Asking too much? The rhetorical role of questions in political discourse

Justine Zhang,[1] **Arthur Spirling,**[2] **Cristian Danescu-Niculescu-Mizil**[1]
[1]Cornell University, [2]New York University
jz727@cornell.edu, arthur.spirling@nyu.edu, cristian@cs.cornell.edu

Abstract

Questions play a prominent role in social interactions, performing rhetorical functions that go beyond that of simple informational exchange. The surface form of a question can signal the intention and background of the person asking it, as well as the nature of their relation with the interlocutor. While the informational nature of questions has been extensively examined in the context of question-answering applications, their rhetorical aspects have been largely understudied.

In this work we introduce an unsupervised methodology for extracting surface motifs that recur in questions, and for grouping them according to their latent rhetorical role. By applying this framework to the setting of question sessions in the UK parliament, we show that the resulting typology encodes key aspects of the political discourse—such as the bifurcation in questioning behavior between government and opposition parties—and reveals new insights into the effects of a legislator's tenure and political career ambitions.

1 Introduction

> "We'd now like to open the floor to shorter speeches disguised as questions..."
>
> – Steve Macone, New Yorker cartoon caption

Why do we ask questions? Perhaps we are seeking factual information that others hold, or maybe we are requesting a favor. Alternatively we could be simply making a rhetorical point, perhaps at the start of an academic paper.

Questions play a prominent role in social interactions (Goffman, 1976), performing a multitude of rhetorical functions that go beyond mere factual information gathering (Kearsley, 1976). While the informational component of questions has been well-studied in the context of question-answering applications, there is relatively little computational work addressing the rhetorical and social role of these basic dialogic units.

One domain where questions have a particularly salient rhetorical role is politics. The ability to question the actions and intentions of governments is a crucial part of democracy (Pitkin, 1967), particularly in parliamentary systems. Consequently, scholars have studied parliamentary questions in detail, in terms of their origins (Chester and Bowring, 1962), their institutionalization (Eggers and Spirling, 2014) and their importance for oversight (Proksch and Slapin, 2011). In particular, the United Kingdom's House of Commons, renowned for theatrical questions periods, has been studied in some depth. However, those accounts are largely qualitative in nature (Bull and Wells, 2012; Bates et al., 2014).

The present work: methodology. In order to approach these problems computationally, we introduce an unsupervised framework to structure the space of questions according to their rhetorical role. First, we identify common ways in which questions are phrased. To this end, we automatically extract these recurring surface forms, or *motifs*, based on the lexico-syntactic structure of the questions posed (Section 4). To capture rhetorical aspects we then group these motifs according to their role, relying on the intuition that this role is encoded in the type of *answer* a question receives. To operationalize this intuition we construct a latent question-answer space in which question motifs triggering similar answers are mapped to the same region (Section 5).

The present work: application. We apply this general framework to the political discourse that occurs during parliamentary question sessions in

the British House of Commons, a new dataset which we make publicly available (Section 3). Our framework extracts intuitive question types ranging from narrow factual queries to pointed criticisms disguised as questions (Section 5, Table 1). We validate our framework by aligning these types with prior understandings of parliamentary proceedings from the political science literature (Section 6). In particular, previous work (Bates et al., 2014) has categorized questions asked in Parliament according to the intentions of the asker (e.g., to help the answerer, or to adversarially put them on the spot); we find a clear, predictive mapping between these expert-coded categories and the induced typology. We further show that the types of questions specific legislators tend to ask vary with whether they are part of the governing or opposition party, consistent with well-established accounts of partisan differences (Cowley, 2002; Spirling and McLean, 2007; Eggers and Spirling, 2014). Concretely, government legislators exhibit a preference for overtly friendly questions, while the opposition slants towards more aggressive question types.

We then apply our methodology to provide new insights into how a legislator's questioning behavior varies with their career trajectory. The pressures faced by legislators at various stages in their career are cross-cutting, and multiple possible hypotheses emerge. Younger, more enthusiastic legislators may be motivated to ask harder-hitting questions, but risk being passed over for future promotion if they are too combative (Cowley, 2002). Older legislators, whose opportunities for promotion are largely behind them and hence have "less to lose", may act more aggressively (Benedetto and Hix, 2007); or simply seek a quiet path to retirement. Viewing each group's behavior through the questions they ask brings evidence for the latter hypothesis that more tenured legislators are more aggressive, even when questioning their own leaders. In this way, their presence in the House of Commons, and their refusal to simply 'keep their heads down', facilitates a core component of democracy.

2 Related Work

Question-answering. Computationally, questions have received considerable attention in the context of question-answering (QA) systems—for a survey see Gupta and Gupta (2012)—with an emphasis on understanding their information need (Harabagiu, 2008). Techniques have been developed to categorize questions based on the nature of these information needs in the context of the TREC QA challenge (Harabagiu et al., 2000), and to identify questions asking for similar information (Shtok et al., 2012; Zhang et al., 2017; Jeon et al., 2005); questions have also been classified by topic (Cao et al., 2010) and quality (Treude et al., 2011; Ravi et al., 2014). In contrast, our work is not concerned with the information need central to QA applications, and instead focuses on the rhetorical aspect of questions.

Question types. To facilitate retrieval of frequently asked questions, Lytinen and Tomuro (2002) manually developed a typology of surface question forms (e.g., 'what'- and 'why'-questions) starting from Lehnerts' conceptual question categories (Lehnert, 1978). Question types were also hand annotated for dialog-act labeling, distinguishing between yes-no, wh-, open-ended and rhetorical questions (Dhillon et al., 2004). To complement this line of work, this paper introduces a completely unsupervised methodology to automatically build a domain-tailored question typology, bypassing the need for human annotation.

Pragmatic dimensions. One important pragmatic dimension of questions that has been previously studied computationally is their level of politeness (Danescu-Niculescu-Mizil et al., 2013; Aubakirova and Bansal, 2016); in the specific context of making requests, politeness was shown to correlate with the social status of the asker. Sachdeva and Kumaraguru (2017) studied another rhetorical aspect by examining linguistic attributes distinguishing serviceable requests addressed to police on social media from general conversation. Previous research has also been directed at identifying rhetorical questions (Bhattasali et al., 2015) and understanding the motivations of their "askers" (Ranganath et al., 2016). Using the relationship between questions and answers, our work examines the rhetorical and social aspect of questions without predefining a pragmatic dimension and without relying on labeled data. We also complement these efforts in analyzing a broader range of situations in which questions may be posed without an information-seeking intent.

Political discourse. Finally, our work contributes to a rapidly growing area of NLP applications to political domains (Monroe et al., 2008; Card et al.,

2016; Gonzalez-Bailon et al., 2010; Niculae et al., 2015; Grimmer and Stewart, 2013; Grimmer et al., 2012; Iyyer et al., 2014b, inter alia). Particularly relevant are applications to discourse in congressional and parliamentary settings (Thomas et al., 2006; Boydstun et al., 2014; Rheault et al., 2016).

3 Data: Parliamentary Question Periods

The bulk of our analysis focuses on the questions asked, and responses given during parliamentary question periods in the British House of Commons. Below, we provide a brief overview of key features of this political system in general, as well as a description of the question period setting.

Parliamentary systems. Legislators in the House of Commons (*Members of Parliament*, henceforth *MPs* or *members*) belong to two main voting and debating *affiliations*: a *government* party which controls the executive and holds a majority of the seats in the chamber, and a set of *opposition* parties.[1] The executive is headed by the Prime Minister (PM) and run by a cabinet of *ministers*, high-ranking government MPs responsible for various departments such as finance and education.

Question periods. The House of Commons holds weekly, moderated *question periods*, in which MPs of all affiliations take turns to ask questions to (and theoretically receive answers from) government ministers for each department regarding their specific domains. Such events are a primary way in which legislators hold senior policy-makers responsible for their decisions. In practice, beyond narrow requests for information about specific policy points, MPs use their questions to critique or praise the government, or to self-promote; indeed, certain sessions, such as Questions to the Prime Minister, have gained renown for their partisan clashes, often fueled by the (mis)handling of a current crisis. The following question, asked to the Prime Minister by an opposition MP about contamination of the meat supply in 2013, encapsulates this odd mix of purposes:

> *"The Prime Minister is rightly shocked by the revelations that many food products contain 100% horse. Does he share my concern that, if tested, many of his answers may contain 100% bull?"*[2]

---

[1] We use *affiliation* to refer broadly to the government and opposition roles, independent of the identity of the current government and opposition parties. In subsequent analysis we only consider the largest, "official" opposition party as the opposition.

[2] MPs almost always address each other in 3rd person.

The moderated, relatively rigid format of questions periods, along with the multifaceted array of underlying incentives and interpersonal relationships, yields a structurally controlled setting with a rich variety of social interactions, taking place in the realm of important policy discussions. This complexity makes question periods a particularly fruitful and consequential setting in which to study questions as social signals, and expand our understanding of their role beyond factual queries.

Dataset description. Our dataset covers question periods from May 1979 to December 2016, encompassing six different Prime Ministers. For each question period, we extract all question-answer pairs, along with the identity of the asker and answerer. Because our focus here is on how questions are posed in a social setting, and not on the subsequent dialogue, we ignore questions which were tabled prior to the session, as well as any followup back-and-forth dialogue between the asker and answerer.

We augment this collection with metadata about each asker and answerer, including their political party, the time when they first took office, and whether they were serving as a minister at a given point in time. Such information is used to validate our methodology and interpret our results in light of the social context in which the questions were asked, described further in Sections 6 and 7.

In total there are 216,894 question-answer pairs in our data, occurring over 4,776 days and 6 prime-ministerships. The questions cover 1,975 different askers, 1,066 different answerers, and a variety of government departments with responsibilities ranging from defense to transport. We make this dataset publicly available, along with the code implementing our methodology, as part of the Cornell Conversational Analysis Toolkit.[3]

4 Question Motifs

The first component of our framework identifies lexico-syntactic phrasing patterns recurring in a collection of questions, which we call *motifs*. Intuitively, motifs constitute wordings commonly used to pose questions. To find motifs in a given collection, we first extract relevant *fragments* from each question. We then group sets of frequently co-occurring fragments into motifs.

---

[3] https://github.com/CornellNLP/Cornell-Conversational-Analysis-Toolkit

Question fragments. Our goal is to find motifs which reflect functional characteristics of questions. Hence, we start by extracting the key *fragments* within a question which encapsulate its functional nature. Following the intuition that the bulk of this functional information is contained in the root of a question's dependency parse along with its outgoing arcs (Iyyer et al., 2014a), we take the fragments of a question to be the root of its parse tree, along with each (root, child) pair. To capture cases when the operational word in the question is not connected to its root (such as "What..."), we also consider the initial unigram and bigram of a question as fragments. The following question has 5 fragments: what, what is, going→*, is←going and going→do.

(1) **What is** the minister **going** to **do** about ... ?

Because our goal is to capture topic-agnostic patterns, we ignore all fragments which contain a noun phrase (NP) or pronoun. NP subtrees are identified based on their outgoing dependencies to the root;[4] in the event that an NP starts with a WH-determiner (WDT), we consider (root, WDT) to be a fragment and drop the remainder of the NP.[5]

Finally, we note that some questions consist of multiple sub-questions ("What does the Minister think [...], *and* why [...]?"). For such questions, we recursively extract fragments from each child subtree in the same manner, starting from their roots.

From fragments to motifs. We define *motifs* as sets of question fragments that frequently co-occur (in at least n questions). We find motifs by applying the apriori algorithm (Agrawal and Srikant, 1994) to find these common itemsets. This results in a collection of motifs $\mathcal{M}$ which correspond to different question phrasings.[6] Examples of motifs are shown in Table 1.

Motifs can identify phrasings to varying degrees of specificity. For example, the singleton motif {what is} corresponds to all questions starting with that bigram, while {what is, going→do} nar-

rows these down to questions also containing the fragment going→do. To model the specificity relation between motifs, we structure $\mathcal{M}$ as a directed acyclic graph where an edge points from a motif m_1 to another motif m_2 if the latter has exactly one more fragment in addition to those in m_1, corresponding to a narrower set of phrasings.

Motif-representation of a question. Finally, a question q contains a motif if it includes all of the fragments comprising that motif. We can hence capture the phrasing of a given question q using the subset of motifs it contains, structured as the subgraph $\mathcal{M}_q \subset \mathcal{M}$ induced by this subset. This directed subgraph represents the question at multiple levels of specificity simultaneously; in particular, the set of sinks (i.e., nodes with outdegree 0; henceforth *sink motifs*) of $\mathcal{M}_q$ is the most fine-grained way to specify the phrasing of q. For example {what is, is←going, going→do} is the only sink motif of the question in example (1); its entire subgraph is shown in Figure 3 in the appendix.

5 Latent Question Types

The second component of our framework structures the space of questions according to their functional roles, thus going beyond the lexico-syntactic representation captured via motifs. The main intuition is that the nature of the answer that a question receives provides a good indication of its intention. Therefore, if two questions are phrased differently but answered in similar ways, the parallels exhibited by their answers should reflect commonalities in the askers' intentions.

To operationalize this intuition, we first construct a latent space based on answers, and then map question motifs (Section 4) to the same space. Using the resultant latent representations, we can then cluster questions with similar rhetorical functions, even if their surface forms are different.

Constructing a space of answers. In line with our focus on functional characterizations, we extract the fragments from each sentence of an *answer*, defined in the same way as question fragments. We then construct a term-document matrix, where terms correspond to answer fragments, and documents correspond to individual answers in the corpus. We filter out infrequent fragments occurring less than n_A times, reweight the rows of this matrix with tf-idf reweighting, and scale to unit norm, producing a fragment-answer matrix $\mathcal{A}$. We perform singular value decomposi-

---

[4]We take as NPs subtrees connected to the root with the following: nsubj, nsubjpass, dobj, iobj, pobj, attr.

[5]In the particular case of the Parliament dataset, removing NPs also removes conventional, partisan address terms (e.g. "my hon. Friend").

[6]In some cases, a pair of motifs almost always co-occurs in the same questions, making them redundant. We treat two motifs m_1 and m_2 as equivalent if, for some probability p, $\Pr(m_1|m_2) > p$ and $\Pr(m_2|m_1) > p$; we keep the smaller of the two as the representative motif, or pick one of them arbitrarily if they are of equal sizes.

tion on $\mathcal{A}$ and obtain a low-rank representation $\mathcal{A} \approx \hat{\mathcal{A}} = U_A S V_A^T$, for some rank d, where rows of U_A correspond to answer fragments and rows of V_A correspond to answers.[7]

Latent projection of question motifs. We can draw a natural correspondence between a question motif m and answer term t if m occurs in a question whose answer contains t. This enables us to compute representations of question motifs in the same space as $\hat{\mathcal{A}}$. Concretely, we construct a motif-question matrix $\mathcal{Q} = (q_{ij})$ where $q_{ij} = 1$ if motif i occurred in question j; we scale rows of $\mathcal{Q}$ to unit norm. To represent $\mathcal{Q}$ in the latent answer space, we solve for $\hat{\mathcal{Q}}$ in $\mathcal{Q} = \hat{\mathcal{Q}} S V_A^T$ as $\hat{\mathcal{Q}} = \mathcal{Q} V_A S^{-1}$, again scaling rows to unit norm. Row i of $\hat{\mathcal{Q}}$ then gives a d-dimensional representation of motif i, denoted $\hat{q}_i$.

Grouping similar questions. Finally, we identify *question types*—broad groups of similar motifs. Intuitively, if two motifs m_i and m_j have vectors q_i and q_j which are close together, they elicit answers that are close in the latent space, so are functionally similar in this sense. We use the K-Means algorithm (Pedregosa et al., 2011) to cluster motif vectors into k clusters; these clusters then constitute the desired set of question types.

To determine the type of a *particular* question q^*, we transform it to a binary vector (q_i^*) where $q_i^* = 1$ if motif i is a sink motif of q^*; using only sink motifs at this stage allows us to characterize a question according to the most specific representation of its phrasing, thus avoiding spurious associations resulting from more general motifs. We scale q^*, project it to the latent space as before, and assign the resultant projection $\hat{q}^*$ to a cluster t, hence determining its type.

Since question motifs and answer fragments have both been mapped to the same latent space (as rows of $\hat{\mathcal{Q}}$ and U_A respectively), we can also assign each answer fragment to a question type. This further facilitates interpretability through characterizing the answers commonly triggered by a particular type of question.

6   Validation

We now apply our general framework to the particular setting of parliamentary question periods, structuring the space of questions posed within these sessions according to their rhetorical function. To validate the induced typology, we quantitatively show that it recovers asker intentions in an expert-coded dataset, and qualitatively aligns with prior findings in the political science literature on parliamentary dynamics.

Question types in Parliament. We apply our motif extraction and question type induction pipeline to the questions in the parliamentary dataset.[8] Over 90% of the questions in the dataset contain at least one of the resulting 2,817 motifs; in subsequent analyses we discard questions without a matching motif. We apply our pipeline to the questions in the parliamentary dataset, and induce a typology of $k = 8$ question types to capture the rich array of questions represented in this space while preserving interpretability.

Table 1 displays extracted types, along with example questions, answers, and motifs.[9] The second author, a political scientist with domain expertise in the UK parliamentary setting, manually investigated each type and provided interpretable labels. For example, in questions of type 4, the asker is aware that his main premise is supported by the minister, and thus will be met with a positive statement backing the thrust of the question; we call this the **agreement** cluster. Types 6 and 7 are much more combative: in type 6 questions the asker explicitly attempts to force the minister to **concede/accept** a point that would undermine some government stance, while type 7 contains **condemnatory** questions that prompt the minister to justify a policy that is self-evidently bad in the eyes of the asker. In contrast, type 2 constitutes tamer **narrow** queries that require the minister to simply report on non-partisan matters of policy. (Extended interpretations in the appendix.)

Quantitative validation. We compare our output to a dataset of 1,256 questions asked to various Prime Ministers labeled by Bates et al. (2014)

---

[7]We experimented with grouping answer fragments into motifs as well, but found that most of the motifs produced were one fragment large. While future work could focus more on understanding consistent phrasings of answers, we note that at least in our chosen corpus, answers are longer and encompass a much greater variation of possible phrasings.

[8]We consider questions to be sentences ending in question marks. If an utterance consists of multiple questions, we extract fragments sets from each question separately, and take the motifs of the utterance to be the union of motifs of each component question. We set $n = 100$, $p = 0.9$, $n_A = 100$ and $d = 25$. The choice of parameters was done via manual inspection of the dataset.

[9]Each type contains a few hundred question motifs and answer fragments.

Type	Question motifs	Answer fragments	Example question-answer pairs
0: Issue update (16,693)	{what, are←taking}, {will←update}	continue→work, met→discuss	Q: **What** steps **are** the Department **taking** to create a system for asylum-seekers? A: We **continue** to **work** with the Department of Education to ensure an equitable [...]
1: Shared concerns (35,954)	{will←take}, {may←urge}	grateful←am, shall←consider	Q: **Will** he **take** steps to support other MPs to employ apprentices? A: I **am grateful** for that suggestion [...]
2: Narrow factual (16,467)	{what←made}, {what←happen, will←happen}	is←considering, have←discussed	Q: **What** representations has the Minister **made** on the future of rural policing [in] Dyfed-Powys? A: The Home Office **is considering** the matter [...]
3: Prompt for comment (16,588)	{what←say, say→to}, {will←tell}	must←say, said→was	Q: **What** has the Prime Minister to **say to** President Reagan for sending troops to Honduras? A: [...] I **must say** that we deplore the reported incursion by Nicaraguan forces [...]
4: Agreement (32,835)	{does←agree, agree→is}, {is→important}	agree→with, agree→completely	Q: **Does** [he] **agree** that one of the best ways to improve the trade balance **is** to continue the Government's strong economic policies? A: I **agree with** my hon. Friend [...]
5: Self promotion (26,351)	{is→aware}, {will←consider}	will←appreciate, am→certain	Q: **Is** my Friend **aware** that members of my parish church are pleased to have received a grant [...] ? A: [My Friend] **will appreciate** the significant performance of parishes up and down the country [...]
6: Concede, accept (31,653)	{will←accept}, {is→not, is→true}	not←accept, not←believe	Q: **Will** [he] **accept** that [the UK exiting the EU] would undermine our security [...]? A: No, I do **not accept** that [...]
7: Condemnatory (21,320)	{can←explain}, {how←justify, can←justify}	knows→well, is→wrong	Q: **Can** the Secretary **explain** why the Government are scrapping child poverty targets? A: The hon. Lady **is wrong** in what she says [...]

Table 1: Question types automatically extracted from the parliamentary question periods, along with representative motifs and question-answer pairs. The number of questions in our dataset assigned to each type is shown in parantheses. Interpretations and more examples in Tables 2 & 3 in the appendix.

(also included in our data distribution). Each question in this data is hand-coded by a domain expert with one of three labels indicating the rhetorical intention of the asker: compared to *standard* questions—denoting straightforward factual queries, *helpful* questions serve as prompts for the PM to talk favorably about their government, while *unanswerable* questions are effectively vehicles for delivering criticisms that the PM cannot respond to. Questions which are *unanswered* by the PM are also labeled. If our framework captures meaningful rhetorical dimensions, we expect a given label to be over-represented in some of our induced types, and under-represented in others.

Even though our clustering of questions is generated in an unsupervised fashion without any guidance from the coded rhetorical roles, we see that several of the types we discover closely align with these annotations. In particular, helpful questions are highly associated with the **agreement** type (constituting 28% of questions of that type compared to 14% over the entire dataset; binomial test $p < 0.01$), reinforcing our interpretation

that this type captures MPs cheerleading their own government. Conversely, unanswerable questions are frequently of the **concede/accept** type (20% in-type vs. 11% overall), while **condemnatory** questions are often unanswered (43% vs. 24% overall), suggesting that questions of these types have an increased tendency to be posed as aggressive criticisms packaged as questions.

We also validate our framework in a prediction setting using these labels, in three binary classification tasks: distinguishing helpful vs. standard, unanswerable vs. standard, and unanswered vs. answered questions. (In each task, we balance the two classes.) To control for asker affiliation effects, we consider only questions asked by government MPs for the helpful task, and opposition questions in the unanswerable and unanswered tasks; we train on questions to Conservative PMs and evaluate on Labour PMs.[10] For each setting, we train logistic regression classifiers; as

A. Log-odds ratios between question type and affiliation   **B. Question type propensities before and after affiliation switches**
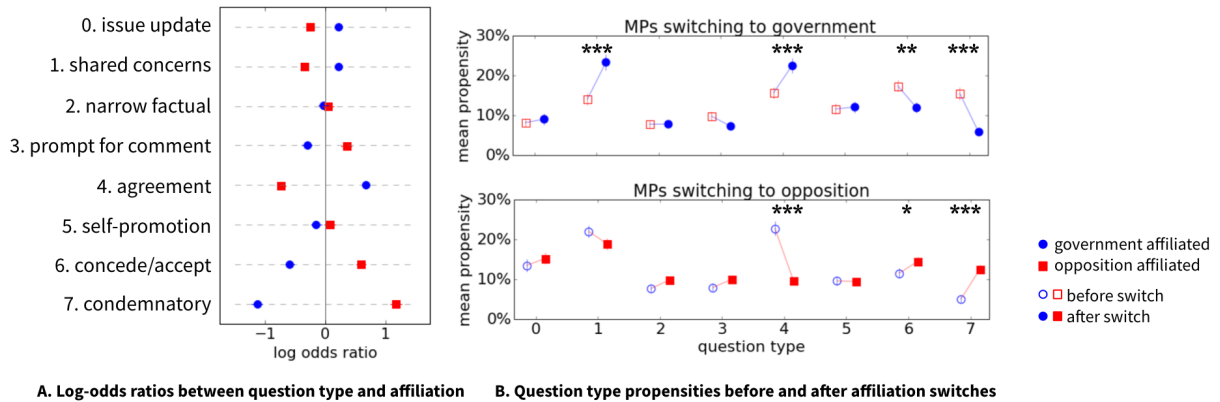
Figure 1: **A**: Log-odds ratios of questions of each type asked by government and opposition MPs, compared to MPs not of the respective affiliation; 95% confidence intervals (sometimes imperceptible) are depicted. **B**: Mean propensities for each question type, for MPs who switch from being in the opposition to being in the government (top) and vice-versa (bottom) after an election. Stars indicate statistically significant differences at the $p < 0.05$ (*), $p < 0.01$ (**) and $p < 0.001$ (***) levels (Wilcoxon test).

features we compare the latent representation of each question to a unigram BOW baseline.[11]

In the *unanswerable* and *unanswered* tasks, we find that the BOW features do not perform significantly better than a random (50%) baseline. However, the latent question features produced by our framework bring additional predictive signal and outperform the baseline when combined with BOW (binomial $p < 0.05$), achieving accuracies of 66% and 62% respectively (compared with 55% and 50% for BOW alone). This suggests that our representation captures useful rhetorical information that, given our train-test split, generalizes across parties. None of the models significantly outperform the random baseline on the *helpful* task, perhaps owing to the small data size.

Qualitative validation: question partisanship. We additionally provide a qualitative validation of our framework by comparing the question-asking activity of government and opposition-affiliated MPs—as viewed through the extracted question types—to well-established characterizations of these affiliations in the political science literature. In particular, prior work has examined the bifurcation in behavior between government and opposition members, in their differing focus on various issues (Louwerse, 2012), and in settings such as roll call votes (Cowley, 2002; Spirling and McLean, 2007; Eggers and Spirling, 2014). Since government MPs are elected on the same party ticket and manifesto, they primarily act to support the government's various policies and bolster the status of their cabinet, seldom airing disagreements publicly. In contrast, opposition members tend to offer trenchant partisan criticism of government policies, seeking to destabilize the government's relationship with its MPs and create negative press in the country at large. In characterizing the question-asking activity of government and opposition MPs, this friendly vs. adversarial behavior should also be reflected in a rhetorical typology of questions.[12]

Concretely, to quantify the relationship between a particular question type t and asker affiliation $\mathcal{P}$, we compute the log-odds ratio of type t questions asked by MPs in $\mathcal{P}$, compared to MPs not in $\mathcal{P}$.[13]

Figure 1A shows the resultant log-odds ratios of each question type for government and opposition members. Notably, we see that **agreement**-type questions are significantly more likely to originate from government than from opposition MPs, while the opposite holds for **concede/accept** and **condemnatory** questions (binomial $p < 10^{-4}$ for each, comparing within-type to overall proportions of questions from an affiliation). No such

---

[11]We used tf-idf reweighting and excluded unigrams occurring less than 5 times.

[12]While we induce the typology over our entire dataset, we perform all subsequent analyses on a filtered subset of 50,152 questions. In particular, we omit utterances with multiple questions—i.e. multiple question marks—to ensure that we don't confound effects arising from different co-occurring question types. Our filtering decisions are also determined by the availability of information about the asker and answerers' roles in Parliament. Further information about these filtering choices can be found in the appendix.

[13]The log-odds values are not symmetric between government and opposition, because they includes questions asked by MPs not in the official opposition.
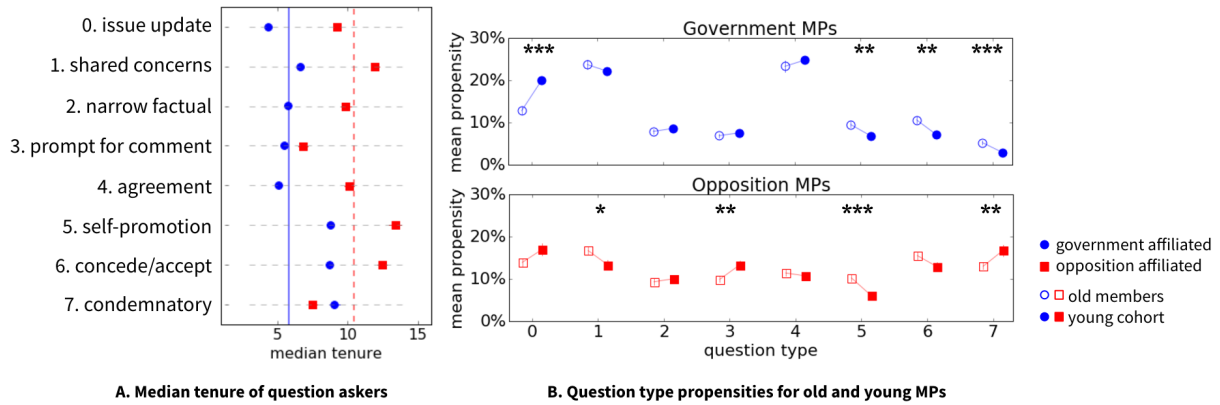
A. Median tenure of question askers

B. Question type propensities for old and young MPs

Figure 2: **A**: Median asker tenures over each question type, for government and opposition askers. Overall median tenures are also shown for reference (solid blue line for government, dashed red line for opposition). **B**: Mean propensities for newly elected MPs during the 1997 and 2010 elections, compared to re-elected MPs in the subsequent parliamentary sitting. Stars indicate statistically significant differences at the $p < 0.05$ (*), $p < 0.01$ (**) and $p < 0.001$ (***) levels (Mann Whitney U test).

slant is exhibited in the **narrow factual** type, further reinforcing the role of such questions as informational queries about relatively non-partisan issues. These results strongly cohere with the "textbook" accounts of parliamentary activity in the literature, as well as our interpretation of these types as bolstering or antagonistic.

Moreover, we find that the *same MP* shifts in her propensity for different question types as her affiliation changes. When a new political party is elected into office, MPs who were previously in the opposition now belong to the government party, and vice versa. Such a switch occurs within our data between the Major and Blair governments (Conservative to Labour, 1997), and between the Brown and Cameron governments (Labour to Conservative, 2010). For both switches, we consider all MPs who asked at least 5 questions both before and after the switch, resulting in 88 members who *became* government MPs and 102 who became opposition MPs. For an MP M we compute $P_{M,t}$, their *propensity* for a question type t, as the proportion of questions they ask which are from t. Comparing $P_{M,t}$ before and after a switch, we replicate the key differences observed above—for instance, we find that former opposition MPs who become government MPs decrease in their propensity for **condemnatory** questions, while newly opposition MPs move in the other direction (Wilcoxon $p < 0.001$, Figure 1B). This suggests that the general trends we observed before are driven by the shift in affiliation, and hence parliamentary role, of *individual* MPs.

7 Career Trajectory Effects

We now apply our framework to gain further insights into the nature of political discourse in Parliament, focusing on how questioning behavior varies with a member's tenure in the institution. As stated in the introduction, two alternative hypotheses arise: younger MPs may be more vigorously critical out of enthusiasm, but are potentially tempered by their stake in future promotion prospects compared to older members (Cowley, 2002, 2012). Alternatively, older MPs who have less at stake in terms of prospects of further promotion may ask more antagonistic questions. Throughout, *young* and *old* refer to tenure—i.e., how many years someone has served as an MP—rather than biological age.

In order to understand the extent to which young or old members contribute a specific type of question, for each question type t we compute the median tenure of askers of each question in t, and compare the median tenures of different question types, for each affiliation (Figure 2A).[14] We see that among both affiliations, more aggressive questions tend to originate more from older members, reflected in significantly higher median tenures (for types 6 in both affiliations, and 7 in government MPs; Mann Whitney U test $p < 0.001$ comparing within-type median tenure with outside-type median tenure); whereas standard **issue update** questions tend to come from younger

---

[14]Median tenures for opposition members are generally higher; winning an election tends to result in more newly-elected and therefore younger MPs (Webb and Farrell, 1999).

members ($p < 0.001$, both affiliations). Notably, the disproportionate aggressiveness of older members manifests even among *government* MPs who direct these questions towards their *own* government. This supports the "less to lose" intuition, offering a rhetorical parallel to previous findings about the increased tendency to vote contrary to party lines from MPs with little chance of ministerial promotion (Benedetto and Hix, 2007).

Interestingly, we find that these differential preferences across member tenure also manifest at a finer granularity than simply less to more aggressive. For instance, younger *opposition* members tend to contribute more **condemnatory** questions compared to older members (Mann Whitney U test $p < 0.01$), who disproportionately favor **concede/accept** questions. While further work is needed to fully explain these differences, we speculate that they are potentially reflective of strategic attempts by younger MPs to signal traits that could facilitate future promotion, such as partisan loyalty (Kam, 2009).

To discount the possibility of these effects being solely driven by a few very prolific young or old MPs, we also consider a setting where type propensities are macroaveraged over MPs. For each affiliation we compare the cohort of younger MPs who are newly voted in at the 1997 and 2010 elections, with older MPs who have been in office prior to the election.[15] We compute the type propensities of these two cohorts over the questions they asked during the *subsequent* parliamentary sitting, and replicate the tenure effects observed previously (Figure 2B). This suggests that these parliamentary career effects reflect behavioral changes at the level of individual MPs, whose incentives evolve over their tenure.

8 Conclusion and Future Work

In this work we introduced an unsupervised framework for structuring the space of questions according to their rhetorical role. We instantiated and validated our approach in the domain of parliamentary question periods, and revealed new interactions between questioning behavior and career trajectories.

We note that our methodology is not tied to a particular domain. It would be interesting to explore its potential in a variety of less structured do-

---

[15]This totals 272 new and 184 old government MPs, and 84 new and 179 old opposition MPs.

mains where questions likewise play a crucial role. For example, examining how interviewers in high-profile media settings (e.g., Frost on Nixon) can use their questions to elicit substantive responses from influential people would aid us in the broader normative goal of holding elites to account, by gaining a better understanding of what and how to ask, and what (not) to accept as an answer.

From a technical standpoint, future work could also augment the representation of questions and answers presently used in our framework, beyond our heuristic of using root arcs without noun phrases. Richer linguistic representations, as well as more judicious ways of weighting different fragments and motifs, could enable us to capture a wider range of possible surface and rhetorical forms, especially in settings where phrasings are potentially less structured by institutional conventions. Additionally, as with most unsupervised methods, our approach is limited by the need to hand-select parameters such as the number of clusters, and manually interpret the typology's output. Having annotations of these corpora could better motivate the methodology and enable further evaluation and interpretation; we hope to encourage such annotation efforts by releasing the dataset.

Inevitably, drawing causal lessons from observational data is difficult. Moving forward, experimental tests of insights gathered through such explorations would enable us to establish causal effects of question-asking rhetoric, perhaps offering prescriptive insights into questioning strategies for objectives such as information-seeking (Dillman, 1978), request-making (Althoff et al., 2014; Mitra and Gilbert, 2014) and persuasion (Tan et al., 2016; Zhang et al., 2016; Wang et al., 2017).

References

Rakesh Agrawal and Ramakrishnan Srikant. 1994. Fast algorithms for mining association rules. In *Proceedings of VLDB*.

Tim Althoff, Cristian Danescu-Niculescu-Mizil, and Dan Jurafsky. 2014. How to ask for a favor: A case study on the success of altruistic requests. In *Proceedings of ICWSM*.

Malika Aubakirova and Mohit Bansal. 2016. Interpreting neural networks to improve politeness comprehension. In *Proceedings of EMNLP*.

Stephen R. Bates, Peter Kerr, Christopher Byrne, and Liam Stanley. 2014. Questions to the Prime Minister: A comparative study of PMQs from Thatcher to Cameron. *Parliamentary Affairs*.

Giacomo Benedetto and Simon Hix. 2007. The Rejected, the Ejected, and the Dejected: Explaining government rebels in the 2001-2005 British House of Commons. *Comparative Political Studies*.

Shohini Bhattasali, Jeremy Cytryn, Elana Feldman, and Joonsuk Park. 2015. Automatic identification of rhetorical questions. In *ACL*.

Amber E. Boydstun, Dallas Card, Justin H. Gross, Philip Resnik, and Noah A. Smith. 2014. Tracking the development of media frames within and across policy issues. In *Proceedings of the APSA*.

Peter Bull and Pam Wells. 2012. Adversarial discourse in Prime Ministers Questions. *Journal of Language and Social Psychology*.

Xin Cao, Gao Cong, Bin Cui, and Christian S Jensen. 2010. A generalized framework of exploring category information for question retrieval in community question answer archives. In *Proceedings of WWW*.

Dallas Card, Justin Gross, Amber Boydstun, and Noah Smith. 2016. Analyzing framing through the casts of characters in the news. In *Proceedings of EMNLP*.

Daniel Chester and Nona Bowring. 1962. *Questions in Parliament*. Clarendon Press, Oxford.

Philip Cowley. 2002. *The Rebels: How Blair Mislaid his Majority*. Politico's Publishing.

Philip Cowley. 2012. Arise, novice leader! The continuing rise of the career politician in Britain. *Politics*.

Cristian Danescu-Niculescu-Mizil, Moritz Sudhof, Dan Jurafsky, Jure Leskovec, and Christopher Potts. 2013. A computational approach to politeness with application to social factors. In *Proceedings of ACL*.

Rajdip Dhillon, Bhagat Sonali, Carvey Hannah, and Shriberg Elizabeth. 2004. Meeting Recorder Project: Dialog Act Labeling Guide. Technical report.

Don A Dillman. 1978. *Mail and telephone surveys: The total design method*. Wiley New York.

Andrew Eggers and Arthur Spirling. 2014. Ministerial responsiveness in westminster systems: Institutional choices and House of Commons debate, 1832–1915. *American Journal of Political Science*.

Erving Goffman. 1976. Replies and responses. *Language in society*.

Sandra Gonzalez-Bailon, Andreas Kaltenbrunner, and Rafael E Banchs. 2010. The structure of political discussion networks: A model for the analysis of online deliberation. *Journal of Information Technology*.

Justin Grimmer, Solomon Messing, and Sean J Westwood. 2012. How words and money cultivate a personal vote: The effect of legislator credit claiming on constituent credit allocation. *American Political Science Review*.

Justin Grimmer and Brandon Stewart. 2013. Text as Data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*.

Poonam Gupta and Vishal Gupta. 2012. A survey of text question answering techniques. *International Journal of Computer Applications*.

Sanda M. Harabagiu. 2008. Questions and intentions. In *Advances in Open Domain Question Answering*.

Sanda M. Harabagiu, Dan I. Moldovan, Marius Paşca, Rada Mihalcea, Mihai Surdeanu, Răzvan Bunescu, Corina R Gîrju, Vasile Rus, and Paul Morărescu. 2000. FALCON: Boosting knowledge for answer engines. In *Proceedings of TREC*.

Mohit Iyyer, Jordan Boyd-Graber, Leonardo Max Batista Claudino, Richard Socher, and Hal Daumé III. 2014a. A neural network for factoid question answering over paragraphs. In *Proceedings of EMNLP*.

Mohit Iyyer, Peter Enns, Jordan Boyd-Graber, and Philip Resnik. 2014b. Political ideology detection using recursive neural networks. In *Proceedings of ACL*.

Jiwoon Jeon, W Bruce Croft, and Joon Ho Lee. 2005. Finding Semantically Similar Questions Based on Their Answers. In *Proceedings of SIGIR*.

Christopher Kam. 2009. *Party Discipline and Parliamentary Politics*. Cambridge University Press, Cambridge.

Greg P. Kearsley. 1976. Questions and question asking in verbal discourse: A cross-disciplinary review. *Journal of Psycholinguistic Research*.

Wendy G. Lehnert. 1978. *The process of question answering: A computer simulation of cognition*. Lawrence Erlbaum Associates.

Tom Louwerse. 2012. Mechanisms of issue congruence: The democratic party mandate. *West European Politics*.

Steven Lytinen and Noriko Tomuro. 2002. The use of question types to match questions in FAQFinder. In *AAAI Spring Symposium on Mining Answers from Texts and Knowledge Bases*.

Tanushree Mitra and Eric Gilbert. 2014. The Language that Gets People to Give: Phrases that Predict Success on Kickstarter. In *Proceedings of CSCW*.

Burt L. Monroe, Michael P. Colaresi, and Kevin M. Quinn. 2008. Fightin' Words: Lexical feature selection and evaluation for identifying the content of political conflict. *Political Analysis*.

Vlad Niculae, Caroline Suen, Justine Zhang, Cristian Danescu-Niculescu-Mizil, and Jure Leskovec. 2015. QUOTUS: The structure of political media coverage as revealed by quoting patterns. In *Proceedings of WWW*.

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*.

Hanna F. Pitkin. 1967. *The Concept of Representation*. University of California Press, Oakland.

Sven-Oliver Proksch and Jonathan Slapin. 2011. Parliamentary questions and oversight in the European Union. *European Journal of Political Research*.

Suhas Ranganath, Xia Hu, Jiliang Tang, Suhang Wang, and Huan Liu. 2016. Identifying rhetorical questions in social media. In *ICWSM*.

Sujith Ravi, Bo Pang, Vibhor Rastogi, and Ravi Kumar. 2014. Great question! Question quality in community Q&A. In *Proceedings of ICWSM*.

Ludovic Rheault, Kaspar Beelen, Christopher Cochrane, and Graeme Hirst. 2016. Measuring emotion in parliamentary debates with automated textual analysis. *PloS One*.

Niharika Sachdeva and Ponnurangam Kumaraguru. 2017. Call for service: Characterizing and modeling police response to serviceable requests on Facebook. In *Proceedings of CSCW*.

Anna Shtok, Gideon Dror, Yoelle Maarek, and Idan Szpektor. 2012. Learning from the past: Answering new questions with past answers. In *Proceedings of WWW*.

Arthur Spirling and Iain McLean. 2007. UK OC OK? Interpreting optimal classification scores for the U.K. House of Commons. *Political Analysis*.

Chenhao Tan, Vlad Niculae, Cristian Danescu-Niculescu-Mizil, and Lillian Lee. 2016. Winning arguments: Interaction dynamics and persuasion strategies in good-faith online discussions. In *Proceedings of WWW*.

Matt Thomas, Bo Pang, and Lillian Lee. 2006. Get Out the Vote: Determining support or opposition from congressional floor-debate transcripts. In *Proceedings of EMNLP*.

Christoph Treude, Ohad Barzilay, and Margaret-Anne Storey. 2011. How do programmers ask and answer questions on the web? In *Proceedings of ICSE*.

Lu Wang, Nick Beauchamp, Sarah Shugars, and Kechen Qin. 2017. Winning on the merits: The joint effects of content and style on debate outcomes. *TACL*.

Paul Webb and David M Farrell. 1999. Party members and ideological change. *Critical Elections: British Parties and Voters in Long-Term Perspective*.

Justine Zhang, Ravi Kumar, Sujith Ravi, and Cristian Danescu-Niculescu-Mizil. 2016. Conversational flow in Oxford-style debates. In *Proceedings of NAACL*.

Wei Emma Zhang, Quan Z. Sheng, Jey Han Lau, and Ermyas Abebe. 2017. Detecting duplicate posts in programming QA communities via latent semantics and association rules. In *Proceedings of WWW*.

Appendix

A.1 Further examples of question types

Tables 2 and 3 provide further examples of representative questios and motifs from each of the eight question types we induce on our dataset of parliamentary question periods. Additionally, we include extended interpretations of each of these types, provided by the second author, a political scientist with domain expertise in the UK parliamentary setting.

A.2 Details about data filtering decisions

Here we provide further details about how we selected the subset of 50,152 questions which was used in the analyses described in Sections 6 & 7. First, we restrict our analysis to the questions which consist of only one question, as opposed to a series of questions (as delimited by multiple sentences ending in question marks; constituting 52% of the data). We omit these multi-question utterances in order to ensure that we don't confound effects arising from different co-occurring question types. Next, we only include questions for which information about the asker and answerers' party affiliations and ministership positions are available (such data is provided consistently from the Blair government onwards). Finally, we omit questions asked by opposition members who are specifically appointed by their party to *shadow* a government minister, and are hence obliged by their appointment to ask more critical questions. Our choice to omit such questions eliminates the possibility that our observed differences in question preference are driven by official appointment.

A.3 Details about the labeled PMQ dataset

Here we provide further details about the size of the labeled Questions to the Prime Minister (PMQ) dataset from Bates et al. (2014) used in the quantitative validation of our typology (Section 6). The dataset contains 931 standard, 186 helpful and 139 unanswerable questions. Additionally, 445 answers to questions are labeled as *answered*, while 305 are labeled as *not answered*; the rest are labeled as *deferred answer*, meaning the Prime Minister did not have the knowledge or capability to provide an answer.

We restrict all analysis done using the unanswered vs. answered labels to *standard* questions, i.e., questions for which the PM has an opportunity to provide a legitimate answer.

The *helpful vs. standard* classification task contains 264 train and 108 test examples; the *unanswerable vs. standard* classification task contains 190 train and 86 test examples; the *unanswered vs. answered* classification task contains 166 train and 84 test examples.

A.4 Example motif subgraph

Figure 3 illustrates a section of the motif DAG $\mathcal{M}$ and highlights (in bold) the subgraph $\mathcal{M}_q$ corresponding to the phrasing of the question in example (1). Nodes that are higher in the graph serve as general representations, and capture similarities between broad sets of phrasings: e.g., {is←going going→do, is←going, what is} groups together example (1) with a question like "When is he going to get a grip on [the scandal]". By projecting such motifs into our latent question-answer space (Section 5) we capture characteristics shared between these phrasings and allow for generalizability. Nodes which are lower in the graph constitute more specific representations, disambiguating between phrasings. In particular, sink motifs serve as the most specific representation of a question, delineating the region of the latent space (and thus the question type) that best captures its phrasing. For instance, this additional specificity allows us to draw contrasts between "**What is** the minister **going** to **do** [about the policy]?" (sink motif: {what is, is←going, going→do}) and the more aggressive "**When is** he **going** to get a grip on [the scandal]?" (sink motif: {when is, is←going}).
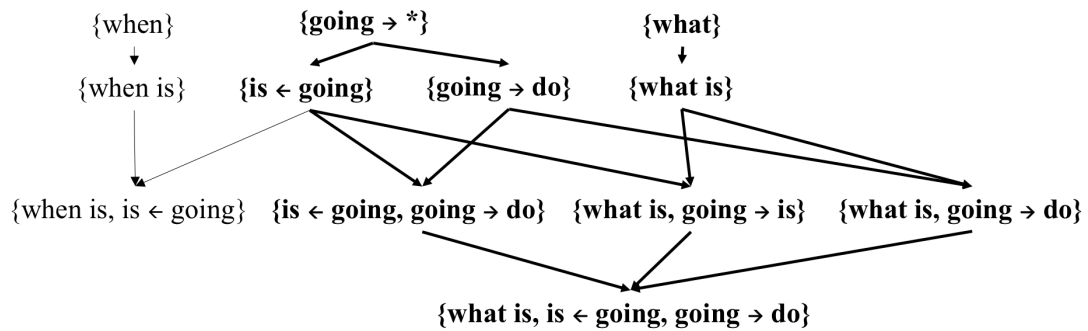
Type	Interpretation and examples
0: Issue update	Requests for information or updates about a current event, issue, or policy. Typically the policy refers to a genuinely 'national' concern, rather than a partisan issue for which the major parties may have differing views. Q: **Will** the Minister also **update** the House on whether any decisions have been made on the post-2014 UK contribution to Afghanistan? Q: **What** more can the Government **do** to **ensure** that each pupil has a single point of contact [for mental health issues] throughout their education? Q: **What** further steps **can** we **take** to resolve [the] terrible situation [in Syria]? Question motifs: {what,are←taking}, {will←update}, {what←do, do→ensure}, {what←take, can←take}
1: Shared concerns	Straightforward factual question with no strong ideological underpinnings; answers are typically vague and involve explaining that the government takes it seriously, will continue to do so and will consult with the relevant stakeholders. Q: **May** I **urge** the Minister to concentrate on tough penalties for people who get involved in alcohol-induced antisocial behaviour? Q: **Will** the Secretary of State **look carefully at** reports that houses built to house the soldiers will block off the rising sun at the summer equinox [at] the Stonehenge? Q: **Will** [my hon. Friend] **raise** [the matter of] product placement **with** people in the film industry [when he meets them]? Question motifs: {will, take→*}, {may←urge}, {will←look, look→at, look→carefully}, {raise→will, raise→with}
2: Narrow factual	Narrow queries about relatively minor policy issues which are either extremely local in nature (perhaps referring to implications for a given constituency) or limited in scope (constituting a small issue within a much broader context of policy). Questions are precise if not especially penetrating. Answers typically explain the ministry response in narrow, non-ideological terms. Q: **What** funding will be **given to** the new postgraduate institute at the Edinburgh College of Dentistry? Q: **Will** the Minister **make** a statement **on** where we are with the website for the [National Health Service appraisals toolkit for general practitioners]? Q: **What will happen** to the French if they are found guilty of infringing the rules of the Commission? Question motifs: {what←made}, {what←happen, will←happen}, {what←given, given→to}, {will←make, make→on}
3: Prompt for comment	Requests for comments, or information, especially on a meeting that has taken place between a minister and constituents, colleagues, or opposite numbers, the contents of which would not normally be immediately accessible to MPs. The asker seeks that the minister clarifies policy where none might presently exist. Q: **Will** the Secretary of State **tell** us which Minister has been appointed to be responsible for green economic growth? Q: **Can** [the Secretary of State] **confirm** whether train platform capacity will **be** a part of the discussions between the city council and Network Rail? Q: **What would** the Prime Minister **say** to a borough council which is considering rejecting Government funding and instead [taxing] my constituents more? Question motifs: {what←say,say→to}, {will←tell}, {can←confirm, confirm→be}, {what←say, would←say}

Table 2: Interpretation for the first four types, with examples of representative questions and motifs.

Type	Interpretation and examples
4: Agreement	Airing a laudatory remark about a policy that the minister and MP clearly already agree on. Often these questions effectively serve as attempts to curry favor with the minister and bolster their (mutual) party. Q: **Is** it not **important** that the Department continues its excellent work [in] building flood defeneces? Q: **Does** [the Secretary of State] **agree with** me that part of protecting Britain's national interests is that Britain should develop relationships with emerging economies? Q: **Does** the Minister **agree** that UK taxpayers **need** to be considered at every single step of the way when it comes to our aid spending? Question motifs: {does←agree, agree→is}, {is→important}, {does←agree, agree→with}, {does←agree, agree→need}
5: Self-promotion	Here, "awareness" is entirely rhetorical: either the minister is aware and agrees or the minister is not aware and will investigate. These questions allow the question asker to be seen to be bringing local concerns to broader attention, but in a way that is more assertive than in type 2 (**narrow factual**). Q: **Has** the right hon. Gentleman **considered** compulsory postal voting , and moving polling day from a Thursday to the weekend? Q: In considering the role of local tribunals , **will** my hon Friend **take** account of the recommendation **in** the Oglesby report? Q: **Will** the Minister **reconsider** his reply to my [colleague], [in light of] the situation [on] school leavers applying for technical positions? Question motifs: {is→aware}, {considered←has}, {will←take, take→in}, {will←reconsider}
6: Concede, accept	Aggressive demand for minister to concede to, or accept, a fault. The premise of such questions is that the minister has been incompetent, or that the government has the wrong policy; these questions do not constitute a genuine attempt to obtain information. Q: **Is** it **not** now completely **true** that the Labour Government are out of touch with gut British instincts? Q: **Will** [the Secretary] **acknowledge** the importance of not completely abandoning the research on sustainable biofuels? Q: **Will** [the Deputy Prime Minister] now **concede** to the House that the Royal Mail was sold off too cheaply? Question motifs: {will←accept}, {is→not, is→true}, {will←acknowledge}, {will←concede}
7: Condemmatory	Similar to type 6 (**concede, accept**) but more aggressive and hectoring in tone, asking the minister to explain themselves and be contrite on the basis of very broad ideological premises that are difficult to answer without 'self-incrimination'. These questions often rely on rhetorical grandstanding and lacks any subtlety or policy detail on which minister can comment precisely. Q: When members of the armed forces are facing a pay freeze, **how can** the Secretary **justify** bonuses to senior offices in the civil service? Q: **Why does** the right hon Gentleman not have an industrial strategy to build that recovery? Q: **Will** the Government now **apologise** for their complacent decision to scrap the future jobs fund, [given that] long-term youth unemployment is rising? Question motifs: {can←explain}, {how←justify, can←justify}, {why does}, {will←apologise}

Table 3: Interpretation for the last four types, with examples of representative questions and motifs.

{when}   **{going → *}**   **{what}**

{when is}   **{is ← going}**   **{going → do}**   **{what is}**

{when is, is ← going}   **{is ← going, going → do}**   **{what is, going → is}**   **{what is, going → do}**

{what is, is ← going, going → do}

(1)   What is the minister going to do about … ?

Figure 3: A section of the motif DAG $\mathcal{M}$ and the subgraph $\mathcal{M}_q$ (in bold) representing the phrasing of the question in example (1). For clarity, redundant and irrelevant nodes and edges are not shown.