

Declarative question acts : two experiments on identification

Citation for published version (APA):

Beun, R. J. (1989). Declarative question acts : two experiments on identification. In *The structure of multimodal dialogue* / ed. M.M. Taylor, F. Néel, D.G. Bouwhuis (pp. 313-321). (Human factors in human technology; Vol. 4). North-Holland Publishing Company.

Document status and date:

Published: 01/01/1989

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Declarative Question Acts: Two Experiments on Identification

R.J. Beun

Introduction

In general it is assumed that questions in natural language are represented by sentences with an interrogative sentence-type¹. In natural dialogues, however, one finds many utterances² with a declarative sentence-type that clearly fulfill the function of a question. Sometimes these utterances have clear cues to indicate this function, e.g. the use of special words (1b) or, as indicated by the question mark in (1c), a rising intonation at the end of the utterance in spoken language:

- | | |
|--|-----------------|
| (1a) "Are you invited to the party?" | (interrogative) |
| (1b) "So, you are invited to the party." | (declarative) |
| (1c) "You are invited to the party?" | (declarative) |

Although we will not consider the exact contribution of the indicators to the function, one can roughly say that the word 'so' in (1b) and the combination of declarative sentence-type and rising intonation in (1c) not only seem to express the question in (1a) but also a speaker's supposition about the answer.

In many cases these clear indicators are absent and the determination of the function seems to come from contextual cues only. The recognition of questions (and of course answers) is of crucial importance for a proper continuation of the dialogue, so we would like to find special characteristics in the utterance which could indicate its question-function even in the absence of contextual cues.

We will use the theoretical framework of Bunt's **dialogue acts** (Bunt, Ch. 4). In doing so we have restricted ourselves to so-called **information dialogues** in which the participants have no other purpose than the exchange of factual information. A dialogue act is determined by the combination of its **semantic content** and **communicative function** (CF). This function stipulates the role of the semantic content in the dialogue and relates it to the speaker's knowledge and goals. The CF of the utterances in (1b) and (1c) for example is a CHECK with semantic content 'INVITED(H,party)' where H is the hearer of the utterance. In terms of knowledge and goals the preconditions of these utterances are:

- (2a) S wants to know whether INVITED(H,party)
 (2b) S suspects (or knows) that H knows whether INVITED(H,party)
 (2c) S suspects that INVITED(H,party)

In this chapter we present an empirical study of questions with a declarative sentence-type in natural dialogues. We carried out two experiments to find out whether subjects are able to recognize the CF of utterances detached from the context. The utterances were obtained in previous dialogue experiments which will not be considered here³ (see Beun, 1985). As a discourse situation we used an information exchange by telephone⁴ concerning arrival and departure times of planes and trains at Schiphol (Amsterdam airport).

The Corpus of Utterances

All utterances investigated in this chapter are **declarative question acts**, i.e., questions, but with a declarative sentence-type. Both complete and elliptical utterances are considered. Since in the case of elliptical utterances it is difficult and sometimes impossible to determine the sentence-type, only those instances with clear interrogative or imperative features were excluded from the corpus.

To determine if an utterance is a question, even in the absence of syntactic indications, the following definition is used:

- (DEF) An utterance U is a declarative question act (DQA) if⁵:
- The sentence-type of the sentence expressed by U is declarative (or if the sentence is elliptical, the sentence-type is at least non-interrogative and non-imperative).
 - The utterance U, uttered by S, is about a topic on which H is the expert.
 - S and H mutually know that H is the expert on the topic⁶.

Note that we have excluded sentences as:

- (3) "I want to know what time it is"
 (4) "Can you tell me what time it is"

The reason is that (3) is not a direct question as far as the literal interpretation is involved: in (3) the literal topic is the goal of the speaker, who is the expert on his/her own goals. (4) is excluded because of its interrogative sentence type.

In many cases where a matrix sentence contains a performative verb in the second person, past tense, the definition is satisfied⁷, e.g.:

- (5) "You said that the plane will leave at ten"
 (6) "You stated that the Germans will win"

For, in the restricted domain we use here, the addressee knows best what he has said, stated, etc. The same can be said about verbs relating to knowledge or opinion in the second person, present or past tense, e.g.:

- (7) "You mean that I have to leave at ten"
 (8) "You think that I'm joking"

Again, the addressee knows best what he means, thinks, etc. In the next sentence, however, it is not clear what the CF is without contextual knowledge:

- (9) "The train leaves at 12.00"

Uttered by an informant in the dialogue experiments the CF can be an answer, uttered by one of the subjects the CF can be a DQA⁸, even without a rising intonation at the end. In fact only 48% (37 out of 77) of the DQA's which appeared in the dialogue experiments had a rising intonation.

The definition of a DQA is largely based on contextual characteristics such as 'topic' and 'mutual knowledge'. We expected however that certain characteristics of the utterances (prosodic or linguistic) could indicate the CF without the use of context. From a linguistic angle one can think of the use of special words such as 'so', 'thus' etc.⁹. Prosodic cues could for instance be given by intonation, accentuation or declination. To find out whether these indicators contribute to the interpretation of the CF we carried out the experiments described in the next section.

The First Experiment

From the Schiphol dialogues we isolated 77 DQA's¹⁰. 47 utterances were excluded from the experiment because the 'DQA indicators' were so clear that it was feared that the subjects would only pay attention to these clear cases to distinguish the CF¹¹. These clear indicators were: 1. rising intonation at the end of the utterance (37 utterances); 2. the use of Dutch particles such as: 'he', 'toch', 'zeker' (13 utterances); and, 3. the use of special verbs as described in the previous section (6 utterances). Some utterances possessed two or more of these indicators. The remaining 30 DQA's were mixed with 24 answers, which were also isolated from the Schiphol dialogues.

The task of the participants was divided into two parts; in one part they had to indicate for each utterance whether it was a question or not (question-task), in the other part they had to say whether the utterance was an answer or not (answer-task)¹². The subjects did not know that in both cases the same utterances were used.

This first experiment was only meant to give a rough indication about the ability of subjects to recognize the CF in general, so only six Dutch subjects (of both sexes, all over 18) took part in the first experiment. They had to write down their responses on paper and they could think about their responses as long as they wanted.

Before we discuss the results we have to introduce some abbreviations. To distinguish the original CF (known from the context) from the responses given by the subjects we will always write the first one in capital letters. The responses of the subjects will be divided into 4 parts:

- (+) question: the utterance is interpreted as a question in the question-task.
 (-) question: the utterance is interpreted as not being a question in the question-task.

- (+) answer: the utterance is interpreted as an answer in the answer-task.
 (-) answer: the utterance is interpreted as not being an answer in the answer-task.

If the subjects agreed in more than 90% of the cases on a question as CF of an utterance (percentage of both (+)questions and (-)answers > 90%), the utterances were called 'Q-utterances'. Conversely if the subjects agreed on an answer in more than 90%, the utterances were called 'A-utterances'. (Note that this has nothing to do with the original CF.)

Responses are called 'inconsistent' if one subject gave the same response in the question- and answer-task.

Results

70% (199 out of 285¹³) of the ANSWERs and 68% (242 out of 351) of the DQA's were correctly identified. (Note that 50% would be purely random.) By 'correctly identified' we mean that, if the utterance was a DQA, the subjects responded: (+)question and (-)answer and vice versa if the utterance was an ANSWER. 11% of the ANSWERs and 12% of the DQA's were identified inconsistently, so there were no striking differences between the ANSWERs and DQA's.

These results do not look very spectacular. When we consider every utterance separately, however, some interesting results appear. Table 23-1 shows the number of utterances as a function of the CF and the percentage of correct responses.

Table 23-1: Number of Utterances as a Function of the Original CF and the Percentage of Correct Responses

correct responses (max=12) \ CF	>90%	>10% and <90%	<10%	Total
DQA	11	17	2	30
ANSWER	8	15	1	24
Total	19	32	3	54

From Table 23-1 it can be seen that in 19 cases subjects can almost perfectly identify the CF, in 3 cases the CF is inverted (ANSWER → Q-utterance, DQA → A-utterance) and in 32 cases the subjects could hardly identify the CF. But how do subjects identify the Q- and A-utterances? In other words, what special characteristics do these utterances have to reveal the CF?

To discover this, all Q- and A-utterances were analysed on intonation patterns but no special differences appeared between these utterances¹⁴. There seemed to be an important difference, however, in the use of special words. Almost all Q-utterances contained particles like 'en' ('and'¹⁵) at the beginning of the utterance, 'dus' ('so'), 'ook' ('also'), etc. These words were missing in the A-utterances. In Table 23-2 the occurrence of special words is shown as a function of the Q- and A-utterances. 'Uncertainty' was expressed by words as 'ongeveer', 'pak em beet', etc. ('about', 'roughly', etc.) or the pause-particle 'uh'.

Table 23-2: Occurrence of indicators as a function of Q- and A-utterances

indicator \ utterance	'dus'	'en'	'ook'	uncertainty
Q	5	6	4	4
A	0	0	0	2

Note that Table 23-2 does not imply that the responses of the subjects were correct, only that they agreed in more than 90% on the interpretation of the CF.

These indicators can also be used in answers, so we had the impression that prosodic aspects also contributed to the interpretation of the CF. In the second experiment, described in the next section, these linguistic indicators were removed from the utterances.

The Second Experiment

In the second experiment we used the following utterances (total 26) from the previous one:

1. all Q-utterances
2. all A-utterances
3. utterances with the following characteristics:
 - a. The appearance of the word 'en' ('and') and 'oh' at the beginning of the utterance.
 - b. The appearance of the pragmatic particles 'dus' ('so'), 'ook' ('also') and 'niet' ('not').
 - c. The appearance of the pause particle 'uh'.
 - d. Mistakes followed by self-repair.
 - e. Words which indicate uncertainty.

Next, these characteristics were removed from the utterances¹⁶ and the 26 original utterances were mixed with the edited ones. Where the utterances had two or more of these characteristics they were removed in arbitrary order. For example, an utterance with three

characteristics would give 6 (=3!) edited versions; we did not include all combinations but only an arbitrary selection, e.g.:

- (10) "And uh, these are all coming/ going to Munich" (original)
 (10a) "And these are all coming/ going to Munich" (edited; 3c)
 (10b) "These are all coming/ going to Munich" (edited; 3a)
 (10c) "These are all going to Munich" (edited; 3d)

Almost the same conditions were applied as in the first experiment; only the number of utterances (total number was 72 of which 46 were edited) differed. In the second experiment 20 subjects took part.

Results

Figure 23-1 shows the difference in (+) answers to utterances with (white rectangle) and without (shaded rectangle) special indicators. Note the considerable variation in the number of utterances per indicator (= number of responses (n) divided by 20). Figure 23-2 shows the same for (+)questions.

In Figure 23-1 we can see for instance that 35% of the utterances with indicator 'en' ('and') at the beginning were interpreted as an answer. The same utterances without 'en' are interpreted as an answer in 55% of the cases. If we turn to Figure 23-2 we can see that in 64% of the cases with 'en' the utterances are interpreted as a question and 51% when the 'en' is removed.

Note that Figures 23-1 and 23-2 do not have to be complements, because responses to the same utterance by the same subject can be (-)question and (-)answer (12%, or 166 out of 1440) or (+)question and (+)answer (10%, or 149 out of 1440).

Discussion

From Table 23-1 it can be seen that the subjects were able to recognize the CF in 35% (19 out of 54) of the cases; in 4% (3 out of 54) the function was inverted. Table 23-2 shows that important indicators to recognize the question function can be the word 'en' at the beginning of the utterance, and the words 'ook' and 'dus'. These words were missing in the utterances which were recognized as answers. When these words were removed from the Q-utterances they never shifted to A-utterances, so prosodic characteristics must be involved in the interpretation.

An important shortcoming in the analysis is that the removal of certain words has the consequence of cutting out certain prosodic aspects as well. We expect to solve this problem in a following experiment in which the utterances will be presented on a screen.

We will now consider some of the indicators separately.

a. 'En' at the beginning of the utterance:

In Figure 23-1, we can see that the removal of the word 'en' has important consequences for the interpretation of the CF. When the 'en' is not removed the subjects prefer the 'question' interpretation (35% (+)answer, 64% (+)question). When the 'en' is cut out

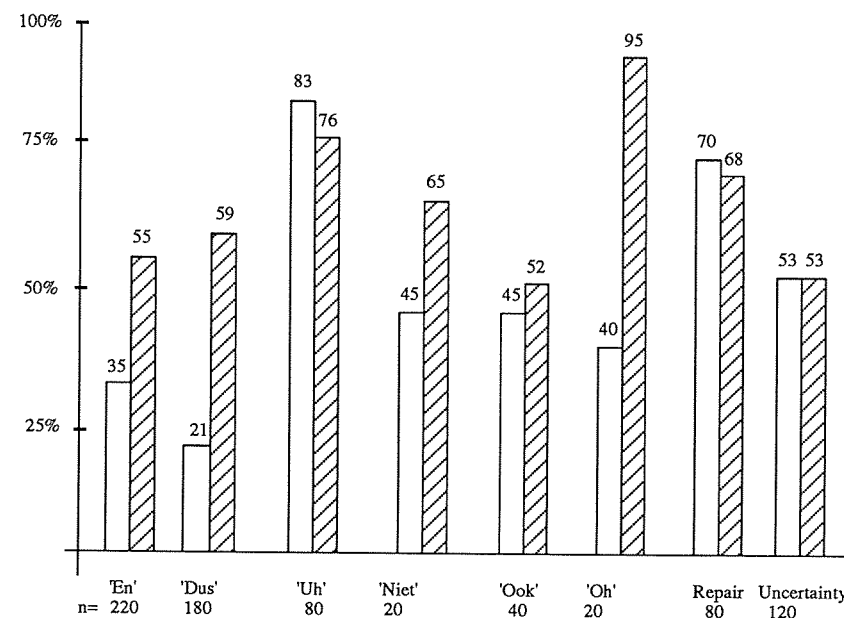


Figure 23-1: The shift of (+)answers as a function of the presence (white rectangle) and the absence (shaded rectangle) of utterance characteristics

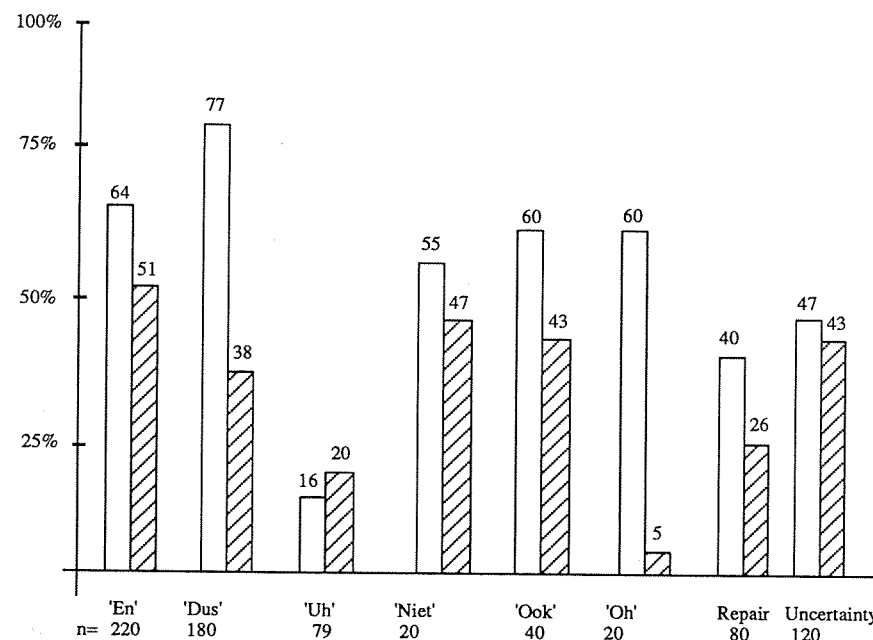


Figure 23-2: The shift of (+)questions as a function of the presence (white rectangle) and the absence (shaded rectangle) of utterance characteristics

the responses shift to an 'answer' interpretation (55% (+)answer, 51% (+)question). Note that the (+)answers and (+)questions are independent because they were given in different tasks. Here it is to be expected that the influence of prosodic aspects will be very small because the 'en' was never accentuated and forms only a small part of the speech signal of the complete utterance.

b. 'Dus':

In this case the shift from question-responses (21% (+)answer, 77% (+)question), when 'dus' is included, to answer-responses (59% (+)answer, 38% (+)question), when 'dus' is removed, is dramatic. For the same reasons as before it can be expected that the prosodic effects will be very small and almost everything will come from the meaning of the word itself.

c. 'Oh':

Again we see a dramatic shift from question to answer responses. Note however that in this case we had only one utterance (number of responses=20) and the CF of the original utterance was an ANSWER. The shift can be explained by the fact that in Dutch the 'oh' at the beginning is often used to express surprise about the content of a previous statement and in many cases is followed by an utterance which asks for an explanation. In this case the 'oh' was uttered because the subject in the dialogue experiment had not expected the previous question from the informant, because he did not pay attention to her. So, the particle had nothing to do with the content of the question.

d. 'Ook' and 'niet':

In this case the shift from question to answer responses, when these words are removed, is less dramatic than in the previous cases. Semantically these words seem to contribute nothing to the CF. An explanation for the shift can possibly be found in the prosodic properties of the words. Both words were accentuated and the removal of the accents could influence the responses of the subjects.

The insertion of repair sequences, the expression of uncertainty and the use of the pause particle 'uh' did not change the responses significantly, so these indicators do not seem to contribute to the interpretation of the CF.

In conclusion it can be said that many DQA's have no rising intonation at the end (in the experimental dialogues only 48%). In some cases subjects can distinguish ANSWERS and DQA's very well without any contextual knowledge. Indicators can be provided by certain pragmatic particles such as 'en' and 'dus' or prosodic characteristics of the utterance. On the other hand the CF of 32 utterances was not identified, so in these cases all cues have to be provided by the context.

Notes

Acknowledgements: I would like to thank Jan-Roelof de Pijper, Kees van Deemter and Frens Dols for their helpful comments on the first draft of this chapter. Special thanks to Harry Bunt and Don Bouwhuis for their co-operation in all the thinking and all the other things they did for me.

- 1 Following Gazdar (1981) I will regard sentence-type as a purely syntactic characteristic of sentences.
- 2 Here an utterance should be taken in the sense of a sentence paired with a context (see e.g. Levinson, 1985 pp.18-19 or Gazdar, 1981).
- 3 A short impression will do no harm, I think. In the experiment twenty four Dutch speaking subjects had to determine by telephone the departure and arrival times of airplanes and trains from an informant at Schiphol airport. The informant was simulated by a well-trained person. Each subject read eight sets of instructions to tell them what kind of information had to be obtained.
- 4 Here we avoid non-verbal aspects of communication which are difficult to measure such as eye-contact, body-movement, etc., so everything can be collected from the original speech signal.
- 5 The interpretation of sentences defined by (DEF) as a question also agrees with the maxim of quantity, for if the addressee knows more about the subject than the speaker (and this is mutually known) the utterance would be superfluous in the case of a statement.
- 6 S and H mutually know that $p = (q)$; S and H know that p and that q'; See also Clark & Marshall (1981).
- 7 In special circumstances one can use these forms to focus certain aspects of an addressee's knowledge or actions to introduce for instance counterexamples, e.g.: "You stated that indirect speech acts can be explained by means of felicity conditions, but how then do you explain the utterance "It's cold inside" as a request to close the window." We did not encounter any of these examples in our restricted domain.
- 8 Strictly speaking a DQA is not a communicative function because it contains syntactic information; this characterization is only maintained for the sake of simplicity.
- 9 In Dutch these particles are widely used in natural dialogues, e.g. 'wel', 'toch', 'dus', 'en', 'ook', etc.
- 10 We had no problems with interruptions because the voices of the informant and the subject were recorded on different tracks. (TEAC A3440 4-channel simul-sync)
- 11 At this moment this is just a supposition and still has to be tested.
- 12 Actually the subjects also had to answer whether the utterance was an 'inform' (Dutch: 'mededeling'). Until now we did not investigate these results, so they are omitted.
- 13 In some cases subjects did not make any decision at all, so the total numbers of ANSWERS and DQA's differ from $12 \times 24 = 288$ and $12 \times 30 = 360$ respectively.
- 14 We have one DQA, however, where all the subjects responded correctly, with a very horizontal declination line. This can be an indication, but because of the lack of other examples any conclusion of this sort seems to be premature.
- 15 The English translations are only meant as a vague indication about the meaning.
- 16 First the utterances were put in the VAX computer and next the special words were removed with the aid of the IPO speech software. This could be done with an accuracy of 25 milliseconds. Some of the utterances (but only very few) sounded a little unnatural; none of the subjects, however, noticed any difference between the original and the edited versions.