Abstract

We discuss similarities between mid-utterance self-correction, which is often seen as a phenomenon that lies outside the scope of theories of dialogue meaning, and other discourse phenomena, and argue that an approach that captures these similarities is desirable. We then provide a sketch of such an approach, using Ginzburg’s KoS formalism, and discuss the implications of including ‘sub-utterance-unit’ phenomena in discourse theories.

1 Introduction

Unlike written language, spoken conversational language is full of what can be described as explicit traces of editing processes, as in the following example:

(1) I was one of the. I was responsible for all the planning and engineering

In this example, the brief silence after one of the (represented here by a full stop) seems to prepare the ‘editing operation’ that is to follow in the form of a partial repetition of material, the result being the ‘cleaned up’ utterance I was responsible for all the planning and engineering, with the fragment I was one of the being struck from the record.

2 Form and Function of Dysfluencies

In this paper we will argue, following much work in the tradition of conversational analysis beginning with (Schegloff et al., 1977), that there are, in fact, strong similarities between self-correction and other discourse phenomena (Section 3), which make an approach that captures these similarities desirable. In contrast to conversation analytic work, however, we actually ground our proposal in a formal model: in Section 4 we sketch such an approach, couched in terms of the KoS formalism (Ginzburg and Cooper, 2004; Purver, 2004; Ginzburg, (forthcoming)). We also discuss there the implications of making such a move for the grammar/parser–discourse interface and for discourse theories in general. Some conclusions are provided in Section 5.

Before coming to this, however, we briefly give some background on speech dysfluencies in the next section and review some of the terminology from the literature.

2 Form and Function of Dysfluencies

In this section we discuss the ‘syntax’ of self-correction, classifications according to the relation of problematic material and replacement, and the kinds of problems that can be corrected with self-correction.

As has often been noted (see e.g. Levelt (1983), and references therein for earlier work), speech dysfluencies follow a fairly predictable pattern. The example in Figure 1 is annotated with the labels introduced by Shriberg (1994) (building on (Levelt, 1983)) for the different elements that can occur in a self-repair.

1From the Switchboard corpus (Godfrey et al., 1992).
2This division of labour also seems to be presupposed by much of the computational work on automatically detecting and repairing dysfluent speech, as expressed e.g. in the following quote from (Heeman and Allen, 1999): “we propose that these tasks [a.o. detecting and correcting speech repairs, the authors] can be done using local context and early in the processing stream.”
3“Although self-initiation and other-initiation of repair are distinct types of possibilities [...] There are quite compelling grounds for seeing self and other-initiations to be related, and for seeing their relatedness to be organized.” (Schegloff et al., 1977)
Of these elements, the editing term is always optional (although some marking, like an extended pause, seems to be always present (McKelvie, 1998)). The relation between reparandum and alteration can be used as the basis of a further classification: if the alteration differs strongly from the reparandum and does not form a coherent unit together with the start, or if alteration and continuation are not present at all, the dysfluency can be classified as an *aborted utterance / fresh start*. Other classes are *repair* (alteration 'replaces' reparandum) and *reformulation* (alteration elaborates on reparandum). The following gives examples for all three classes:

(2) a. \{ *I mean* \} [ *I, + I, / there are a lot, + there are so many* ] different songs,
   b. [ *We were + I was* ] lucky too that I only have one brother.
   c. at that point, [ *it, + the warehouse* ] was over across the road

Within the class of repairs, finally, a further distinction can be made (Levelt, 1983) into *appropriateness-repairs* that replace material that is deemed inappropriate by the speaker given the message she wants to express (or has become so, after a change in the speaker's intentions), and *error-repairs*, where the material is erroneous.

### 3 From Other to Self

Figure 2 shows (constructed) examples of 'normal' discourse correction (a), two uses of clarification requests (b & c), correction within a turn (d), other-correction mid-utterance (e), and two examples of self-correction as discussed above (f & g). The first four examples clearly are instances of phenomena within the scope of discourse theories.

**What about the final two?**

...
Our task, then, is to develop a formal model that can capture the similarities exhibited by self-initiated within-utterance repair and other-initiated cross-utterance repair, without neglecting the important characteristics that differentiate them. To this we turn now.

4 A Model of Other- and Self-Repair

4.1 KCRT: A Theory of Inter-Utterance, Other-Initiated Repair

For concreteness we take as our starting point the theory of CRification developed in (Ginzburg and Cooper, 2004; Purver, 2004; Ginzburg, (forthcoming)) (henceforth Kos CR Theory (KCRT)). This theory attempts to explain a.o. the coherence of CRs/corrections such as the following:

7 a. A: Did Bo leave? B: Bo? (= Who do you mean ‘Bo’? or Are you asking if BO left?)

b. A: Did Bo phone? B: You mean Mo.

c. A: Should we... B: leave? (= Is ‘leave’ the word to be said after ‘we’? )

How to analyze examples like (7c) is actually only mentioned in passing in (Purver, 2004), given certain formal difficulties it involves, not least of which is parsing an incomplete utterance.

The main features of KCRT are:

Initialization: Utterances are kept track of in a contextual attribute PENDING (cf. the G/DU bifurcation in PTT (Poesio and Traum, 1997).) in the immediate aftermath of the speech event. Given a presupposition that \( u \) is the most recent speech event and that \( T_u \) is a grammatical type that classifies \( u \), a record of the form \[ \text{sit} = u \]

of type \( \text{sit-type} = T_u \)

LocProp (locationary proposition), gets added to PENDING.

Contextual/phonological instantiation: In so far as A’s information state \( IS_0 \) enables her to...
fully instantiate the contextual parameters specified in \( T_u \), and \( T_u.phon \) is uniquely specified, \( \text{sit} = u \), \( \text{sit-type} = T_u \), can trigger an illocutionary update of IS\(_0\) (i.e., a new move is added to MOVES—an assertion, query etc.)

**CR/Correction coherence:** Failure to fully instantiate contextual parameters or recognize phonological types triggers CRification. This involves accommodation of questions into context by means of Clarification Context Update Rules (CCURs). Each CCUR specifies an accommodated MaxQUD built up from a sub-utterance \( u_1 \) of the target utterance, the maximal element of PENDING, MaxPending. Common to all CCURs is a license to follow up MaxPending with an utterance whose qud-update is copropositional with MaxQud\(^8\); either a CR which differs from MaxQud at most in its domain, or a correction—a proposition that instantiates MaxQud. The CCURs differ primarily in the question whose accommodation into QUD they give rise to. (8) is a simplified formulation of one CCUR, (9)-(11) provide a specification of the MaxQud instantiation of other CCURs:

**Parameter identification:**

\[
\begin{align*}
\text{Input:} & \quad \text{Spkr} : \text{Ind} \\
& \quad \text{MaxPending} : \text{LocProp} \\
& \quad u_0 \in \text{MaxPending}.\text{sit.constits} \\
\text{Output:} & \quad \text{MaxQUD} = \text{What did spkr mean by } u_0? \\
& \quad \text{LatestMove} : \text{LocProp} \\
& \quad \text{c1: CoProp(LatestMove.cont,MaxQUD)}
\end{align*}
\]

**Parameter focussing:** raises as \( \text{MaxQud} \)
\[
\lambda x.\text{MaxPending}.\text{content}(u_1.\text{content} \rightarrow x)
\]

**Utterance repetition:** raises as \( \text{MaxQud} \)
\[
\lambda x.\text{Utter}(A,u_1.x) \quad (\text{What did } A \text{ utter in } u_1? \quad \text{“What did you say?”})
\]

**Utterance prediction:** raises as \( \text{MaxQud} \)
\[
\lambda x.\text{UtterAfter}(A,u_1.x) \quad (\text{What will } A \text{ utter after } u_1? \quad \text{“What were you going to say?”})
\]

**Answers:** Accepting an answer to a CR/correction gives rise to an modified MaxPending via **Contextual/phonological instantiation:** (in the case of content-related CRs (corrections): the contextual assignment of \( u \) is extended (replaced by a substitute); in the case of phonological CRs this applies to \( T_u.phon \)).

**Speaker/hearer asymmetry:** Speakers cannot self-CR because their own utterance is downdated from PENDING following successful contextual parameter instantiation (which always applies to a speaker’s own utterance). Hence, the different contextual possibilities, exemplified in (4) and (5).

**CR accommodation:** If \( A \) utters \( u \) and \( B \) follows up with a CR/correction, \( A \) accommodates the MaxQud B accommodated and \( \text{sit} = u \), \( \text{sit-type} = T_u \) becomes MaxPending.

### 4.2 Extending KCRT to Self-Initiated Mid-Utterance Repair

How do we extend this model to mid-utterance self and other correction? As things stand, there are two things that prevent KCRT from accounting for self-repair: (1) all CR/corrections are forced to occur after complete utterances, and (2) CR/corrections can only be posed by others (given that the speaker downdates PENDING immediately). Let us take up each of these issues in turn.

The first move we make is indeed to extend PENDING to incorporate utterances that are in progress, and hence, incompletely specified semantically and phonologically. Conceptually this is a natural step to make. Formally and methodologically this is a rather big step, as it presupposes the use of a grammar which can associate types word by word (or minimally constituent by constituent), e.g., in Categorial Grammar, Dynamic Syntax, (Steedman, 2000; Kempson et al., 2000). It raises a variety of issues with which we cannot deal in the current paper: monotonicity, nature of incremental denotations etc.

For our current purposes, the decisions we need to make can be stated independently of the specific grammatical formalism used, modulo the fact that as in the KCRT work, we need to assume that grammatical types specify a feature/label/field \( \text{CONSTITS} \) which keeps track of all not just immediate constituents of a given speech event (gram-
matical type). The main assumptions we are forced to make concern where pending instantiation and contextual instantiation occurs, and more generally, the testing of the fit between the speech events and the types assigned to them. We assume that this takes place incrementally, say word by word.

The incrementalization of PENDING has good consequences, as well as certain seemingly undesirable ones. On the positive side, since PENDING now includes also incomplete utterances, we can now account also for CRs/other corrections that occur mid-utterance, dispreferred as they might be (Schegloff et al., 1977). One such corpus example is (12a). The constructed (12b) shows that in such contexts the same ambiguities are maintained as in cross-utterance cases exemplified above:

\(12\) a. A: *There are subsistence farmers that...* B: *There are what?* (attested example from the Potsdam Pentomino Corpus)

b. A: *Did Bo... (no pause) B: Bo? (= Who do you mean ‘Bo’? or Are you asking something about BO?) A: I mean Mo/Yeah, Mo’s partner.*

On the other hand, without saying more, it will overgenerate in precisely the way we were trying to avoid, given (4) and (5). We can block this via a route any dialogue theory has to go through in any case: moves such as acceptances involve obligatory turn change. For this reason KCRT already keeps track of speaker/addressee roles, while underspecifying these where the turn is up for grabs (as e.g. following the posing of a query.). So the CCURs we specified above will now carry information that ensures that the various interpolated utterances do indeed involve a turn change.

This in turn means that simply enlarging the scope of what goes into PENDING has not offered a route to characterize the potential for mid-utterance self correction. But this is probably inevitable: while there may be some cases such as (12) involving other participants, self-correction in mid-utterance (and elsewhere) involves, as we discussed earlier, the presence of an editing phrase (EditP) (encompassing also extended silences.). What we need to do, therefore, is to provide a means for licensing EditPs. This is simple to do: all we need to say is that an EditP can be interpolated essentially at any point, or more precisely, at any point where PENDING is non-empty. (13) is an informal such specification. It enforces turn continuity and the non-inclusion of the EditP in PENDING:

\(13\) Edit Move Update Rule:

\[
\text{Input: } \left[ \begin{array}{l} \text{Spkr} : \text{Ind} \\ \text{MaxPending} : \text{LocProp} \end{array} \right] \\
\text{Output: } \left[ \begin{array}{l} \text{Spkr} = \text{Input.spkr} : \text{Ind} \\
\text{Pending} = \text{Input.MaxPending} : \text{LocProp} \\
\text{LatestMove} = \text{Edit} (\text{Spkr,MaxPending}) \end{array} \right]
\]

The output state this brings us to is a state where PENDING contains repairable material and the LatestMove is an EditP. Now we can specify coherent Self/Other corrections in a manner akin, though not identical to (8)-(11). We will assume the following as a tentative characterization, though clearly it is not exhaustive:

\(14\) ... u0... EditP u1 (= Spkr meant to utter u1)

\(15\) ... u0... EditP u0’? (= Did Spkr mean to utter u0?)

\(16\) A: ... u0... \{um, uh\} u1 (= Spkr meant u1 to be the next word after u0)

We sketch here only a rule that will capture (14) and (15). The URs in (17) take as input a state where the LatestMove is an EditP and specify a new state in which the MaxQUD is *What did spkr mean to utter at u0?* and where the new utterance has to be an instantiation of MaxQud (propositional or polar question):

\(17\) Utterance identification:

\[
\text{Input: } \left[ \begin{array}{l} \text{Spkr} : \text{Ind} \\
\text{MaxPending} : \text{LocProp} \\
\text{LatestMove} = \text{Edit} (\text{Spkr,MaxPending}) \\
\text{u0} \in \text{MaxPending.sit.constits} \end{array} \right] \\
\text{Output: } \left[ \begin{array}{l} \text{MaxQUD} = \text{What did spkr mean to say at u0?} \\
\text{LatestMove} : \text{LocProp} \\
\text{c2} : \text{InstPropQ(LatestMove.cont,MaxQUD)} \end{array} \right]
\]

With this machinery in hand, we can now consider some examples:

1. Self-correction mid-utterance:

\(18\) A: *Peter, no Paul quit.*

1.a After utterance of ‘Peter’: in A’s FACTS (shared assumptions etc—whatever underwrites presuppositions) the presuppositions that the most recent speech event is u0 (‘Peter’), classified by a
type T\textsubscript{u0}; PENDING gets updated with the following record:
\[
\begin{align*}
\text{sit} &= u0; \\
\text{Sit-Type} &= \text{‘Utterance whose first word} \\
&\text{is Peter; involves reference to p...’}
\end{align*}
\]

1.b This allows for an EditP to be interpolated: LatestMove = Edit(A, MaxPending).

1.c This allows for utterance identification: MaxQUD = What did spkr mean to say at u0?; LatestMove: Assert(A, MeanUtter(A, ‘Paul’))

1.d Accepting this gives rise to an application of Contextual/phonological instantiation: PENDING is modified to the following record:
\[
\begin{align*}
\text{sit} &= u1; \\
\text{Sit-Type} &= \text{‘Utterance whose first word} \\
&\text{is Paul; involves reference to p...’}
\end{align*}
\]

1.e Note: the utterance u0 is still in the information state, though not as a component of PENDING—PENDING was originally initialized due to the presence in FACTS of the proposition that the most recent speech event is u0 (‘Peter’), classified by a type T\textsubscript{u0}. Hence, anaphoric possibilities to this utterance are not eliminated.

2. Self-correction after utterance:


Same procedure as in 1., initiated with the completed utterance as MaxPending.

3. Other-correction, indirect:

(20) A: (1) Peter is not coming.
   B: Peter? (in ‘indirect correction’ reading)
   A: Oh, sorry, I meant Paul.

In consequence of B’s utterance A applies CR accommodation, which makes What did A mean by ‘Peter’ MaxQud and (1) MaxPending. Applying Contextual/phonological instantiation after A’s correction leads to a modification in (1).

4. Other-correction, direct:

(21) A: (a) Peter is not coming.
   B: (b) No, (c) Peter is, Paul isn’t.

This is simply a disagreement at the illocutionary level: A’s assertion pushes ?Coming(peter) to MaxQud but not to FACTS, giving rise to the discussion which B initiates. If A accepts B’s assertion (c) will be added to FACTS, whereas ?Coming(peter) gets downdated from QUD.

5 Conclusions

In this paper we have related self- and other-initiated repair. We have argued, following a long but unformalized tradition in Conversation Analysis, that the two processes bear significant similarities: a problem is detected with an utterance, this is signalled, and then the problem is addressed and repaired, leaving the incriminated material with a special status, but within the discourse context. We provide a unified account: a single repository, PENDING carries CR/correct-able material within and across utterances. Consequently, a single set of rules regulate the up- and downdating of PENDING, as well as the modification of its elements by answers to CRs or corrections, regardless of whether the utterances that are in progress or completed. Different rules trigger within and cross-utterance CRs/corrections, but that is as should be, as the form and content of these differ, as we have shown.

Acknowledgements This work was carried out while the first author was a senior visitor at the University of Potsdam within the DEAWU project. The work was supported by the EU (Marie Curie Programme) and Deutsche Forschungsgemeinschaft (Emmy Noether Programme).

References


