. This comparison adds an additional layer of validating whether the patterns in the data remain systematic across these two time periods. Finally, for ease of analysis, I collapse the dependent variable into a binary indicator of whether a particular dyad experienced any disagreement on any resolutions during a topic-year period. I conduct this analysis using the Caret package in R (Kuhn, 2016).

The results provide further validation that speech patterns add useful information to understanding foreign policy alignment. I present results from two models: a model including all variables except for speech, and the model presented above in Equation 2. Note that this is an extremely difficult baseline for performance: the null model is, in part, driven by Ideal Point differences estimated by (Bailey, Strezhnev and Voeten, 2015), which are derived from UN General Assembly votes themselves. We
would expect, therefore, that these measures would already provide an extremely good prediction for how countries vote on those same UN General Assembly resolutions. However, the results indicate that based on two metrics, precision-recall and the area under the Receiver-Operator curve, the addition of speech patterns adds to the performance of the model. The cross-validation results (10-fold, held-out samples) show similar patterns. Cross-validation results and test-set performance are provided in Figure 7 and Figure 8. This provides further evidence that there is useful information being communicated by speech; moreover, it is also inconsistent with the claim that the model presented in the previous section is overfitted, since in that case we would expect a drop in predictive capability (Colaresi and Mahmood, 2017, p.199). Thus, the addition of this test provides a useful complement to the analysis presented above, and adds evidence to the argument that statements of opinion like those at the UN General Assembly do in fact communicate some useful patterns of information about the alignment of foreign policy interests.

![Cross-validated ROC comparison](image)

**Figure 7:** Area under the ROC curve, cross-validated in training set (95% confidence intervals)
Figure 8: ROC curves and Precision-recall curves, test set.
Conclusion and Discussion

In summary, the convention wisdom that speech patterns contain no meaningful information about interests—measured by patterns of voting at the UN General Assembly (Bailey, Strezhnev and Voeten, 2015)—is not supported by the data. Rather, I present a theory which states that leaders can use speech strategically to communicate their alignment with other nations, where the ability to identify friends and foes drives the incentive structure which produces different patterns mapping speech to behavior. This has implications for the study of diplomacy and cooperation, both at the United Nations and in international relations more broadly. Specifically, while the United Nations is a vehicle for coordinating foreign policy (Abbott and Snidal, 1998), and it can reduce foreign policy costs with respect to things like military intervention (Thompson, 2009; Chapman, 2011), its role as a vehicle to communicate the alignment of foreign policy interests more generally has not yet been understood. Particularly on the most polarizing, controversial issues in international politics—which are also, coincidentally, the issues that international institutions tend to take on (Abbott and Snidal, 1998; Fortna, 2008)—this role of institutions in facilitating public diplomacy should be of interest to the scholarly community.

There are several limitations to this study. First, it relies on a measurement of ideal points derived from text, which is an extremely difficult inferential process. A substantial literature has been devoted to how best to estimate this information (Budge, Robertson and Hearl, 1987; Laver, Benoit and Garry, 2003; Slapin and Proksch, 2008; Monroe, Colaresi and Quinn, 2008; Elff, 2013), as well as both the prospects and limitations of text-based methods (Grimmer and Stewart, 2013; Haddi, Liu and Shi, 2013). Thus, there is no single, proven method of extracting an objectively best estimate for how countries align in their speech. Other observable behaviors—such as the pattern of grouped statements at the UN General Assembly, or variation in which countries speak most on which issues—which are less reliant on latent measurements of alignment can be of benefit to a more thorough assessment of the theory. Second, while UN General Assembly votes have been

<sup>25</sup>Chiba and Fang (2014) is an example of work that moves in this direction with respect to the UN. Other promising work on the topic of speech specifically includes Ramsay (2011); Bils and Spaniel (2017).
long established in the literature as being one possible measure of state interests and behavior (Bailey, Strezhnev and Voeten, 2015), the scope of this paper remains within the confines of the United Nations General Assembly. It does not, for example, speak to scholars who may believe that the entirety of the UN--votes, speech, laws, and all--is simply epiphenomenal or irrelevant to international relations as a whole. Thus, there remains additional work to be done in linking these speech patterns to behaviors outside the UN. Finally, the nature of the hypotheses in the paper, which are conditional and require extensive interaction terms, make them difficult to test reliably without fragility. This makes it all the more important to use alternative measures in concert with the ones presented here, in order to understand the nature of speech delivery at the UN.

However, this study does demonstrate that scholars interested in diplomacy and international institutions should take speech more seriously in international politics, both as a measure of interests and as a vehicle for diplomacy. The theoretical intractabilities which have surrounded the study of speech in the past can be overcome with an alignment-based approach, where the incentives and costs of giving particular types of speech are driven not by the absolute content of speech, but rather its relative content. I show that the most controversial issues--where it is easier to discern whether one is aligning with a friend or foe--tend to incentivize the most informative speech; in this way, for the study of diplomacy, controversy is ultimately king. Overall, the types of speech most common at the UN--expressions of sentiment and opinions--have to date gone largely overlooked in the literature. The prevalence of this type of speech, be it at institutions or in media outlets, provides a new domain for scholars of international diplomacy and politics to understand how countries communicate the underlying constellation of alignment in foreign policy interests.
References


Greene, Kevin, Baekkwan Park and Michael Colaresi. forthcoming. “Machine learning human rights and wrongs: How the successes and failures of supervised learning algorithms can inform the debate about information effects.” *Political Analysis*.


Hayes, Danny and Matt Guardino. 2013. *Influence from abroad: foreign voices, the media, and US public opinion*. Cambridge University Press.


contributions from Jed Wing, Steve Weston, Andre Williams, Chris Keefer, Allan Engelhardt, Tony Cooper, Zachary Mayer, Brenton Kenkel, the R Core Team, Michael Benesty, Reynald Lescarbeau, Andrew Ziem, Luca Scrucca, Yuan Tang, Can Candan, and Tyler Hunt.

URL: https://CRAN.R-project.org/package=caret


