Query Responses

Anonymous

ABSTRACT

In this article we consider the phenomenon of answering a query with a query. Although such answers are common, no large scale, corpus-based characterization exists, with the exception of clarification requests. After briefly reviewing different theoretical approaches on this subject, we present a corpus study of query responses in the British National Corpus and develop a taxonomy for query responses. We point at a variety of response categories that have not been formalized in previous dialogue work, particularly those relevant to adversarial interaction. We show that different response categories have significantly different rates of subsequent answer provision. We provide a formal analysis of the response categories in the framework of KoS.

1 INTRODUCTION

Responding to a query with a query is a common occurrence, representing on a rough estimate more than 20% of all responses to queries found in the British National Corpus (BNC).1 Research on dialogue has long recognized the existence of such responses. However, with the exception of one of its subclasses – albeit a highly substantial one – the class of query responses has not been characterized empirically in previous work.

The class that has been studied in some detail are Clarification Requests (thereafter referred to as CRs) (see e.g., Purver et al. 2001; Rodriguez and Schlangen 2004; Rieser and Moore 2005). However, in the spoken part of the BNC, using SCoRE (Purver 2001), we found 9,279 ?/? cross-turn sequences, whereas 41,041 ?/. cross-turn sequences, so the ?/? pairs constitute 22.61%.

Keywords: questions, query responses, corpus study, KoS

Journal of Language Modelling Vol i^2, No e^{i\pi} (1970), pp. 1–50
CRs can be triggered by any utterance, interrogative or otherwise. Researchers on the semantics and pragmatics of questions (see e.g., Carlson 1983; Wiśniewski 1995) have been aware for many years of the existence of one class of query responses – responses that express questions dependent in some sense on the question they respond to, as in (1a,b). This led Carlson to propose (1d) as a sufficient condition for a query response (cf. (1a,c)).

(1)

a. A: Who murdered Smith? B: Who was in town?
b. A: Who is going to win the race? B: Who is going to participate?
c. Who killed Smith depends on who was in town at the time.
d. q2 can be used to respond to q1 if q1 depends on q2.

How to define question dependence is an important issue if the criterion in (1d) is to have much substance. A number of proposals concerning dependence have been made in the literature, for instance Ginzburg (2012) offers the definition in (2):

(2) \( q_1 \) depends on \( q_2 \) iff any proposition \( p \) such that \( p \) resolves \( q_2 \), also satisfies \( p \) entails \( r \) such that \( r \) is about \( q_1 \). (Ginzburg 2012, (61b), p. 57)

For Ginzburg, this notion of dependence is an agent–relative notion, given the agent–relativity of the relation resolves. An arguably more open-ended view is taken by Roberts (1996), who suggests that a query move \( m \) is relevant in a context where \( q \) is the question under discussion if \( m \) is part of a strategy to answer \( q \) (Roberts 1996, p. 17). In similar fashion, Larsson (2002) and Asher and Lascarides (2003) argue that the proper characterization of query response is pragmatically based. Asher and Lascarides (2003) propose to characterize non-CR query responses by means of the rhetorical relation of question elaboration (Q-Elab) with stress on the plan-oriented relation between the

---

2 The agent–relativity of the relation resolves is argued for in great detail in (Ginzburg (1995)). resolves is the answerhood notion implicated in examples such as ‘…knows where she is’ and ‘…knows who came to the talk’, which is, arguably, relativized by agent goals and background knowledge.
initial question and the question expressed by the response. Q-Elab might be informally summarized as follows:

(3) If Q-Elab(α, β) holds between an utterance α uttered by A, where g is a goal associated by convention with utterances of the type α, and the question β uttered by B, then any answer to β must elaborate a plan to achieve g.

The relation of Q-Elab, motivated by interaction in cooperative settings, is vulnerable to examples such as those in (4). There is a reading of (4a) that can be characterized by using dependence (What I like depends on what you like), but it can also be used simply as a coherent retort. (4b) could possibly be used in political debate without necessarily involving any attempt to discover an answer to the first question

   b. A: What is Sarkozy going to do about it? B: What is Papan-dreou?

In the field of the logic of questions we can mention approaches proposed within Inferential Erotetic Logic (IEL) (Wiśniewski 1995, 2013) and inquisitive semantics (INQ) (Groenendijk 2009; Groenendijk and Roelofsen 2011). Although INQ and IEL represent different approaches to questions itself, both frameworks share a similar treatment of question dependency. In IEL, the central notion used to express dependency between questions is erotetic implication. Erotetic implication is a semantic relation between a question, Q, a (possibly empty) set of declarative well formed formulae, X, and a question, Q1. Intuitively, erotetic implication ensures the following: (i) if Q has a true direct answer and X consists of truths, then Q1 has a true direct answer as well ‘transmission of soundness\(^3\) and truth into soundness’ – cf. Wiśniewski 2003, p. 401), and (ii) each direct answer to Q1, if true, and if all elements of X are true, narrows down the class at which a true direct answer to Q can be found ‘open-minded cognitive usefulness’ – cf. Wiśniewski 2003, p. 402).

In the framework of inquisitive semantics the dependency relation has been analysed in terms of compliance. Roughly speaking, INQ treats

\(^3\) A question Q is sound iff it has a true direct answer (with respect to the underlying semantics).
questions as sets of possibilities or, in other words, as an issue to be resolved. The intuition behind the notion of compliance is to provide a criterion to “judge whether a certain conversational move makes a significant contribution to resolving a given issue” (Groenendijk and Roelofsen 2011, p. 167).

Other question generation mechanisms in a broadly dialogical context have been noticed in the literature. One such notion is askability. The intuition behind askability relates to the issue – when is it reasonable to (publicly) ask a question? Peliš and Majer (2010), applying a dynamic epistemic logic of questions combined with a public announcements’ logic for modelling communicative interaction and knowledge revision during this process, propose three conditions that have to be met by an agent in order to ask a question within a group of agents: (i) the answer is not known to the agent posing the question (non-triviality); (ii) each direct answer is considered as possible by the agent (admissibility); and (iii) at least one of the direct answers must be the right one in a given context (context condition).

Van Kuppevelt (1995) proposes topicality as the general organizing principle in discourse. The topic (for a discourse unit) is provided by an explicit or implicit question. van Kuppevelt does not consider simple question–query response pairs, but rather speaks about discourse units. However, the relation between such units is determined by the relation between aforementioned topic-providing questions. From the current perspective the most interesting is the notion of subtopic-constituting sub-question:

(5) An explicit or implicit question $Q_p$ is a subtopic-constituting subquestion if it is asked as the result of an unsatisfactory answer $A_{p-n}$ to a preceding question $Q_{p-n}$ with the purpose of completing $A_{p-n}$ to a satisfactory answer to $Q_{p-n}$. (van Kuppevelt 1995, p. 125)

Graesser et al. (1992) propose four question generation mechanisms for natural settings (especially in educational contexts). The first group consists of knowledge deficit questions. The other three groups are: common ground questions, social coordination questions and conversation-control questions. Common ground questions, like ‘Are we working on the third problem?’ or ‘Did you mean the independent variable?’, are asked to check whether knowledge is shared
between dialogue participants. Social coordination questions relate to different roles of dialogue participants, such as in student–teacher conversations. Social coordination questions are requests for permission to perform a certain action or might be treated as indirect request for the the addressee to perform such an action (e.g., ‘Could you graph these numbers?’, ‘Can we take a break now?’). Conversation-control questions, as it is indicated by their name, aim at manipulating the flow of a dialogue or the attention of its participants (e.g., ‘Can I ask you a question?’).

How many kinds of query responses are there and what aspects of context or agents’ information states are needed to characterise these? In order to better understand the nature of query response, we ran a corpus study on one large balanced corpus, the British National Corpus (BNC), and several smaller, more domain specific corpora, a selection from CHILDES (parent/child interaction; MacWhinney 2000), AMEX (interactions in the travel domain; Kowtko and Price 1989), and BEE (tutor/student interaction; Rosé et al. 1999). The results we obtained, discussed in Section 3 of this paper, show that, apart from CRs, dependent questions are indeed by far the largest class of query responses. However, they reveal also the existence of a number of response categories, characterisable neither as dependent questions nor as plan supporting responses. To exemplify, these include:

- a class akin to what Conversation Analysts refer to as *counters* (Schegloff 2007) – responses that attempt to foist on the conversation a distinct issue from the current discourse topic and
- *situation-relevant responses* – responses that ignore the current topic but address the situation it concerns.

Just as wide coverage is an important goal for any computational theory of sentential grammar (tempered by some notion of ‘strong generative capacity’, i.e., attaining this in a principled way), the same goal *mutatis mutandis* applies to theories of dialogue; their corresponding aim is to characterise in a principled way the relevance or coherence as a wide range of utterance sequences. Attaining wide coverage for the particular case of the response space of a query naturally has significant practical importance for dialogue management and the design of user interfaces. Beyond that general goal, a better understanding of e.g., *counters* and *situation-relevant responses* is important for adver-
serial interaction (courtroom, interrogation, argumentation, certain games).

The rest of the paper is structured as follows: in Section 2 we present the taxonomy underlying our corpus study; Sections 3 and 4 describe the results, with issues concerning annotation reliability discussed in section 5; in Section 6 we show how to analyse the relevance of each of the response categories emergent in the corpus study. We do this in terms of information state transitions of two interlocutors participating in a dialogue, using the dialogue framework, KoS. We conclude with a brief cross-theoretical evaluation of potential analyses of the various response classes and with possibilities for future work.

2 A CORPUS–BASED TAXONOMY OF QUERY RESPONSES

In this section we present a corpus based taxonomy of query responses. It was designed on the basis of 1,051 examples of query – query-response pairs obtained from BNC. Initially, examples were obtained by using the search engine SCoRE (Purver 2001). Subsequently, cross talk and tag questions were eliminated manually. The annotation was performed by the first author; we discuss the reliability of this annotation in Section 5. In what follows we describe and exemplify each class of the resulting taxonomy. To make the description clearer we will use $q_1$ for the initial question posed and $q_2$ for a question given as a response to $q_1$. The taxonomy is focused on the function of $q_2$ in a dialogue.

2.1 Clarification requests (CR)

Clarification requests are all query responses that concern the content or form of $q_1$ that was not completely understood. This class contains intended content queries (see example 6a), repetition requests (example 6b) and relevance clarifications (example 6c). In this article we will not consider this class in detail, mainly because of existing, detailed work on this subject (see e.g., Purver 2006; Ginzburg 2012).

---

KoS is a toponym – the name of an island in the Dodecanese archipelago – bearing a loose connection to conversation oriented semantics (Ginzburg 2012, p. 2).
Query Responses

(6) a. A: What’s Hamlet about?
   B: Hamlet? [KPW, 945–946]

b. A: Why are you in?
   B: What?
   A: Why are you in? [KPT, 469–471]

c. A: Is he knocked out?
   B: What do you mean? [KDN, 3170–3171]

2.2 Dependent questions (DP)

By a dependent question we understand \( q_2 \) where a dependency statement as in (1d; see page 2) could be assumed to be true. The following examples illustrate this:

(7) a. A: Do you want me to \(<\text{pause}>\) push it round?
   B: Is it really disturbing you? [FM1, 679–680]
   (cf. Whether I want you to push it depends on whether it really disturbs you.)

b. A: Any other questions?
   B: Are you accepting questions on the statement of faith at this point? [F85, 70–71]
   (cf. Whether more questions exist depends on whether you are accepting questions on the statement of faith at this point.)

c. A: Does anybody want to buy an Amstrad? \(<\text{pause}>\)
   B: Are you giving it away? [KB0, 3343–3344]
   (cf. Whether anybody wants to buy an Amstrad depends on whether you are giving it away.)

2.3 ‘How should I answer this?’ questions (FORM)

This class consists of query responses addressing the issue of the way the answer to \( q_1 \) should be given. In other words, whether the answer to \( q_1 \) will be satisfactory to A depends on \( q_2 \). This relation between \( q_1 \) and \( q_2 \) might be noticed in the following examples. Consider (8a). The way B answers A’s question in this case will be dictated by A’s answer to \( q_2 \) – whether or not A wants to know details point by point.

---

\( ^5 \)This notation indicates the BNC file (KPW) together with the sentence numbers (945–946).
Anonymous

(8)  a. A: Okay then, Hannah, what, what happened in your group?
    B: Right, do you want me to go through every point? [K75, 220–221]

    b. A: Where’s that one then?
    B: Erm, you know Albert Square? [KBC, 506–507]

    c. A: Another thing I found out today was do we know where our main supplier of our coffee is.
       Any guesses?
       B: Which country? [G3U, 251–253]

2.4 Requests for underlying motivation (MOTIV)

In the case of requests for underlying motivation $q_2$ addresses the issue of the motivation underlying asking $q_1$. Whether an answer to $q_1$ will be provided depends on the answer to $q_2$ (i.e. providing proper reasons for asking $q_1$). In this aspect this class differs from the previous ones, where we may assume that an agent wishes to provide an answer to $q_1$. Most of query responses of this kind are of the form ‘Why?’ (32 out of 41 gathered examples, see example 9a), but also other formulations were observed (9 out of 41, see examples (9b) and (9c)). Most direct questions of this kind are: What’s it got to do with you?; What’s it to you?; Is that any of your business?; What’s that gotta do with anything?

(9)  a. A: What’s the matter?
    B: Why? [HDM, 470–471]

    b. A: Out, how much money have you got in the building society?
    B: What’s it got to do with you? [KBM, 2086–2087]

    c. A: Just what the fucking do you think you’re doing?
    B: Is that any of your business? [KDA, 1308–1309]

2.5 ‘I don’t want to answer your question’ (NO ANSW)

The role of query responses of this class is to signal that an agent does not want, at least at a given stage of the conversation, to provide an answer to $q_1$. Instead of answering $q_1$ the agent provides $q_2$ and attempts to “turn the table” on the original querier, as exemplified in examples (10 a–b).

(10)  a. A: Yeah what was your answer?
     B: What was yours? [KP3, 636–637]
Query Responses

b. A: come on Stacey get on with it (pause) can you move up a bit?
   B: What? (unclear) why didn’t you pull the bench out? [KCG, 378–379]

c. A: What about my fifty p?
   B: Fucking hell, where’s my tenner? [KDA, 3527–3528]

d. A: Why is it recording me?
   B: Well why not? [KSS, 43–44]

2.6 Indirect responses (IND)

This class consists of query responses, which provide (indirectly) an answer to $q_1$. Interestingly, providing an answer to $q_2$ is not necessary in this case. Consider (11a). By asking the question *Do you know how old this sweater is?*, B clearly suggests that the answer to A’s question is negative. Moreover, B does not expect to obtain an answer to his/her question. This may also be observed in examples (11b) (*of course I am Gemini*) and (11c) (*no, my job is not safe*).

(11) a. A: Is that an early Christmas present, that sweater?
   B: Do you know how old this sweater is? [HM4, 7–8]

   b. A: Are you Gemini?
   B: Well if I’m two days away from your, what do you think?
   [KPA, 3603–3604]

   c. A: Is your job safe?
   B: Well, whose job’s safe? [G5L, 130–131]

Another means of providing indirect answers can also be observed in the corpus data. These are cases in which by asking $q_2$ an agent already presupposes the answer to $q_1$. (12a) illustrates this – we note that a positive answer to $q_1$ is presupposed in B’s question. A similar situation can be observed in examples (12b) (*no, I have not tasted this*) and (12c) (*I will help you*).

(12) a. A: I’ve got to do the washing up?
   B: Shall I, shall I come and help you? [KPU, 1861–1862]

   b. A: have you tasted this?
   B: are they nice? [KPY, 653–654]

   c. A: Will you help with the *<pause>* the paint tonight?
   B: What can I do? [KE4, 3263–3264]
Anonymous

2.7 ‘I ignore your question’ (IGNORE)

The final class in the taxonomy involves cases where \( q_2 \) does not address \( q_1 \), but is, nonetheless, related to the situation associated with \( q_1 \). This is evident in example (13c). A and B are playing Monopoly. A asks a question, which is ignored by B. It is not that B does not want to answer A’s question and that’s why he/she asks \( q_2 \). Rather, B ignores \( q_1 \) and asks a question related to the situation (in this case the board game).

(13) a. A: Well do you wanna go down and have a look at that now? \(<\text{pause}>\) While there’s workmen there?
   B: Why haven’t they finished yet? [KCF, 617–619]

b. A: Just one car is it there?
   B: Why is there no parking there? \(<\text{unclear}>\) [KP1, 7882–7883]

c. A: I’ve got Mayfair \(<\text{pause}>\) Piccadilly, Fleet Street and Regent Street, but I never got a set did I?
   B: Mum, how much, how much do you want for Fleet Street? [KCH, 1503–1504]

2.8 Summary

In this section a corpus based taxonomy of query responses was presented. Seven classes of query-responses were described. The classification is focused on the function of \( q_2 \) (question given as a response) to \( q_1 \) (initial question). In what follows we present a corpus study that led to the classification. First, a study using the BNC is discussed, then the class distribution over specific genres is presented. Subsequently, we consider the issue of annotation reliability.

3 RESULTS

As we noted, this study used a sample of 1,051 query – query response pairs from the BNC. The procedure for obtaining the sample was the following. First the search engine SCoRE was used on the whole spoken part of the BNC using as the search string: \? $ | ? $\(^6\). Following

\(^6\) ‘? $’ expression would match any sentence/turn with a question mark at the end and the pipe character matches the break between sentences/turns. For
this, the search results were checked manually. The collected sample covers a wide range of dialogue domains, like interviews, radio and TV broadcasts, tutorials, meetings, training sessions or medical consultations (blocks D, F, G, H, J, K of the BNC). The summary of dialogue domains for the sample is presented in Table 1.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Frequency</th>
<th>% of the Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>free conversation</td>
<td>940</td>
<td>89.44</td>
</tr>
<tr>
<td>educational context (lesson, tutorial, training)</td>
<td>36</td>
<td>3.43</td>
</tr>
<tr>
<td>meeting (public meeting, seminar, conference)</td>
<td>27</td>
<td>2.57</td>
</tr>
<tr>
<td>radio broadcast</td>
<td>25</td>
<td>2.38</td>
</tr>
<tr>
<td>interview</td>
<td>15</td>
<td>1.43</td>
</tr>
<tr>
<td>medical consultation</td>
<td>4</td>
<td>0.38</td>
</tr>
<tr>
<td>TV broadcast</td>
<td>4</td>
<td>0.38</td>
</tr>
<tr>
<td>Total</td>
<td>1,051</td>
<td>100</td>
</tr>
</tbody>
</table>

The sample was classified and annotated by the first author with tags presented in Table 2.

<table>
<thead>
<tr>
<th>Tag</th>
<th>query-response type</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>clarification requests</td>
</tr>
<tr>
<td>DP</td>
<td>dependent questions</td>
</tr>
<tr>
<td>FORM</td>
<td>questions considering the way of answering q1</td>
</tr>
<tr>
<td>MOTIV</td>
<td>questions about the underlying motivations behind asking q1</td>
</tr>
<tr>
<td>NO ANSW</td>
<td>questions aimed at avoiding answering q1</td>
</tr>
<tr>
<td>IND</td>
<td>questions with a presupposed answer</td>
</tr>
<tr>
<td>IGNORE</td>
<td>questions ignoring q1</td>
</tr>
</tbody>
</table>

To guide the classification process we used the following questions:

1. (CR) Is q2 a query about something not completely understood in q1?
2. (DP) Is it the case that the answer to q1 depends on the answer to q2?

more details about the SCoRE syntax see http://www.dcs.qmul.ac.uk/imc/ds/score/help.html
Table 3:
Frequency of query–query response categories in the BNC. The parenthesized percentage is the category’s percentage of the sample that excludes CRs.

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>% of the Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>832</td>
<td>79.16</td>
</tr>
<tr>
<td>DP</td>
<td>108</td>
<td>10.28 (49.31)</td>
</tr>
<tr>
<td>MOTIV</td>
<td>41</td>
<td>3.90 (18.72)</td>
</tr>
<tr>
<td>NO ANSW</td>
<td>26</td>
<td>2.47 (11.87)</td>
</tr>
<tr>
<td>FORM</td>
<td>16</td>
<td>1.52 (7.31)</td>
</tr>
<tr>
<td>IND</td>
<td>22</td>
<td>2.09 (10.05)</td>
</tr>
<tr>
<td>IGNORE</td>
<td>6</td>
<td>0.57 (2.74)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1,051 (219)</strong></td>
<td>100</td>
</tr>
</tbody>
</table>

3. (MOTIV) Does q2 address the motivation underlying asking q1?
4. (NO ANSW) Is it the case that q2 enables the speaker to avoid answering q1 while attempting to force the other speaker to answer q2 first?
5. (FORM) Is it the case that the way the answer to q1 will be given depends on the answer to q2?
6. (IND) Is it the case that q2 is rhetorical and in this sense it does not need to be answered and provides (indirectly) an answer to q1?
7. (IGNORE) Does q2 relate to the situation described by q1?

The results of the performed classification are presented in Table 3. The parenthesized percentage is the category’s percentage of the sample that excludes CRs.

Putting aside CRs, the largest class is DP. What is perhaps striking is the relatively large frequency of adversarial responses (the classes MOTIV, NO ANSW, IGNORE).

We also compared which query categories lead to a subsequent answer, either about q2 or about q1. Bearing in mind that our taxonomy is focused on the function of q2 in a dialogue we would expect the following results.

**DP** Answering q2 should lead to answering q1. The figures for q1 and q2 should be similar.

**FORM** Whether the answer to q1 will be satisfactory to A depends on q2. q2 addresses only the form of answer to q1, hence is somewhat less important than with DP – the number of answers to q1 could be higher than for q2.
MOTIV Whether an answer to q1 will be provided depends on a satisfactory answer to q2. The numbers for q1 and q2 should be comparable, though q1 may be somewhat smaller.

NO ANSW Instead of answering q1 the agent provides q2 and attempts to “turn the table” on the original querier. The original querier is pressured to answer q2 and put q1 aside. Hence, the numbers for q1 should be significantly smaller than for q2.

IND q2 provides (indirectly) an answer to q1. Providing an answer to q2 is not necessary in this case, so the numbers should be low here. q1 is answered by q2 definitionally.

IGNORE q2’s speaker shows lack of interest in q1, but since q2 relates to the situation associated with q1, there is some expectation that q2 be responded to. Thus the numbers for q1 should be significantly smaller than for q2. Moreover, the numbers for q2 should also be rather low (asking q2 is not very cooperative).

The results of the data analysis are presented in Table 4. They are in line with the discussed intuitions of the taxonomy.

4 CLASS DISTRIBUTION OVER SPECIFIC GENRES

We conducted our study in the BNC since it is a general corpus with a variety of domains and genres. However, we wanted also to check how the classes are distributed in more genre specific corpora. To do this we decided to study the following corpora:

- The Child Language Data Exchange System (CHILDES; MacWhinney 2000), which contains adult-child conversations.
- The Basic Electricity and Electronics Corpus (BEE; Rosé et al. 1999), which contains tutorial dialogues from electronics courses.
Table 5: Frequency of query – query response categories (CHILDES). The parenthesized percentage is the category's percentage of the sample that excludes CRs.

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>% of the Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>319</td>
<td>88.12</td>
</tr>
<tr>
<td>DP</td>
<td>11</td>
<td>3.04 (25.58)</td>
</tr>
<tr>
<td>MOTIV</td>
<td>2</td>
<td>0.55 (4.65)</td>
</tr>
<tr>
<td>NO ANSW</td>
<td>5</td>
<td>1.38 (11.63)</td>
</tr>
<tr>
<td>FORM</td>
<td>3</td>
<td>0.83 (6.98)</td>
</tr>
<tr>
<td>IND</td>
<td>5</td>
<td>1.38 (11.63)</td>
</tr>
<tr>
<td>IGNORE</td>
<td>17</td>
<td>4.70 (39.53)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>362 (43)</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

Table 6: Frequency of query – query response categories (BEE). The parenthesized percentage is the category's percentage of the sample that excludes CRs.

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>% of the Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>10</td>
<td>22.22</td>
</tr>
<tr>
<td>DP</td>
<td>28</td>
<td>62.22 (80)</td>
</tr>
<tr>
<td>NO ANSW</td>
<td>6</td>
<td>13.33 (17.14)</td>
</tr>
<tr>
<td>IGNORE</td>
<td>1</td>
<td>2.22 (2.86)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>45 (35)</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

Table 7: Frequency of query – query response categories (AMEX). The parenthesized percentage is the category's percentage of the sample that excludes CRs.

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>% of the Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>1</td>
<td>12.5</td>
</tr>
<tr>
<td>DP</td>
<td>7</td>
<td>87.5 (100)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>8 (7)</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

As is apparent, the DP class is the second largest class in the

- The SRI/CMU American Express dialogues (AMEX; Kowtko and Price 1989), which contains conversations with travel agents.

As with the BNC study, data was initially obtained by using the search engine SCoRE. Subsequently, cross talk and tag questions were eliminated manually. The annotation was then performed by the first author. 362 examples were obtained from the sample of the CHILDES corpus (files bates, belfast, manchester/anne); 45 examples were obtained from the whole BEE corpus and 8 from the whole AMEX corpus (the low numbers for BEE and AMEX are caused by the significantly smaller size of these corpora in comparison to BNC and CHILDES). The results of the classification applied to these corpora are presented in Tables 5, 6 and 7. The parenthesized percentage is the category's percentage of the sample that excludes CRs.
CHILDES corpus and is the largest class in the task oriented dialogues obtained from the BEE and AMEX corpora. As for the adversarial classes (MOTIV, NO ANSW, IGNORE), these are very rare in task oriented dialogues. One exception is the NO ANSW class in the case of BEE corpus. Here the percentage of NO ANSW questions is even higher than in the BNC and CHILDES. This type of query response is used in a teaching context to encourage a student to provide his/her answer to the teacher’s question (e.g., Student: can you remind me which colors mean what on the different resistors?; Tutor: Is that the first thing you need to know? [log-stud31]). When it comes to the CHILDES corpus, a large percentage of IGNORE query responses was observed – in all examples it was a child, who used this kind of query response. One can also note that for NO ANSW, FORM and QA, the frequency is similar for the CHILDES and the BNC corpora. The summary of the classes’ distribution in the BNC, CHILDES, AMEX and BEE is presented in Figure 1.

We also compared which query categories lead to a subsequent
answer, either about $q_2$ or about $q_1$. The results are presented in Table 8. In terms of answer analysis, task oriented dialogues are interesting in the context of the DP query response class. For all observed examples, an answer was provided to $q_2$ and to $q_1$. The NO ANSW category also behaves in line with the observations for BNC. We can observe the fulfilment of certain of our predictions in the case of the CHILDES corpus. When it comes to interaction with children neither maintaining attention nor topic continuity are a given, and this can be observed in the data. In the case of DP questions we still have a high number of answers provided for $q_2$, but the number of answers provided to $q_1$ is relatively low. As for the IGNORE class, our prediction in general was that the number of answers provided for $q_2$ should be low (since the behaviour it represents is not very cooperative). However, in the CHILDES case we observe a high number of provided answers. In our sample it was a child who posed the IGNORE query response, and this offers the basis for an explanation of the results: child-adult conversation generally demands of an adult to provide answers to a child’s questions, even if the question somehow deviates from the topic of the conversation.

5 ANNOTATION RELIABILITY

As we mentioned above, the annotation process was performed by the first author. In order to check the reliability of the classification process, inter- and intra-annotator study was performed.

For the inter-annotator study a sample of 100 randomly chosen
examples of query – query response (retrieved from all four analysed corpora) was used. The distribution of the classes was in line with the distribution observed by the primary annotator: CR: 31 examples; DP: 32 examples; MOTIV: 11 examples; NO ANSW: 8 examples; FORM: 5 examples; IND: 7 examples; IGNORE: 6 examples.

All the examples were supplemented with a context. The guideline for annotators (see appendix) contained explanations of all the classes and examples of question-responses of each category. Also the OTHER category was included. The instruction was to annotate the query response to the first question in each example. The control sample was annotated by four annotators (three experienced linguists and one a logician with moderate experience in corpus annotation).

The reliability of the annotation was evaluated using $\kappa$ (Carletta 1996), established by using the R statistical software (R Core Team 2013; version 3.1.2) with the irr package (Gamer et al. 2012). The interpretation of the kappa values is based on that of Viera and Garrett (2005).

The Fleiss $\kappa$ for all five annotators was 0.64 (i.e. substantial) with agreement 51% over 100 cases. The agreements between the main and other annotators were all substantial:

1. main and second annotator: $\kappa = 0.67$ with 73% agreement;
2. main and third annotator: $\kappa = 0.65$ with 72% agreement;
3. main and fourth annotator: $\kappa = 0.61$ with 69% agreement;
4. main and fifth annotator: $\kappa = 0.63$ with 70% agreement.

When it comes to detailed analysis of the annotation we will start with the OTHER category. The annotators were given the option of using this category and this option was not used frequently (from a sample of 100 cases): annotator 1: 0, annotator 2: 3, annotator 4: 6, annotator 4: 8, annotator 5: 0.

When we take a closer look on the disagreements between the main and the fourth annotator (the lowest agreement observed) we can notice that the most problematic cases were DP vs. IGNORE (5 cases). This fact is quite surprising since these categories are rather distant. This suggests most plausibly that the fourth annotator had misunderstood the category IGNORE).
Anonymous

An analysis of the cases which involved the most disagreement suggests that there are not infrequently cases which involve genuine ambiguities given the intentional nature of many of the dialogical relations tying together queries – these naturally enough get exacerbated for annotators required to make decisions in a largely context independent manner.

Thus, (14a) was annotated as DP by two annotators, as IGNORE by two annotators, and as NO ANSw by one annotator; in the context of adult/child interaction, IGNORE is possibly more likely – the child observing the same situation as the parent but ignoring politeness in trying to impose the issue momentarily captivating her own interest; at the same time a DP reading is potentially plausible given the plausibility of the assumption *What a fireman does with his axe depends on where his axe is*. In a similar fashion, (14b) was annotated as DP by three annotators and by two annotators as NO ANSw. Both are potentially plausible classifications: in a cooperative setting (as in a dinner enjoyed by a couple in a restaurant) DP is more appropriate (*What Norrine will have as a starter depends on what Chris wants*), whereas in a more adversarial setting the query response can simply be a means of avoiding the initial question.

(14) a. between DP and IGNORE: Parent: what does a fireman do with his axe? Child: where’s his axe is?
   b. between DP and NO ANSw: Chris: What would you like to have start with? Norrine: What do you want?

In light of this, we hypothesize that a more satisfying account of annotator reliability for this task would involve developing an annotation model, which accommodates ambiguity in annotation, as for instance in work on the basis of crowdsourced labels, as pioneered in *(Passonneau and Carpenter 2013)*. This constitutes work we hope to perform in the future.

For the intra-annotator study another control sample of 100 examples was randomly chosen from the gathered data. The distribution of the classes was similar to the one in the first control sample. In this case the agreement of the coding between the first annotation and the one obtained in the study was substantial ($\kappa = 0.78$).
6 MODELING QUERY RESPONSE CATEGORIES IN KOS

6.1 Dialogue Gameboards, Conversational Rules, and Dialogical Relevance

We offer a formal explication of the coherence that underlies the various different types of query responses within the framework of KoS (Ginzburg 2012). We offer here an analogy to formal syntax. When one discovers a class of constructions $C$ in need of analysis, one means of showing that a given formalism $F$ is adequate involves showing that $F$’s weak (strong) generative capacity properly includes the string set (analysis trees etc) corresponding to $C$. Within dialogue similar desiderata exist, where constructions are replaced by pairs (or longer sequences of) coherent utterances. We seek to show that KoS’s notion of coherence properly includes the class of queries and their questions responses, a demonstration that to the best of our knowledge has not hitherto been attempted for any dialogue formalism.

KoS is a framework for dialogue formulated using Type Theory

---

7 An anonymous reviewer for this journal cautions us about the analogy between syntactic grammaticality and dialogue coherence. We agree that the analogy with syntax should not be exaggerated. There are differences. But the analogy between syntactic ungrammaticality and dialogical incoherence is not entirely far fetched: if one says something incoherent, one could be adjudged to be linguistically incompetent, just like with ungrammaticality. With the latter one can use repair mechanisms to fix ungrammaticality (‘I know who did Mary like, I mean who she liked.’), just as with the former (A: Who came yesterday? B: I’m having a coffee. A: Did you hear what I said? B: Oh sorry um yes, no I have no idea.). Of course, given the possibility of interpreting incoherence as intended irrelevance, one can often draw that as a possible inference, but grammaticality errors also potentially push us to expect repair, to view the other speaker as momentarily confused etc. The same reviewer points out that weak/strong generative capacity are not necessarily notions to be held as some kind of ideal for scientific explanation. Whatever one thinks of those notions, we think that they were, nonetheless, useful in stimulating syntactic research in the 60s to 80s in trying to figure out which constructions stretch a given formalism to its limit (e.g., phrase structure grammars and cross-serial dependencies.). Similar considerations apply at present mutatis mutandis to formal dialogue theory. We return to this issue and how it relates to cross-theoretical comparison in Section 7.

8 Ginzburg (2010, 2012) sketched such a characterization for the entire class of queries and their responses, though without a detailed corpus study.
Anonymous

with Records (TTR; Cooper 2005, 2012; Cooper and Ginzburg 2015). It provides formal underpinnings for the information state approach to dialogue management (Larsson and Traum 2003) and underlies dialogue systems such as GoDiS (Larsson 2002) and CLARIE (Purver 2006). On the approach developed in KoS, there is actually no single context – instead of a single context, analysis is formulated at a level of information states, one per conversational participant. The type of such information states is given in (15a). We leave the structure of the private part unanalyzed here, as with one exception all our characterizations do not make reference to this; for one approach to private, see (Larsson 2002). The dialogue gameboard represents information that arises from publicized interactions. Its structure is given in (15b) – the spkr, addr fields allow one to track turn ownership, Facts represents conversationally shared assumptions, Pending and Moves represent respectively moves that are in the process of/have been grounded, QUD tracks the questions currently under discussion: 9

(15) a. TotalInformationState (TIS) = def
\[
\text{dialoguegameboard : DGBTType}
\]
\[
\text{private : Private}
\]

b. DGBTType = def
\[
\text{spkr: Ind}
\]
\[
\text{addr: Ind}
\]
\[
\text{utt-time: Time}
\]
\[
\text{c-utt : addressing(spkr,addr,utt-time)}
\]
\[
\text{Facts : Set Proposition)}
\]
\[
\text{Pending : list(loc(utionary)prop(osition))}
\]
\[
\text{Moves : list(loc(utionary)prop(osition))}
\]
\[
\text{QUD : poset(Question)}
\]

A dialogue gameboard c1 will be a record r1 such that (16a) holds; by definition this means that: 10 (i) the set of labels of r1 needs

9 The motivation for Pending and the type loc(utionary)prop(osition) is explained in Section 6.4.

10 For a more detailed discussion and exemplification see Cooper and Ginzburg (2015), section 2.2.
to be a superset of the set of labels of DGBT and (ii) for each judgement constituent of DGBT \( l_k : T_k \), the value \( r1 \) gets for that label, denoted by \( r1 \cdot l_k \), it is the case that \( r1 \cdot l_k : T_k \). Thus, concretely in this case, \( r1 \) should have the make up in (16b) and the constraints in (16c) need to be met:

(16) a. \( r1 : \text{DGBT} \)
   
   b. 
   \[
   \begin{align*}
   \text{spkr} &= A \\
   \text{addr} &= B \\
   \text{utt-time} &= t1 \\
   \text{c-utt} &= \text{p}_{\text{utt}}(A,B,t1) \\
   \text{Facts} &= \text{cg1} \\
   \text{Pending} &= \langle p1, \ldots, pk \rangle \\
   \text{Moves} &= \langle m1, \ldots, mk \rangle \\
   \text{QUD} &= Q
   \end{align*}
   \]

   c. A: Ind, B: IND, t1: TIME, \( \text{p}_{\text{utt}}(A,B,t1) : \text{addressing}(A,B,t1) \),
   
   \( \text{cg1} : \text{Set(Proposition)} \), \( \langle p1, \ldots, pk \rangle : \text{list(locutionaryProposition)} \),
   
   \( \langle m1, \ldots, mk \rangle : \text{list(locutionaryProposition)} \), Q : \text{poset(Question)}

The basic units of change are mappings between dialogue gameboards that specify how one gameboard configuration can be modified into another on the basis of dialogue moves. We call a mapping between DGBT types a conversational rule. The types specifying its domain and its range we dub, respectively, the preconditions and the effects, both of which are subtypes of DGBT. We explain shortly how this allows one to capture the coherence of responses.

We start by specifying how a question becomes established as in the DGB. The rule in (17) says that given a question \( q \) and \( \text{ASK}(A,B,q) \) being the LatestMove, one can update QUD with \( q \) as the maximal element of QUD (henceforth, a QUD–maximal element or Max-QUD, the “discourse topic”).

\[11\]

Here, as in the rest of the paper, we make use of manifest fields (Coquand et al. (2003)). A manifest field \( [\ell = a : T] \) is a convenient notation for \( [\ell : T_a] \) where \( T_a \) is a singleton type whose only witness is \( a \). Singleton types are introduced by the clauses in (17).

1. If \( a : T \) then \( T_a \) is a type.
2. \( b : T_a \) iff \( b = a \)
In order to avoid the prolixity exemplified in (17), the rules in this paper employ a number of abbreviatory conventions. First, instead of specifying the full value of the list Moves, we usually record merely its first member, which we call ‘LatestMove’. Second the preconditions can be written as a merge of two record types $DGBTy^<-\land_{merge} PreCondSpec$, one of which $DGBTy^<$ is a subtype of DGBTy and therefore represents predictable information common to all conversational rules; $PreCondSpec$ represents information specific to the preconditions of this particular interaction type. Similarly, the effects can be written as a merge of two record types $DGBTy^0 \land_{merge} ChangePrecondSpec$, where $DGBTy^0$ is a supertype of the preconditions and $ChangePrecondSpec$ represents those aspects of the preconditions that have changed.\textsuperscript{12} So we can abbreviate (17) as (18b):

\textsuperscript{12}This procedure is described in much more general terms using the operation of asymmetric merge in (Cooper (2016)), who shows the use of this operation for a wide range of semantic uses.
We can exemplify how this rule works. Assume (19a) to be a record that satisfies the preconditions of the type (18b), in other words it is a record which is of the type of the type assigned to ‘pre’ in (17) or in abbreviated form in (18b). Hence it constitutes the appropriate context for Ask QUD–incrementation. The output of that rule is (19b):

We also assume an analogue of (18b) for assertion, given in (20). In an interactive setting A asserting $p$ raises the issue $p$? for B—she can then either decide to discuss this issue (as a consequence of the rule QSPEC introduced below as (23)) or accept it as positively resolved (as a consequence of the rule (21)):
An obvious complement to QUD incrementation is a principle controlling QUD downdate. Since QUD consists of questions that are unresolved relative to FACTS, QUD downdate is formulated simultaneously with FACTS update: when \( p \) is added to FACTS, one needs to verify for all existing elements of QUD that they are not resolved by the new value of FACTS. This joint process of FACTS update / QUD downdate is formulated in (21): given an acceptance or confirmation of \( p \) by B, \( p \) can be unioned into FACTS, whereas QUD is modified by the function NonResolve. NonResolve is a function that maps a partially ordered set of questions \( \text{poset}(q) \) and a set of propositions \( P \) to a partially ordered set of questions \( \text{poset}^r(q) \) which is identical to \( \text{poset}(q) \) modulo those questions in \( \text{poset}(q) \) resolved by members of \( P \).

(21) \text{Fact Update/ QUD Downdate } = def \\
\[ \begin{align*}
\text{pre} & : \text{LatestMove} = \text{Accept(spkr,addr,p)} : \text{IllocProp} \\
\text{effects} & : \text{qud} = \langle p?,\text{pre.qud}\rangle : \text{poset(Question)}
\end{align*} \]

a. \text{NonResolve } = def \\
\[ \begin{align*}
B : \text{Ind} \\
F : \text{set(Prop)} \\
Q : \text{poset(InfoStruc)}
\end{align*} \]
\[ \begin{align*}
Q' : \text{poset(InfoStruc)} \\
c1 : Q' \subset r.Q \\
c2 : \forall q_0 \in Q' \exists f \in F \\
\text{Resolve}(f,q_0,q)
\end{align*} \]

With this in hand, we can now turn to explaining how dialogical relevance gets handled in KoS. Pretheoretically, relevance relates an utterance \( u \) to an information state \( I \) just in case there is a way to successfully update \( I \) with \( u \). Ginzburg (2012) defines two notions of
Query Responses

relevance, a simpler one at the level of moves, i.e. illocutionary contents of utterances, and a somewhat more complex one at the level of utterances. For expository simplicity, we restrict attention here to the former and refer the reader to (Ginzburg 2010, 2012) for the more complex notion.

The basic concept introduced here is contextual \( m(ove) \)-coherence defined in (22a) as applying to \( m_1 \) and \( dgb_0 \) just in case there is a conversational rule \( c_1 \) which maps \( dgb_0 \) to \( dgb_1 \) and such that \( dgb_1 \)'s LatestMove value is \( m_1 \). **Pairwise M(ove)-Coherence**, defined in (22b), applies to a pair of moves \( m_1, m_2 \) if \( m_1 \) is \( M\text{-Coherent} \) relative to some DGB \( dgb_0 \) and there is a sequence of updates leading from LatestMove being \( m_1 \) to LatestMove being \( m_2 \). Finally, **Sequential M(ove)-Coherence**, defined in (22c) applies to a sequence of moves \( m_1, \ldots, m_n \) just in case each successive pair of moves are **Pairwise M\text{-Coherent}**: (22)

a. **M(ove)-Coherence**: Given a set of conversational rules \( \mathcal{C} \) and a dialogue gameboard \( dgb_0 : DGBType \), a move \( m_1 : IllocProp \) is \( m(ove)_{\mathcal{C} dgb_0} \text{-coherent} \) iff there exists
   (i) \( dgb_1 : DGBType, c_1 \in \mathcal{C} \) such that \( c_1(dgb_0) = dgb_1 \) and
   (ii) \( dgb_1, \text{Latest Move} = m_1 \).

b. **Pairwise M(ove)-Coherence**: Given a set of conversational rules \( \mathcal{C} \) two moves \( m_1, m_2 \) are \( m(ove)_{\mathcal{C} dgb_0} \text{-pairwise-coherent} \) iff
   (i) there exists \( dgb_0 : DGBType \) such that \( m_1 \) is \( m(ove)_{\mathcal{C} dgb_0} \text{-coherent} \) and
   (ii) there exist \( dgb_i, c_i, (1 \leq i \leq k - 1, dgb_i : DGBType, c_i \in \mathcal{C}) \) such that \( c_{i+1}(dgb_i) = dgb_{i+1} \) and \( dgb_i, \text{Latest Move} = m_1 \), whereas \( dgb_k, \text{Latest Move} = m_2 \).

c. **Sequential M(ove)-Coherence**: A sequence of moves \( m_1, \ldots, m_n \) is \( m_{\mathcal{C}} \text{-coherent} \) iff for any \( 1 \leq i \) \( m_i, m_{i+1} \) are \( m_{\mathcal{C}} \text{-pairwise-coherent} \).

6.2 Question Accepting Responses

6.2.1 The class \( \text{DP} \)

We start by characterizing the moves in which the responder B accepts question \( q_1 \) as an issue to be resolved. The potential for \( \text{DP} \) responses is explicated on the basis of QSPEC, the conversational rule in (23a). This rule characterizes the contextual background of reactive queries and
Anonymous

assertions. It specifies that if $q$ is QUD–maximal, then subsequent to this either conversational participant may make a move constrained to be $q$–specific, conveying either a proposition $p$ which is a partial answer to $q$ ($p$ is about $q$) or a question $q_1$ on which $q$ depends, as the relation is defined in (23c); one possible definition of dependence is given in (23d); intuitively the idea is that if $q$ is dependent on $q_1$, then once one knows an answer that resolves $q_1$, some information about $q$ (viz. a partial answer) becomes available. This originates in Ginzburg (2012), whence formal characterizations of aboutness and resolvedness can be found.\footnote{We notate the underspecification of the turn holder as TurnUnderspec, an abbreviation for the following specification which gets unified together with the rest of the rule:}

\begin{align*}
\text{(23)} & \quad \text{a. QSPEC} \\
& \quad \quad \text{pre} : [\text{qud} = \langle q, Q \rangle ; \text{poset(} \text{Question} \rangle] \\
& \quad \quad \text{effects} : \text{TurnUnderspec} \land \text{merge} \\
& \quad \quad r : \text{Question} \lor \text{Prop} \\
& \quad \quad R : \text{IllocRel} \\
& \quad \quad \text{LatestMove} = R(\text{spkr},\text{addr},r) : \text{IllocProp} \\
& \quad \quad c1 : \text{Qspecific} (r,q) \\
\text{b. } q\text{-specific utterance: an utterance whose content is either a} \\
& \quad \quad \text{proposition } p \text{ About } q \text{ or a question } q_1 \text{ on which } q \text{ Depends} \\
\text{c. } q_1 \text{ depends on } q_2 \text{ iff any proposition } p \text{ such that } p \text{ resolves } q_2, \\
& \quad \quad \text{also satisfies that for some } r \text{ } p \text{ entails } r \text{ such that } r \text{ is about } q_1. \\
\end{align*}

Other characterizations of dependency are conceivable and could replace the one given here. For now, we illustrate how dependent responses emerge as relevant responses: in (24) A asks $q_1$, responded to by B with a dependent question response $q_2$. A answers $q_2$, which gets accepted by B, leading to an answer to $q_1$: 

\begin{align*}
\begin{aligned}
\text{PrevAud} &= \{ \text{pre.spkr,pre.addr} \} : \text{Set(Ind)} \\
\text{spkr} & : \text{Ind} \\
\text{c1} & : \text{member}\text{(spkr, PrevAud)} \\
\text{addr} & : \text{Ind} \\
\text{c2} & : \text{member}\text{(addr, PrevAud)} \\
\end{aligned}
\land \text{addr} \neq \text{spkr}
\end{align*}

[ 26 ]
### Query Responses

(24)

A(1): Who should we invite for tomorrow?

B(2): Who will agree to come?

A(3): Helen and Jelle and Fran and maybe Sunil.

B(4): (a) I see. (b) So, Jelle I think.

A(5): OK.

(25)

<table>
<thead>
<tr>
<th>Utt.</th>
<th>DGB Update (Conditions)</th>
<th>Rule</th>
</tr>
</thead>
</table>
| initial | MOVES = {}  
QUD = {}  
FACTS = cg1 | |
| 1 | LatestMove := Ask(A,B,q1)  
QUD := (q1) | Ask QUD–incrementation QSPEC |
| 2 | LatestMove := Ask(B,A,q2)  
Influence(q2,q1)  
QUD := (q2,q1) | Assert QUD–incrementation QSPEC |
| 3 | LatestMove := Assert(A,B,p2)  
(Acta(p2,q2))  
QUD := (p2?,q2,q1) | Assert QUD–incrementation Accept |
| 4a | LatestMove := Accept(B,A,p2)  
FACTS := cg1 ∪ {p2}  
QUD := (q1) | Fact update/QUD downdate QSPEC |
| 4b | LatestMove := Accept(B,A,p1)  
(Acta(p1,q1))  
QUD := (p1?,q0) | Assert QUD–incrementation Accept |
| 5 | LatestMove := Accept(A,B,p1)  
FACTS := cg1 ∪ {p1,p2}  
QUD := (q1) | Fact update/QUD downdate |

### 6.2.2 The class QA

This class consists of query responses, where q2 is posed rhetorically, and which provide (indirectly) an answer to q1. In other words, q2 is posed in a context where an answer that resolves q2 can be assumed to be in FACTS – the repository of shared assumptions in the DGB, and, moreover, this answer entails a (resolving) answer to q1. Handling this class does not involve any additional conversational rules; it requires solely two independently needed additions to the setup described hitherto, a mechanism for rhetorical interpretation of ques-
tions and a means of accommodating *indirect* answers:

1. **Rhetorical interpretation of interrogatives**: a rhetorical use arises when an interrogative $q_1$ is used in a context where the DGB contains a fact $f$ that resolves $q_1$. There are, in fact, two possible ways this can be satisfied: either the question has been discussed and a resolving answer provided; alternatively, certain answers are default values for such uses – a negative universal for wh–question (‘Who cares?’, ‘Who knows?’) one of the polar values for polar–questions (‘Do I care?’, ‘Is the Pope catholic?’). One possible treatment is proposed in Ginzburg (2012), §8.3.5: given a context in which a proposition $p$ resolving a question $q$ is presupposed one postulates a root construction that assigns a clause denoting a question $q$ the force of a reassertion of $p$, where $p$, a proposition resolving $q$, is a presupposition satisfied by the context.

2. **Indirect answers**: we need to allow $q$-specificity to include propositions that are *indirectly* about $q$. An explicit account of indirectness would take us too far afield here, but see e.g., (Asher and Lascarides 1998), (Ginzburg 2012, §8.3.3–8.3.5) for discussion relating to questions in dialogue.

We exemplify how this works in (26): A asks the question $p_0$? (‘Is B’s job safe?’). B responds with a reassertion of a question $q_1$ (‘Whose job is safe?’), which in this context reasserts the proposition $p_1$ ‘No one’s job is safe.’, which in particular resolves the question $p_0$?. This explains *inter alia* why there is no need for A to respond to B’s question.

(26)  

**A:** Is your job safe?

**B:** Whose job’s safe?

<table>
<thead>
<tr>
<th>Utterance</th>
<th>DGB Update (Conditions)</th>
<th>Rule</th>
</tr>
</thead>
</table>
| initial   | MOVES = $\langle \rangle$  
\nQUD = $\langle \rangle$  
\nFACTS = cg1 |  |
| 1         | LatestMove := Ask(A,B,p0?)  
\nQUD := $\langle p0? \rangle$ | Ask QUD-incrementation  
\nQSPEC  
\nResolve(p1,q1) |
| 2         | LatestMove := ReAssert(B,A,p1) |  |
Query Responses

We mentioned previously a subclass of QA – query responses where \( q_2 \) presupposes an answer that resolves \( q_1 \). We do not offer a detailed analysis of this subclass here, but they could, for instance, be accommodated by a slight adjustment of q-specificity which licensed responses whose content semantically presupposed a proposition \( p \) about \( q_1 \).

6.2.3 The class FORM

This class consists of question-responses addressing the issue of the way the answer to \( q_1 \) should be given. The class FORM raises interesting issues since it seems to be the sole class whose coherence intrinsically involves reasoning by the responder about the original querier’s unpublicized intentions. One possible explication would be in terms similar to the relation \( Q\text{-ELAB} \) (Asher and Lascarides 2003). Perhaps the simplest way to do this in the current setting would be, following Larsson (2002), to widen the definition of q-specificity so that it is relative to an information state which provides a notion of the agent’s plan, decomposed into a sequence of questions to be resolved:

(27) \( u \) is q-specific relative to information state \( I \): an utterance whose content is either a proposition \( p \) About \( q \) or a question \( q_1 \) on which \( q \) Depends or a question \( q'_1 \) which is a component of \( I \).plan

One could try and collapse DP and FORM. One reason not to do this is precisely that the former does not intrinsically involve reasoning about intentions and so, in principle, its coherence should be easier to compute.

6.3 Adversarial query responses

Adversarial query responses are challenging for most semantic theories of questions for reasons we discuss below. Common to all three classes is a lack of acceptance of \( q_1 \) as an issue to be discussed. In MOTIV–type responses the need/desirability to discuss \( q_1 \) is explicitly posed, in NO ANS–type responses there is an implicature that \( q_1 \) is of lesser importance/urgency than \( q_2 \), whereas for IGNORE–type responses there is an implicature that \( q_1 \) as such will not be addressed.

One commonality between MOTIV and NO ANSW worth noting is that in both cases \( q_1 \) actually needs to be added to QUD at the outset.
Anonymous

One might think that a consequence of a responder’s failure to accept \( q \) for discussion is that \( q \) will only resurface if explicitly reposed. There is evidence, however, that actually \( q \) remains in a conversational participant’s QUD even when not initially adopted, its very posing makes it temporarily DGB available. In (28), where move (2) could involve either a MOTIV query (2a), or a NO ANSW query (2b), the original question has definitely not been reposed and yet B still has the option to address it, which he should be unable to do if it is not added to his gameboard before (28(2)).

(28) A: Who are you meeting next week?
   B(2): (2a) What’s in it for you? / (2b) Who are you meeting next week?
   A: I’m curious.
   B: Aha.
   A: Whatever.
   B: Oh, OK, Jill.

We turn to a discussion of the coherence of each class, starting with MOTIV and NO ANSW, leaving IGNORE for later given a certain additional complexity it embodies.

6.3.1 The class MOTIV

MOTIV utterances are an instance of metadiscursive interaction—interaction about what should or should not be discussed at a given point in a conversation, as exemplified by utterances such as (29):

(29)

   a. I don’t know.
   b. Do we need to talk about this now?
   c. I don’t wish to discuss this now.

A natural way to analyze such utterances is along the lines of a rule akin to QSPEC given in (23): \( q \) being MaxQUD gives (the responder) B the right to follow up with an utterance specific to the issue we could paraphrase informally as \(?WishDiscuss(B,q)\).\footnote{We are formulating this rule asymmetrically with respect to the interlocuters, in contrast to QSPEC, since A posing \( q_1 \) means that A keeping the turn and uttering (29b,c) would be somewhat incoherent; the status of (29a) as a fol-
Query Responses

\[ 'QUD = \{ \text{max} = \{ ? \text{WishDiscuss}(B, q1), q1 \}, Q \} ' \]

indicates that both ?WishDiscuss(B, q1) and q1 are maximal in QUD, unordered with respect to each other. The motivation for this latter is the need to integrate q1 in context, as per (28) above.

(30) MetaDiscussing q1

\[
\begin{align*}
\text{pre:} & \quad \text{QUD} = \langle q1, Q \rangle; \text{poset(Question)} \\
& \quad \text{spkr} = \text{pre.addr}: \text{Ind} \\
& \quad \text{addr} = \text{pre.spkr}: \text{Ind} \\
& \quad \text{r} : \text{Question} \lor \text{Prop} \\
& \quad \text{R: IlocRel} \\
\text{effects:} & \quad \text{Moves} = \langle \text{R(spkr, addr, r)} \rangle \uplus \text{pre.Moves}: \text{list(IlocProp)} \\
& \quad c1 : \text{Qspecific}(\text{R(spkr, addr, r)}, ? \text{WishDiscuss(spkr, pre.maxqud)}) \\
& \quad \text{qud} = \langle \text{max} = \{ ? \text{WishDiscuss(spkr, q1), q1 \}, Q \rangle: \text{poset(Question)} \\
\end{align*}
\]

In case information is accepted indicating negative resolution of ?WishDiscuss(B, q1), then q1 may be down-dated from QUD. This involves a minor modification of the Fact Update/QUD Downdate rule (see (21) above.).

We exemplify (30) in two ways. First, with a variant of (28), where B’s rejection of a question leads to the down-dating of q1; then, with a very similar analysis of a MOTIV query response.

low up to q1 is somewhat different: in the commonest case, where a query is posed because the querier does not know the answer, (29a) is redundant and somewhat infelicitous. In cases where q1 is uttered in the spirit of ‘Here is an interesting issue to discuss’, (i) seems acceptable:

(i) I don’t know.

Whether this should be taken to imply that (29a) and (29b,c) should be licensed by distinct mechanisms is an issue we will not try to resolve here.

15 All that this involves is a modification of the function NonResolve which fixes the value of QUD after the fact update: in its new definition it maps a poset of questions poset(q) and a set of propositions P to a poset of questions poset’(q) which is identical to poset(q) modulo those questions in poset(q) resolved by members of P, as well as those questions q for whom ?WishDiscuss(q) is negatively resolved.
We start with (31): A poses q1; B does not wish to discuss q1, hence she accommodates ?WishDiscuss(B, q1) into QUD and offers an utterance concerning this issue. A accepts B’s assertion, so using the new version of fact-update/qud-downdate q1 can be downdated and either conversationalist could introduce a new topic, as in (31):

(31)

A(1): Who are you meeting next week?
B(2): No comment.
A(3): I see.
A/B(4): What are you doing tomorrow?

<table>
<thead>
<tr>
<th>Utt.</th>
<th>DGB Update (Conditions)</th>
<th>Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>initial</td>
<td>MOVES = {}</td>
<td>Ask QUD–incrementation Discussing u?</td>
</tr>
<tr>
<td></td>
<td>QUD = {}</td>
<td>Assertion QUD–incrementation Accept</td>
</tr>
<tr>
<td></td>
<td>FACTS = cg1</td>
<td>Fact update/QUD downdate</td>
</tr>
<tr>
<td>1</td>
<td>LatestMove := Ask(A,B,q1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>QUD := {q1}</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>LatestMove := { Assert(B,A,p1) }</td>
<td></td>
</tr>
<tr>
<td></td>
<td>QUD := { p1?, &gt; ?WishDiscuss(q1), q1 }</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>LatestMove := { Assert(B,A,p1) }</td>
<td></td>
</tr>
<tr>
<td></td>
<td>QUD := {}</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FACTS := cg1 ∪ {p1}</td>
<td></td>
</tr>
</tbody>
</table>

We suggest that a dialogue like (32) works in a similar way: A’s answer to B’s question (32(2)) can satisfy B, which will lead to the question ?WishDiscuss(B, q1) being positively resolved, enabling B to

\[
\begin{align*}
\text{pre :} & \quad p : \text{Prop} \\
\text{LatestMove = Accept(spkr,addr,p)} \\
\text{qud = \{p?,pre.qud\}; poset(}\text{Question}) \\
\text{effects :} & \quad \text{facts = pre.facts }\cup \{p\}; \text{Set(Prop)} \\
& \quad \text{qud = NonResolve(pre.qud,facts).Q'} \\
& \quad : \text{Poset(}\text{Question})
\end{align*}
\]

\[
\text{NonResolve} = _{\text{def}}
\]

\[
\begin{align*}
\text{B : Ind} \\
\text{r : (} \\
& \quad \text{F : set(Prop)} \\
& \quad \text{Q : poset(}\text{Question})
\end{align*}
\]

\[
\begin{align*}
\text{Q'} : & \quad \text{poset(InfoStruct)} \\
\text{c1 :} & \quad Q' \subseteq r.Q \\
\text{c2 :} & \quad \forall q_0 \in Q' \lnot \exists f \in F \\
\text{resolve}(f,q_0,q) \\
\text{\lor resolve}(f,?WishDiscuss(r,B,q_0,q))
\end{align*}
\]
downdate it from her QUD and address the question (32(1)). B not being satisfied with A’s answer is entirely similar to (31) mutatis mutandis:

(32)

A(1): Who are you meeting next week?  
B(2): Why?  
A(3): I need to know which refreshments to buy.

<table>
<thead>
<tr>
<th>Utt.</th>
<th>DGB Update (Conditions)</th>
<th>Rule</th>
</tr>
</thead>
</table>
| initial | MOVES = {}  
QUD = {}  
FACTS = cg1 | |
| 1 | LatestMove := Ask(A,B,q1)  
QUD := (q1) | |
| 2 | LatestMove := (Ask(B,A,q2))  
QUD := (q2 > ?WishDiscuss(B,q1), q1) | Ask QUD–incrementation  
Discussing u? |
| 3 | LatestMove := Assert(A,B,p1)  
(QUD := (p1,q2))  
FACTS := cg1 ∪ {p1}  
QUD := (q0) | Assert QUD–incrementation  
QSPEC |
| 4a | LatestMove := Accept(B,A,p1) | Fact update/QUD downdate |

6.3.2 The Class NO ANSW

NO ANSW–queries can be analyzed in a fairly similar fashion. The main challenge such queries pose is to consider the coherence relation between q1 and q2. In contrast to IGNORE, where it seems like there is little that need connect the two questions, save for some reference to the situation associated with q1, for NO ANSW the questions seem to need a fairly tight link. A tentative characterization of this link is the following: q1 and q2 are not dependent on each other, but instead there exists a third question, q3, such that q3 depends on q1 and q3 depends on q2. The rationale behind this characterization is that by responding with q2 B provides (a) an issue that is not unconnected with q1, but (b) it is informationally not subservient to q1. Hence, given that q3 is (or can be accommodated to be) the general topic under discussion, q2 has an arguable case to being at least as discussion worthy as q1.16

16 An anonymous reviewer for this journal point out the following exchange
Anonymous

(33) a. \( q_1 = \) what do you (B) like? \( q_2 = \) what do you (A) like? \( q_3 = \) Who likes what?

b. \( q_1 = \) Why should we buy that scanner? \( q_2 = \) Why should we not buy that scanner? ; \( q_3 = \) Should we buy that scanner?

Based on this we define the relation of being *unifiably coherent*:

(34) Given \( q_1, q_2 : \) Question \( q_1 \) and \( q_2 \) are unifiably coherent iff

1. Neither \( q_1, \) nor \( q_2 \) depend on the other: \( \neg \text{Depend}(q_2,q_1) \)
   \& \( \neg \text{Depend}(q_1,q_2) \)

2. There exists \( q_3 : \) Question which depends on both \( q_1 \) and \( q_2 : \text{Depend}(q_3,q_1) \) \& \( \text{Depend}(q_3,q_2) \)

The potential for making such queries can be captured by the conversational rule in (35). Given that \( q_1 \) is MaxQUD, the responder may respond with \( q_2 \), assuming it to be unifiably coherent with \( q_1 \). The immediate effect of this is to update QUD with the issue ?WishDiscuss(B,q1).

(35) Challenging \( q_1 \)

as problematic for our taxonomy, suggesting that it is ‘fully coherent given the sequel but the pair does not seem to fit any of the schemes’:

(i) A: Are you coming on Friday?
   B: Did you ever consider quarks?
   A. No.
   B. Well you should for your work and Friday there will be a lecture that is just right for you. I may be there myself.

Actually, we would suggest that this example would be classified as a NO ANS by the annotation criteria we offer (since B views A’s question as less important to consider than his and one could eliminate B’s answer at the end without affecting coherence.). Nonetheless, it calls into question our formalized definition for NO ANS in that it is not clear that the \( q_2 \) and \( q_1 \) are *unifiably coherent*. One might use this (constructed) example to argue for weakening the unifiable coherence clause. At the same time, it seems likely that B’s response would initially be viewed as incoherent by A and this should be reflected by e.g., response time, frowning etc.
Query Responses

\[
\text{pre} : \begin{bmatrix}
\text{qud} &= \langle q1, Q \rangle ; \text{poset}(\text{Question}) \\
\text{spkr} &= \text{pre}.\text{addr} : \text{Ind} \\
\text{addr} &= \text{pre}.\text{spkr} : \text{Ind} \\
q2 &= \text{Question} \\
\text{Moves} &= \langle \text{Ask}(\text{spkr}, \text{addr}, q2) \rangle \oplus \text{pre}.\text{Moves} : \text{list}(\text{IllocProp}) \\
\text{c1} &= \text{Unifiably coherent}(q1, q2) \\
\text{qud} &= \langle \text{max} = \{ ?\text{WishDiscuss}(B, q1) \}, Q \rangle ; \text{poset}(\text{Question}) \end{bmatrix}
\]

In (36)\textsuperscript{17} A asks q1, B responds with q2 that unifies coherently with q1 via, for example, the issue q3 = ‘Should they wait?’. A responds to q2 and then B’s second utterance can be understood as addressing q2. If A accepts (4), q2 can be downdated and, consequently q1 and ?WishDiscuss(B,q1) as well – q1 has been resolved as well, and hence ?WishDiscuss(B,q1) could be taken to be resolved as well.\textsuperscript{18}

(36) 
A(1): Why won’t they wait?
B(2): Why should they?
A(3): I waited.
B(4): They have lives of their own.

6.4 DGB divergence: Ignore and Clarification requests

Both clarification requests (CRs) and IGNORE type responses involve reasoning that requires reference to two DGBs. CRs arise due to a mis-

\textsuperscript{17}Inspired by the BNC example:

Eddie: But it’s something, something in you, you have to rush don’t they? Why won’t they wait? Unknown: Why should they? Eddie: Why should they? Unknown: No, why should they? Eddie: I have Unknown: Take the rest of it Unknown: (unclear) Eddie: pleasure spending Unknown: (unclear) Unknown: No why, they’ve got lives of their own Eddie: Well Sally: let them live it, don’t want saving for the children, no, they don’t want nothing Eddie: Well Unknown: They’ve had far more than what we’ve ever had (BNC, KCF 3584–3596)

\textsuperscript{18}A general principle linking the downdating of ?WishDiscuss(B,q0) once q0 has been downdated should be introduced, though we will not do so here.
match that occurs between what the speaker assumes his interlocutor’s linguistic/contextual knowledge is and what it actually is; consequently, in the immediate aftermath of such an utterance – before the mismatch becomes manifest, the speaker updates his IS with the query he posed and the addressee updates hers with the clarification question she calculated.

Similarly, in the case of IGNOREs the initial speaker updates his IS with the query he posed and, ignoring this, the addressee updates hers with the situationally relevant question she has decided to pose.

6.4.1 Clarification Requests

We start by discussing CRs since they have been studied in great detail, see (Ginzburg and Cooper 2004; Schlangen 2004; Purver 2006; Ginzburg et al. 2014); we will summarize briefly the most detailed account we are aware of, that provided in Ginzburg (2012). This will provide tools enabling us to analyze IGNORE-type responses.

Integrating clarification interaction into the DGB involves two modifications to the representations we have been using so far. One minor modification, drawing on an early insight of Conversation Analysis (Schegloff 2007), is that repair can involve ‘putting aside’ an utterance for a while, a while during which the utterance is repaired. That in itself can be effected without further ado by adding further structure to the DGB, specifically the field we call Pending. ‘Putting the utterance aside’ raises the issue of what is it that we are ‘putting aside’. In other words, how do we represent the utterance? The requisite information needs to be such that it enables the original speaker to interpret and recognize the coherence of the range of possible clarification queries that the original addressee might make. Ginzburg (2012) offers detailed arguments on this issue, including considerations of the phonological/syntactic parallelism exhibited between CRs and their antecedents and the existence of CRs whose function is to request repetition of (parts of) an utterance. Taken together with the obvious need for Pending to include values for the contextual parameters specified by the utterance type, Ginzburg concludes that the type of Pending combines tokens of the utterance, its parts, and of the constituents of the content with the utterance type associated with the utterance. An entity that fits this specification is the locutionary proposition defined by the utterance. A locutionary proposition is a proposition whose sit-
Query Responses

...ational component is an utterance situation, typed as in (37a) and will have the form in (37b):

(37) a. \( \text{LocProp} = \text{def} \left[ \begin{array}{c} \text{sit} : \text{Sign} \\ \text{sit-type} : \text{RecType} \end{array} \right] \)

b. \( \left[ \begin{array}{c} \text{sit} = u \\ \text{sit-type} = T_u \end{array} \right] \)

Here \( T_u \) is a grammatical type for classifying \( u \) that emerges during the process of parsing \( u \). It can be identified with a sign in the sense of Head Driven Phrase Structure Grammar (HPSG) Pollard and Sag (1994).

How then can one characterize the relevance of CRs in this set up? Corpus studies of CRs Purver et al. (2001); Rodriguez and Schlangen (2004); Rieser and Moore (2005) indicate that the subject matter of CRs is, in practice, restricted to three classes: CRs requesting repetition, CRs requesting confirmation, and CRs in which query the intended content of a sub-utterance. This means that the potential for CRs can be modelled in terms of a small number of schemas (Clarification Context Update Rules (CCURs)) of the form: “if \( u \) is the maximal element of Pending (\( \text{MaxPending} \)) and \( u_0 \) is a constituent of \( u \), add the clarification question \( \text{CQ}^i(u_0) \) into \( \text{QUD} \).”, where \( \text{CQ}^i(u_0) \) is one of the three types of clarification question (repetition, confirmation, intended content) specified with respect to \( u_0 \).

(38) is a simplified formulation of one CCUR, Parameter identification, which allows \( B \) to raise the issue about \( A \)’s sub-utterance \( u_0 \): what did \( A \) mean by \( u_0 \)?:

(38) Parameter identification:

\[
\left[ \begin{array}{c} \text{Spkr} : \text{Ind} \\ \text{pre} : \begin{array}{c} \text{MaxPending} : \text{LocProp} \\ u_0 \in \text{MaxPending.sit.constits} \end{array} \\ \text{effects} : \begin{array}{c} \text{MaxQUD} = \lambda x \text{Mean}(A,u_0,x) : \text{Question} \\ \text{LatestMove} : \text{LocProp} \\ c1 : \text{CoPropositional(LatestMove.cont,MaxQUD)} \end{array} \end{array} \right]
\]

Here CoPropositionality for two questions means that, modulo their domain, the questions involve similar answers: for instance
'Whether Bo left', 'Who left', and 'Which student left' (assuming Bo is a student.) are all co-propositional. More generally, the definition is given in (39):

\[(39) \text{ Two utterances } u_0 \text{ and } u_1 \text{ are co-propositional iff the questions } q_0 \text{ and } q_1 \text{ they contribute to QUD are co-propositional.} \]

a. qud-contrib(m0.cont) is m0.cont if m0.cont : Question
b. qud-contrib(m0.cont) is ?m0.cont if m0.cont : Prop

19Recall from the assertion protocol that asserting p introduces p? into QUD.
Query Responses

Speech event: \( u_0 \)
\( u_1, u_2, u_3 \)

\[ T_{u_0} = \begin{cases} \text{phon: } \text{did bo leave} \\
\text{cat: } S[+\text{root}] \\
\text{constits:} \\
\text{u1: aux} \\
u2: \text{NP} \\
u3: \text{VP} \\
\text{dgb-params:} \\
\text{spkr: Ind} \\
b: \text{Ind} \\
\text{cont = Ask(spkr,?Leave(b)): IllocProp} \end{cases} \]

Speaker’s witnesses for dgb-params:
\[ w_A = \begin{bmatrix} \text{spkr = A} \\
b = b_0 \end{bmatrix} \]

Addressee’s witnesses for dgb-params:
\[ w_B = \begin{bmatrix} \text{spkr = A} \end{bmatrix} \]

Speaker’s DGB update:
\[ \text{LatestMove} = \begin{bmatrix} \text{sit = u0} \cup w_A \\
\text{sit-type = } T_{u_0} \end{bmatrix} \]
\[ \text{MaxQUD} = \text{?Leave(b0)} \]

Addressee’s DGB update:
\[ \text{MaxPending} = \begin{bmatrix} \text{sit = u0} \cup w_B \\
\text{sit-type = } T_{u_0} \end{bmatrix} \]
\[ \text{MaxQUD} = \lambda x.\text{Mean}(A,u2,x) \]

\( \lambda x.\text{Mean}(A,u2,x) \), and the locutionary proposition associated with \( u_0 \) which B has constructed remains in Pending.

6.4.2 The class IGNORE

The final class we consider are IGNORE-type responses. Such responses implicate that \( q_1 \) will not be addressed, somewhat analogously to the classic Gricean floutings of relevance (A: Bob is an embarrassment B: It’s very hot in here). Nonetheless, the effect such responses have is different from Gricean floutings since these responses are situationally relevant, which appears to minimize significantly the potential impoliteness associated with ignoring \( q_1 \). We think the difference between these two cases should be experimentally testable (e.g. response times
for Gricean floutings should be significantly larger than for IGNOREs.)

The conversational rule we propose both allows the potential for \(q_2\) and captures the implicature concerning \(q_1\)’s being ignored. The formulation of such a rule presupposes a notion of relevance between the content of an utterance \((q_2)\) and the current context. We assume here the notion of relevance we mentioned in section 6.1 and define irrelevance as failure of relevance: for an utterance \(u\) being Irrelevant to an information state \(I\) amounts to: there is no way to successfully update \(I\) with \(u\). At the same time we assume that \(q_2\) being situationally relevant means that the open proposition component of \(q_2\) is of the form \(p_2(\ldots a \ldots)\) with \(a\) being in the situation which concerns \(q_1\).

This involves positing a conversational rule along the lines of (41) – given that (the content of) MaxPending – the most recent utterance, as yet ungrounded, hence maximal in Pending – is irrelevant to the DGB but situationally relevant to \(q_2\), one can make MaxPending into LatestMove while updating Facts with the fact that the speaker of MaxPending does not wish to discuss MaxQUD:

\[
(41) \quad \text{Ignoring questions}
\]

\[
\begin{align*}
&\text{pre : } q_2 = (G1) \left[ \begin{array}{l}
sit = s \\
sit-type = \left[ c : p_2(a) \right]
\end{array} \right] : \text{Question} \\
&\text{dgb = MaxQUD = q1 : Question} \\
&\text{maxpending}^{\text{content}} = q_2 : \text{Question} \\
&\text{c: IrRelevant(maxpending}^{\text{content}},\text{dgb}) \\
&\text{effects : } \\
&\quad \text{LatestMove = pre.maxpending : LocProp} \\
&\quad \text{Facts = pre.Facts} \cup \\
&\quad \left\{ \neg \text{WishDiscuss(pre.spkr,pre.maxqud)} \right\}.
\end{align*}
\]

Note that this does not make the unwillingness to discuss be the content of the offending utterance; it is merely an inference. Still this in-
Query Responses

ference will allow MaxQUD to be downdated, via fact update/question downdate, as was discussed with respect to MOTIV moves and the rule MetaDiscussing q1. We exemplify this with respect to (42).

(42)  A: Is there just one car there?
      B: Why is there no parking there?

As we noted earlier, given the contextual mismatch involved, in order to describe such dialogues one needs to consider the dialogue on the basis of two distinct DGBs. One possible evolution of A’s DGB is this: A utters q1, which becomes MaxQUD; he then encounters B’s response; A applies the rule Ignoring questions, which leads to q1’s downdate, q2 becomes MaxQUD.

(43)

<table>
<thead>
<tr>
<th>Utt.</th>
<th>DGB Update (Conditions)</th>
<th>Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>initial</td>
<td>MOVES = {}</td>
<td>Ask QUD–incrementation</td>
</tr>
<tr>
<td></td>
<td>QUD = {}</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FACTS = cg1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>LatestMove := Ask(A,B,q1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>QUD := (q1)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>LatestMove := ( Ask(B,A,q2) )</td>
<td>Ignoring questions</td>
</tr>
<tr>
<td></td>
<td>FACTS := FACTS \ ¬ WishDiscuss(B,q1)</td>
<td>FACTS update/QUD downdate</td>
</tr>
<tr>
<td></td>
<td>QUD := ( )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>QUD := ( q2 )</td>
<td>Ask QUD–incrementation</td>
</tr>
</tbody>
</table>

To the extent B wishes to ignore A’s utterance, we do not need any additional machinery, save for a general principle needed for a variety of other not necessarily linguistic events (e.g., in case one of the participants A burps, spits, or farts) – pretense that an event was not perceived. Assuming this, a possible evolution of B’s DGB is as in (44): B pretends that A’s utterance u1 did not take place, she utters q2, which relates to the situation A and B are jointly perceiving; q2 becomes MaxQUD:

(44)

<table>
<thead>
<tr>
<th>Utt.</th>
<th>DGB Update (Conditions)</th>
<th>Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>initial</td>
<td>MOVES = {}</td>
<td>Ask QUD–incrementation</td>
</tr>
<tr>
<td></td>
<td>QUD = {}</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FACTS = cg1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>LatestMove := ( Ask(B,A,q2) )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>QUD := ( q2 )</td>
<td>Ask QUD–incrementation</td>
</tr>
</tbody>
</table>
Table 9:

Increasing complexity of reasoning needed to accommodate query responses

<table>
<thead>
<tr>
<th>Query response type</th>
<th>Information state complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>DP, IND</td>
<td>QUD, dependence relation</td>
</tr>
<tr>
<td>FORM</td>
<td>QUD, parametrised dep. relation</td>
</tr>
<tr>
<td>MOTIV</td>
<td>QUD as poset</td>
</tr>
<tr>
<td>NO ANSW</td>
<td>unif-coh relation, QUD as poset</td>
</tr>
<tr>
<td>CR, IGNORE</td>
<td>QUD, Pending, DGB split</td>
</tr>
<tr>
<td></td>
<td>non-semantic coherence</td>
</tr>
</tbody>
</table>

6.5 Summary

In this section we have shown how to characterize the relevance of the range of possible query responses $q_2$ to an initial query $q_1$ using DGB–based dynamics. The relevance of dependent questions is characterized in terms of QUD and the dependence relation, a relation defined on pairs of questions; QA and IND use the same contextual set up (plus mechanisms independently needed for accommodating rhetorical uses of interrogatives and indirect/presupposed answers); accommodating FORM involves reasoning similar to DP, but requires making reference to the issues constituting an interlocutor’s plan; MOTIV and NO ANSW involve postulating additional conversational rules that make reference to the issue of whether $q_2$’s speaker wishes to discuss $q_1$, leaving this and $q_1$ as issues simultaneously under discussion, hence this makes crucial use of QUD being a partially ordered set; NO ANSW also involves computing an additional coherence relation ‘unifiable coherence’ that needs to relate $q_1$ and $q_2$; clarification requests and IGNORE both require making reference to distinct DGBs for the two participants, make use of an additional buffer for ungrounded utterances, Pending, and involve coherence relations defined at the level of utterances, not merely $q_1$ and $q_2$.

7 CONCLUSIONS

The article provides the first comprehensive, empirically-based study of query responses to queries. One interesting finding here is the existence of a number of classes of adversarial responses, that involve the rejection/ignoring of the original query. Indeed, in such cases the original query is rarely responded to in subsequent interaction. We designed our taxonomy based on data from the BNC since it is a gen-
eral corpus with a variety of domains and genres, but have also shown that our classification works well in a number of more specific genres and domains, which display quite different distributions of query responses. We have proposed qualitative, domain specific explanations for the variation displayed by these distributions.

On the theoretical side, we have provided a comprehensive, information state dynamics-based characterisation of the relevance of the entire range of query response types. Our account uses the KoS framework for representing dialogue information states and its component of information arising from publicized interaction, the dialogue game board (DGB). This enables us to offer a pre–theoretical sketch of the expressive complexity of the different classes of query response types, ranging from dependent questions and IND, which, assuming a semantic relation of question dependence, can be accommodated in a fairly vanilla query/response set up, through MOTIV and NO ANSW, which intrinsically require the dynamic question repository QUD to be a partially–ordered set, through IGNORE and clarification requests, which require distinct information DGBs for the two participants, make use of an additional buffer for ungrounded utterances, Pending, and involve coherence relations defined at the level of utterances, not merely $q_1$ and $q_2$.

What are the more general theoretical implications of this characterization? We believe that they offer concrete desiderata for semantic theories, more specifically for the nature of conversational context. We offer brief remarks relative to frameworks that have put forward theories of question responses, as discussed in Section 1.

Some account of question dependence can be developed by any theory of questions which supplies notions of exhaustive and partial answerhood, though it is clear that providing a more detailed empirical and theoretical account of this notion than we have given here is an important task.

Relations like MOTIV and NO ANSW require structure within context since they need maintaining several questions simultaneously accessible to the participants. This constitutes a challenge for views of contexts in terms of stacks. Such a view has been made prominent in Roberts (1996) view of QUD. It can also be found, for instance in the discourse model of Farkas and Roelofsen (2011), where a discourse context $X$ is identified as a pair $<M,T>$, where $M$ is a Kripke model
and \( T \) is a stack of sentences, those sentences that have been uttered so far.

The problem for stacks can be defused by adopting distinct structure, for instance a partial order. Nonetheless, for these accounts and most other existing views of context, context is an entity shared by the conversational participants. This was also the case for the view of discourse structure in earlier work in SDRT (e.g., Asher and Lascarides 1998, 2003). In more recent work (e.g., Lascarides and Asher 2009) SDRT adopts a view advocated in KoS and also in the framework of PTT (Poesio and Traum 1998) that associates a distinct contextual entity with each conversational participant.

Given this, it seems like a framework like SDRT has the potential for developing an account of question relations like IGNORE and CR which require context to ‘diverge’ across participants. There is one important caveat – we have argued that the notion of relevance that underpins both these question relations must make reference to non-semantic information. By contrast, in SDRT the semantics/pragmatics interface has no access to linguistic form, but only to a partial description of the content that is derived from linguistic form. This has been argued to be necessary to ensure the decidability of SDRT’s glue logic (see e.g., Asher and Lascarides 2003, p. 77).

In closing, we note two questions raised by our account. The coherence follows in some cases on the basis of quite general conversational rules (e.g., QSPEC and MetaDiscussing q1) and in other cases on the basis of rather specific – though domain independent – rules (e.g. Ignoring questions). An obvious theoretical issue is whether one can attain similar coverage on the basis of more “general” rules allied with some other very general pragmatic principles. A converse question is whether investigation of specific genres will lead to the need for genre–specific conversational rules for certain classes of question relations.
In this study we are interested in the phenomenon of answering a query with a query.

In the attached file you will find 100 samples of question–question–response pairs (with added context). In what follows, we will use \( q_1 \) for the initial question posed and \( q_2 \) for a question given as a response to \( q_1 \).

In the annotation file we use the following conventions:

- an utterance which is \( q_1 \) is written in *italics*;
- an utterance which is \( q_2 \) is written with **bold font**;
- a context is marked with grey coloured font.

The task is to ascribe each \( q_2 \) into one of the following categories:

<table>
<thead>
<tr>
<th>Tag</th>
<th>question-answer type</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>clarification requests</td>
</tr>
<tr>
<td>DP</td>
<td>dependent questions</td>
</tr>
<tr>
<td>FORM</td>
<td>questions considering the way of answering ( q_1 )</td>
</tr>
<tr>
<td>MOTIV</td>
<td>questions about the underlying motivations behind asking ( q_1 )</td>
</tr>
<tr>
<td>NO ANSW</td>
<td>questions aimed at avoiding answering ( q_1 )</td>
</tr>
<tr>
<td>QA</td>
<td>questions providing an answer to ( q_1 ) or questions with a presupposed answer</td>
</tr>
<tr>
<td>IGNORE</td>
<td>questions ignoring ( q_1 )</td>
</tr>
</tbody>
</table>

During the classification process, please use the following questions:

1. Is \( q_2 \) a query about something not completely understood in \( q_1 \)?
   *If yes, then it is CR*
   
   **example:**
   
   A: Why are you in?
   B: What?

2. Is it the case that the answer to \( q_1 \) depends on the answer to \( q_2 \)?
   *If yes, then it is DP*
   
   **example:**
   
   A: Do you want me to `<pause>` push it round?
B: Is it really disturbing you? [FM1, 679–680]  
(cf. Whether I want you to push it depends on whether it really disturbs you.)

3. Does \( q_2 \) address the motivation underlying asking \( q_1 \)?  
   *If yes, then it is MOTIV*
   
   example:
   
   A: What’s the matter?  
   B: Why? [HDM, 470–471]

4. Is it the case that \( q_2 \) enables the speaker to avoid answering \( q_1 \) while attempting to force the other speaker to answer \( q_2 \) first?  
   *If yes, then it is NO ANSW*
   
   example:
   
   A: Why is it recording me?  
   B: Well why not? [KSS, 43–44]

5. Is it the case that the way the answer to \( q_1 \) will be given depends on the answer to \( q_2 \)?  
   *If yes, then it is FORM*
   
   example:
   
   A: Okay then, Hannah, what, what happened in your group?  
   B: Right, do you want me to go through every point? [K75, 220–221]

6. Is it the case that \( q_2 \) is rhetorical and in this sense it does not need to be answered and provides (indirectly) an answer to \( q_1 \)?  
   *If yes, then it is IND*
   
   example:
   
   A: Are you Gemini?  
   B: Well if I’m two days away from your, what do you think? [KPA, 3603–3604]  
   (cf. of course I am Gemini)

7. Is it the case that by asking \( q_2 \) an agent already presupposes the answer to \( q_1 \).  
   *If yes, then it is IND*
   
   example:
   
   A: have you tasted this?  
   B: are they nice? [KPY, 653–654]  
   (cf. no, I have not tasted this)

8. Does \( q_2 \) ignore \( q_1 \) but at the same time is related to the situation described by \( q_1 \)?  
   *If yes, then it is IGNORE*
   
   example:
   
   A: Just one car is it there?  
   B: Why is there no parking there? <unclear> [KP1, 7882–7883]  
   *(A asks a question, which is ignored by B. It is not that B does not want to*
Query Responses

answer A’s question and that’s why he/she asks q2. Rather, B ignores q1 and asks a question related to the situation.

9. If none of the above fits, please mark q2 as OTHER.

REFERENCES


Thierry Coquand, Randy Pollack, and Makoto Takeyama (2003), A Logical Framework with Dependent Types, *Fundamenta Informaticae*, 20:1–21.

Donka Farkas and Floris Roelofsen (2011), Polarity particles in an inquisitive discourse model, Manuscript, University of California at Santa Cruz and ILLC, University of Amsterdam.


Anonymous


This work is licensed under the Creative Commons Attribution 3.0 Unported License. http://creativecommons.org/licenses/by/3.0/