

The Keystone Role Played by Questions in Debate

Zlata Kikteva¹, Kamila Gorska², Wassiliki Siskou^{1,3}

Annette Hautli-Janisz¹, Chris Reed²

¹University of Passau, Germany

²University of Dundee, UK

³University of Konstanz, Germany

Abstract

Building on the recent results of a study into the roles that are played by questions in argumentative dialogue (Hautli-Janisz et al., 2022a), we expand the analysis to investigate a newly released corpus that constitutes the largest extant corpus of closely annotated debate. Questions play a critical role in driving dialogical discourse forward; in combative or critical discursive environments, they not only provide a range of discourse management techniques, they also scaffold the semantic structure of the positions that interlocutors develop. The boundaries, however, between providing substantive answers to questions, merely responding to questions, and evading questions entirely, are fuzzy and the way in which answers, responses and evasions affect the subsequent development of dialogue and argumentation structure are poorly understood. In this paper, we explore how questions have ramifications on the large-scale structure of a debate using as our substrate the BBC television programme *Question Time*, the foremost topical debate show in the UK. Analysis of the data demonstrates not only that questioning plays a particularly prominent role in such debate, but also that its reverberations can reverberate through a discourse.

1 Introduction

Whether employed to garner votes, determine policy, air grievances or test theories, debates rely upon questioning; questions provide the driving force, a rhythmic sparking of the engine of debate. They do not merely provide impetus for a moment or a turn, however: they can have far-reaching ramifications, scaffolding large structures of subsequent discursive interaction. We intuitively recognise this keystone role of questions, when, for example, we criticise politicians for evading questions, or are frustrated when our own are ignored. (The US 2016 presidential debates offer a good example, wherein Clinton was roundly criticised by the press

at both ends of the political spectrum for failing to give direct answers to the questions she was asked). Our intuitions, however, have heretofore not been matched by a theoretical understanding. How do different types of questions lead to different interactional and argumentative structures? What patterns of ramifications do different classes of reactions to those questions have? How can questions be used strategically to open up and limit the space of information and outcome?

For the purpose of exploring these questions in naturally occurring argumentation at scale, we investigate a new dataset, QT30 (Hautli-Janisz et al., 2022b), comprising analysis of thirty episodes of broadcast topical debate from the UK BBC television programme *Question Time*. This is three times the size of the previously largest dataset and offers a rich environment in which to test the theories from Hautli-Janisz et al. (2022a). By comparing the analysis of QT30 with that of the earlier datasets, we can establish a robust grounding for our conclusions about the role that *information-seeking*, *assertive* and *rhetorical questions* play across domains of argumentation. We also show that the *response space* across these three question types remains consistent, i.e., whether questions are ignored, reacted to, responded to, or in fact answered.

2 Related work

Categorising the function of questions has been a notoriously tricky issue, not only in theoretical but also in computational approaches to discourse. The field is therefore confronted with “a trade-off between usefulness and ease or consistency of coding” (Carletta et al., 1997, p. 15). The majority of approaches represent and model information-seeking (or so-called ‘canonical’) questions, as for instance in FriendsQA (Yang and Choi, 2019). In the MapTask coding scheme (Carletta et al., 1997), rhetorical questions do not form a discrete category. The Switchboard corpus (Calhoun et al., 2010) en-

codes information-seeking questions according to their syntactic property (e.g. yes-no-questions, wh-questions), non-information seeking questions are summed up in the category ‘rhetorical questions’. Stivers and Enfield (2010) propose seven different ‘social actions’ of questions, among them ‘Request for Information’, ‘Outloud’ (“delivered to no one in particular”), and ‘Rhetorical Question’ (“Questions that may seek a response but do not seek an answer”).

In terms of characterising the response space, Stivers and Enfield (2010) use the four categories of ‘Non-response’, ‘Non-answer response’, ‘Answer’, and ‘Can’t determine’ (p. 2624) for spontaneous, naturally occurring conversation, showing some overlap with the six-way distinction of Berninger and Garvey (1981) with ‘Possible answers’, ‘Indirect answers’, ‘Confessions of Ignorance’, ‘Clarification requests’, ‘Evasive replies’ and ‘Miscellaneous’ (targeted at nursery school child conversation). Łupkowski and Ginzburg (2016) propose a taxonomy for query responses, bearing some resemblance to Berninger and Garvey (1981). For the task-based interactions in MapTask, the BEE corpora and informal conversations in the BNC corpus, Ginzburg et al. (2019) propose a two-way distinction of responses into ‘Answers’ and ‘Non-answers’, with more fine-grained categories subsumed under them.

Our goal in this paper is to combine the strengths of these approaches whilst maintaining a simplicity in taxonomy that enables us to focus specifically upon the impact of questions in argumentative discourse. For our investigation, we pair structures of question-answering with an analysis of how argumentation is invoked or ‘anchored’ in dialogue. It is only in this combination that we are able to derive insights into how these two realms of structures impact surrounding discourse. The only framework that allows us to investigate both realms of discourse in parallel is Inference Anchoring Theory (IAT) (Budzynska et al., 2014, 2016), a pragmatically-driven account of dialogical argumentation.

3 Inference Anchoring Theory and the question/response space

The central motivation of Inference Anchoring Theory (IAT) (Budzynska et al., 2014, 2016) is to provide a theoretical framework for analysing dialogical argumentation. As such, IAT sets out to answer

the question of where argumentation *comes from* in dialogical interaction and acts as a theory-neutral scaffolding that integrates different communicative structures, namely *dialogue structure*, *argument structure* (including inference understood in the logical manner) and *illocutionary forces* such as asserting, suggesting or promising, to anchor argument structure in dialogue structure. One of the primary loci of argumentation is precisely in the asking of, and responding to, questions of particular sorts, and it is upon these questions that we focus here, rather than questions of clarification (Purver et al., 2001, 2003); of recovery from non-understanding (Bohus and Rudnicky, 2005); or in many cases of personal information (Sakai et al., 2018), inter alia. An example of how questions can lead to argumentative structure is shown in (1) in which Question Time moderator Fiona Bruce uses questions as a tool that allows her to forward and control the debate:

- (1) a. Fiona Bruce [00:48:45] [...] *Nadhim, should the UK be doing more to help find a long lasting solution? Because who knows how long this particular ceasefire will last.*
- b. Nadhim Zahawi [00:49:05] *The simple answer is yes, and we have been doing. The Foreign Secretary Dominic Raab and the US Secretary of State Antony Blinken have worked tirelessly, one, to deescalate the situation... Obviously President Biden has made a huge difference, as has our Prime Minister to urge both sides...*
- c. Fiona Bruce [00:49:30] *But there is no peace process at the moment. Is there anything the UK can do with the international community to try and foster a situation where that can start again?*

In (1-a) Fiona Bruce asks one of the panel members a question about a long-lasting resolution of the Israeli–Palestinian conflict. She then supports her question by expressing her concerns regarding the current ceasefire (*‘Because who knows how long this particular ceasefire will last’*). Panel member Nadhim Zahawi answers Fiona Bruce’s question directly with ‘yes’ in (1-b) and further elaborates his answer with *‘we have been doing [more]’*. He then proceeds to support his claim by mentioning

previous work on the de-escalation of the conflict by Foreign Secretary Dominic Raab and the US Secretary of State Antony Blinken as well as President Biden and the UK Prime Minister. However, by mentioning the past work of the UK and US politicians, he avoids going into detail regarding the future steps of the UK in the conflict. In (1-c) Fiona Bruce picks up on that and steers the conversation back to the current Israeli–Palestinian conflict by stating that ‘there is no peace process at the moment’. She then uses this as a reason to ask Nadhim Zahawi another and more specific question on whether the UK can play a role in restarting the peace talks, pressing him to give a more relevant answer. This second question leads to further discussion, which is not illustrated in this example.

In the following we go through the components of analysis that are needed to investigate the role of questions in argumentative dialogue.

3.1 Propositional relations

Argumentative structures are relations between propositions; core IAT assumes three different relations that are designed to capture argumentative structure in dialogue:

Inference (Support, ‘Default Inference’, RA, green node) Holds between propositions when one (or more) proposition is used to provide a reason to accept another proposition.

Conflict (Attack, ‘Default Conflict’, CA, red node) Holds between two propositions when one proposition is used to provide an incompatible alternative to another proposition.

Rephrase (Rephrase, ‘Default Rephrase’, MA, orange node) Holds between two propositions when one proposition is used to rephrase, refine or generalise another proposition. Rephrases also hold between questions and answers.

3.2 Illocutionary acts of questioning

Illocutionary relations anchor propositions in locutions and capture the communicative intent of the speaker in uttering a locution or a pair of locutions. We instantiate illocutionary relations with the concept of illocutionary forces of Searle (1969) and Searle and Vanderveken (1985), however we use a simplified version in which there are no commissives or expressives and just two simple types of directives (questions and challenges). For QT30, we use a set of nine illocutionary connections that

either anchor propositions in locutions (as with ‘Asserting’) or propositional relations in the dialogical structure (as with ‘Arguing’, ‘Disagreeing’ and ‘Restating’ in the case of ‘Default Inference’, ‘Default Conflict’ and ‘Default Rephrase’, respectively). The subset we require for the purposes of this paper are summarised briefly below.

- **Assertive Questioning** Speaker S communicates information and at the same time asks for confirmation/rejection from their interlocutor.
- **Pure Questioning** Speaker S is seeking information or asking for the opinion of their interlocutors.
- **Rhetorical Questioning** Speaker S is expressing an opinion in the form of an interrogative.
- **Restating** Speaker S rephrases a previous claim. Anchors a ‘Default Rephrase’.
- **Default Illocuting** Used to anchor a ‘Default Rephrase’ which captures an answer to a question, i.e., the answer instantiates (parts of) the question.

The graph on the right-hand side of Figure 1 provides the IAT analysis for Example (1), produced with OVA+ (Online Visualisation of Argument – <http://ova.arg.tech/>), an open-source online interface for the analysis of argumentation in dialogues (Janier et al., 2014). OVA+ allows for a representation of the argumentative structure of a text as a directed graph.¹ For zooming out of the discourse and getting a ‘bird’s-eye view’ of an entire debate, there are visualisation tools that form a part of the Argument Web (Reed et al., 2017): the left-hand side of Figure 1, for example, is produced using *ArgNav* (argnav.arg.tech/).

The graph on the right-hand side of Figure (1) shows the different components of analysis for Example (1). In this graph, the right side shows the dialogical structure (blue boxes on the right), the left shows the propositional structure (blue boxes on the left). They are connected with the illocutionary acts (yellow) in the middle. Argument relations are between propositions on the left.

The exchange starts with Fiona Bruce questioning Nadhim Zahawi regarding the UK’s further

¹The IAT annotation guidelines are here: <http://www.arg.tech/f/IATannotationguidelines.pdf>.

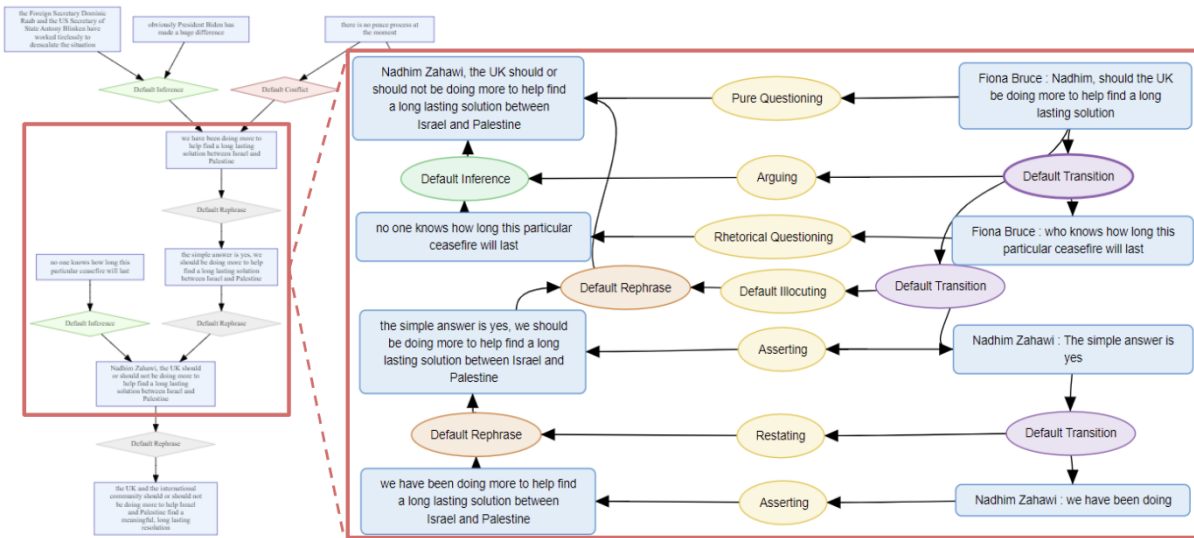


Figure 1: The IAT diagram of Example (1) (<https://www.aifdb.org/argview/21415>) on the right features locutions (blue nodes on the right-hand side), propositions (blue nodes on the left-hand side), illocutionary relations (yellow nodes in the middle), dialogical relations (purple nodes on the right) and propositional relations – ‘Default Inference’ (green) and ‘Default Rephrase’ (orange). The diagram on the left features propositions only and the argument relations between them. It captures the larger argumentative exchange of which Example (1) is part of.

help in de-escalation of the Israeli–Palestinian conflict (*‘Should the UK be doing more to help find a long-lasting solution?’* – ‘Pure Questioning’) and then by way of a rhetorical question she brings up her concerns regarding the longevity of the current ceasefire (*‘Who knows how long this particular ceasefire will last?’* – ‘Rhetorical Questioning’). Nadhim Zahawi then answers her initial question in (1-b) with *‘the simple answer is yes’*. The relation between propositions is ‘Default Rephrase’ which is anchored via ‘Default Illocuting’, indicating that this is an answer to a question. He then proceeds to explain his answer with *‘we have been doing [more]’*. Discourse moves like that are analysed as rephrases, which are indicated with ‘Default Rephrase’ between two propositions and anchored in the dialogical structure on the right with ‘Restating’. The graph on the left-hand side of Figure 1 shows the complete exchange between Fiona Bruce and Nadhim Zahawi from the Example (1), but only in terms of the argumentative structure. It illustrates how one question starts an elaborate conversation with intricate argumentative structures including those of inferences, rephrases and conflicts.

3.3 The response space

The context succeeding a question is generally called ‘the response space’, with a number of cod-

ing schemes having worked on a partition or characterisation of that space. As mentioned before, IAT pursues an approach of underspecification – in the case of responses this means that we differentiate only four types of responses:

- **Ignored** Similar to the ‘Ignore’ type of query response in Ginzburg et al. (2019) and the Non-response category in Stivers and Enfield (2010) where the “person did nothing in response” (p. 2624) we can identify discourse moves that entirely ignore the question and relate to discourse material before the question.
- **Reactions** The question is not ignored – it is reacted to, but is ‘left hanging in the air’ and the reaction rather contributes to the ongoing discourse. Reactions are identified in IAT through the existence of an outgoing ‘Default Transition’ (TA – purple transition between blue boxes on the left in Figure 1) from the question locution to the succeeding locution. There is no propositional relation between the proposition succeeding the question and the question itself.
- **Responses** are understood to be discourse moves directly related to a question without answering it. This is the case for rhetorical and assertive questions which can be re-

sponded to, but not answered. ‘Responses’ in IAT are captured via a ‘Default Transition’ following the question and anchoring an illocutionary structure (YA – yellow relations), e.g., due to the fact that there is an argumentative relation between the question and the succeeding proposition.

- **Answers** Discourse moves that provide answers to a question by delivering propositional information that instantiates the lambda expression (in the case of a wh-question) or resolves the disjunction (in the case of a polar and alternative question). In IAT answers are captured via a Default Rephrase between answer and question which is anchored via Default Illocuting in the transition between them.

4 Data

The combined dataset underlying our investigation consists of thirty six corpora in four argumentative genres: election debate, moral debate, public deliberation and topical debate. All corpora are annotated with full IAT. In total, our dataset consists of 414,726 words (tokens), and 51,993 illocutionary structures.

For the ‘Election Debate’ genre, we use the US2016tv² corpus (Visser et al., 2020), which includes the transcripts of all televised debates in the US 2016 Presidential Election. The ‘Election Debate’ dataset has 39,694 words, and 6,570 illocutionary structures and an inter-annotator agreement of Cohen’s $\kappa = 0.61$ (substantial agreement) (Visser et al., 2020).

For the genre ‘Moral Debate’, we include two sub-corpora. MM2012³, a collection of transcripts of BBC Radio 4’s ‘The Moral Maze’, a series of debates about moral dilemmas and BBC2017⁴, which consists of a special issue of ‘The Moral Maze’ and a related TV debate about the morality of abortion. The ‘Moral Debate’ dataset has 60,273 words, and 5,488 illocutionary structures.

For ‘Public Deliberation’, we include three sub-corpora. USCD2011⁵ consists of several transcripts of public meetings and hearings from departments of transportation (DOTs) across the

USA (Lawrence et al., 2017), (Konat et al., 2016). DEDD2019⁶ is a set of two experimental 4-party deliberations on whether or not to allow fracking in Germany. UKDD2019⁷ is a public deliberation organised by the Royal Society of Edinburgh in the context of an inquiry about the future of energy in Scotland. The ‘Public Deliberation’ dataset has 34,759 words and 3,931 illocutionary structures.

The corpora in the genres ‘Public Deliberation’, ‘Moral Debate’ and ‘Election Debate’ were annotated by small teams of two to ten analysts. These three genres provide background and context for our main study which focuses upon the genre of ‘Topical Debate’, for which we use the QT30 corpus (Hautli-Janisz et al., 2022b).⁸ This unique dataset includes transcripts of 30 episodes aired between June 2020 and November 2021 of BBC’s Question Time, one of the most viewed political talk shows in the UK. The corpus consists of more than 29 hours of transcribed broadcast debate and comprises 280,000 words and 36,004 illocutionary structures by more than 400 participants, making it the largest corpus of analysed broadcast topical debate published to date. The annotation was conducted by 38 students of linguistics, philosophy, literature and computer science across Europe. Inter-annotator agreement for QT30 is at a Combined Argument Similarity Score (CASS) (Duthie et al., 2016) of 0.56, signaling moderate agreement.

Table 1 shows that in terms of the total number of illocutionary structures, the ‘Topical Debate’ corpus is the largest dataset and almost five times the size of the ‘Election Debate’ corpus. ‘Asserting’ is consistently the most frequent illocutionary connection across the four argumentative genres, making up almost 60% of all illocutionary connections, followed by ‘Arguing’ with around 20%. The total number of questions of all types in the QT30 corpus is comparable to the number of questions across the other corpora, i.e. 4.71% and 4.75% respectively. This number is slightly lower than the number of questions in the Switchboard corpus (8%) (Jurafsky et al., 1997), but higher than in the Penn Discourse Treebank 3.0, where questions make up only 0.2% of all explicit and implicit discourse relations (Webber et al., 2017). In the following, we discuss question-answering and argumentative structures in more detail.

²<http://corpora.aifdb.org/US2016tv>

³<http://corpora.aifdb.org/mm2012>

⁴<http://corpora.aifdb.org/BBCaaaTV5RADIO>

⁵<http://corpora.aifdb.org/cd>

⁶<http://corpora.aifdb.org/DEDD2019>

⁷<http://corpora.aifdb.org/UKDD2019>

⁸<http://corpora.aifdb.org/qt30>

	Election Debate	Moral Debate	Public Deliberation	Topical Debate	Total
Sum Non-questioning	6289 (95.72%)	5243 (95.54%)	3760 (95.65%)	32506 (94.66%)	47798 (95.28%)
Asserting	4218 (64.20%)	2953 (53.81%)	2245 (57.11%)	19224 (55.98%)	28640 (57.09%)
Arguing	1523 (23.18%)	1280 (23.32%)	879 (22.36%)	5497 (16.01%)	9179 (18.30%)
Agreeing	36 (0.55%)	180 (3.28%)	65 (1.65%)	346 (1.01%)	627 (1.25%)
Disagreeing	153 (2.33%)	368 (6.71%)	151 (3.84%)	1259 (3.67%)	1931 (3.85%)
Restating	243 (3.70%)	189 (3.44%)	376 (9.56%)	4082 (11.88%)	4890 (9.75%)
Default Illocuting	93 (1.42%)	243 (4.43%)	35 (0.89%)	1969 (5.73%)	2340 (4.66%)
Challenging	23(0.35%)	30 (0.55%)	9 (0.23%)	129 (0.38%)	191 (0.38%)
Sum Questioning	281 (4.28%)	245 (5.01%)	171 (4.35%)	1671 (4.71%)	2368 (4.72%)
Pure Questioning	166 (2.53%)	113 (2.06%)	52 (1.32%)	1201 (3.50%)	1532 (3.05%)
Rhetorical Questioning	30 (0.46%)	39 (0.71%)	20 (0.51%)	224 (0.65%)	313 (0.62%)
Assertive Questioning	85 (1.29%)	93 (1.69%)	99 (2.52%)	246 (0.72%)	523 (1.04%)
Total	6570	5488	3931	34177	50166

Table 1: Distribution of illocutionary structures across genres of argumentative discourse.

5 The role of questions in broadcast topical debate

5.1 Types of questioning

‘Pure Questioning’ makes up 3.5% of illocutionary structures in the QT30 corpus (‘Topical Debate’), in comparison to 2.07% across previous corpora. Assertive questions are substantially more unusual in QT30 by comparison to the other corpora - 0.72% in QT30 vs aggregated 1.83%, constituting barely a third of the frequency. The proportion of pure (information-seeking) questions makes up the majority across the total number of questions of all types in both corpora, but to a marked extent in QT30 – 72% in the QT30 corpus, and 44% across other corpora. This increase in instances of ‘Pure Questioning’ is expected due to the nature of a topical debate, in which explanations are sought after by the audience members. In addition to that, on multiple occasions, the moderator is forced to rephrase the questions in order to urge panel members to react to those they have previously attempted to avoid.

5.2 The response space

Table 2 shows that questions (of any type) are rarely left ignored, with 83%-89% eliciting some type of reaction, over half of which are responses, with around 58% of all responses being answers. If we compare question types based on their response pattern, ‘Pure Questioning’ is distinct in that it has the highest frequency of answers: 39% of ‘Pure Questioning’ are answered while the other two question types, ‘Assertive Questioning’ and ‘Rhetorical Questioning’, are answered 14% and 13% of the time, respectively. These observations are exactly as our intuitions would suggest: assertive and

	Topical Debate
Pure Questioning	1201
~Ignored (no outgoing TA)	135
~with Reactions (existence of outgoing TA)	1066
~~with Responses (anchor a YA)	679
~~~with Answers (the content is MA)	468
Same speaker	53
Different speaker	406
Same and different speakers	9
<b>Assertive Questioning</b>	<b>246</b>
~Ignored (no outgoing TA)	43
~with Reactions (existence of outgoing TA)	203
~~with Responses (anchor a YA)	132
~~~with Answers (the content is MA)	34
Same speaker	7
Different speaker	27
Same and different speakers	0
Rhetorical Questioning	224
~Ignored (no outgoing TA)	24
~with Reactions (existence of outgoing TA)	200
~~with Responses (anchor a YA)	111
~~~with Answers (the content is MA)	29
Same speaker	24
Different speaker	5
Same and different speakers	0
<b>Questions of all types</b>	<b>1671</b>
~Ignored (no outgoing TA)	202
~with Reactions (existence of outgoing TA)	1469
~~with Responses (anchor a YA)	922
~~~with Answers (the content is MA)	531
Same speaker	84
Different speaker	438
Same and different speakers	9

Table 2: Analysis of Question-Response patterns (individual question types)

rhetorical questions by their very nature do not typically invite an answer, as part of the role they play is to convey information or provide an opinion. In terms of speaker patterns, a large proportion of pure and assertive questions are answered by a different speaker (87% and 79% respectively), in contrast to rhetorical questions where 83% were answered by

	Topical Debate
Inferences	5675
Premise is a question	104
Pure Questioning	35
Assertive Questioning	22
Rhetorical Questioning	47
Conclusion is a question	235
Pure Questioning	106
Assertive Questioning	63
Rhetorical Questioning	66
As immediate response to questions	132
Pure Questioning	48
Assertive Questioning	36
Rhetorical Questioning	48
As eventual response to questions	627
Pure Questioning	477
Assertive Questioning	70
Rhetorical Questioning	80
Conflicts	1240
Antecedent is a question	31
Pure Questioning	4
Assertive Questioning	17
Rhetorical Questioning	10
Consequent is a question	52
Pure Questioning	17
Assertive Questioning	24
Rhetorical Questioning	10
As immediate response to questions	44
Pure Questioning	12
Assertive Questioning	22
Rhetorical Questioning	10
As eventual response to questions	158
Pure Questioning	107
Assertive Questioning	27
Rhetorical Questioning	24

Table 3: Analysis of Question-Argument patterns

the same speaker. This is to be expected as rhetorical questions typically do not seek an answer from another interlocutor, being used rather as rhetorical tools, with the speaker asking and answering their own question as a way of bringing their point forward.

5.3 The argumentative impact of questions

Our analysis in Table 3 is divided into the two sides of the argumentative coin: inference (support) and conflict (attack). For each, we look first at the overall frequencies of the antecedent and consequent (for inferences, the premise, and conclusion) and the proportion of them that are realised by questions (‘the premise/conclusion/antecedent/consequent is a question’). This allows us to answer the question of whether (particular types of) questions are stereotypically associated with particular local roles in argumentative structures.

In the next step, we look at cases where infer-

ences and conflicts are direct responses to questions (‘As immediate response to’), i.e., their length of the chain of dialogical or argumentative structures between the question and an inference or conflict that follows it is exactly one. We then also identify how often inferences and conflicts occur in response to a question at any number of dialogically relevant steps later (‘As eventual response to’); i.e. where the chain is of length greater than one. At each step in the chain, there must, in the IAT analysis, be both dialogical relevance (i.e. a chain of TAs and locutions from the locution constituting the initial question) and propositional relevance (i.e. a chain of inferences, RAs, conflicts, CAs and/or rephrases, MAs from the propositional content of the initial question).

Table 3 indicates that questions are used as elements of both inferences and conflicts in about 3.1% of all of the structures. More specifically, questions are encountered as either antecedents or consequents in 3.3% of all conflicts in the topical debate data, which is slightly more frequently than questions used as either a premise or conclusion in inferences (3%). In the case of inferences, questions of all types are two times more frequent in a role of a conclusion rather than a premise with 2.1% and 0.9% respectively.

Closer inspection reveals particular roles that questions of different types play in the debate. ‘Pure Questioning’ make up almost half of all questions used as the conclusion in the inferential structures with the rest fairly equally split between assertive questions and rhetorical ones. This kind of use of pure questions is illustrated in Example (1-a) where Fiona Bruce uses her concern regarding the current ceasefire as a reason to inquire about UK’s further involvement in the peace talks. We hypothesize that speakers often give a reason as to why they are asking a particular question in order to justify it. A need for justification might be stronger in the case of ‘Pure Questioning’ since the speakers ask those questions with an intention of getting answers.

With respect to premises expressed by questions, it is most often ‘Rhetorical Questioning’ (45%) that is used (compared to 34% for ‘Pure Questioning’ and 21% for ‘Assertive Questioning’). As already discussed in Section 5.2, this type of question is often employed in a manner that allows speakers to make certain points or put forward ideas. The use of rhetorical questions as premises further supports

this theory: questions are used as a sophisticated linguistic means to express an opinion and to draw conclusions.

In terms of conflict, ‘Assertive Questioning’ is most commonly encountered in both antecedents and consequents (55% and 46% respectively). This suggests that when using assertive questions in the role of an antecedent (the source of a conflict), speakers tend to soften their disagreement via the use of assertive questions instead of expressing their position via an assertion. Frequent use of assertive questions as the consequent (the target) of conflict implies that this type of question provokes more disagreement than other types.

The number of inferences and conflicts used in responses (either immediate or eventual) to questions of all types indicates their particular role in the development of complex argumentative structures. Almost 7% of all inferential constructions are elicited by questions; in the case of conflicts, the number is slightly higher at 8%. These percentages are more indicative when put into perspective by the fact that questions make up less than 5% of all illocutionary structures.

‘Pure Questioning’ is by far the most common type of question making up around 72% of questions of all types. Surprisingly, it does not elicit a higher share of immediate responses than ‘Assertive Questioning’ or ‘Rhetorical Questioning’, neither for inferences nor conflicts. Only between 27% and 36% of all immediate responses were directed at pure questions. ‘Assertive Questioning’ on the other hand elicits the most immediate responses which are conflicts (making up 50% of responses). This supports previous observations regarding the role assertive questions play in provoking disagreement.

About 5.5% of inferences and 6.3% of conflicts are elements in chains of responses to questions. It is ‘Pure Questioning’ that elicits the majority of responses, with 76% and 82% in case of inferences and conflicts, respectively. This illustrates how pure questions are more than just information-seeking instruments. They trigger extended discussions that can be traced through a number of argumentative moves including those of inference and conflict. The diagram on the left in Figure 1 illustrates how one question can elicit a chain of responses including two linked arguments and a conflict followed up by another inference.

6 Conclusion

Our analysis here extends upon and refines the findings reported in Hautli-Janisz et al. (2022a), which illustrates the strategic role questions play in formulating various argumentative structures across genres of debate. For the first time, we are able to inspect the role of questions at scale, in the largest annotated corpus of dialogical argumentation, QT30 (Hautli-Janisz et al., 2022b). The results reported here scale up previous work in two distinct ways. First of all, there is explicit annotation of relationships diachronically through each hour, allowing tracking of ramifications and consequences of questions long after they have been uttered. Secondly, we also scale up the robustness of analysis by considering more debates – thereby covering more speakers (so we are at less risk of merely picking up on speaker idiosyncrasy), and more themes (so we are less at risk of merely picking up on thematic specificity). Some of the observations presented in the analysis are to be expected – that pure questions elicit more answers and that rhetorical questions are usually answered by the same speaker. Some are a little less straightforward – that assertive questions and rhetorical questions often present material that is used to support further argumentation, whereas the contents of pure questions typically end up being the conclusions of arguments. That assertive and rhetorical questions are used to provide an evidential basis for further reasoning further supports the hypothesis that such illocutions provide a rhetorical means to lower the interpersonal cost of making an assertion – a gambit to reduce dialogical risk. Sometimes, though, this gambit fails, as suggested by the fact that assertive questions in particular very often lie at the heart of conflict. The most striking new insight from this analysis is that across both inferential and conflict structures by which argumentation is driven forward, almost a quarter are triggered, either immediately or eventually, by questioning.

Questions are important to debate and reasoning: from journalistic punditry, from personal experience, from scholarly study across philosophy of science to rhetoric, from practical political training; wherever humans debate, the role of questions and question-answering is placed front and centre. This work, however, is the first time that it has become possible to quantify that role, and pin down exactly how it is, and to what extent, questions are keystone in the development of reasoning in debate.

Acknowledgements

The work reported on in this paper was funded by the Deutsche Forschungsgemeinschaft (DFG – German Research Foundation) under Germany’s Excellence Strategy – EXC-2035/1 – 390681379, the VolkswagenStiftung under grant Az. 98544 ‘Deliberation Laboratory’.

References

- Ginger Berninger and Catherine Garvey. 1981. Relevant replies to questions: Answers versus evasions. *Journal of Psycholinguistic Research*, 10:403–420.
- Dan Bohus and Alexander I. Rudnicky. 2005. [Sorry and I didn’t catch that! - an investigation of non-understanding errors and recovery strategies](#). In *Proceedings of the 6th SIGdial Workshop on Discourse and Dialogue*, pages 128–143, Lisbon, Portugal. Special Interest Group on Discourse and Dialogue (SIGdial).
- Katarzyna Budzynska, Mathilde Janier, Juyeon Kang, Chris Reed, Patrick Saint-Dizier, Manfred Stede, and Olena Yaskorska. 2014. Towards argument mining from dialogue. In *Proceedings of the Fifth International Conference on Computational Models of Argument (COMMA 2014)*, pages 185–196. IOS Press.
- Katarzyna Budzynska, Mathilde Janier, Chris Reed, and Patrick Saint Dizier. 2016. Theoretical foundations for illocutionary structure parsing. *Argument & Computation*, 7(1):91–108.
- Sasha Calhoun, Jean Carletta, Jason M. Brenier, Neil Mayo, Dan Jurafsky, Mark Steedman, and David Beaver. 2010. The next-format switchboard corpus: a rich resource for investigating the syntax, semantics, pragmatics and prosody of dialogue. *Language Resources and Evaluation*, 44(4):387–419.
- Jean Carletta, Amy Isard, Stephen Isard, Jacqueline C. Kowtko, Gwyneth Doherty-Sneddon, and Anne H. Anderson. 1997. The reliability of a dialogue structure coding scheme. *Computational Linguistics*, 23(1):13–31.
- Rory Duthie, John Lawrence, Katarzyna Budzynska, and Chris Reed. 2016. [The CASS technique for evaluating the performance of argument mining](#). In *Proceedings of the Third Workshop on Argument Mining (ArgMining2016)*, pages 40–49, Berlin, Germany. Association for Computational Linguistics.
- Jonathan Ginzburg, Zulipiye Yusupujiang, Chuyuan Li, Kexin Ren, and Pawel Lupkowski. 2019. Characterizing the response space of questions: a corpus study for English and Polish. In *Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue*, pages 320–330.
- Annette Hautli-Janisz, Katarzyna Budzynska, Conor McKillop, Brian Plüss, Valentin Gold, and Chris Reed. 2022a. Questions in argumentative dialogue. *Journal of Pragmatics*, 188:56–79.
- Annette Hautli-Janisz, Zlata Kikiteva, Wassiliki Siskou, Kamila Gorska, and Chris Reed. 2022b. QT30: A corpus of argument and conflict in broadcast debate. In *Proceedings of the Thirteenth International Conference on Language Resources and Evaluation (LREC 2022)*. European Language Resources Association (ELRA).
- Mathilde Janier, John Lawrence, and Chris Reed. 2014. OVA+: An argument analysis interface. In *Computational Models of Argument: Proceedings of COMMA*, volume 266, pages 463–464.
- Daniel Jurafsky, Elizabeth Shriberg, and Debra Biasca. 1997. Switchboard swbd-damsl shallow-discourse-function annotation – coders manual, draft 13. Technical report, University of Colorado Institute of Cognitive Science.
- Barbara Konat, John Lawrence, Joonsuk Park, Katarzyna Budzynska, and Chris Reed. 2016. [A corpus of argument networks: Using graph properties to analyse divisive issues](#). In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC’16)*, pages 3899–3906, Portorož, Slovenia. European Language Resources Association (ELRA).
- John Lawrence, Mark Snaith, Barbara Konat, Katarzyna Budzynska, and Chris Reed. 2017. Debating technology for dialogical argument: Sensemaking, engagement, and analytics. *ACM Trans. Internet Technol.*, 17(3):24:1–24:23.
- Matthew Purver, Jonathan Ginzburg, and Patrick Healey. 2001. On the means for clarification in dialogue. In *Proceedings of the Second SIGdial Workshop on Discourse and Dialogue*.
- Matthew Purver, Patrick G.T. Healey, James King, Jonathan Ginzburg, and Greg J. Mills. 2003. [Answering clarification questions](#). In *Proceedings of the Fourth SIGdial Workshop of Discourse and Dialogue*, pages 23–33.
- Chris Reed, Katarzyna Budzynska, Rory Duthie, Mathilde Janier, Barbara Konat, John Lawrence, Alison Pease, and Mark Snaith. 2017. The argument web: An online ecosystem of tools, systems and services for argumentation. *Philosophy & Technology*, 30(2):137–160.
- Kazuki Sakai, Ryuichiro Higashinaka, Yuichiro Yoshikawa, Hiroshi Ishiguro, and Junji Tomita. 2018. [Introduction method for argumentative dialogue using paired question-answering interchange about personality](#). In *Proceedings of the 19th Annual SIGdial Meeting on Discourse and Dialogue*, pages 70–79, Melbourne, Australia. Association for Computational Linguistics.

- John Searle. 1969. *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.
- John Searle and Daniel Vanderveken. 1985. *Foundations of Illocutionary Logic*. Cambridge: Cambridge University Press.
- Tanya Stivers and N.J. Enfield. 2010. [A coding scheme for question-response sequences in conversation](#). *Journal of Pragmatics*, 42(10):2620 – 2626. Question-Response Sequences in Conversation across Ten Languages.
- Jacky Visser, Barbara Konat, Rory Duthie, Marcin Koszowy, Katarzyna Budzynska, and Chris Reed. 2020. Argumentation in the 2016 us presidential elections: annotated corpora of television debates and social media reaction. *Language Resources and Evaluation*, 54:123–154.
- Bonnie Webber, Rashmi Prasad, Alan Lee, and Aravind Joshi. 2017. [The Penn Discourse Treebank 3.0 annotation manual](#). Technical report, LDC.
- Zhengzhe Yang and Jinho D. Choi. 2019. [FriendsQA: Open-domain question answering on TV show transcripts](#). In *Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue*, pages 188–197, Stockholm, Sweden. Association for Computational Linguistics.
- Paweł Łupkowski and Jonathan Ginzburg. 2016. Query responses. *Journal of Language Modeling*, 4:245–292.