Self-addressed questions in disfluencies

Jonathan Ginzburg[1], Raquel Fernández[2] & David Schlangen[3]

[1] Laboratoire de Linguistique Formelle (LLF) and CLILLAC-ARP and LabEx-EFL,
Université Paris-Diderot, Sorbonne Paris Cité
[2] Institute for Logic, Language & Computation, University of Amsterdam
[3] Faculty of Linguistics and Literary Studies, Bielefeld University

Abstract

The paper considers self-addressed queries – queries speakers address to themselves in the aftermath of a filled pause. We study their distribution in the BNC and show that such queries show signs of sensitivity to the syntactic/semantic type of the sub-utterance they follow. We offer a formal model that explains the coherence of such queries.

1. Introduction

How to characterize the context associated with hesitations? For production, [3] claimed that fillers like 'uh' and 'um' should be treated as words with different distributions ('uh' more for short pauses,' um' for long pauses) and with discourse functions intended by the speaker. ([10] proposes a related account for Swedish glottalized filled pauses. Clark and Fox-Tree's hypothesis has more recently been strongly disputed for distributional differences [12] as well as with respect to speakers' intentions [4].

In this paper we consider a phenomenon that occurs in the aftermath of a filled pause, namely self-addressed queries, exemplified in ((1)):

(1) a. Carol 133 Well it's (pause) it's (pause) er (pause) what's his name? Bernard Matthews' turkey roast. (BNC, KBJ)

b. They're pretty ... um, how can I describe the Finns? They're quite an unusual crowd actually.

http://www. guardian.co.uk/sport/2010/sep/10/small-talk-steve-backley-interview

The question we investigate is whether such queries are essentially reflexive or show signs of sensitivity to the (the syntactic/semantic type of) the sub-utterance they follow.

Section 2 describes a corpus study we ran on the BNC to investigate this issue. The study demonstrates clearly a strong effect, with distinct distributions clustering around a small number of triggering contexts. After brief discussion of the results in section 3, section 4 provides a formal model in which we analyze the coherence of such moves, as part of an account of what we call *forwards-looking disfluences* – disfluencies where the moment of interruption is followed by a completion of the utterance which is delayed by a filled or unfilled pause (hesitation) or a repetition of a previously uttered part of the utterance (repetitions).

Conclusions and further work are provided in section 5.

2. Corpus study

We ran a corpus study on the BNC, using the search engine SCoRE ([14]) to search for all self-addressed queries. We searched using the pattern 'noun preceding 'er' or 'erm' preceding a wh word, adjacent to a verb.'. This yielded 692 hits, from this we manually selected all self-addressed queries, resulting in a corpus of 83 queries.

Representative examples are in (2) and the distribution is summarized in Table 1. Tables 2–6 provide a detailed summary of queries found, relative to triggering

(2) a. (**anticipating an N'**:) on top of the erm (pause) what do you call it?

b. (**anticipating a locative NP**:) No, we went out on Sat , er Sunday to erm (pause) where did we go?

c. (**anticipating an NP complement**:) He can't get any money se (pause) so so he can't get erm (pause) what do you call it?

d. (**anticipating a person–denoting NP**:) But you see somebody I think it was erm what's his name?

e. (**anticipating a person–denoting NP**: with erm, who was it who went bust?

f. (**anticipating a predicative phrase**: she's erm (pause) what is she, Indian or something?

Table 1: *Distribution of Self addressed questions in disfluencies in the British National Corpus*

categorial context	questions found
pre NP; prer _or verb _or NP and _	42
det _	20
locative prep _	12
be _	5
say _	4
Total self addressed questions	**83**

Table 2: *Distribution of Self addressed questions in pre NP context*

what's his/her name?	19
what do they/you call him/her/it?	13
who was it/the woman?	3
what's the other one?	3
what did you/I say?	2
what did it mention?	2
Total	**42**

Table 3: *Distribution of Self addressed questions in post Det context*

what do/did they/you call it/that/them?	14
what's it called?	2
what is it?	3
what am I looking for?	1
Total	**20**

Table 4: *Distribution of Self addressed questions in post Loc Prep context*

where is it?	3
where do they call that?	2
what's the name of the street/address?	2
what do they call X?	2
where do we go?	1
where did it say now?	1
what is it?	1
Total	**12**

Table 5: *Distribution of Self addressed questions in post-copular context.*

what is she/it?	3
what's the word I want?	1
what do you call it?	1
Total	**5**

Table 6: Distribution of Self addressed questions in post 'say' context

what did X say?	3
where did I get the number?	1
Total	**4**

3. Discussion

Table 1 indicates that self-addressed queries occur in a highly restricted set of contexts, above all where an NP is anticipated and after 'the'. Moreover, the distribution of such queries across these contexts varies manifestly: the anticipated NP contexts involve predominantly a search for a name or for how the person/thing is called with some 'who'-questions as well, whereas the post 'the' contexts only allow 'what' questions, predominantly of the form 'what does X call Y'; anticipated location NP contexts predominantly involve 'where' questions. The final two classes identified are somewhat smaller, so generalizations there are less robust – nonetheless, the anticipated predicative phrase and post 'say' context involve seem to involve quite distinct distributions from the other classes mentioned above.

4. Forward Looking Dysflencies in a dialogue model

4.1. Dialogue GameBoards

We start by providing background on the dialogue framework we use here, namely KoS (see e.g. [9, 8]). On the approach developed in KoS, there is actually no single context – instead of a single context, analysis is formulated at a level of information states, one per conversational participant. The dialogue gameboard represents information that arises from publicized interactions. Its structure is given in (3) – the *spkr,addr* fields allow one to track turn ownership, *Facts* represents conversationally shared assumptions, *Pending* and *Moves* represent respectively moves that are in the process of/have been grounded, *QUD* tracks the questions currently under discussion, though not simply questions qua semantic objects, but pairs of entities which we call *InfoStrucs*: a question and an antecedent sub-utterance.[1] This latter entity provides a partial specification of the focal (sub)utterance, and hence it is dubbed the *focus establishing constituent* (FEC) (cf. *parallel element* in higher order unification-based approaches to ellipsis resolution e.g. [6].)[2]

$$(3) \quad \text{DGBType} =_{def} \begin{bmatrix} \text{spkr: Ind} \\ \text{addr: Ind} \\ \text{utt-time : Time} \\ \text{c-utt : addressing(spkr,addr,utt-time)} \\ \text{Facts : Set(Proposition)} \\ \text{Pending : list(locutionary Proposition)} \\ \text{Moves : list(locutionary Proposition)} \\ \text{QUD : poset(Infostruc)} \end{bmatrix}$$

The basic units of change are mappings between dialogue gameboards that specify how one gameboard configuration can be modified into another on the basis of dialogue moves. We call a mapping between DGB types a *conversational rule*. The types specifying its domain and its range we dub, respectively, the *preconditions* and the *effects*, both of which are supertypes of DGBType.

Examples of such rules, needed to analyze querying and assertion interaction are given in (4). Rule (4-a) says that given a question q and ASK(A,B,q) being the LatestMove, one can update QUD with q as QUD–maximal. QSPEC is what characterizes the contextual background of reactive queries and assertions. (4-b) says that if q is QUD–maximal, then subsequent to this either conversational participant may make a move constrained to be q–specific (i.e. either About or Influencing q).[3]

$$(4) \quad a. \quad \text{Ask QUD–incrementation}$$
$$\begin{bmatrix} \text{pre} & : & \begin{bmatrix} \text{I} & : & \text{InfoStruc} \\ \text{LatestMove} = \\ \text{Ask(spkr,addr,I.q)} & : & \text{IllocProp} \end{bmatrix} \\ \text{effects} & : & \begin{bmatrix} \text{qud} = \langle \text{I.q,pre.qud} \rangle & : & \text{poset(InfoStruc)} \end{bmatrix} \end{bmatrix}$$

---

[1] Extensive motivation for this can be found in [5, 8], based primarily on semantic and syntactic parallelism in non-sentential utterances such as short answers, sluicing, and various other fragments.

[2] Thus, the FEC in the QUD associated with a wh-query will be the wh-phrase utterance, the FEC in the QUD emerging from a quantificational utterance will be the QNP utterance, whereas the FEC in a QUD accommodated in a clarification context will be the sub-utterance under clarification.

[3] We notate the underspecification of the turn holder as 'TurnUnderspec', an abbreviation for the following specification which gets unified together with the rest of the rule:

$$\begin{bmatrix} \text{PrevAud} = \{ \text{pre.spkr,pre.addr} \} & : & \text{Set(Ind)} \\ \text{spkr} & : & \text{Ind} \\ \text{c1} & : & \text{member(spkr, PrevAud)} \\ \text{addr} & : & \text{Ind} \\ \text{c2} & : & \text{member(addr, PrevAud)} \\ & & \wedge \text{ addr} \neq \text{spkr} \end{bmatrix}$$

(4) b. QSpec

$$\begin{bmatrix} \text{pre:} \begin{bmatrix} \text{qud} = \langle \text{i, I} \rangle : \text{poset(InfoStruc)} \end{bmatrix} \\ \text{effects : TurnUnderspec} \\ \wedge_{merge} \begin{bmatrix} \text{r : AbSemObj} \\ \text{R: IllocRel} \\ \text{LatestMove} = \text{R(spkr,addr,r) : IllocProp} \\ \text{c1 : Qspecific(r,i.q)} \end{bmatrix} \end{bmatrix}$$

4.2. Forwards-looking disfluencies and self-addressed queries

Our starting point is the account developed within the KoS framework for Clarification Requests (see e.g. [13, 7]): in the aftermath of an utterance u a variety of questions concerning u and definable from u and its grammatical type become available to the addressee of the utterance. These questions regulate the subject matter and ellipsis potential of CRs concerning u and generally have a short lifespan in context. We argue that disfluencies can and should be subsumed within a similar account, a point that goes back to [16]: in both cases (i) material is presented publicly, (ii) a problem with some of the material is detected and signalled (= there is a 'moment of interruption'); (iii) the problem is addressed and repaired leaving (iv) the incriminated material with a special status, but within the discourse context. Concretely for disfluencies – as the utterance unfolds incrementally questions can be pushed on to QUD about what has happened so far, as with Backwards Looking Disfluencies (BLDs) (e.g. *what did the speaker mean with subutterance u1?*) or what is still to come, as with Forwards Looking Disfluencies (FLDs) (e.g. *what word does the speaker mean to utter after sub-utterance u2?*).

We specify FLDs with the update rule in (5) – given a context where the LatestMove is a forward looking editing phrase by A, the next speaker – underspecified between the current one and the addressee – may address the issue of what A intended to say next by providing a co-propositional utterance:[4,5]

(5) Forward Looking Utterance Rule:

$$\begin{bmatrix} \text{preconds :} \begin{bmatrix} \text{spkr : Ind} \\ \text{addr : Ind} \\ \text{pending} = \langle \text{p0,rest} \rangle : \text{list(LocProp)} \\ \text{u0 : LocProp} \\ \text{c1 : member(u0, p0.sit.constits)} \\ \text{LatestMove}^{content} = \text{FLDEdit(spkr,u0) : IllocProp} \end{bmatrix} \\ \text{effects : TurnUnderspec} \wedge_{merge} \\ \begin{bmatrix} \text{MaxQud} = \\ \begin{bmatrix} \text{q} = \lambda x \text{ MeanNextUtt(pre.spkr,pre.u0,x)} \\ \text{fec} = \{\} \end{bmatrix} : \text{InfoStruc} \\ \text{LatestMove : LocProp} \\ \text{c2: Copropositional(LatestMove}^{content},\text{MaxQud)} \end{bmatrix} \end{bmatrix}$$

─────────────

[4]   This rule is inspired in part by Purver's rule for *fillers*, (91), p. 92, ([15]). Given that our rule leaves the turn ownership unspecified we unify FLDs with fillers.

[5]   CoPropositionality for two questions means that, modulo their domain, the questions involve similar answers. For instance 'Whether Bo left', 'Who left', and 'Which student left' (assuming Bo is a student) are all co-propositional. In the current context co-propositionality amounts to: either a CR which differs from MaxQud at most in terms of its domain, or a correction – a proposition that instantiates MaxQud.

Rule (5) differs from its BLD analogue, in two ways. First, in that the preconditions involves the LatestMove having as its content what we describe as an FLDEdit move, which we elucidate somewhat shortly. Words like 'uh', 'thee' will be assumed to have such a force, hence the utterance of such a word is a prerequisite for an FLD. A second difference concerns parallelism: for BLDs it is intuitive that parallelism exists between reparandum and alteration (with certain caveats), given that one is replacing one sub-utterance with another that is essentially of the same type. However, for FLDs there is no such intuition—what is taking place is a search for the word after the reparandum, which has no reason to be parallel to the reparandum. Hence in our rule (5), the FEC is specified as the empty set. To make this explicit, we assume that 'uh' could be analyzed by means of the lexical entry in (6):

(6)

$$\begin{bmatrix} \text{phon :} \texttt{ uh} \\ \text{cat} = \textit{interjection} : \text{syncat} \\ \text{dgb-params :} \begin{bmatrix} \text{spkr : IND} \\ \text{addr : IND} \\ \text{MaxPending : LocProp} \\ \text{u0 : LocProp} \\ \text{c1: member(u0, MaxPending.sit.constits)} \\ \text{rest : address(spkr,addr,MaxPending)} \end{bmatrix} \\ \text{cont} = \begin{bmatrix} \text{c1 : FLDEdit(spkr,addr,MaxPending)} \end{bmatrix} : \text{Prop} \end{bmatrix}$$

We demonstrate how to analyze (7):

(7)   A: Show flights arriving in uh Boston. [18]

After A utters $u0$= 'in', she interjects 'uh', thereby expressing FLDEdit(A,B,'in'). This triggers the Forward Looking Utterance rule with MaxQud.q = λx MeanNextUtt(A,'in',x). 'Boston' can then be interpreted as answering this question, with resolution based on the rule used to interpret (elliptical) short answers.

Similar analyses can be provided for (8). Here instead of 'uh' we have lengthened versions of 'the' and 'a' respectively, which express FLDEdit moves:

(8)   a.   And also the- the dog was old. [2]
       b.   A vertical line to a- to a black disk [11]

Let us return to consider what the predicate 'FLDEdit' amounts to from a semantic point of view. Intuitively, (9) should be understood as 'A wants to say something to B after $u0$, but is having difficulty (so this will take a bit of time)':

(9)   FLDEdit(A,B,u0)

This means we could unpack (9) in a number of ways, most obviously by making explicit the utterance-to-be-produced $u1$, representing this roughly as in (10):

(10) $\exists u1[\text{After}(u1,u0) \wedge \text{Want}(A,\text{Utter}(A,B,u1))]$

Moving on finally to (dysfluent) self addressed queries of the kind described in section 2, on our account such queries are licensed because these questions are co-propositional with the issue 'what did A mean to say after $u0$'.

Self addressed queries also highlight another feature of KoS's dialogue semantics: the fact that a speaker can straightforwardly answer their own question, indeed in these cases the speaker is the "addressee" of the query. Such cases get handled easily in KoS because turn taking is abstracted away from querying: the conversational rule QSpec, introduced earlier as (4-b), allows either conversationalist to take the turn given the QUD-maximality of q. This contrasts with a view of querying derived from Speech Act Theory (e.g. [17]) still widely as- sumed (see e.g. [1]), where there is very tight link to intentional categories of 2-person dialogue ('. . . Speaker wants Hearer to provide an answer . . . Speaker does not know the answer . . . ').

5. Conclusions

In this paper we offer the first detailed corpus study of self-addressed queries that occur in the aftermath of filled pauses. We show that such queries show marked signs of sensitivity to the (the syntactic/semantic type of) the sub-utterance $u0$ they follow. We then offer a formal model from which the possibility for such queries follows directly. An obvious next step is to study differences between the distribution we found in the BNC and that occurring in other languages, as surface syntax in particular of NPs seems to be a significant factor, as does predicational structure.

6. Acknowledgements

7. References

[1] N. Asher and A. Lascarides, *Logics of Conversation*. Cambridge: Cambridge University Press, 2003.

[2] J. Besser and J. Alexandersson, "A comprehensive disfluency model for multi-party interaction", in Proceedings of SigDial 8, pp. 182–189, 2007.

[3] H. Clark and J. FoxTree, "Using uh and um in spontaneous speech", *Cognition*, vol. 84, pp. 73–111, 2002.

[4] M. Corley and O. W. Stewart, "Hesitation disfluencies in spontaneous speech: The meaning of 'um'", *Language and Linguistics Compass*, vol. 2, no. 4, pp. 589–602, 2008.

[5] R. Fernández, "Non-sentential utterances in dialogue: Classification, resolution and use", Ph.D. dissertation, King's College, London, 2006.

[6] C. Gardent and M. Kohlhase, "Computing parallelism in discourse," in *Proc. IJCAI*, pp. 1016–1021, 1997.

[7] J. Ginzburg, "Situation semantics: from indexicality to metacommunicative interaction", in *The Handbook of Semantics*, K. von Heusinger, C. Maierborn, and P. Portner, Eds. Walter de Gruyter, 2011.

[8] J. Ginzburg, *The Interactive Stance: Meaning for Conversation*. Oxford: Oxford University Press, 2012.

[9] J. Ginzburg and R. Fernández, "Computational models of dialogue", in *Handbook of Computational Linguistics and Natural Language*, A. Clark, C. Fox, and S. Lappin, Eds. Oxford: Blackwell, 2010.

[10] M. Horne, "Attitude reports in spontaneous dialogue: Uncertainty, politeness and filled pauses", in *From Quantification to Conversation*, L. Borin and S. Larsson, Eds. Gothenburg: Gothenburg University, pp. 309–318, 2008.

[11] W. J. Levelt, "Monitoring and self-repair in speech", *Cognition*, vol. 14, no. 4, pp. 41–104, 1983.

[12] D. C. O'Connell and S. Kowal, "Uh and Um Revisited: Are They Interjections for Signaling Delay?" *Journal of Psycholinguistic Research*, vol. 34, no. 6, pp. 555–576, 2005.

[13] M. Purver, "Clarie: Handling clarification requests in a dialogue system", *Research on Language & Computation*, vol. 4, no. 2, pp. 259–288, 2006.

[14] M. Purver, "Score: A tool for searching the bnc", King's College, London, Tech. Rep. TR-01-07, 2001.

[15] M. Purver, "The theory and use of clarification in dialogue", Ph.D. dissertation, King's College, London, 2004.

[16] E. Schegloff, G. Jefferson, and H. Sacks, "The preference for self-correction in the organization of repair in conversation", Language, vol. 53, pp. 361–382, 1977.

[17] J. Searle, *Speech Acts*. Cambridge: Cambridge University Press, 1969.

[18] E. E. Shriberg, "Preliminaries to a theory of speech disfluencies", Ph.D. dissertation, University of California at Berkeley, Berkeley, USA, 1994.