"And then what was the question again?"
- TEXT AND TALK IN STANDARDISED INTERVIEWS

Hanne-Pernille Stax, University of Southern Denmark
Institute of Language & Communication, Campusvej 55, DK-5230 Odense M, Denmark

Key Words: Interviewer-Respondent Interaction, Comprehension of Questions, Question-Response Contiguity, Standardisation / (Standardization).

This study is part of a Ph.D.-project in Language and Communication about standardised questionnaire interviews as a form of discourse between oral and written communication. This paper will focus on how (or if) respondents discriminate between pre-scripted and locally generated talk in standardised interviews.

The project is inspired by conversation analysis (CA), and so is the perspective of this paper's approach to the theme of the AAPOR 2000 panel: 'Improving Respondent Comprehension of Questions'. In the following I will present 1: Research question, 2: Research design, 3: Preliminary results and 4: Possible implications

1. RESEARCH QUESTION

When questionnaires are enacted into interviews it is of crucial importance that the recorded answers are both valid and comparable: That is, 1) Answers must be produced BY the respondents, and 2) Answers must be produced AS responses to the questions as they are formulated in the pre-scripted questionnaire.

In short, the recorded answers should represent the respondents reply to the questionnaire text. And they should do so even if interviewer and respondent engage in non-scripted interaction about a question before a recordable answer is arrived at (Fowler & Mangione, 1990).

Fig. 1

I:	**Question**	*(enacted from prescripted text)*
(I/+ R:	*Talk*	*(non-script))*
R:	**Response**	*(recorded in prescripted text)*

In survey research the questionnaire TEXT is often seen as THE 'medium of communication' between researcher and respondent. And questionnaire pre-tests are often carried out to identify and solve problems in the wording of this TEXT (DeMaio et al, 1993; Oksenberg et al, 1991; Tourangeau, 1984).

Thus the interviewer is conceived to be a mere mediator, through which the words of the pre-scripted questionnaire text are presented to the respondent, and through which the words of the respondent are transformed and recorded as written responses in the questionnaire.

However, in actual standardised interviews not all of the interviewers talk is 'enacted', that is read aloud, from the pre-scripted questionnaire text. And when respondents and interviewers sometimes engage in lengthy interactions before arriving at a recordable response, it is not always evident whether this response is responding to the pre-scripted questionnaire question or to the immediately preceding talk.

Recent CA studies of the interaction in standardised interviews have focused on the role of the interviewer as more than a mere mediator between researcher and respondent (Houtkoop-Steenstra, 1995, 2000; Schaeffer, 1991; Schaeffer & Maynard 1996; Suchman & Jordan, 1990). And these studies have shown that interactional tensions can emerge when questionnaire texts are enacted into talk-in-interview-interaction.

CA studies of verbal interaction do not focus on WHAT participants understand, but on HOW they understand, and HOW they display their understandings to each other (Schegloff 1984, 1992; Pomerantz, 1988). Instead of asking the classical pre-test question 'What does it do to a particular text to be transformed into talk-in-interaction?' CA studies of interview interaction ask: 'What does it do to talk-in-interaction to be based upon pre-scripted texts?' Or in, more operational terms, 'how do participants in pre-scripted interactions orient to – or deal with - this scriptedness'?

The main focus of this paper is not on HOW scripted talk deviates from ordinary, non-scripted interaction, but on how participants COMPREHEND and ORIENT to the special interactional framework of scripted interactions. The central research question is: 'Do respondents in standardised interviews discriminate between enacted questionnaire text and locally generated talk?' Thus in this paper, the panel theme "respondent comprehension of questions" is approached with a focus on respondent comprehension of questions AS pre-scripted, non-negotiable items.

2. RESEARCH DESIGN

The data for this study is a set of 11 authentic, audio-recorded, standardised telephone interviews from a major Danish attitude survey. They were randomly selected among interviews to be conducted by three

experienced interviewers at the Danish National Institute for Social Research. Each interview consists of app. 80 question items and their duration ranges from 15 to 30 minutes. These audio-recorded interviews are transcribed for sequential microanalysis according to a modified version of the transcription conventions developed by Gail Jefferson (Jefferson, 1984).

Applying sequential micro-analysis in order to answer the question: "Do respondents discriminate between enacted questionnaire TEXT and locally produced TALK?" I will present two ways in which respondents display their orientation to the special interactional framework of the standardised interview :
1) Respondent's request for repetition of a pre-scripted question after some non-scripted interaction. And
2) Respondent's use of a pre-scripted response scale while the local interaction makes another type of response more sequentially fitted.

2.1. DISPLAYING UNDERSTANDING

But before proceeding to this, a brief note on the concept of 'display of understanding' is in place: A basic assumption of sequential micro-analysis of talk-in-interaction is the view on human conduct. As formulated by Pomerantz & Fehr (1997: 69): "conduct is produced and understood as responsive to the immediate, local contingencies of interaction." Consequently the actual production, placement and packaging of each contiguous turn in an interaction, displays a lot about how the speaker of this turn has understood the immediately preceding turns of talk.

And as the interaction unfolds, each turn will not only display how the present speaker has understood the preceding turn. In each turn the speaker will also display whether he accepts the understandings of prior talk, which were displayed by the other speaker(s) in the preceding turns.

Fig. 2. Translated excerpt from interview no. 21

	A:	.hh and we begin by observing that eh (.) you are male who is born in nineteen hundred and (0.6)	
→	B:	seventy eight	(answer to ↑)
	A:	seventy eight yes	(acknowledgement
		and how far did you...	+ new question)

In Figure 2 speaker B produces a turn, which relates to the preceding turn, as an answer to a question. This turn displays (to A and to us) that speaker B has understood the preceding turn as a question, addressed at speaker B and projecting an answer in the following turn.

Within the third turn speaker A then proceeds to a new question, thus displaying that speaker A is treating

the preceding turn – the answer – as a relevant and adequate response to A's preceding turn. Thus in the third turn speaker A displays (to speaker B and to us) that speaker B's displayed understanding of the first turn - as a question - is acceptable.

In this paper I will focus on whether respondents comprehend - and display comprehension of - enacted survey questions AS pre-scripted, non-negotiable items.

3.1. REQUESTS FOR QUESTION REPETITION

A first example of respondent's displayed orientation to the questionnaire text is cited in the paper title. The phrase "And then what was the question again?" is a translation from a sequence where a respondent requests a repetition of a question after some non-scripted 'clarification talk' about words or implications of a question. Figure 3 is a diagram of the general format of sequences where some non-scripted clarification talk after the first production of a question is followed by a request for a repetition of the question.

Fig. 3: Request for repetition of question

	I:	Question item enacted	(script)
	R/+I:	Clarification talk	(non-script)
→	R:	Request for repetition:	(non-script)
		{ 'og hvad var så sporgsmålet?'}	
		'And then what was the question again?'	
⌈	I:	Question item enacted	(script)
⌊	R:	Response	(script)

I will point to two somewhat conflicting observations on this example. In incorporating a "then" {"så"} in the request for repetition (→ Fig. 3), the respondent displays that the intervening talk HAS broken new ground. The "then" marks that the question to be re-produced will not be treated simply as a re-positioned recycle of a prior turn. In pointing to the immediately preceding non-scripted talk as consequential for the next move, the respondent is not simply back-stepping the preceding talk. She is markedly requesting the question to be repeated as a 'next-positioned' question (Jefferson: 1980). That is, as positioned AFTER some non-scripted interaction, which will inform how the repeated question will be heard and responded to.

On the other hand in requesting a repetition of 'the question' (rather than say: 'what do you want to know?' or "and then what about it?'') the respondent also displays a comprehension of the question as a non-negotiable text. She displays that she is treating the intervening talk as an insertion into a pre-planned exchange of scripted questions and responses and not as a negotiation towards a modified version of the question. In initiating a resumption of the exchange of

pre-scripted questions and responses, she displays that she orients to the task of producing a response to the 'question proper' – and not to any other locally negotiated version. Similar interactional work is done in other 'post clarification' requests for repetition. E.g: "ehh then can't I just hear the question one more time?"

In these sequences a model relationship ([in Fig. 3) between response and question proper is accomplished regarding substance as well as sequentiality, re-establishing a planned contiguity between pre-scripted question and response. A contiguity, which is challenged, when the two are separated by non-scripted, intervening talk.

I am aware of the ambiguity of this analysis. On the one hand I claim that the respondent is displaying that she orients to the non-scripted interaction as consequential for the sequence in which the response will be produced. On the other hand I claim that she is displaying an orientation to the status of the questionnaire question as pre-scripted and non-negotiable. But this is exactly the point I want to make.

The pre-scripted interview requires that the respondent will distinguish between pre-scripted questionnaire text and locally produced talk. In sequences where a respondent produces a request for repetition of a question after some non-scripted clarification, the respondent is displaying an awareness of this implication of standardisation.

3.1.1. OCCURRENCES

Since answers in standardised interviews must respond to the questionnaire questions, one could expect that requests for question repetition would recurrently be found after non-scripted clarification talk. But in fact very few respondents make this type of display of 'doing being a competent respondent.' Thus very few respondents display an orientation to non-scripted clarification talk as 'time out' from the pre-planned exchange of pre-scripted questions and responses.

About one third of the question-response sequences in the material for this study take more than five new speaker turns to complete. But out of these 260 sequences, there are only three instances where a respondent requests a repetition of a question after some non-scripted interaction. And this might indicate that the essential contiguity between questionnaire text and recorded response is at risk.

3.2. FORMALLY/LOCALLY FITTED RESPONSE

A more covert display of orientation to the questionnaire TEXT appears in respondent's production of formally fitted responses where another type of response would be more fitted to the local interaction.

In standardised interviews not only the questions but also the responses are in fact pre-scripted. And some respondents display comprehension of this condition, e.g. by responding with "I agree" or "Disagree" rather than "Yes" or "No" to a statement like: "Also racist groups should have the right to assemble and speak their case"

In the questionnaire behind the data for this study there are 2 series of app. 20 question items, which are formulated as statements with a four-point 'agree-disagree' rating scale. They are interesting for a number of reasons:

First: With 40 items using the same rating scale, it will be possible to see whether any development - or learning of the pre-scripted format - is going on in the interviews.

Second: Being formulated as statements these questionnaire items make OTHER response types than the format fitted "agree" or "disagree" sequentially relevant – like "yes" or "no", a modification, a counter-statement or another statement.

And last: When a rating scale is presented at the beginning of a series of statements there is a planned distance between the presentation of the rating task and the respondent's production of a response.

This distance means that in order to produce formally relevant responses to questionnaire statements respondents must repeatedly switch from the local context of each enacted statement to the distant context of the rating scale.

Fig. 4: Formally / Locally Fitted Responses (a-d)

Scale – response distance

I: "Would you say that you agree or disagree with the following statements?"

I: "Also racist groups should have the right to assemble and speak their case"

R: a: "I agree"
b: "Yes"
c: "Well they shouldn't become violent"
d: "Yes, I agree"

I: Statement 2
R: Response
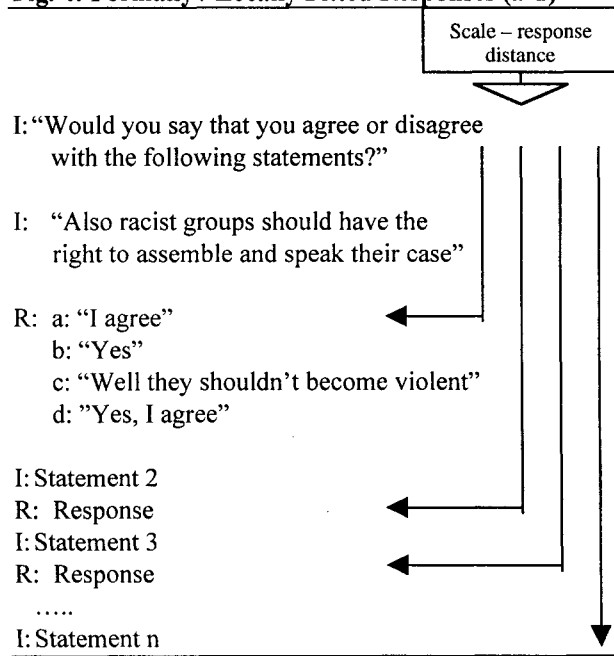I: Statement 3
R: Response
.....
I: Statement n

Figure 4 is a diagram of interviewer–respondent interaction involved in completing a series of questionnaire statements, which are preceded by a general introduction, displaying (some of) the response options on the pre-scripted rating scale.

A possible place to detect respondent's orientation to the fact, that both questions and response options are pre-scripted and non-negotiable, are 'first responses' produced immediately after each statement.

3 out of 4 of these responses display an orientation to the pre-scripted scale by including one of the formally relevant response words "agree" or "disagree".

But 1 out of 5 responses include "yes" or "no" instead. And about half as many include neither "agree" or "disagree" nor "yes" or "no".

Thus the formal preference for responding with "agree" or "disagree", is - recurrently - overruled by what seems to be a structural preference for producing a locally fitted response – that is "yes" or "no".

Now one out of every five responses might not make a strong case However, the argument for a presence of a structural preference for responding with "yes" or "no" is substantiated by another finding: In responses, which include "agree" or "disagree", these words are recurrently preceded by "yes" or "no".

In such two-component responses like "yes - I agree", or "no – partly disagree", the first component is locally fitted to the immediately preceding talk and the second is formally fitted to the pre-scripted text.

Here an observation from conversation analysis is of relevance: When two questions are produced within the same turn, then next speaker tends to answer the last question first, thereby preserving maximum contiguity between question and response (Sacks, 1987).

When respondents in standardised interviews are first presented with an overall pre-scripted rating scale, and then a particular statement, they are actually presented with two response tasks: 1) to produce a relevant next utterance to a statement and 2) to answer the implicit question: "Do you agree or disagree?"

The two-component responses display orientation to both tasks. And in placing "yes" or "no" BEFORE the formatted response options, the respondents display an orientation to preserving contiguity where it can be preserved – that is on the level of local talk.

This means that in order to produce formally fitted responses (here: agree / disagree), respondents are required to overrule a built in preference for contiguity on the local level of the interaction. This does not make the response task easier.

3.2.1. DISTRIBUTION

In the data at hand, most responses are formally fitted. Thus most respondents seem to have a good comprehension of the implications of standardisation. However the number of format fitted answers are quite unequally distributed among the respondents, varying from 100% - to zero.

In some interviews (3 out of 11) the vast majority of statements are responded to in a format fitted way. If some of the first statements generate a "yes" or "no" (and they recurrently do), then the interviewers probe the response by enacting relevant options from the scale. And after the first few questions these respondents start adhering to the scale options.

But in some interviews (3 out of 11) this does not seem to work, and the respondents continue to produce unformatted answers throughout the series. In these interviews the interviewer's probes are not treated as requests for revised responses to the enacted questionnaire questions. They are treated as next-positioned 'new' questions.

Fig. 5

I: People should only be licensed to get help from the state, if they are entirely unable to support them selves

R: (.h 1.6) hm yes that=that is obvious, then of course they must have help hhhh of course they must
   (15 lines left out)

I: So eh do you agree or disagree with that?

R: I agree with that they must have help

Translation, simplified

These respondents display no orientation to the frame question: "Do you agree or disagree", NOR to fact that only responses from the pre-scripted scale are acceptable. Thus the relationship between the produced answer and the question proper is diffused and unclear as in the example in figure 5.

Summing up: In agree-disagree items there are two competing preferences regarding response format: A formal preference for a format-fitted response containing "agree" or "disagree", and a structural preference for a locally fitted "yes" or "no".

When respondents respond to most statements in the formally preferred format, they display an orientation to the fact, that even their own talk is pre-scripted. Some respondents need some instruction from the interviewer's probes to overrule the preference for producing locally fitted responses. These respondents learn the format in the course of the interview.

However to some respondents this implicit instruction is inadequate. They continue to produce locally fitted answers to each of the statements in the series, and show no orientation to the fact that the format disallows recording of these answers.

4: POSSIBLE IMPLICATIONS

The initial question in this study was 'Do respondents discriminate between text and talk in survey interviews?' And the preliminary answer will have to be 'Some respondents do and others don't.'

Whether this finding is problematic depends on the answer to another question: 'What are the implications of an imperfect discrimination between text and talk in standardised interviews?'

Some respondents display no comprehension of the special limitations of scripted interaction. This may result in lengthy interactions about response format, which are time consuming (and thus costly), and which might produce answers with a diffused relationship to the questionnaire questions.

Moreover respondents might not comprehend that the interviewers do discriminate between formatted responses elicited through probing and comments or modifications produced in immediate response to the question proper. That is: Some respondents might well believe that both their (format fitted) responses to probes AND other qualifying remarks will be recorded. Thus the recorded answers might not reflect the respondents positions.

In standardised interviews some of the talk is scripted and planned ahead of the interview, some is not. Only the proportion of the actual talk, which adheres to the plan, is taken into account when the data is used. Respondent's comprehension of this implication of standardisation is essential to the validity and reliability of the data.

When respondents do not orient to the pre-scriptedness of the interaction, then the interviewers have but ONE 'legal' resource at hand: To produce probes, which only implicitly display the desired response format. In my material this is not an adequate resource to solve the problem of orientation when respondents do not know the game.

Standardised interviewing presupposes that respondents comprehend the implications of standardisation (Clark & Schober, 1992) - but this assumption does not always seem to be realistic. Thus additional, conversational recourses for interviewers to make the text 'visible' to all respondents must be developed, or - at least - allowed.

5: REFERENCES

Clark, H. H. & M. F. Schober, 1992: 'Asking Questions and Influencing Answers.' In: Tanur, J. M. (ed.): **Questions about Questions**. New York: Russel Sage Foundation.

DeMaio, T. N. et al, 1993: Protocol for Pretesting Demographic Surveys at the Census Bureau. Census Bureau report, June 28, 1993, Washington, D. C.

Fowler, F. J. & T. W Mangione, 1990: **Standardized Interviewing: Minimizing Interviewer-Related Error**. Newbury Park, CA: Sage Publications

Houtkoop-Steenstra, H. 1995: 'Meeting Both Ends' In: ten Have, P. & G. Psathas (eds.): **Situated Order**. Washington, D. C: University Press of America.

--- 2000: **Interaction and the standardized survey interview; The living questionnaire**. Cambridge University Press

Jefferson, G. 1980: 'The Abominable *Ne?* In: Schröder, P. & H Steger: **Dialogforschung**. Düsseldorf: Pädagogisher Verlag Schwann.

--- 1984: **Caricature versus detail**. Tilburg Papers on Language and Literature No. 231, University of Tilburg, NL.

Oksenberg, L. et al 1991: 'New strategies for pretesting survey questions' In: **Journal of Official Statistics, 7**

Pomerantz, A. 1988: 'Offering a Candidate Answer.' In: **Communication Monographs**, 55: 360-73.

Pomerantz, A. & B. J. Fehr, 1997: 'Conversation analysis: An approach to the study of social action as sense making practices.' In: Van Dijk, T.A. (ed.): **Discourse as social interaction**. London: Sage Publications.

Sacks, H. 1987: ''On the preference for agreement and contiguity in sequences in conversation.' In: Button, G. & J. R. E. (eds.): **Talk and Social Organisation**. Clevedon: Multilingual Matters.

Schaeffer, N. C. 1991: 'Conversation with a purpose – or conversation?' In: Beimer et al (eds.): **Measurement Errors in Surveys**. John Wiley & Son

Schaeffer, N. C. & D. W. Maynard, 1996: 'From Paradigm to Prototype and Back Again'. In: Schwarz, N. & S. Sudman (eds.): **Answering Questions**. Jossey-Bass Publishers, San Francisco

Schegloff, A. E. 1984: 'On some questions and ambiguities in conversation.' In: J. M. Atkinson & J. Heritage (eds.): **Structures of Social Action**. Cambridge: Cambridge University Press.

--- 1992: 'Repair after Next Turn: The last structurally provided defence of intersubjectivity in conversation.' In: **American Journal of Sociology, 97, 5**

Schober, M. & F. Conrad, 1997: 'Does conversational interviewing reduce survey measurement error?' In: **Public Opinion Quarterly**. Vol. 61

Suchman, L. (1987): **Plans and situated Action, The problem of human-machine interaction**. Cambridge University Press, Cambridge.

Suchman, L & B. Jordan, 1990: 'Interactional Troubles in Face-to-Face Survey Interviews.' In: **Journal of the American Statistical Association**. V 85, No 409

Tourangeau, R. 1984: 'Cognitive science and survey methods' In: Jabine et al: **Cognitive aspects of survey methodology**. Washington, D. C: National Academy Press