Institute for Research in Cognitive Science

Yes/No Questions and Answers in the Map Task Corpus

Deborah Rossen-Knill
Beverly Spejewski
Beth Ann Hockey
Stephen Isard
Matthew Stone

University of Pennsylvania
3401 Walnut Street, Suite 400A
Philadelphia, PA 19104-6228

July 1997

Site of the NSF Science and Technology Center for Research in Cognitive Science

IRCS Report 97--11

Yes/No Questions and Answers in the Map Task Corpus

Deborah Rossen-Knill[†] **Beverly Spejewski**[‡] **Beth Ann Hockey**[‡] **Stephen Isard**[§] **Matthew Stone**[‡]

[†]**Phila. College of Textiles & Science** [‡] **University of Pennsylvania** [§]**University of Edinburgh**

Philadelphia PA 19144-5497 **Philadelphia PA 19104-6228** **Edinburgh EH1 1HN**

{drossen, spejewsk, beth, matthew}@linc.cis.upenn.edu stepheni@cstr.ed.ac.uk

Address correspondence to:

Deborah Rossen-Knill

The Learning Center

Philadelphia College of Textiles & Science

Henry Ave and School House Ln

Philadelphia, PA 19144

Acknowledgements. We would like to thank the Institute for Research in Cognitive Science at the University of Pennsylvania for supporting D. Rossen-Knill, M. Stone, and S. Isard during this project.

Abstract

We analyze question-answer pairs in a variety of ways, for three different kinds of yes/no questions. We find that the classification of yes/no questions described in (Carletta et al., 1995) for the Edinburgh map task corpus correlates well with whether a response will be a bare *yes* or *no*, a *yes* or *no* plus additional speech, or just speech without an overt *yes* or *no*. Correlation with responses described as "direct" or "indirect" is less good. We also find that the strength of a question's expectation for a YES response correlates with the move type, the form of the response, and lexical *yes* choices; and that the move type correlates with the form of the question and with turn-taking schema.

Introduction

While it may at first seem fairly straightforward that the answer to a yes/no question should be *yes* or *no*, or some variant of these terms, it has often been pointed out that acceptable responses may not contain such a term (e.g., (Stenström, 1984; Green and Carberry, 1992; Green and Carberry, 1994; Stubbs, 1983; Ginzburg, 1995). Green and Carberry's (1994) example illustrates a potential sequence:

(1) a Q: I need a ride to the mall.

b Are you going shopping tonight?

c R: [no]

d My car's not running.

e The timing belt is broken.

The square brackets indicate that R does not utter *no*, although the response communicates a negative answer to the question. This happens fairly frequently: in our data, 27% of the responses to yes/no questions are similar to (1) in that they do not contain any overt *yes* or *no* term. In addition, we find that when overt *yes* or *no* terms do appear in responses, they are accompanied by additional speech 41% of the time.[1] If we are to describe, interpret or model dialogue involving yes/no questions and answers, we need to be able to distinguish the different kinds of responses to yes/no questions. Example (1) raises several relevant questions. When is it appropriate to produce answers that don't contain an overt *yes* or *no* form? When is it sufficient to produce just an overt *yes* or *no*, and When is it appropriate to produce both an overt polar term and other relevant speech?

In this paper, we address these questions for three different kinds of yes/no questions, based on analysis of natural language data from the Edinburgh map task corpus (Anderson et al., 1991). We analyze the whole of a response in relation to its question to determine if it communicates a YES, NO or NON-COMMITTAL meaning. We then separate the response into one of three categories of form: just affirmative or negative terms, which we refer to as *bare yes/no*; polar terms plus additional speech, which we refer to as *yes/no+stuff*; and speech without overt polar terms, referred to as *stuff*.

The data studied here leads us to two major observations. First, the form of answer correlates with the function of the question in dialogue as given by the MOVE TYPE of the question (Carletta et al., 1995). Second, the form of the answer and the *yes* term used to answer a question correlate with the overall likelihood of a given question type producing a YES—a quality we refer to as the YES-expectation. In addition, we find that there are strong patterns in the form the question takes for each move type, and in the turn-taking sequence

Align:
"Right?"
"Okay? Can you do that?"

Check:
"So you don't have a graveyard?"
"Walk below the springs?"

Query-yes/no:
"Do you have a fenced meadow?"
"Are we going to go below the picket fence?"

Figure 1: Examples of Moves

in which the move is embedded. The explanations we offer for these findings come mainly from the domain of pragmatics, including such notions as lexical variation, conversational roles, and turn-taking.

We begin with a description of the data used for this study. We then look at patterns in the answer and how they relate to the move type and YES-expectation, patterns in the form of the question, and finally how the move type relates to its surrounding dialogue.

Description of Dialogues Studied

We analyze eight dialogues averaging seven minutes each of Glaswegian English from the Edinburgh map task corpus ((Anderson et al., 1991), files eaq1c1-eaq1c8). For these dialogues, one person has a route marked on a map; the two people converse in order to enable the other person to mark the same route on her own map. Each resulting dialogue has been previously coded for dialogue structure using the system described by Carletta et al. (1995). Their coding scheme has three levels of structure: the conversational game, conversational moves, and specific move types. The conversational game consists of an initiating utterance and all utterances following it until the purpose of the game is fulfilled. An initiating utterance and its responses are each classified as particular types of CONVERSATIONAL MOVES based on their purposes. The initiating utterances of interest for our study are yes/no questions. In this study we restrict our attention to questions without negation; negated questions introduce a further level of complexity, which could be better investigated once we have an understanding of the more straightforward cases.

Carletta et al. (1995) classify yes/no questions into three different move types: ALIGN, CHECK and QUERY-YES/NO. An align "checks the attention or agreement of the partner, or his readiness for the next move"; a check "requests the partner to confirm information that the checker has some reason to believe, but is not entirely sure about," and a query-yes/no is any other question that takes a *yes* or *no* answer (Carletta et al., 1995). Examples of each move type are given in figure 1.

In addition to the coding of move types done by the Edinburgh group, we have coded for other features

Polar answer:
Q: And you loop round about a bakery? <check>
R: Yeah.

Polar + STUFF:
Q: Have you got a white mountain? <qyn>
R: Yes, I have a white mountain to my north.

STUFF:
Q: Do you meet the savannah? <qyn>
R: Ehm, when I go round the collapsed shelter my savannah is to my right.

Figure 2: Answer forms

of the answers. We indicate whether or not an answer contains any overt *yes* or *no* terms (polar terms), which particular *yes* or *no* lexical items it contains, and whether or not it includes any other material (see figure 2).[2] We also code for whether the answer as a whole conveys a YES or a NO, or a meaning that does not appear to commit to either. For example, (2)(a) shows a YES, (b) a NO, and (c) a non-commital answer to query-yes/no questions. Furthermore, we code for broad syntactic features of the form of the question itself. In addition to coding for features of the question-answer game, we also identify the speakers before and after the game, and whether those moves relate to the game or not.

(2) a A: Antelopes?
B: Yes, I've got some antelopes right in the centre.

b B: Have you got anything down that side?
A: Not in that corner.

c A: You want to go right up north, past it
B: Past the top of it?
A: Right up above it.

Data Analysis

YES/NO meanings of the answers

We begin our analysis by looking at the YES/NO meaning of the answers to see if there are differences among the move types. We classify answers as having a YES meaning, a NO meaning, or not apparently committing to either (figure 3). For aligns the data show a strong bias toward producing responses that mean YES. The only answers other than a YES are non-committal. With checks we see a majority of YES answers but also a small number of NO answers. Query-yes/no's get YES answers about half the time, with plenty of NO

	NO	YES	NON-COMMITTAL
align	0%	89%	11%
check	11%	80%	8%
q-y/n	30%	55%	15%

Figure 3: Percent of answers having each meaning type

answers also given.

This data does suggest a strong correlation between answer meaning and move type, and essentially defines YES-expectation. Aligns have the highest proportion of YES answers and thus have the highest YES-expectation. Because checks have primarily YES answers but also some NO's, they are considered to have a somewhat lesser but still apparent YES-expectation. Query-yes/no's do not show a significant expectation for a YES answer. Our notion of having a YES-expectation is the same as the notion of a question being *conducive* (Bolinger, 1957; Stenström, 1984) or *biased* (Pope, 1976) towards a positive answer. As subsequent data will show, the YES-expectation strongly correlates with a number of features of the answers given to these three different question moves.

A Three-Way Analysis of Answers

In this section we investigate the forms of the answers given to the three kinds of question move types. We first discuss a previous analysis of answer forms, and then propose a new analysis.

Background: Direct vs. Indirect Answers In Green and Carberry (1992; 1994), responses to questions are broken down into direct and indirect components. An answer consisting of an overt polar term is categorized as direct, and an answer without overt polar terms but which commits to a YES or NO meaning as indirect.

Because people produce and understand answers that consist only of indirect responses, Green and Carberry investigate ways for a computational system to generate answers which consist only of indirect responses. Given this purpose, the overriding relationship between an indirect and direct NO, or an indirect and direct YES, is one of redundancy. When, for example, an indirect NO is indicated in an answer, it renders the direct *no* term unnecessary. Green and Carberry offer the following composed example and explanation to demonstrate the workings of their generator (1994) 195-196:

(3) a Q: You went to the party, didn't you?

b R: [no]

c [The baby sitter could not sit.]

d The baby sitter was sick.

In response to the question Q in (3), the generator initially produces the full response, (b–d), including both direct (b) and indirect (c-d) parts. Next, the response is pruned, if possible, starting from the end: non-inferrable parts are retained; inferrable parts are deleted. Since (d) cannot be inferred from (c), it is retained. However, as (c) may be inferred from (d), and (b) from (c) and (d), neither (b) nor (d) need to appear in the final response.

The direct/indirect split works well for Green and Carberry's purpose of producing informative answers that do not contain redundancies. However, as an examination of spoken data forces one to recognize, eliminating redundancy does not seem to be an overriding a goal in human-human interaction. Responses regularly include not only the direct portion of a response, but also an indirect portion, as Green and Carberry point out. To some extent, their model does account for these occurences: if the direct portion of the response were not inferrable from the indirect portion, then both the direct and indirect would co-exist in the generated response. However, this approach does not account for answers in which the direct is inferrable from the indirect, as in these attested cases:

(4) a Q: Well do you have an Apache camp?

b R: Yeah.

c Got an Apache camp (map task corpus)

(5) a Q: Have you got a big white mountain?

b R: Mmhmm,

c it's up a bit. (map task corpus)

In example 4, since (c) repeats the content of (a), (b) is inferrable from (c), yet both coexist in the actual response. Similarly, in example (5), the use of "it" presupposes the existence of a mountain, making (b) unnecessary. Once again, however, (b) and (c) co-exist in the actual response. The inability of Green and Carberry's model to account for examples such as (4) and (5) seems to result from their focus on redundancy in the relationship between the direct and indirect portions of the response. Our data suggest that a more fruitful explanation stems from (1) viewing the portions of the response in their collective relationship to the function of the question as a MOVE TYPE, and (2) analyzing responses along two dimensions rather than one: whether or not they contain an overt *yes* or *no* term, and whether there is other material or not.

Results Using an analysis that considers both the presence of *yes/no* lexical terms and the presence of *stuff* yields three categories of answers: bare *yes/no*, *yes/no* plus *stuff*, and just *stuff*.[3] The usefulness of the three-way split is shown most clearly by a comparison of two tables classifying response types. The

	NO: % of NO's		YES: % of YES's		NON-COMMITTAL
	direct	indirect	direct	indirect	% of total
align	−(0)	−(0)	100%(40)	0%(0)	11%(5)
check	88%(7)	13%(1)	81%(46)	19%(11)	8%(6)
q-y/n	77%(17)	23%(5)	70%(28)	30%(12)	15%(11)

Figure 4: Breakdowns of responses into direct and indirect

	NO			YES			NON-
	bare	*no+stuff*	*stuff*	bare	*yes+stuff*	*stuff*	COMMITTAL
align	− (0)	− (0)	− (0)	90% (36)	10% (4)	0% (0)	11% (5)
check	0% (0)	88% (7)	13% (1)	37% (21)	44% (25)	19% (11)	8% (6)
q-y/n	50% (11)	27% (6)	23% (5)	35% (14)	35% (14)	30% (12)	15% (11)

Figure 5: Breakdowns of responses into three content forms

first table results from the categorization of responses into direct and indirect categories. In figure 4 we separate YES, NO and non-commital answers, showing how many are given as direct and how many as indirect for each of the kinds of question types we consider. (Non-commital answers will all be indirect, and these are given as percent of total answers.) Our overall results here for indirect vs. direct checks and query-yes/no's are virtually identical to Stenström's (1984), although she does not break hers down by YES/NO answers in the same way, so we cannot compare those values. However, she uses dialogue from a variety of conversational contexts, and so it is encouraging that we found similar overall values for our particular context.

The breakdown in figure 4 shows a clear difference between aligns and the other two categories, but the differences between checks and query-yes/no's are more subtle. Aligns are not answered by NO's at all, while the checks and query-yes/no's both have a large number of direct NO's plus some indirect NO's. Aligns also do not have any indirect YES answers, while the other two do. Besides not distinguishing checks and query-yes/no's very clearly, this classification also does not clearly distinguish the YES and the NO answers within those move types.

In figure 5, we have teased apart the direct category into those that consist only of a polar *yes/no* term, and those that also include *stuff*[4]. While there does not appear to be much difference in YES responses for checks and query-yes/no's (though this will be discussed further in the section on distinctions based on *stuff*), we do see a clearer difference between checks and query-yes/no's in the way that NO answers are formed. Aligns do not have NO answers. Checks do not have bare NO answers; rather they have a high percentage of NO answers conveyed through an overt polar term and additional speech, and show some possibility for a NO communicated only through *stuff*. Query-yes/no's have responses of all types. We also see now that the YES and NO answers are clearly distinguished within each move type. Overall we see that,

while our 43% proportion of bare *yes/no* answers is much higher than the <10% found by (Richards, 1977), our results agree with his finding that *yes/no+stuff* and *stuff* make up the majority of answers. In addition, we found values similar to Stenström's for overall bare *yes/no* answers.

Discussion We can explain these results by appealing to the purpose of each move type and the strength of its YES-expectation. The purpose of an align is to check that the conversation is proceeding smoothly. This offers the hearer the chance to either indicate that all is well, as the normal case, or to point out that the conversation is in need of some repair that the speaker is not aware of. All of the "non-committal" answers for aligns in our data serve to initiate a repair, by requesting clarifying information. Although there is some chance of an answer other than YES to an align, the speaker expects the hearer to respond with YES—a strong YES-expectation—or to initiate a repair sequence. As suggested in figures 4 and 5, even when a YES answer is not valid, the response is never an overt or even apparent NO. This suggests that, rather than being a true yes/no question, aligns may be a more restricted yes-only question, where the responses that are available are YES or *initiate-repair*.

A check does seem to be a true yes/no question, in that it gets all three possible answer meanings. However, there are no bare NO answers in our data, which we explain by the the fairly high *yes*-expectation of checks. It seems that, since a NO is not really expected, speakers feel compelled to provide information about the unexpected NO, presumably for reasons of clarity and politeness (Brown and Levinson, 1987). That there are fewer non-committal responses for checks than aligns likely occurs because speakers can produce NO responses to checks, and are therefore not dependent on non-committals to communicate a lack of agreement.

Finally, query-yes/no's as a class do not have an intrinsic expectation of a YES or a NO—although there may be such expectations for individual occurrences due to context (Stubbs, 1983). For this move type, we see the greatest number of bare NO's because in this case NO is as informative, cooperative (Grice, 1989) and perhaps polite as YES. Put differently, the query-yes/no move does not require *stuff* in a negative response in the way an align and a check do, though *stuff* may be present.

Distinctions based on forms of *stuff*

As we noted earlier, while the three-way split as presented in figure 5 indicates a clear difference in NO answers to the three moves, it does not indicate a strong difference in YES-responses to aligns and checks. However, a closer look at the *stuff* does reveal differences.

Results Building on Green and Carberry's classification of *stuff* types for their generator, we have developed a new classification based on our data. Different kinds of *stuff* can have different intrinsic YES or

REP : Repetition of key part of the question
Q: Do you have an Apache camp?
R: Yeah. Got an Apache camp.

CONT : Continuation on to another subject
A: And then west.
B: Underneath the elephants?
A: Mmhmm, and then north.

EM : Emphasis (*"Yes, I do"*)

POL : Politeness (*"Sorry! Right."*)

LM : Limited Commitment (*"I think so"*, etc.)

ANS : Answer-other-Q: the person answers what they think is actually the relevant question
Q: Is that on the site of the plane crash?
R: Uh-huh, I've got that. I've got a site of plane crash.

ELAB : Accept and elaborate on the answer
Q: Underneath them?
R: Mmhmm. Until you just get past them.

Figure 6: Some kinds of *stuff*

	align	check	query-yn
ANS	–	–	1(4%)
CONT	1	14(39%)	4(15%)
ELAB	1	12(33%)	16(61%)
EM	–	–	2(8%)
LM	–	1(3%)	1(4%)
POL	2	–	–
REP	–	22(61%)	6(23%)
#answers	4	36	26

Figure 7: Total occurrences for YES answers with *stuff*

NO contents, or none at all. For example, continuing on to a new topic or elaborating on the proposition imply YES, whereas repetition and emphasis have no inherent YES or NO content–if we consider negation markers not to be part of the repetition–but depend on context to communicate a strong YES or NO. Combined with an overt polar term, repetition and emphasis function to strengthen and/or clarify (as well as communicate) the meaning of the overt term; without any overt polar term, they are solely responsible for communicating an unequivocal YES or NO. The particular kinds of *stuff* that occur for different move types gives us a further way to distinguish the moves. We list in figure 6 the kinds of *stuff* that we found in YES answers.[5] (There were not enough occurrences of *stuff* in NO answers to analyze them.)

In figure 7 we indicate for each move how often a particular kind of *stuff* occurrs in a YES answer. For answers containing more than one kind of *stuff*, each kind is counted separately, so that the total number of

answers given may not equal the sum of the column. We see a much higher use of Repetition in the YESes given to checks than to the other moves.[6] We also see a greater proportion of Continuations for checks. For Elaborations, we see the reverse, with a greater use for query-yes/no's. However, a more careful look reveals that this difference lies only in the plain *stuff* answers; for *yes+stuff* answers, both have the same rate (42-44%).

Discussion We attribute these kinds of results to the different roles that the moves play. Since an align is a general probe on the status of the conversation without reference to any particular piece of information, we would not expect a lot of discussion here for YES answers. With the check, a speaker seeks verification of his beliefs, whereas with a query-yes/no a speaker seeks to form beliefs. Because the check expresses some doubt, the check responder may be more likely to repeat the proposition (i.e. use Repetition) as a means of confirming it more carefully than just a YES would, since Repetition diminishes the chance for the speaker to erroneously have in mind a different proposition.

It is not surprising that there are Elaborations for both checks and query-yes/no's, since in both cases the speakers hope to reach a state of shared beliefs, and language is often ambiguous and does not map directly onto beliefs. A speaker may add more information in the hopes of increasing the chance that both conversants hold the same beliefs. It is not entirely clear why there are more Elaborations for query-yes/no's, or why the difference occurs only in bare *stuff* answers.

The greater number of Continuations for checks is attributed to the roles of the moves in turn-taking, as will be discussed in a later section.

Lexical choice of YES/NO terms

We also analyze differences in lexical *yes/no* terms chosen in the response to the different moves. Affirmatives in our data include *yes*, *yeah*, *uh-huh*, *right*, *okay*, *mm-hmm*, and combinations of these. In answers meaning NO, only *no* appears as an overt polar term in this data.

Results Figure 8 indicates how often each lexical term occurred in response to each move type[7]. From this, we can identify for each move type the most common YES term(s) used as the response: *right* for aligns, *uh-huh* and *yeah* for checks, and *yes* for query-yes/no's. We can also identify the most common use for each lexical term: the most common use for *right* and *okay* is to answer an align; for *uh-huh* and *yeah*, to answer a check; and for *yes*, to answer query-yes/no's.

Discussion We explain the lexical differences by proposing a correlation between YES-expectation and the semantic/pragmatic content of the lexical terms. *Yes*, for example, seems to involve a stronger affirmation than *okay* does. In figure 8, the *Yes* terms are listed in order of increasing strength. With this arrangement,

	align	check	query-yn
Okay	7	1	0
Right	19	11	2
Mmhmm	7	9	5
Uh-huh	6	15	7
Yeah	1	15	5
Yes	0	1	11
No	0	7	17

Figure 8: Number of occurrences of each form

	Decl	Interr	Phrase	Word
align	1(2%)	–	1(2%)	43(96%)
check	28(40%)	4(6%)	37(53%)	1(1%)
qyn	4(5%)	49(67%)	19(26%)	–

Figure 9: The forms of the questions

for each kind of move, the histogram of *yes* terms falls off fairly smoothly from its peak, but each move has its peak in a different place. We conclude from the data that the stronger the YES-expectation of the question, the weaker are the affirmative forms that tend to be used in answering it. Aligns, which have a strong YES-expectation, have a high likelihood of being answered with weak terms such as *okay* and *right*; checks have a somewhat weaker YES-expectation, and tend to be answered by somewhat stronger terms, such as *uh-huh* or *yeah*; and query-yes/no's generally have no YES-expectation, and the answers to these questions cluster closer to *yes*.

The Form of the Question

Our approach has been to investigate the response as it relates to the question, and not just the response in isolation. For this reason, we also analyze the form of the question itself. Figure 9 indicates whether the question consists of declarative syntactic form (e.g. "I'm going east underneath the bakery?"), an interrogative form ("Is that to the right?"), a non-sentential phrase ("To my right?"), or just a *yes/no* word ("Right?"). We cannot assume the elliptical phrases correspond to any sentence structure, and so we rely on the full forms to give us insight here. What we see is that very few checks have an interrogative form, and very few query-yes-no's have a declarative form. Nearly all the aligns use just a *yes*-term.

From this we claim that the prototypical form for an align is a *yes*-term, for a check a declarative form, and for a query-yes/no an interrogative form. It appears that the weaker the YES-expectation, the more likely the question is to take the syntactic form of an interrogative. This fits with Stenström's (1984) claim that the interrogative form expresses a weaker assumption than a question posed in declarative syntax. Interestingly, even though the move types are defined in a purely functional way, there are clear correlations between the

	Interrupt?	Not
align	2(4%)	43(96%)
check	44(63%)	26(37%)
qyn	19(26%)	54(74%)

Figure 10: Whether the question is likely to be an interruption

	Ask/old	Ask/new	Ans/new
align	–	97%(36)	3%(1)
check	33%(22)	54%(36)	13%(9)
qyn	12%(8)	82%(55)	6%(4)

Figure 11: Who says what after the answer

move type and the syntactic form of the question.

The context of the question and answer

We now back away slightly from the question-answer pair itself and look at the way the pair is related to its context, which in our case means the utterances before and after it. We are interested here in whether the different kinds of questions have different statuses in the turn-taking structure, and also whether the YES-expectation has an effect on the dialogue after the answer is given.

Results Figure 10 indicates whether it seems that the question is an interruption in the turn-taking sequence. It is NOT considered an interruption if the previous utterance (1) is uttered by the asker (2) is a question for which the current asker requests clarification, or (3), is a simple polar term (an acknowledgement or simple answer). The typical case where the question may be an unexpected turn is when the previous utterance is performed by the person other than the asker, and when it is simply gave contentful information. This table suggests that checks have the greatest potential to be interruptions in a turn-taking sequence.

Next we look at what happens right after the answer to a yes-no question.[8] First, in Figure 11, we look at who takes the next turn, and whether that turn moves on to new information (new) or continues with the same issue as the question (old). For instance, an acknowledgement or follow-up question will be *old*; a new instruction will be *new*. If there is an acknowledgement and then new information by the same speaker, this is counted as *new*. Here we again see a stronger match between aligns and query-yes/no's than either of these with checks. Both aligns and query-yes/no's typically end with the asker moving on to new information. With a check the asker frequently offers a simple acknowledgement or else asks a follow-up question (Ask/old), or the responder answers the question and then moves directly on to something new (Ans/new).

Now we move to a more specific question about the move after the answer: figure 12 indicates whether

	Acknowlg	Other	None
align	5(14%)	–	32(86%)
check	25(37%)	22(33%)	20(30%)
qyn	42(63%)	8(12%)	17(25%)

Figure 12: Responses to the answer

the answer to the question is acknowledged by the asker. There can be a direct acknowledgement (typically an "okay"), some other response that addresses the answer, such as a follow-up question, or no verbal response to the answer. Overall, we get a response rate of 42%, which is similar to what Stenström (1984) found overall (41%). In fact, our value, which emerges from a situation in which the participants are partially concealed from each other, is midway between the values she found for face-to-face and telephone conversation (36%/50%). Looking at the separate moves, we see an increase in direct acknowledgement as we move from aligns to checks to query-yes/no's, and a decrease in the lack of any response. In other words, the greater the YES-expectation, the less the answer is acknowledged.

Discussion In the data relating to the context in which the question-answer pair is embedded, we notice some data that correlate with YES-expectation, i.e. whether the answer is acknowledged, and some that don't, i.e. the data about who takes a turn when. This is the first data we have seen in which the move type has an effect but the YES-expectation does not. In both tables dealing with whose turn it is (figures 10 and 11), the aligns look more similar to the query-yes/no's than to the checks. We suggest that there is a conversational function to each move that explains these results (Sacks et al., 1974), and that it is not due to the YES-expectation.

Based on the patterns in the tables, we propose turn-taking schema for each move type. We have seen that aligns overwhelmingly get *yes* answers, the answers are rarely acknowledged, the speaker is not interrupting, and the asker continues to new information after the typical brief *yes* response. From this, we propose that the typical turn-taking structure of an align is that the asker already has the floor and expects a simple confirmation to the question, after which she will reclaim the turn, as schematized below.

align-scheme
(A's turn)
A: align
B: YES
A: start new game

A query-yes/no is similar to an align in all respects except that it typically does get acknowledged. We

thus propose a similar turn-taking scheme for query-yes/no's. We suggest that the elements that query-yes/no's have in common with aligns are due to their same turn-taking structure, while the differences–the answer form and whether an acknowledgement is given–are due to differences in YES-expectation.

qyn-scheme
(A's turn)
A: query-yes/no
B: answer
A: acknowledge; start new game

Checks are different from the other two. A check typically occurs after the other person has given information, without any overt indication that it is the asker's turn to speak. Also, the person who initiates a check is less likely to move on to new information. These indicate to us that a check is an interruption of the normal turn-taking process more often than are the other move types, as (Schiffrin, 1994) has suggested. A check interaction schematically looks like this:

check-scheme
(B's turn)
A: check
B: YES
A: (Acknowledge)
B: Continue her turn

The overall picture gained here is that aligns and query-yes/no's tend to be cases in which the asker has the turn, asks a question, gets an answer, and then continues her turn. A check appears to commonly be an interruption in the turn-taking process, so that after the question is answered, the turn goes back to the person who had it before the interruption.

This view provides an explanation for some of the differences in context we see between aligns and query-yes/no's on the one hand and checks on the other in what comes before and after the answer. Since aligns and query-yes/no's have similar turn-taking structures, they should exhibit similarities in what comes before and after the question game, and these should be different from what comes before and after a check.

Turn-taking sequences may also be able to explain one of the differences in the kinds of *stuff* found in answers, indicated in figure 7. We saw that Continuations are more likely to occur in the answer to a check.

This is presumably because the answerer perceives that it will still be her turn after the answer, and so she may decide to move right on to her the rest of her turn, although she may also allow the asker the chance to acknowledge or follow up.

Conclusion

Understanding what kinds of answers are to be expected for specific kinds of yes-no questions is an important kind of information that should be available to systems for modelling or interpreting dialogue. Our results and analysis contribute to the understanding of this domain in several ways by:

- Providing a new three-way classification of answers, and the notion of YES-expectation
- Demonstrating that the form of the question (i.e., declarative, interogative,) correlates with the MOVE TYPE of the question
- Showing that the YES-expectation of a question correlates with
  - the class of the response
  - the lexical form of the response, and
  - the turn-taking pattern of the dialogue

It is apparent that a number of different kinds of information contribute to the formation of responses to yes-no questions. We have suggested that these include what type of question was asked, and therefore what the speaker's expectations are about the answer, lexical knowledge about the strength of various *yes*-terms, and who is understood to be due a turn. We assume that there are also a number of other contextual factors that were not detailed here. For instance, we expect that a person answering a question will have to weigh the relative importance of politeness versus the need for the other person to be clear about what the answer is, and that there may be differences based on the situation or the question in whether explanation of an answer is appropriate. We also expect that the nature of the task the speakers are involved in may affect some dimensions of question asking and answering. We have tried to show that even without including all the possible considerations a speaker may have in formulating an answer to a yes-no question, there are some fairly strong patterns in answers. These derive from the kind of question that was asked, and our analysis explains these while still allowing some variability based on individuals or situations.

Our results are useful for generation in dialogue systems and for speech recognition. Our results suggest a simple strategy based on the intended answer and the move type of the user's question for selecting what polar lexical items to use and what kind of *stuff*, if any, to produce in a response, as well as what form of question to produce, based on the intended move. It would be straightforward to follow this strategy in the

generation component of a dialogue system because first, move types of the users' questions can be identified using prosody (Taylor et al., 1996) as well as form, and second, many dialogue systems already incorporate similar kinds of decision networks for choosing forms based on contextual considerations (e.g.(Bateman, 1995; O'Donnell, 1996)).

In speech recognition systems, having tighter constraints on the language model produces better performance. Being able to predict the form of answers from the move type of question they answer provides a valuble source of additional constraints that can be incorporated into such models.

The work done in this paper was based on one set of data in a particular kind of context. We are encouraged by how many different factors converge on a particular analysis, based on the roles of move types and their inherent YES-expectation, and by how well some of our numbers correlate with similar measures given by Stenström (1984), who looks at a variety of different dialogue contexts. This provides us with the hope that the results here will be generalizable in at least some ways to other dialogue situations than the one-on-one, goal-oriented experimental situation in which our dialogues occur.

Notes

[1]Throughout the paper, *yes* and *no* in italics refer to overt polar terms communicating affirmative and negative meanings (e.g., yes, yeah, right, no); YES and NO in caps refer to meanings.

[2]Terms which convey YES or NO meanings but cannot stand alone (e.g., do, not) are not considered overt polar terms. They are categorized as part of *stuff*.

[3]There were cases in which it was not clear whether an align was answered; those cases are not included here.

[4]The *stuff* typically occurs after the polar term, though not always.

[5]In addition to the *stuff* in figure 6, we have identified other *stuff* in the NO and NON-COMMITTAL categories (see also the papers by Carberry and Green), and are in the process of categorizing and defining it.

[6]We did see a number of cases in which an align was answered by repeating the question in the affirmative; e.g. "Right?" "Right." However, these are counted as bare-YES answers, since we count "right" as a *yes*-variant.

[7]When more than one overt lexical term is used, we only count the first one in figure 8.

[8]Figures 11 and 12 only include questions which are answered, and not ones where the responder asks for clarification instead of answering.

References

A. H. Anderson, M. Bader, E. G. Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, J. Kowtko, J. McAllister, J. Miller, C. Sotillo, H. Thompson, and R. Weinert. 1991. The HCRC map task data. *Language and Speech*, 34(4):351–366.

J. A. Bateman. 1995. KPML: the KOMET-Penman (multilingual) development environment. Technical report, GMD, Darmstadt. Release 0.8.

D. L. Bolinger. 1957. *Interrogative structures of American English*. University of Alabama Press.

P. Brown and S. C. Levinson. 1987. *Politeness: Some Universals in Language Usage*. Cambridge University Press.

J. Carletta, A. Isard, S. Isard, J. Kowtko, A. Newlands, G. Doherty-Sneddon, and A. Anderson. 1995. HCRC dialogue structure coding manual. Technical report, Association for Computational Linguistics.

J. Ginzburg. 1995. Resolving questions, I. *Language and Philosophy*, 18(5):459–527.

N. Green and S. Carberry. 1992. Conversational implicatures in indirect replies. In *ACL*.

N. Green and S. Carberry. 1994. A hybrid reasoning model for indirect answers. In *ACL*.

H. P. Grice. 1989. *Logic and Conversation*. Harvard University Press.

M. O'Donnell. 1996. Input specification in the WAG sentence generation system. In *INLG 8*, pages 41–50.

E. Pope. 1976. *Questions and Answers in English*. Mouton, The Hague.

J. C. Richards. 1977. Answers to *yes/no*-questions. *English Language Teaching*, 31:136–141.

H. Sacks, E. Schegloff, and G. Jefferson. 1974. A simplest systematics for the organization of turn-taking in conversation. *Language*, 50:696–735.

D. Schiffrin. 1994. *Approaches to Discourse*. Blackwell.

A. Stenström. 1984. *Questions and Responses in English Conversation*. Lund Studies in English 68. CWK Gleerup, Malmo.

M. Stubbs. 1983. *Discourse Analysis: The Sociolinguistic Analysis of Natural Language*. University of Chicago Press.

P. Taylor, H. Shimodaira, S. Isard, S. King, and J. Kowtko. 1996. Using prosodic information to constrain language models for spoken dialogue. In *ICSLP 96*.