

# When a nod is as good as a word: Form–function relationships between questions and their responses

D. V. M. BISHOP

*MRC Cognition and Brain Sciences Unit, Cambridge*

J. CHAN

*University of Manchester*

J. HARTLEY

*University of Sheffield*

F. WEIR

*University of Manchester*

## ADDRESS FOR CORRESPONDENCE

D. V. M. Bishop, Department of Experimental Psychology, South Parks Road, Oxford OX1 3UD,  
United Kingdom. Email: dorothy.bishop@psy.ox.ac.uk

## ABSTRACT

It is well established that syntactic form and communicative function do not always correspond: for instance, a syntactic question might function as a request for information (“did you see the play?”) or a request for acknowledgment when there is no doubt about polarity of the response (“there’s a tower at Blackpool, isn’t there?”). Using data from a corpus of 18 child–adult conversations, we distinguished adult utterances that solicited information from those soliciting acknowledgment (i.e., where the response was predictable, and the utterance served a predominantly social function). Both types of utterance were usually responded to by children, but the form of response differed according to the communicative function of the utterance. Nonverbal and prosodic responses (e.g., nods or “mmh”) were significantly more likely to occur in response to utterances soliciting acknowledgment than in response to yes/no questions that solicited information. There were consistent form–function relationships for responses as well as for soliciting utterances. Nonverbal nods and headshakes were not functionally equivalent to verbal “yes” and “no.”

Those interested in discourse readily appreciate the need to distinguish syntactic form and communicative function. Thus, the syntactic question “Would you mind opening the window?” functions as a (polite) command. The declarative form of “I don’t suppose you’d like to come to dinner with me” acts as an invitation. A tag question such as “It’s got salt in it, hasn’t it?” could be a request for information by an uncertain speaker (perhaps a forgetful cook) or a statement of opinion by a speaker who has no doubt of the truth of the proposi-

tion. To understand what makes some sequences of utterances acceptable as a conversation and others not, we must attend to the functional level: for instance, if we asked someone "Would you mind opening the window?" we would not expect them simply to reply "No [I wouldn't mind]" without taking any action. We expect action rather than information, and so if someone says "Of course," accompanied by the appropriate action, this seems entirely acceptable; we would not think that this meant "Of course I mind" but rather "Of course I will comply with your request." Research on communicative functions has been concerned with the questions of how far specific syntactic forms are correlated with particular communicative functions, and how nonsyntactic cues (e.g., intonation, awareness of the speaker's knowledge) are used to disambiguate the speaker's meaning.

Quirk, Greenbaum, Leech, and Svartik (1972) addressed the first of these questions in a syntactic classification of English question forms.

- wh*-questions, formed by using interrogative words, such as *who*, *why*, and *how*, in initial position.
2. yes/no questions, formed by subject-auxiliary inversion (e.g., "Are you coming tonight?").
  3. alternative questions, which offer a set of possible answers as options and may resemble yes/no questions (e.g., "Do you want eggs or bread?") or *wh*-questions (e.g., "Where are you going, Africa or Asia?") in their form.
  4. tag questions, formed by auxiliary + pronoun or auxiliary + negative + pronoun appended to a statement (e.g., "You are coming, aren't you?").
  5. declarative questions, which use the syntactic word order of a declarative statement but with questioning intonation (e.g., "John is coming as well?").

Quirk et al. noted that these question forms differ in their presuppositions and in the expectations they convey as to the response. For tag questions and declarative questions, they argued that one can predict the polarity of the anticipated answer (yes or no). Such utterances function to establish the existing common ground between conversational partners (see Clark & Brennan, 1991). There is no uncertainty to be resolved; the interlocutor is simply invited to agree with the self-evident truth of an assertion. The expected response to yes/no questions, *wh*-questions, and alternative questions is typically more uncertain.

Using this taxonomy, we may draw a simple distinction between question forms that solicit acknowledgment, where the polarity of the answer is entirely predictable, and those that solicit information, where the speaker has some uncertainty (which is reduced when the answer is supplied). Quirk et al. (1972) maintained that usually questions with the syntactic form of tag questions and declarative questions will be of the former type, and the other syntactic forms of the latter. However, there is little quantitative data on how reliable syntactic form is as a cue to communicative function or how syntactic information is combined with other sources of information, such as intonation and mutual knowledge, when determining a speaker's intention.

The distinction between utterances soliciting acknowledgment and those so-

liciting information entails that the same form of words used in response to a question may serve different communicative functions according to the context. "Yes" may serve to provide information in response to a yes/no question such as "Have you got any brothers and sisters?", but it may also serve merely to acknowledge the interlocutor's contribution in response to a tag question such as "That's nice, isn't it?" Whereas the former type of response plays a role in topic development, the latter serves more of a social function, keeping a conversation running smoothly by providing mutual evidence of the participants' engagement in the interaction. Although the range of possible syntactic forms used in response to questions is limited, there are different ways of expressing agreement and disagreement: for example, agreement may be expressed by nodding or by saying "yes" or "mmm," thus raising the question of whether these different forms serve different functions.

Two related questions are of interest. First, do information- and acknowledgment-soliciting utterances differ in their success in soliciting a response? Because a response to an acknowledgment-soliciting utterance is not crucial for topic development, we might expect such responses to be regarded as more optional than responses to information-soliciting utterances. Second, do information- and acknowledgment-soliciting utterances differ in the types of response they solicit? Clearly, the nature of a response is heavily constrained by the grammatical form of an utterance, and so the response to a *wh*-question will be different to the response to a yes/no question. However, if we exclude *wh*-questions and alternative questions, we find that the remaining grammatical question forms may be regarded as binary solicitations: that is, they anticipate "yes" or "no" as a response. For these solicitations, the response may be verbal ("yes" or "no"), prosodic ("mmh," "uh huh," etc.), or gestural (a nod or shake of the head). Summarizing research on responses that serve a purely social acknowledging function ("backchannel" or "minimal" responses), Reid (1995) noted that all of these forms may be used for this purpose, but as far as we have been able to establish, no research has been conducted to compare the use of these different forms of response in response to different communicative acts. One prediction we might make is that the verbal forms "yes" and "no" would be more likely to be used when a response provides information than when it provides acknowledgment. The reasoning behind this prediction is that the verbal forms are the most salient: gestural responses are fairly unambiguous (nod vs. shake) but are easy to miss if the conversational partner is not watching; prosodic responses are less likely to be missed but may be more ambiguous, with a less clear differentiation between positive and negative polarity. Although it is possible to express positive or negative meaning more or less unambiguously (as when [ʔəʔə] is used to mean "no"), in other cases the intonation of "mmh" or "uh huh" may be all that distinguishes a positive from a negative interpretation.

Our interest in this topic developed in the course of studying children with language impairments. Some of these children had communicative problems that were much more apparent in an open-ended conversational setting than in a formal test situation (Adams & Bishop, 1989); we wanted to investigate this difference further by studying how children responded to different types of con-

versational solicitation. However, we found that, although authors such as Quirk had documented the different functions of particular question forms, much less was known about the types of responses these forms typically elicited, and we could find no work on children's ability to distinguish and respond to soliciting utterances with different functions. Thus, we embarked on this study to investigate how normally developing children responded to conversational solicitations with different functions.

## METHOD

### *Participants*

This analysis is based on conversations between an unfamiliar adult and 18 children (9 boys, 9 girls), aged 7 and 8 years, who were attending regular classes at English primary schools in Lancashire and Yorkshire. All of the children were native English speakers; some used features of local dialect, such as producing "the" as [tʰ], but in no case were dialect variations large enough to affect intelligibility, either for the adults conversing with the children or for those who rated the conversations. A total of 6 children conversed with adult S, a speech-language pathologist (JH), and 12 conversed with adult L, a linguist who had no experience of therapeutic work with children (FW). Psychological test data confirmed that this was a representative sample of children of this age: expressed in *z* scores, mean nonverbal reasoning, as assessed by Raven's Coloured Matrices (Raven, Court, & Raven, 1986), was 0.24 (*SD* = 0.66); mean expressive language, as indexed by the Clinical Evaluation of Language Fundamentals-Revised, Repeating Sentences subtest (Semel, Wiig, & Secord, 1980), was 0.6 (*SD* = 1.12); and mean receptive vocabulary, as measured by British Picture Vocabulary Scale (Dunn, Dunn, Whetton, & Pintilie, 1982), was -0.04 (*SD* = 1.04).

### *Conversational data*

The conversations were gathered in a semistructured format, originally used by Adams and Bishop (1989), in which photographs of familiar events (e.g., a visit to the doctor, a trip to the seaside) were used to initiate specific topics of conversation, with the child then being encouraged to recount similar personal experiences. The adults were asked to avoid using demonstration questions to which the answer was already known; rather, the idea was to elicit some natural conversation from the child about topics of common interest. No attempt was made to control the number or type of questions asked because of concern that this could lead to an unnatural and stilted interaction, with findings that would have little relation to everyday conversation (Duncan & Fiske, 1985). Each conversation lasted from 7 to 15 minutes and was recorded on both audio- and videotape. The tape recorder was a walkman-sized Aiwa machine with directional microphone, using standard TDK low-noise high-output tapes. Conversations were transcribed from the audiotape using the system shown in Appendix 1; the transcript was then supplemented by relevant nonverbal information from

the video. Communicative gestures such as nods and headshakes were noted on the transcript, but no attempt was made to code other aspects of nonverbal communication, such as eye contact or facial expression. The number of turns per conversation ranged from 133 to 215 ( $M = 186$ ), the number of adult utterances ranged from 138 to 237 ( $M = 201$ ), and the number of child utterances ranged from 73 to 232 ( $M = 161$ ). For the adult, on average, 1.6% of the utterances were unintelligible because of overlapping speech and 0.9% were unintelligible for other reasons (e.g., background noise, soft speech). For the child, on average, 2.9% of the utterances were unintelligible because of overlapping speech and 4.8% were unintelligible for other reasons.

### *Coding of communicative functions*

The system for coding communicative functions was developed over a period of five years and was influenced by a range of sources, including researchers concerned with normal and disordered language development (Dore, 1977a, 1977b; Fey, 1986) and discourse analysts such as Coulthard (1977). No existing system proved entirely suitable for our purposes. Most of the coding schemes developed for handling child language were targeted at relatively young children engaged in toy play, and, as Chapman (1981) pointed out, little information is available on the reliability of existing coding systems.

The unit of analysis that we adopted was the utterance. We defined an utterance as a unit, in terms of its form and function, which begins with a fresh start and continues until the next fresh start is perceived. We found that a satisfactory definition of an utterance boundary required one to listen beyond a given stretch of talk to see whether the following talk belonged with it, rather than solely relying on prosodic or syntactic features of a stretch of talk. Take, for instance, an utterance such as "That's a nice one, isn't it?"; if one were to stop the tape after "one," the utterance would sound complete, both syntactically and in terms of forming a complete tone group. However, when listening on, the syntactic, semantic, and prosodic features of the tag that follows make it clear that it is not a fresh start but rather part of the preceding utterance.

Judges were asked to use all of the information available (syntactic, semantic, prosodic, and communicative function) to determine whether a stretch of speech made a fresh start; although this instruction may seem vague, raters had little difficulty in using native speaker intuition to make this judgment. We obtained acceptable agreement between the trained transcribers in using this definition: pairwise agreement for three transcribers rating 100 intelligible turns ranged from 88 to 95%.

*Turn-handling symbols.* The distinction between utterances that solicited information and those that solicited acknowledgment was incorporated into the coding we used to mark turn-handling functions of utterances at the transcription stage (summarized in Appendix 1). The terminating symbol (?) was used where the coder judged that the speaker had finished speaking and was directing another speaker to take a turn which would provide information. If the interlocutor did not take a turn at this point, there would be a sense of an expectation unful-

filled. The content of the response was not a foregone conclusion, although one response might be more probable than another.

The symbol (~) was used to terminate utterances where it was judged that the turn had passed to the interlocutor with the expectation that he or she would provide acknowledgment. Often, such utterances would have the syntactic form of a question, but unlike utterances marked (?) they left no uncertainty to be resolved. The judgment was that the speaker anticipated a response of given polarity and would have been surprised if any other response had occurred.

Inevitably there are cases where the decision to code (~) or (?) is a matter for debate. Though coders might be uncertain about the speaker's mental state, it is also the case that the dividing line between these two classes is a fuzzy one; it is more realistic to treat soliciting utterances as varying quantitatively on a dimension of the extent to which a speaker is judged to predict a response of given polarity. In general, the instructions were to use the (~) code if the expectation of a "yes" or "no" answer was so strong that a response other than the expected polarity could be expected to engender surprise in the speaker. Thus, for an utterance coded (?) there might be a much stronger probability of one response rather than another, but both possibilities could reasonably be entertained. In deciding whether to code (?) or (~), transcribers were instructed to take into account all available information, including their awareness of the speaker's knowledge and prosodic and grammatical cues, without relying exclusively on any one source of information. However, the response (or lack of response) that followed the utterance was not taken into account in making this judgment. This was important because we wanted our coding system to be able to detect cases where the child's response did not meet the expectations of a solicitation.

The symbol (.) was used to terminate tone groups which were deemed neutral in turn-handling function: that is, where it was judged that no response was being solicited from the interlocutor, but where the interlocutor had the option of taking up the turn because the speaker did not indicate a clear intention to retain the turn.

Other turn-handling marks, which do not concern us here, include the following: (,) is used to denote a speaker's intention to continue with the turn; (/) is used to denote an incomplete utterance terminated by broken-off speech; and (≈) is used to denote ambiguous turn-handling.

Transcribers were trained to treat turn-handling marks as different from punctuation marks. After training, agreement between raters in judging whether the utterance terminated in (?), (~), or another turn-handling mark was 87.5% (Cohen's [1960]  $\kappa = .73$ ) for two conversations totaling 344 utterances.

*Types of communicative function.* Each complete utterance was coded for communicative functions, as shown in Table 1. It is important to note that terms such as "question" or "statement," when used in this context, refer to communicative functions and not to syntactic forms. As can be seen from Table 1, there are correspondences between the turn-handling codes and the codes for communicative functions. For utterances terminating in (?), three functions are recognized:

Table 1. *Codes used for communicative functions of utterances with topical content<sup>a</sup>*

	Topic development	Topic maintenance
Solicits information: transcribed terminating in (?)	q: question	rc1: request for clarification ch: challenge
Solicits acknowledgment: transcribed terminating in (~)	s: statement	sum: summary re: restatement ev: evaluation
Neutral, does not solicit a response: transcribed terminating in (.)		

<sup>a</sup>Where consecutive utterances in a turn serve the same communicative function (i.e., they function as a single discourse "move"), they each receive the appropriate code, with + preceding the code for the second and subsequent utterances.

question, request for clarification, and challenge. As shown in examples (1) through (3), a question (q) develops the topic with a request for information.

- (1) 1.1 A: have you got a 'camera at home? q  
2.1 C: yeah.
- (2) 33.2 A: ^did it happen at 'night time, or during the 'day? q  
34.1 C: night time.
- (3) 61.2 A: what was your 'favourite thing about 'america? q  
62.1 C: the desert.

A request for clarification (rc1) backtracks to clarify something that is unclear. In example (4), the adult asks a succession of questions to establish what the child means by "sort of a bridge" in utterance 12.3, establishing that what he really means is a pier.

- (4) 12.3 C: ^we only went on, um, (sniffs) sort of 'bridge, on the pacific 'beach. ... (C gestures)
- 13.1 A: bridge? rc1  
what 'sort of a bridge? +rc1
- 14.1 C: d you know those 'things, at the end there s a kind of 'funfair~
- 15.1 A: = the pier? rc1
- 16.1 C: yeah!.
- 17.1 A: so it s like a 'bridge, but only joined at one 'end? rc1
- 18.1 C: yeah.
- 19.1 A: = it goes off into the 'sea? rc1
- 20.1 C: yeah.

A challenge (ch) backtracks to query the truth value of what the interlocutor has said. In example (5), the adult has no doubt about what C has said, but she doubts its veracity.

- (5) 21.2 C: and we saw a big red 'shark.
- 22.1 A: a shark!? ch
- 22.2 at blackpool!? +ch

Challenges were rare in this corpus and will not be discussed any further.

Utterances terminating in (∼) could be assigned the function of statement, summary, evaluation, restatement, or request for acknowledgment. A statement (s) develops the topic by providing unsolicited information, as in turn 22.1 in example (6).

- (6) 20.1 A: where did you go on 'holiday?  
21.1 C: to blackpool. 1.0  
22.1 A: there s a really big 'tower there, is nt there∼ s

Note that the (∼) turn-handling mark is used rather than (?) because the transcriber judged that there was no doubt in the speaker's mind about the existence of the tower.

A summary (sum) summarizes what the interlocutor has said, either integrating the content of several prior conversational turns or making inferred information explicit. The summary does not move the conversation forward but rather helps draw together what has gone before. It is often used to round off a subject before moving on to a new branch of a topic, as shown in example (7).

- (7) 68.3 C: we even saw a 'ground 'squirrel.  
69.1 A: what does 'that look like?  
70.1 C: = desert ground squirrel.  
70.2 C: ^ it's just sort of like um, a 'skunk, and, ... (C gestures)  
70.3 C: <um> i 'think it ... did nt have a 'tail. ...  
70.4 C: and it just 'scattered across, so i did nt get a very good ...  
'view of it.  
71.1 A: = right.  
71.2 A: does it look a bit like a 'mouse, or a 'rat, or som[ething]?  
72.1 C: [yeah.] (C nods)  
73.1 A: how 'big is it?  
74.1 C: about 'that big. ... (C demonstrates)  
75.1 A: ri:ght.  
75.2 A: oh,  
75.3 A: so it was 'almost like a squirrel, without a 'tail then∼ sum  
76.1 C: yeah. ...

The decision to code (sum) rather than (rcl) was, in effect, made at the transcription stage. If the transcriber had judged there to be any uncertainty in the speaker's mind about the polarity of the expected response, then this would have been transcribed terminating in (?), and the (rcl) code would have been used.

An evaluation (ev) evaluates the affective significance of what has been said, as in example (8).

- (8) 9.2 A: 'where did you go? ...  
10.1 C: llandudno.  
11.1 A: it s 'nice there, is nt it∼ ev

A restatement (re) maintains the topic by restating what the interlocutor has said, as in example (9). The judgment is that the utterance serves to express



interest in the topic and keep the conversation going. If the speaker had been querying or challenging what the other had said, then the (rcl) or (ch) code would have been used instead.

- (9) 20.1 A: where did you go on 'holiday?  
 21.1 C: to blackpool.  
 22.1 A: **blackpool~** re  
 23.1 C: mm.

For this analysis we excluded from this category cases where a "yes" or "no" answer to a question by the child was followed by a "yes" or "no" by the adult, spoken with a questioning intonation.

A request for acknowledgment (rack) serves solely to solicit acknowledgment from the interlocutor and has no topical content. This class of utterances will not be considered further here because they have little scope for syntactic variation and virtually always take the form of isolated tags, as in example (10).

- (10) 9.2 A: 'where did you go?  
 10.1 C: llandudno.  
 11.2 A: 'did you~ rack

The codes of (s), (re), (sum), and (ev) could also all be applied to utterances terminating in (.): that is, those judged to be neutral in turn-handling function. For instance, in example (8), if at 11.1 A had said "oh, that's nice" with a falling intonation, this would have been transcribed terminating in (.) but still would have been coded as an (ev).

Each soliciting utterance type had a corresponding code for the anticipated type of response. Thus, the predicted response to a question (q) was an answer (a), the predicted response to a request for clarification (rcl) was a clarification (cl), and the predicted response to any utterance soliciting acknowledgment was an acknowledgment (ack).

We will not discuss any further the other codes used in our system, as these are not of direct relevance to this article. These include codes for utterances concerned with behavior regulation (e.g., commands) and utterances that were judged to have no topical content but merely served to keep the conversation going (e.g., unsolicited backchannel utterances, such as "uh huh").

A training manual with numerous examples and exercises was developed, and after prolonged training using these materials, the first two authors achieved 87.2% agreement (Cohen's  $\kappa = .86$ ) in using the full set of 23 codes for communicative functions to code five pretranscribed conversations totaling 897 utterances. (These conversations were not part of the data set analyzed here.) Agreement was also computed for the more stringent condition, where both transcription and coding were carried out independently by two coders. Here, agreement for coding communicative function of two conversations totaling 362 utterances was 74% (Cohen's  $\kappa = .69$ ). A complete description of the coding system would require far more space than is available here, but a sample of coded transcript is shown in Appendix 2 for illustrative purposes.

### *Coding responses to different types of soliciting utterances*

In our corpus of conversations, it was overwhelmingly the adult who led the conversation and who thus asked virtually all the questions. Analysis was restricted to adult utterances that had been coded as (q), (rcl), (s), (sum), (ev), or (re). As can be seen from the examples, in some cases a communicative function was seen as extending over more than one utterance. In these cases, the first utterance was given the conventional code, and the subsequent ones were coded with a (+) to indicate that they were being treated together as a single unit of analysis: in the terminology of Coulthard (1977), they constituted a single discourse move.

Children's responses to these adult moves were classified as follows:

no response, despite the opportunity to respond; given in two situations: (a) when the adult paused for at least 1 second, waiting for a response that was not forthcoming, and then continued, and (b) when the child took up the turn without responding to the adult's contribution (e.g., by making an unrelated comment or continuing with a narrative account that ignored an intervening solicitation)

2. yes or no (including "yeah")
3. prosodic response ("mmh," "uhuh," etc.)
4. nonverbal response (nod, headshake, shrug)
5. other verbal response (e.g., response to *wh*-question, "don't know")

Utterances that were unintelligible, incomplete, or could not be adequately coded because of overlap with the adult utterance were excluded from consideration.

## RESULTS

### *Probability of responses to adult solicitations*

Table 2 shows the frequency with which children responded to the adult utterance types shown in Table 1. The category "solicits acknowledgment" includes (s), (sum), (ev), (re), and (rack). These were transcribed terminating in (~). The category "neutral" includes (s), (sum), (ev), and (re). These were transcribed terminating in (.).

The rates of responding to information-soliciting utterances were extremely high. They were slightly lower for acknowledgment-soliciting utterances, but nevertheless the great majority of these utterances were responded to.

Using the method of log-linear analysis described in Appendix 3 to take into account the different totals from which proportions were computed, there was a highly significant difference between the three types of adult utterances, in terms of their success in eliciting a response. Information-soliciting versus acknowledgment-soliciting utterances yielded  $z = 6.16$ ; information-soliciting versus neutral utterances yielded  $z = 13.4$ ; and acknowledgment-soliciting versus neutral utterances yielded  $z = 7.37$ .

Table 2. *Mean (SD) number of adult soliciting and neutral utterances per conversation, with frequency of response to each type*

	<i>N</i> adult utterances	Proportion responded to <sup>a</sup>	
		Raw	Weighted <sup>b</sup>
Soliciting information (?)	53.0 (11.61)	.979 (.075)	.966 (.028)
Soliciting acknowledgment (~)	14.8 (5.48)	.889 (.113)	.861 (.101)
Neutral (.) <sup>c</sup>	8.6 (3.42)	.510 (.282)	.533 (.199)

<sup>a</sup>Excludes adult utterances where there was no opportunity to respond because the adult continued with the turn within 1 second.

<sup>b</sup>See Appendix 3 for explanation.

<sup>c</sup>Utterances coded as responses to acknowledgment-soliciting and neutral utterances were "yes" or "no," "I see," "I know," prosodic responses, or non-verbal equivalents.

Table 3. *Mean (SD) proportions of different response types to adult soliciting utterances (excluding wh- and alternative questions)*

	Soliciting information		Soliciting acknowledgment		<i>z</i>
	Raw	Weighted	Raw	Weighted	
Nonverbal response as proportion of all responses	131 (.132)	195 (.113)	.289 (.257)	.351 (.176)	6.41
Prosodic response as proportion of all vocal responses (prosodic + verbal)	.044 (.050)	.069 (.053)	.246 (.243)	.331 (.207)	7.56

### *Type of responses to binary solicitations*

Analysis of children's responses is restricted to those syntactic types that can elicit a verbal response of "yes" or "no," a prosodic response such as "mm," or a nonverbal response such as a nod or headshake (i.e., binary solicitations). These are utterances with the syntactic form of yes/no questions, tag questions, or full or elliptical declarative questions.

Table 3 shows the proportion of responses that were nonverbal to information-soliciting and acknowledgment-soliciting utterances. Log-linear analysis showed a highly significant difference between the two types of soliciting utterances. Nonverbal responses were made significantly more frequently to acknowledgment-soliciting utterances.

Vocal responses were next subdivided into verbal (yes/no) responses and prosodic responses. The two types of soliciting utterance again differed significantly. A higher proportion of verbal responses were made to information-soliciting utterances. Thus, we can conclude that prosodic and nonverbal responses are seldom used to provide information and are much more likely to be used to denote acknowledgment.

Before we can accept this result with confidence, however, we need to consider that it might be an artifact of transcriber bias. This could arise if, for instance, the decision to code a soliciting utterance as (?) or (~) were unwittingly influenced by the response that was given (or by a failure to observe a response). To test this possibility, four conversations were selected at random, and the transcripts were printed with the adults' soliciting and neutral utterances marked in bold. Another researcher was trained in the transcription conventions and then asked to give a turn-handling mark to the utterances marked in bold, listening to the original tape recording and stopping the tape immediately at the end of each adult utterance. Cases where the transcriber inadvertently heard the start of the next utterance were excluded from consideration. Agreement for coding the marked adult utterances as (?), (.), or (~) was 93.1% ( $\kappa = .839$ ,  $N$  observations = 245). Restricting consideration to binary solicitations, agreement for coding these as (?) or (~) was 90.4% ( $\kappa = .691$ ,  $N$  observations = 125). Interrater agreement on turn-handling functions thus remained high and well above chance when no information was available about the response given.

### *Syntactic form and yes/no solicitations*

The different pattern of responding seen for these two types of utterance raises the question of how far children are able to rely on syntactic form as a reliable cue to utterance function. The relationships noted by Quirk et al. (1972) agree with our native speaker intuition about the presuppositions conveyed by different syntactic types, but it is generally accepted that form-function correspondences are far from perfect. One can readily generate counterinstances where, for instance, a tag or declarative question may be used even when the response is not a foregone conclusion: for example, an anxious spouse might ask "You remembered to buy the vegetables?" or "You remembered to buy the vegetables, didn't you?" Although nonsyntactic sources of information usually disambiguate meaning, misunderstanding can sometimes arise when a speaker is interpreted as soliciting acknowledgment but is in fact seeking information, as in the following real-life example:

- A: I'm hoping to give a talk at Bangor.  
B: Your paper has been accepted.  
A: Has it?  
B: No, I was asking you.

Conversational partners presumably use multiple sources of information to infer a speaker's intended meaning, including knowledge of what the speaker knows, which allows them to infer whether the speaker lacks critical information. In addition, prosodic and nonverbal cues may help disambiguate different communicative functions. It was not feasible in this study to investigate all possible cues to communicative function, but we were able to consider the extent to which syntactic form was a reliable source of information.

The final utterance in each adult move was classified as follows: *wh*-question; yes/no question (i.e., with auxiliary inversion); alternative question; full declarative question (i.e., declarative with questioning intonation); short (elliptical) de-

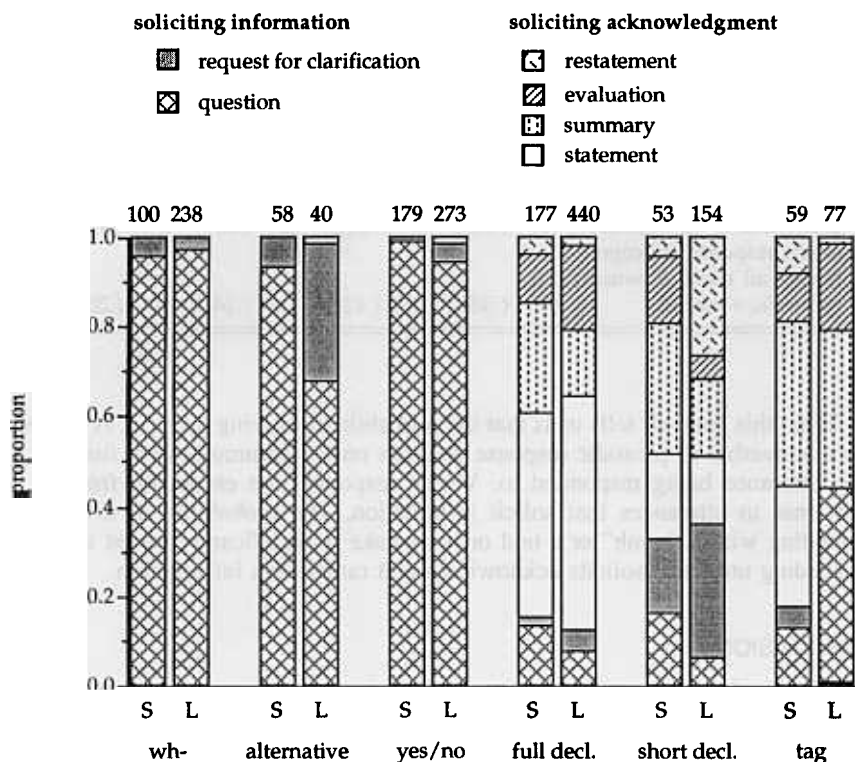


Figure 1. Proportions of utterances of a given grammatical form that have a particular function. S and L denote the two adults who participated in conversations. The numbers at the top of each bar denote the total number of utterances of each grammatical type (from 6 conversations for adult S and from 12 conversations for adult L).

clarative question (e.g., “mum and dad?”); and tag question. Figure 1 shows the distribution of syntactic classes of adult utterances in relation to communicative function. It is apparent that, although *wh*-questions, yes/no questions, and alternative questions were almost invariably used to solicit information, the situation is less clear-cut for tag questions and declarative questions, which were used predominantly for soliciting acknowledgment but on occasion were also used to solicit information. We therefore carried out another analysis to see whether the different pattern of responding to information-soliciting versus acknowledgment-soliciting utterances still obtained when attention was restricted to grammatical forms that could take either function (i.e., full or short declarative questions and tag questions). Results of the analysis are shown in Table 4. Even though the absolute number of soliciting utterances was greatly reduced, the pattern remained the same. Nonverbal responses were significantly more common to acknowledgment-soliciting utterances than to information-soliciting utterances, and prosodic responses formed a higher proportion of the vocal responses to acknowledgment- versus information-soliciting utterances.

Table 4. *Mean (SD) proportions of different response types to adult soliciting utterances with grammatical form of declarative or tag question*

	Soliciting information		Soliciting acknowledgment		z
	Actual	Weighted	Actual	Weighted	
Nonverbal response as proportion of all responses	.207 (.261)	.267 (.186)	.283 (.254)	.343 (.175)	2.04
Prosodic response as proportion of all vocal responses (prosodic + verbal)	.057 (.086)	.132 (.126)	.257 (.241)	.343 (.209)	3.60

What this analysis tells us is that the probability of giving a verbal as opposed to a nonverbal or prosodic response depends on the communicative function of the utterance being responded to. Verbal responses are especially frequent in response to utterances that solicit information. The probability of a child responding with a “mmh” or a nod or headshake is significantly raised when the preceding utterance solicits acknowledgment rather than information.

## DISCUSSION

Responses that appear semantically equivalent – namely saying “yes,” nodding, or saying “mmh” – are used differentially to serve specific communicative functions. The effect is probabilistic rather than absolute, and the children we studied tended predominantly to use verbal responses for all functions. Nevertheless, when nonverbal gestures and prosodic responses were observed, they were more likely to be used to acknowledge than to provide information. Furthermore, this did not merely reflect a simple matching of response type to syntactic form of the soliciting utterance: the differentiation could still be seen when attention was restricted to declarative and tag questions, which could serve both kinds of function. Our analysis does not allow us to specify what other cues children rely on to identify the communicative function of soliciting utterances, but we can speculate as to what these might be. One source of information is an awareness of the speaker’s mental state, which enables the child to deduce whether the speaker already possesses information mentioned. Children as young as 4 years of age are able to appreciate that other people may have different knowledge, desires, and beliefs from themselves and to apply their skills in social cognition to interpret what others say to them (Lewis & Mitchell, 1994). Vocal cues are also important: prosodic features such as pauses and intonation help us estimate a speaker’s knowledge about, commitment to, and confidence in what he or she is saying (Brennan & Williams, 1995). Little is known about the development of the ability to use vocal cues. Another source of information all too often overlooked in linguistic accounts of pragmatics is nonverbal behavior. Our study shows that nonverbal communication plays a distinctive role in signaling different functions when responding to a solicitation. It seems likely that

cues such as facial expression, eye contact, and nodding while speaking might also be used to signal the function of soliciting utterances.

It is important to note the limitations of this study, which focused on the interactions between school-age children and an unfamiliar adult in a standardized setting. Our aim was to encourage informal conversation between an adult and child, with the talk centering on past experiences and future plans rather than on the here and now. In this context, talk concerned with behavior regulation (commands and suggestions) is seldom observed. In a nursery school setting, Dore (1977b) found that 11% of *wh*-questions in a corpus of conversations functioned as directives (e.g., "How about doing x?" or "Why don't you do x?"). Clearly, it would be dangerous to conclude that the communicative functions of *wh*-, yes/no, and alternative question forms are always as predictable as they appeared to be in the setting that we used.

It would be of interest to study interactions between different kinds of dyads, to investigate how much the cognitive maturity, gender, and relative social status of participants influence the patterns of results. We know that young children (aged around 4 years) make heavy use of nonverbal responses when interacting with adults (Adams & Bishop, 1989; Doherty-Sneddon & Kent, 1996), whereas in the sample seen here, such responses were in the minority. This raises the question of whether the differential use of verbal, prosodic, and nonverbal response types is something that emerges as the child's communicative skills mature.

## APPENDIX 1

Below is a summary of main transcription conventions used in this study.

### *Layout*

Turns and utterances are numbered consecutively. The turn is identified by the number before decimal point, and the utterances within turn is identified by the number after it. Lowercase is used throughout.

### *Tone group-ending markers*

- continuation of turn
- neutral end of utterance
- turn passed to interlocutor to provide information
- turn passed to interlocutor without soliciting information (i.e., solicits acknowledgment or continuation by the interlocutor)
- ambiguous transition point (usually because of unintelligibility or overlap)

### *Delivery of speech – prosodic and phonetic features*

- placed before stressed word
- disrupted speech
- sound lengthening
- animation

### *Placement and timing*

- [] overlap
- = fast turn take-up
- ^ fast turn continuation
- ... short untimed interval of less than 1 second
- 1.5 timed interval in half seconds

### *Transcription doubt*

- < > encloses best guess at unclear speech
- <x> uninterpreted syllable(s)

## APPENDIX 2

Below is a sample of the transcript, coded for communicative acts.

- |      |    |   |      |
|------|----|---|------|
| 1.1  | A: | have you ever taken a photo like 'that? ... | q    |
| 2.1  | C: | yeah. (C nods)                              | a    |
| 3.1  | A: | yeah~ ...                                   | re   |
| 3.2  |    | right,                                      | rf   |
| 3.3  |    | where d you reckon 'they are?               | q    |
| 4.1  | C: | seaside.                                    |      |
| 5.1  | A: | right.                                      | rf   |
| 5.2  |    | have 'you ever been to the seaside?         | q    |
| 6.1  | C: | (nods)                                      | a    |
| 7.1  | A: | who have you 'been with?                    | q    |
| 8.1  | C: | mum and dad, and my 'brother.               | a    |
| 9.1  | A: | ri:ght.                                     | rf   |
| 9.2  |    | 'where did you go? ...                      | q    |
| 10.2 | C: | llandudno.                                  | a    |
| 11.1 | A: | oh:, did you!~ ...                          | rack |
| 11.2 |    | it s 'nice there. ...                       | ev   |
| 11.3 |    | my 'brother used to live there for a while. |      |
| 11.4 |    | yea:h.                                      | rf   |
| 11.5 |    | ^ did you 'enjoy yourself?                  | q    |
| 12.1 | C: | yeah.                                       |      |
| 12.2 |    | ^ we <was> stayed in a 'hotel.              |      |
| 13.1 | A: | oh:, were you~ ...                          | rack |
| 13.2 |    | right.                                      | rf   |

## APPENDIX 3

### *Statistical methods*

Comparisons of proportions are problematic when the proportions are based on different numbers of responses for different subjects or when floor and ceiling effects occur. Using the weighted empirical logit transform makes an allowance for the fact that proportions based on a large number of observations are likely to be more accurate than those based on only a small number of observations (Cox & Snell, 1989). For each subject, where  $n$



is the number of responses and  $r$  is the number of responses in a particular category, we computed the empirical log-odds score,  $L = \log [(r + .5)/(n - r + .5)]$ , its associated variance,  $U = [(n + 1) \cdot (n + 2)]/[n \cdot (r + 1) \cdot (n - r + 1)]$ , and a weighting term,  $W$ , which is equivalent to  $1/U$ . For testing paired comparisons between conditions  $i$  and  $j$ , we computed a difference score,  $d = L_i - L_j$ , and a combined weight,  $w_{ij} = 1/(U_i + U_j)$ . We can then test the significance of  $z = \Sigma w_{ij} \cdot d / \sqrt{w_{ij}}$ .

For between-subjects comparisons between independent groups 1 and 2, we computed:

$$z = [(\Sigma w_{1j} L_{1j} / \Sigma w_{1j}) - (\Sigma w_{2j} L_{2j} / \Sigma w_{2j})] / \sqrt{[(1 / \Sigma w_{1j}) + (1 / \Sigma w_{2j})]}.$$

Another consideration is that, where responses can be categorized into mutually exclusive categories, proportions of different categories are not independent. For instance, if the proportion of cases of nonverbal response is more than .5, then the proportion of other response types must sum to less than .5. To avoid spuriously inflating the number of significant results, one can test independent proportions by testing one response category and then removing that category from consideration before testing the next. For example, if a child had a total of 25 coded responses, of which 5 were nonverbal responses and 10 were prosodic responses, we would first do tests on the nonverbal response category, computing the log-odds score based on the proportion  $5/25 = .20$  (i.e., as a proportion of *all* responses). We would then remove the nonverbal response category, so that tests of the prosodic response category would be based on the remaining categories, giving the proportion  $10/20 = .5$ .

## ACKNOWLEDGMENTS

We extend thanks to Cheryl James and Sonia Bishop for assisting with grammatical coding, transcribing, and reliability checks and to Rik Henson and Ian Nimmo-Smith for statistical advice. We also express our gratitude to the staff and pupils of the following Huddersfield primary schools, who generously assisted with the research: Crow Lane Primary, Beech Junior, Holmfirth Primary, Marsden County Junior, and Kaye's First School. This work was supported by Medical Research Council grant number G8924454N.

## REFERENCES

- Adams, C., & Bishop, D. V. M. (1989). Conversational characteristics of children with semantic-pragmatic disorder. I. Exchange structure, turntaking, repairs and cohesion. *British Journal of Disorders of Communication*, 24, 211-239.
- Brennan, S. E., & Williams, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language*, 34, 383-398.
- Chapman, R. S. (1981). Exploring children's communicative intents. In J. F. Miller (Ed.), *Assessing language production in children* (pp. 111-136). Baltimore: University Park Press.
- Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127-149). Washington: American Psychological Association.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.
- Coulthard, M. (1977). *An introduction to discourse analysis*. London: Longman.
- Cox, D. R., & Snell, E. J. (1989). *Analysis of binary data* (2nd ed.). London: Chapman and Hall.

- Doherty-Sneddon, G., & Kent, G. (1996). Visual signals and the communication abilities of children. *Journal of Child Psychology and Psychiatry*, 37, 949–959.
- Dore, J. (1977a). Children's illocutionary acts. In R. O. Freedle (Ed.), *Discourse production and comprehension* (pp. 227–244). Norwood, NJ: Ablex.
- (1977b). "Oh them sherif": A pragmatic analysis of children's responses to questions. In S. Ervin-Tripp & C. Mitchell-Kernan (Eds.), *Child discourse* (pp. 139–164). New York: Academic.
- Duncan, S., & Fiske, D. W. (1985). Structure, strategy, and research. In S. Duncan & D. W. Fiske (Eds.), *Interaction structure and strategy* (pp. 294–327). Cambridge: Cambridge University Press.
- Dunn, M. P., Dunn, M. L., Whetton, C., & Pintilie, D. (1982). *British Picture Vocabulary Scale*. Windsor: NFER Nelson.
- Fey, M. (1986). *Language intervention with young children*. Boston: College-Hill Press.
- Lewis, C., & Mitchell, P. (1994). *Children's early understanding of mind: Origins and development*. Hove: Erlbaum.
- Quirk, R., Greenbaum, S., Leech, G., & Svartik, J. (1972). *A grammar of contemporary English*. London: Longman.
- Raven, J. C., Court, J. H., & Raven, J. (1986). *Raven's Progressive Matrices and Raven's Coloured Matrices*. London: H. K. Lewis.
- Reid, J. (1995). A study of gender differences in minimal responses. *Journal of Pragmatics*, 24, 489–512.
- Semel, E. M., Wiig, E. H., & Secord, W. (1980). *Clinical Evaluation of Language Fundamentals-Revised*. San Antonio, TX: Psychological Corporation.