Dialogue and Discourse 2(1) (2011) 279–311 doi: 10.5087/dad.2011.111

On Incrementality in Dialogue: Evidence from Compound Contributions*

Christine Howes C.HOWES@QMUL.AC.UK
School of Electronic Engineering and Computer Science
Queen Mary University of London
Mile End Road, London E1 4NS, UK

Matthew Purver M.PURVER@QMUL.AC.UK
Queen Mary University of London

Patrick G. T. Healey PAT.HEALEY@EECS.QMUL.AC.UK
Queen Mary University of London

Gregory J. Mills GJMILLS@STANFORD.EDU
Department of Psychology
Stanford University
Stanford, CA 94305, USA

Eleni Gregoromichelaki ELENI.GREGOR@KCL.AC.UK
Philosophy Department
King's College London
Strand, London WC2R 2LS, UK

Editor: Hannes Rieser and David Schlangen

Abstract

Spoken contributions in dialogue often continue or complete earlier contributions by either the same or a different speaker. These *compound contributions* (CCs) thus provide a natural context for investigations of incremental processing in dialogue.

We present a corpus study which confirms that CCs are a key dialogue phenomenon: almost 20% of contributions fit our general definition of CCs, with nearly 3% being the cross-person case most often studied. The results suggest that processing is word-by-word incremental, as splits can occur within syntactic 'constituents'; however, some systematic differences between same- and cross-person cases indicate important dialogue-specific pragmatic effects. An experimental study then investigates these effects by artificially introducing CCs into multi-party text dialogue. Results suggest that CCs affect people's expectations about who will speak next and whether other participants have formed a coalition or 'party'.

Together, these studies suggest that CCs require an incremental processing mechanism that can provide a resource for constructing linguistic constituents that span multiple contributions and multiple participants. They also suggest the need to model higher-level dialogue units that have consequences for the organisation of turn-taking and for the development of a shared context.

Keywords: Compound contributions; corpus study; party formation; dialogue; incrementality

*. This research was carried out under the *Dynamics of Conversational Dialogue* project, funded by the UK ESRC (RES-062-23-0962) with EPSRC support through the *DiET (Dialogue Experimentation Toolkit)* project (EP/D057426/1). We also thank Ruth Kempson and Arash Eshghi for many useful discussions.

 Submitted 1/10; Accepted 4/11; Published online 5/11

1. Introduction

Compound contributions (*CCs*) – spoken dialogue contributions that continue or complete an earlier contribution,[1] see e.g. (1) – have been claimed to occur regularly in dialogue, especially according to the Conversation Analysis (CA) literature, where specific types of compound contributions have been studied under a variety of names, including completions and joint productions (see section 2).

(1) **Daughter:** Oh here dad, a good way to get those corners out
Dad: is to stick yer finger inside.
Daughter: well, thats one way. *[from Lerner (1991)]*

CCs are of interest to dialogue theorists as they provide evidence about how contributions can cohere with each other at multiple levels – syntactic, semantic and pragmatic (though of course they are not the only way). They also indicate the radical context-dependency of conversational contributions, which can, in general, be highly elliptical without disrupting the flow of the dialogue. CCs are a dramatic illustration of this: speakers must rely on the dynamics of the unfolding context (linguistic and extra-linguistic) in order to guarantee successful processing and production.

As early as 1967, in his series of Lectures on Conversation, Sacks (1992) noted that the existence of CCs supports the (now largely accepted) thesis that language in dialogue is processed incrementally:

> Such a fact as that persons go about finishing incomplete sentences of others with syntactically coherent parts would seem to constitute direct evidence of their analysing an utterance syntactically in its course. . . (Sacks, 1992, p651)

However, we argue here that the evidence from CCs goes further; they show that not just processing (parsing), but also production (generation) must be incremental; and that because of the variation in CCs, this must also be at a finer-grained level than is often assumed (see also Ferreira, 1996; Guhe, 2007).

Compound contributions that are split across speakers also present a canonical example of participant coordination in dialogue (here we call these *cross-person CCs* to distinguish them from the *same-person* cases where the original speaker later continues his own contribution – see below). The ability of one participant to continue another interlocutor's contribution coherently, both at the syntactic and semantic level, implies that speaker and hearer can be highly coordinated in terms of processing and production. The initial speaker must be able to switch to the role of hearer, processing and integrating the continuation of their contribution, whereas the initial hearer must be monitoring the grammar and content of what they are being offered closely enough that they can take over and continue in a way that respects the constraints set up by the first contribution. This switch is particularly obvious in those cases where the initial hearers continuation is not the same as that which the original speaker would have provided, as in (1, 2).

1. These terms will be defined in detail below.

(2) **BMA:** She got compensation
Just like that
Because what she had in her suitcase
PM: was Grade A.

[from comedy news quiz Have I got news for you, *s35 ep1]*

There is evidence that such constraints are respected across speaker and hearer in compound contributions (see e.g. Gregoromichelaki et al., 2009). In both Finnish (which has a rich inflectional morphology), and Japanese (a verb-final language), cross-person CCs within a single clause conform to the strict syntactic constraints of the language, despite the change in speaker (Helasvuo, 2004; Hayashi, 1999; Lerner and Takagi, 1999).

These observations have important theoretical implications. Firstly, the grammar and semantics employed by the interlocutors must be able to license and interpret chunks much smaller than the usual sentential or propositional units. Moreover, the possibility of role switches while syntactic/semantic dependencies are pending suggests direct involvement of the grammar in the parsing and production processes, or, at least, a very tight coupling between those processes and the grammar and intermediate representations being used (see Gargett et al., 2009). Indeed, Poesio and Rieser (2010) claim that "[c]ollaborative completions ... are among the strongest evidence yet for the argument that dialogue requires *coordination* even at the sub-sentential level" (italics original).

From a psycholinguistic point of view, the phenomenon of CCs is compatible with mechanistic approaches as exemplified by the Interactive Alignment model of Pickering and Garrod (2004), which claims that it should be as easy to complete someone else's sentence as one's own (p186). According to this model, speaker and listener ought to be interchangeable at any point. This is also the stance taken by the grammatical framework of Dynamic Syntax (DS: Kempson et al., 2001; Cann et al., 2005). In DS, parsing and production are taken to employ the same mechanisms, leading to a prediction that CCs ought to be strikingly natural (Purver et al., 2006). However, continuation by another speaker is sometimes taken to involve guessing or preempting the other interlocutor's intended content.[2] It has therefore been claimed that a full account of CCs requires a complete model of pragmatics that can handle intention recognition and formation. Indeed, Poesio and Rieser (2010) propose sentence completions as the testing ground of competing claims about coordination i.e. whether it is best explained with an intentional model like Clark's (1996) or with a model based on simpler alignment models like Pickering and Garrod's (2004). They conclude that a model which includes modelling of intentions better captures the data, though see (Gregoromichelaki et al., 2011) for an alternative argument.

For computational models of dialogue, compound contributions pose a challenge. While Poesio and Rieser (2010) and Purver et al. (2006) provide general foundational models for various aspects of CCs, there are many questions that remain if automatic processing of naturally occurring dialogues is ever to be completely realised. A computational dialogue system must be able to identify CCs, match up their two (or more) parts (which may not necessarily be adjacent), integrate them into some suitable syntactic and/or semantic representation, and determine the overall pragmatic contribution to the dialogue context. CCs also have implications for the organisation of *turn-taking* in such models (see e.g. Sacks et al., 1974), as regards what conditions (if any) allow or prevent successful turn transfer.

2. Note that this says nothing about whether such a continuation is the same as the initial speaker's intended continuation. For cases where this cannot be the case see (Gregoromichelaki et al., 2011), as well as (1, 2) above.

From an organisational point of view, it has been claimed that turn-taking operates not on individual conversational participants, but on ‘parties’ (Schegloff, 1995). For example, a couple talking to a third person may organise their turns as if they are one ‘party’, rather than two separate individuals. Lerner (1991) speculates that cross-person compound contributions can clarify the formation of such parties, as they reveal a relationship between syntactic mechanisms and social organisation. He claims that this provides evidence of one way in which syntax can be used to organise participants into “groups”.

Of course, our earlier definition of compound contributions begs several questions; most importantly, what do we mean by a ‘dialogue contribution’? Our use of this term, and the related notion of a turn, can be best explained by reference to a short extract of dialogue taken from the British National Corpus (3).

(3)	1.	A:	I were gonna say, they wash [[better than]]	
	2.	J:	[[But I’ve had]]	
	3.	A:	velvet.	
	4.	J:	I’ve had to take them up.	
	5.		Cos they were, they were gonna be miles too long.	
	6.		And I’ve not even took them out the thing.	
	7.		They said he’d swap them if they didn’t fit.	
	8.	A:	[[Ah they do!]]	
	9.	J:	[[And he]] <pause><unclear>.	
	10.	A:	Where d’ya get them from Joyce?	
	11.	J:	I got them from that er	
	12.	B:	Top Marks.	
	13.	J:	that shop.	*[BNC KB2 4134-4146]*

In our usage, each of the transcribed lines (1-13) is a *contribution*. Our use of *contribution* is intended to correspond to Clark’s (1996) “a *contribution* to discourse – [is] a signal successfully understood” (p227).[3] With transcribed corpus text, of course, it is not always possible to determine whether contributions have been successfully understood, as we have no access to non-verbal signals (such as nodding). We therefore take contributions to be stretches of talk bounded by a change in speaker, a significant pause, or the end of the sentence, and assume that in most cases the transcribers’ decision to split the text into separate lines indicate some (e.g. prosodic) cues to suggest that the line has been successfully understood, i.e. treated as a contribution. Thus, whilst contributions can be single words (as in line 3) or backchannels (e.g. ‘mm’), or complete syntactic sentences (e.g. line 4), they can also be partial sentences (e.g. the incomplete sentences at lines 1, 2 and 11 and the fragments at lines 3, 12 and 13). Note however, that single words in longer contributions (e.g. ‘they’ at the start of line 7) do not count as contributions in their own right.

Compound contributions can now be defined as *single* syntactic or semantic (propositional) units built across multiple contributions, which could be provided by one speaker or several. The exchange in lines 11-13 provides two examples. J’s contribution ‘I got them from that er’ starts a sentence, which B’s contribution ‘Top Marks’ (the name of a shop) completes. This counts as a

3. Note that Clark uses contribution to refer to both “the joint act of . . . completing the signal and its joint construal” and for the interlocutors “participatory act, his *part* of that joint act, as when we speak of Roger’s contribution to the discourse.” We use contribution in this second sense only.

compound contribution under our definition. J then also completes her own contribution (with 'that shop') at line 13, and this also counts as a (same-person) compound contribution, as it is spread across multiple contributions (in this case, with intervening material). Note that even though the short extract in (3) also exhibits many other conversational *tying techniques* (Sacks, 1992), such as a question and answer (lines 10-11), and the use of pronouns linked to referents previously introduced in the dialogue, our focus here is not on all pragmatic dependencies between turns.

It should be noted, however, that this definition depends on the protocol used by the corpus transcribers; and with the BNC, this can lead to possibly undesirable segmentation of stretches of talk into multiple "contributions". The insistence on linear ordering means that cases of interruption of one speaker by another will always result in an apparent speaker change, even if the interruption consists only of non-verbal noises (e.g. coughing) or is entirely overlapping – see e.g. lines 1-3 (overlapping material is shown in the examples with square brackets aligned to the material with which it overlaps). J's interruption in line 2 overlaps with A's speech, but forces A's sentence to be transcribed as two lines (1 and 3). These count as separate contributions under our definition, giving a compound contribution: A begins her contribution 'I were gonna say, they wash better than', which she completes in line 3 with 'velvet'. In many cases this may be the correct analysis – in Clark's usage, overlap *can* signal understanding (I might not need you to syntactically or semantically finish your sentence to accept it as a valid contribution to the discourse). In this case, though, it may be that lines 1 and 3 were intended (and processed) as one single contribution – to avoid possibly misleading conclusions we therefore report CC figures both including and excluding such cases (see section 4). Note, however, that these concerns only apply to same-person CCs and not to cross-person CCs.

Whe also define a notion of *turn* here as all talk to the next change of speaker; the contributions by J in lines 4-7 would therefore be classified as a single *turn*. We will use this notion below to distinguish CCs which span multiple turns from those spanning multiple contributions within a single turn. Even a backchannel or overlapping material, such as line 2 (which completely overlaps with the end of line 1) counts as a change of speaker (and thus separate turns) here.

Analysis of CCs, when they can or cannot occur, and what effects they have on the coordination of agents in dialogue, is therefore an area of interest not only for conversation analysts wishing to characterise systematic interactions in dialogue, but also for linguists trying to formulate grammars of dialogue, psychologists and sociolinguists interested in alignment mechanisms and social interaction, and those interested in building automatic dialogue processing systems. In this paper we present and examine empirical corpus data and an experimental manipulation of CCs, in order to shed light on some of the questions and controversies around this phenomenon.

2. Related Work

Most previous work on what we call CCs has examined specific sub-cases, generally of the cross-person type, and have referred to these variously as *collaborative turn sequences* (Lerner, 1996, 2004), *collaborative completions* (Clark, 1996; Poesio and Rieser, 2010), *co-constructions* (Sacks, 1992), *joint productions* (Helasvuo, 2004), *co-participant completions* (Hayashi 1999, Lerner and Takagi 1999), *collaborative productions* (Szczepek, 2000a) and *anticipatory completions* (Fox, 2007) amongst others (with some differences of emphasis in the different terms). Here we discuss some of this work.

2.1 Conversation Analysis

Anticipatory Completions Lerner (1991) identifies various structures typical of CCs which contain characteristic split points. One group of these are 'compound' *turn-constructional units* (TCUs), which are structures that include an initial constituent that hearers can identify as introducing some later final component. Examples include the IF X-THEN Y, WHEN X-THEN Y and INSTEAD OF X-Y constructions (4).

(4) **A:** Before that then if they were ill
G: They get nothing. *[BNC H5H 110-111]*

Other cues for potential *anticipatory completions* include quotation markers (e.g. SHE SAID), parenthetical inserts and lists, as well as non-syntactic cues such as contrast stress or prefaced disagreements. Another important category that he identifies is *terminal item* completions, which involve completing the final one or two lexical items of an interlocutor's utterance at projectable locations of the current speaker's turn ending (possibly involving overlap).

Opportunistic Cases Although Lerner focuses on these projectable turn completions, he also mentions that CCs can occur at other points such as "intra-turn silence", laugh tokens and hesitations, for example in cases of a stalled word search. All these cases he terms *opportunistic completions* (5).

(5) **A:** Well I do know last week thet=uh Al was certainly very ⟨pause 0.5⟩
B: pissed off *[Lerner (1996), p260]*

As he makes no claims regarding the frequency of such devices for CCs, it would be interesting to know how common these are, especially as studies on CCs in Japanese (Hayashi, 1999) show that although CCs do occur, compound TCUs do not play as prominent a role as in English. It should be noted, however, that Lerner's definitions are not intended to be mutually exclusive.

Expansions vs. Completions Other classifications of CCs often distinguish between *expansions* and *completions* (Ono and Thompson, 1993). Expansions are continuations which add, e.g., an adjunct, to an already complete syntactic element (6, 7).

(6) **T:** It'll be an E sharp.
G: Which will of course just be played as an F. *[BNC G3V 262-263]*

(7) **M:** yep dr goes everyones happy
N: except the dr *[DiET SU1 4213-4214]*

Completions involve the addition of syntactic material which is required to make the whole compound contribution (syntactically) complete (5, 8).

(8) **A:** . . . and then we looked along one deck, we were high up, and down below there were rows of, rows of lifeboats in case you see
B: There was an accident.
A: of an accident *[BNC HDK 63-65]*

Importantly, though we consider both expansions and completions to be CCs according to our terminology, we distinguish between the two types by considering the completeness or otherwise of the first part of the CC. Thus while there might be arguments for restricting the definition of a CC to only the completion type, we are also interested in comparing the relative distributions of the different sub-types.

In terms of frequency, the only estimate we are aware of in the CA literature is Szczepek (2000a), who found approximately 200 cross-person CCs in 40 hours of English conversation (there is no mention of the number of sentences or turns this equates to), of which 75% are completions.

As briefly outlined above, CA analyses of CCs tend to focus on their sequential implications in particular cases. These analyses provide clear examples of cross person co-ordination, however, it is unclear how representative they are (with the exception of Szczepek (2000a), who offers limited figures). Additionally, as the emphasis in the CA literature on CCs is in identifying their organisational consequences for the unfolding dialogue (which can range from indicating understanding to highlighting differences of opinion (Szczepek, 2000b)), they leave open the question of where a speaker switch may occur.

2.2 Linguistic Models

Purver et al. (2006) present a grammatical model for compound contributions, using an inherently incremental grammar formalism, Dynamic Syntax (Kempson et al., 2001; Cann et al., 2005). This model shows how syntactic and semantic processing can be accounted for no matter where the split point occurs; however, as their interest is in grammatical processing, they give no account of any higher-level inferences which may be required. Poesio and Rieser (2010) present a general model for *collaborative completions* (a subclass of cross-person CCs) based in the PTT framework, using an incremental LTAG-based grammar and an information-state-based approach to context modelling. While many parts of their model are compatible with a simple alignment-based communication model like Pickering and Garrod's (2004), they see intention recognition as crucial to dialogue management. They conclude that an intention-based model, like Clark's (1996), is more suitable. Their primary concern is to show how such a model can account for the hearer's ability to infer a suitable continuation, but their use of an incremental interpretation method also allows an explanation of the low-level utterance processing required. Nevertheless, the use of an essentially head-driven grammar formalism suggests that some syntactic split points ought to be more problematic than others.

2.3 Corpus Studies

Skuplik (1999) collected data from German two-party task-oriented dialogue, and annotated for cross-person compound contribution phenomena. She found that *expansions* (cases where the part before the split point can be considered already complete, as described above) were more common than *completions* (where the first part is syntactically or semantically incomplete as it stands), with 72 expansions (57%) and 54 completion CCs (43%) in her corpus. This contrasts with the data reported by Szczepek (2000a), detailed above. There are several possible reasons for this contrast; for example, there may simply be a difference in the distributions of CCs in different languages, or between experimentally controlled task-oriented dialogue (which Skuplik (1999) focused on) and casual conversational dialogue. Additionally, there may be issues with the classification schemes used. For example, Szczepek (2000a) did not include what she calls *appendor questions* in her data, which could also be argued to be expansion CCs. The corpus study here should shed some light on some of these possible sources of disagreement.

Rühlemann (2007) uses corpus analysis on the BNC to examine a subset of *expansion* CCs, *sentence relatives* of one's own or another's turn (6, 9).

(9) **A:** profit for the group is a hundred and ninety thousand pounds.
B: Which is superb. *[BNC FUK 2460-2461]*

He found that *sentence relatives* are slightly more likely to be same-person than cross-person, with a total of 104 (55%) of 190 being same-person cases. This contrasts with Tao and McCarthy (2001) who found 96% of their corpus sample were same-person; however, this discrepancy can be attributed to the fact that they were measuring different things: Tao and McCarthy (2001) included all non-restrictive (*'which'*) relative clauses in their analysis, thus excluding restrictive readings, and including cases which were intra-sentential and thus would not count as CCs in our terminology (see section 3.1). In fact, Rühlemann (2007) also excluded intra-turn cases where the sentence relative was annotated as a separate sentence but there was no intervening material; our definition would include these.

In addition, de Ruiter and van Dienst (in preparation) are also in the process of studying cross-person *completions* and their effect on the progressivity of dialogue turns; however no results are available to us at this point in time. Notably, the definition used by de Ruiter and van Dienst (pc) only includes those completions where the additional material combines with the incomplete first part of the CC such that neither part could be considered complete without the other. In our view, this excludes a number of interesting cases; not only expansion type CCs, but also those in which the continuation does not finish in a complete way (including, for example, CCs which spread over more than two parts).

2.4 Dialogue Models

Skantze and Schlangen (2009) and Buß et al. (2010) present incremental dialogue systems (for limited domains) which can deal with some kinds of same-person compound contribution, allowing the system or user to provide mid-sentence backchannels, and/or resume with sentence completion if interrupted. Some related empirical work regarding the issue of turn-switch addressed here is also presented by Schlangen (2006) but the emphasis there centers mostly on prosodic rather than grammar/theory-based factors.

For cross-person CCs, the only system we are aware of is that presented in DeVault et al. (2009) in which the system is able to generate a completion to a user's input based on the semantic representation it has built up so far. Due to the limited domain of possible semantic interpretations, the system is able to produce terminal item completions, once the possible interpretations have been sufficiently narrowed down. It does not, therefore, produce the range of CCs seen in naturally occurring human dialogue (including expansions as discussed above); we hope that empirical data such as that presented here can be used in constructing such systems and evaluating whether they achieve DeVault et al.'s stated aim of enabling virtual agents to display natural conversational behaviour.

3. General Methods

3.1 Terminology

In this paper, as our interest is general, we use the term **compound contributions** (*CCs*) to cover all instances where more than one dialogue contribution combine to form a (intuitively propositional) unit – whether the contributions are by the same or different speakers. We therefore use the term **split point** to refer to the point at which the compound contribution is split (rather than e.g. *transition point* which is associated with a speaker change). Cases where the speaker does change across the split point are called **cross-person** CCs; otherwise we call them **same-person** CCs.

As not all cases will lead to complete propositions, and not all will be split over exactly two contributions, we also avoid terms like *first-half*, *second-half* and *completion*: instead the contributions on either side of a split point will be referred to as the **antecedent** and the **continuation**. In cases where an compound contribution has more than one split point, some portions may therefore act as the continuation for one split point, and the antecedent for the next. We can then talk about **completeness** of each portion independently, with the traditional completion/expansion distinction corresponding to completeness (or otherwise) of the antecedent. See the sub-section on annotation scheme in section 4.1 for details of how completeness is assessed.

3.2 Questions

Questions about frequencies and distributions are addressed in the corpus study (section 4); these lead to others about the effects on the ongoing dialogue, which are examined in the experimental manipulation (section 5).

General Our first interest is in the general statistics regarding CCs: how often do they occur? When they do, do they usually fall into the specific categories (with specific preferred split points) examined by e.g. Lerner (1991), or can the split point be anywhere? What effects do CCs then have on the ongoing dialogue? Do same- and cross- person CCs have different effects? Specifically, do CCs have a bearing on 'party'-formation in Schegloff's (1995) sense, as Lerner (1991) claims?

Same- vs cross-person We are also interested in the balance between same- and cross-person CCs. Some grammatical formalisms (Purver et al., 2006) and psycholinguistic models (Pickering and Garrod, 2004) predict that CCs should be equally natural in both same- and cross- person conditions – is this the case? What are the similarities and differences between same- and cross-person cases?

Completeness For a grammatical treatment of CCs, as well as for implementing parsing/production mechanisms for their processing, we need to know about the likely completeness of antecedent and continuation (for example, if they are always complete in their own right, a standard head-driven grammar may be suitable; if not, something more fundamentally incremental may be required). In addition, CA analyses of dialogue phenomena predict that compound contributions should preferably occur at turn-transfer points that are foreseeable by the participants. Complete syntactic units serve this purpose from this point of view and lack of such completeness will seem to weaken this general claim.

We therefore ask how often antecedents and continuations are themselves complete. For antecedents, we are more interested in whether they *end* in a way that seems complete as they may have started irregularly due to overlap or another CC (**end-complete**); for continuations, whether they *start* in such a way – they may not get finished for some other reason, but we want to know if they would be complete if they do get finished (**start-complete**).[4] These notions are by no means entirely clear cut (as pointed out by an anonymous reviewer there is much debate on whether e.g. adverbial adjuncts and semantic roles are necessary in a sentence) and Schegloff (1996) concedes that his definitions are both arguable and not fully specified, although conversational participants do orient themselves to points of *possible* completion. In practice, however, in most cases there was a high level of agreement between annotators on what constitutes syntactic or semantic completeness.[5] We also look at the syntactic and lexical categories which occur either side of the split point. We are interested to know whether there are different effects on the unfolding dialogue from CCs with complete and incomplete antecedent contributions, and whether the position of the split point has an effect.

Repair and Overlap Finally, we look at how often the continuation of an CC involves explicit *repair* (repetition, reformulation, modification or replacement) of antecedent material. Any grammar of dialogue or computational system will need to be able to identify where this takes place, and we therefore also look at how such repair depends on antecedent completeness and the type of split point.

As our focus is on CCs, note that our use of `repair` refers only to those cases where the 'end' of the antecedent (immediately preceding the split point) is explicitly repeated or reframed at the start of the continuation. An example can be seen in (12), where the last word of the antecedent is repeated in the continuation. Repairs at other points in the s-unit or turn are not taken into consideration.[6]

4. Study 1: Corpus Study

4.1 Materials and Procedure

For this exercise we used the portion of the BNC (Burnard, 2000) annotated by Fernández and Ginzburg (2002), chosen to maintain a balance between what the BNC defines as context-governed dialogue (tutorials, meetings, doctor's appointments etc.) and demographic dialogue (casual unplanned conversations). This portion comprises 11,469 *s-units* – roughly equivalent to sentences[7] – taken from 200-turn sections of 53 separate dialogues.

The BNC transcripts are already annotated for overlapping speech, for non-verbal noises (laughter, coughing etc.) and for significant pauses. Punctuation is included, based on the original audio and the transcribers' judgements; as the audio is not available, we allowed annotators to use punc-

4. The notion of end-completeness that we are trying to capture is the CA notion of *endings* as outlined in Schegloff (1996); "for any TCU we can ask . . . does it end with an ending, i.e., does it come to a recognizable possible completion – syntactic, prosodic and action/pragmatic." Likewise his *beginnings* for our start-completeness; "Turn constructional units – and turns – can start with a "beginning" or with something which is hearably *not* a beginning."
5. See the sub-section on annotation scheme in section 4.1 for operational details, and table 2 for kappa agreement scores between annotators.
6. Consequently, our use of `repair` should be understood not as capturing all instances of repair but only as indexing the frequency with which these specific aspects of the contribution are repaired.
7. The BNC is annotated into *s-units*, defined as "sentence-like divisions of a text", and *utterances*, defined as "stretches of speech usually preceded and followed by silence or by a change of speaker". Utterances may consist of many s-units; s-units may not extend across utterance boundaries. While s-units are therefore often equivalent to complete syntactic sentences, or complete functional units such as bare fragments or one-word utterances, they need not be: they may be divided by interrupting or overlapping material from another speaker.

tuation where it aided interpretation. The BNC transcription protocol divides the transcript into sentence-like units (*"s-units"*) as well as speaker turns (*"utterances"* – see footnote 7), where utterances may contain several s-units from the same speaker. We annotated at the level of individual s-units, to allow self-continuations within a turn to be examined; we are therefore taking the BNC's *s-unit* to correspond to our notion of dialogue *contribution*, and the BNC's *utterance* as our notion of *turn*.

The BNC forces speaker turns to be presented in linear order, which is vital if we are to accurately assess whether turns are continuations of one another; however, this has a side-effect of forcing long turns to appear as several shorter turns when interrupted by intervening backchannels. We will discuss this further below.

Tag	Value	Explanation
`end-complete`	y/n	For all s-units: does this s-unit end in such a way as to yield a complete proposition or speech act?
`continues`	s-unit ID	For all s-units: does this s-unit continue the proposition or speech act of a previous s-unit? If so, which one?
`repairs`	number of words	For continuations: does the start of this continuation explicitly repair words from the end of the antecedent? If so, how many?
`start-complete`	y/n	For continuations: does this continuation start in such a way as to be able to stand alone as a complete proposition or speech act?

Table 1: Annotation Tags

Annotation Scheme The initial stage of manual annotation involved four tags: `end-complete`, `continues`, `repairs` and `start-complete` – these are explained in Table 1 above. S-units which somehow *require* continuation (whether they receive it or not) are therefore those marked `end-complete=n`; s-units which act as continuations are those marked with non-empty `continues` tags; and their antecedents are the values of those `continues` tags. Further specific information about the syntactic or lexical nature of antecedent or continuation could then be extracted (semi-) automatically, using the BNC transcript and part-of-speech annotations.

(3)

			end-complete	continues	repairs	start-complete
1.	A:	I were gonna say, they wash [[better than]]	n			
2.	J:	[[But I've had]]	n			
3.	A:	velvet.	y	1		n
4.	J:	I've had to take them up.	y	2	3	y
5.		Cos they were, they were gonna be miles too long.	y	4		n
6.		And I've not even took them out the thing.	y	5		n
7.		They said he'd swap them if they didn't fit.	y			
8.	A:	[[Ah they do!]]	y			
9.	J:	[[And he]] <pause><unclear>.	y	7		n
10.	A:	Where d'ya get them from Joyce?	y			
11.	J:	I got them from that er	n			
12.	B:	Top Marks.	y	11		n
13.	J:	that shop.	y	11	1	n

Returning to the extract in (3), repeated here, we can see how these tags are applied in practice. Note that all s-units have an `end-complete` tag whilst only those that are judged to continue some prior contribution have any other tags. The reason for judging end-completeness rather than whether the s-unit constitutes a complete proposition or speech act in its own right, is due to both the fragmentary nature of dialogue and the transcription practices of the BNC, which, as already discussed, may break up a syntactic sentence into several s-units due to overlapping material etc. Whether an s-unit *ends* in a potentially complete way is therefore independent of whether it *starts* in one. For the `continues` tag, the value is the line number which this s-unit is judged to continue (i.e. the line number of the antecedent); lines 12 and 13, for example, are both judged to be a continuation of line 11. The `repair` tag takes as its value (if it has one) the number of words from the end of the antecedent which are repeated, reformulated, modified or replaced at the start of the continuation. Line 4 has a `repair` value of 3, because the continuation repeats the three words from the end of line 2 (which is the antecedent) – *'I've had'*.[8] Finally, the `start-complete` tag (also only applied to continuations) indicates whether the contribution starts in a way that it might be the beginning of a complete sentence (even though it may not itself be complete). Continuations starting with *and/or/but/because* etc. are always tagged as `start-complete=n`, as can be seen in lines 5, 6 and 9.

Inter-Annotator Agreement In some cases, it is not easy to identify whether a fragment is a continuation or not, or what its antecedent is – see e.g. (10), where G's second contribution could be seen as continuing either his own prior utterance, or A's intervening contribution:

(10) **G:** Well a chain locker is where all the spare chain used to like coil up
A: So it ⟨unclear⟩ came in and it went round

8. *'I've'* is counted as two words as a contraction of *'I have'*.

G: round the barrel about three times round the barrel then right down into the chain locker but if you kept, let it ride what we used to call let it ride well ⟨unclear⟩ well now it get so big then you have to run it all off cos you had one lever, that's what you had and the steam valve could have all steamed. *[BNC H5G 174:176]*

Similar issues also arise in judgements of completeness, as it is not always obvious if a contribution is syntactically or semantically `end-` and/or `start-complete`. We therefore assessed inter-annotator agreement between the three authors who acted as annotators. First, all three annotated one dialogue independently, then compared results and discussed differences. They then annotated 3 further dialogues independently and agreement was measured; kappa statistics (Carletta, 1996) are shown in Table 2 below.

	BNC Dialogue Code		
Tag	KND	KBG	KB0
`end-complete`	.86-.92	.80-1.0	.73-.90
`continues` (y/n)	.81-.89	.76-.85	.77-.89
`continues` (ant)	.82-.90	.74-.85	.76-.86
`repairs`	1.0-1.0	.55-.81	1.0-1.0
`start-complete`	.59	.68	.62

Table 2: Inter-Annotator κ statistic (min-max)

With the exception of the `repairs` tag for one annotator pair for one dialogue and the `start-complete` tags, all are above 0.7; the low figure in the `repair` category results from a few disagreements in a dialogue with only a very small number of `repairs` instances. The `start-complete` kappa figures, between the two annotators who completed this task, are around 0.6 suggesting that this measure may be less easy to determine. The remaining dialogues were then divided evenly between the three annotators.

4.2 Results and Discussion

The 11,469 s-units annotated yielded 2,231 CCs, of which 1,902 were same-person and 329 cross-person cases; 112 examples involved an explicit repair by the continuation of the antecedent. The data come from the full range of dialogues; all dialogues had at least three same-person cases, though 5 of the 53 dialogues had no cross-person CCs. The mean number of same-person CCs is 35.89 per dialogue (standard deviation 22.46). For cross-person CCs the mean was 6.21 per dialogue (s.d. 5.69).

Within- and cross-turn cases Same-person CCs are much more common than cross-person; however, many of these same-person cases (around 44%) are self-continuations within a single speaker turn (such as those between lines 4 and 5 in (3)). As explained in section 3.2, we consider same-person cases to be interesting in their own right. From a processing/psycholinguistic point of view, we would like to know whether such split points occur in the same places in cross-person CCs as in same-person CCs. However, there are certainly arguments for considering CCs within a turn as single contributions, and including them when comparing the frequency or nature of same- and cross-person CCs may give an unfair comparison, as cross-person CCs can only occur at speaker turn boundaries.

In addition, some apparently cross-turn cases (around 17%) may in fact only appear as such due to the BNC transcription protocol, which forces speaker turns to be strictly linearly ordered. A sentence from a single speaker which is interrupted by material from another speaker will be transcribed as two separate turns – even if the intervening material is non-verbal (e.g. a cough) and/or entirely overlaps with the original sentence rather than actually interrupting its flow (as seen in (3) lines 1-3). In the tables and results below, we therefore present same-person CC figures both including all cases, and excluding those cases which are either within-turn or separated only by non-verbal or overlapping material. We label these figures as *all* and *cross-turn* respectively.

person:	Same-				Cross-	
	all		cross-turn		(all)	
	N	%	N	%	N	%
overlapping	0	*0*	0	*0*	18	*5*
adjacent	840	*44*	0	*0*	262	*80*
sep. by overlap	320	*17*	0	*0*	10	*3*
sep. by backchnl	460	*24*	456	*63*	17	*5*
sep. by 1 s-unit	239	*13*	229	*32*	16	*5*
sep. by 2 s-units	31	*2*	31	*4*	4	*1*
sep. by 3 s-units	5	*0*	3	*0*	1	*0*
sep. by 4 s-units	4	*0*	4	*1*	0	*0*
sep. by 5 s-units	1	*0*	1	*0*	0	*0*
sep. by 6 s-units	2	*0*	2	*0*	1	*0*
Total	1902		726		329	

Table 3: Antecedent/continuation separation

General Observations Looking at cross-turn cases, even excluding those within-turn and overlapping cases discussed above, there are over twice as many same-person CCs (726) as cross-person CCs (329). Many CCs have at least one s-unit intervening between the antecedent and continuation (see Table 3). In same-person cases, once we have excluded the within-turn CCs described above, this must in fact always be the case (see, for example, lines 11 and 13 in (3), where the contribution at line 12 means that the antecedent (line 11) and continuation (line 13) are non-adjacent); the intervening material is usually a backchannel (63% of remaining cases) or a single other s-unit (32%, often e.g. a clarification question), but two intervening s-units are possible (4%) with up to six being seen. In cross-person cases, 88% are adjacent or separated only by overlapping material, but again up to six intervening s-units were seen, with a single s-unit most common (10%, in half of which the intervening s-unit was a backchannel).

Many compound contributions have more than two separate contributions. In same-person cases, a CC can be split over as many as thirteen individual s-units; although such extreme cases occur generally within one-sided dialogues such as tutorials, many multi-split cases are also seen in general conversation. Only 63% of cases consisted of only two s-units. Antecedents can also receive more than one competing continuation (as in (3), where line 11 is continued in both lines 12 and 13), although this is rare: two continuations are seen in 2% of cases.

CA Categories We searched for examples which match CA categories (Lerner, 1991; Rühlemann, 2007) by looking for particular lexical items on either side of the split point. This search was per-formed in two stages: a loose (very high recall but low precision) automatic matching followed by manual checking to remove false positives (although some counts may still be slight over-estimates). For Lerner's (1996) *opportunistic* cases, we looked for filled pauses (*'er/erm'* etc.) or pauses explicitly annotated in the transcript ('`<pause>`'), so counts in this case may be underestimates if short pauses were not transcribed. We also chose some other broad categories based on our observations of the most common cases. Results are shown in Table 4 (where the || token represents the split point).[9]

person:	Same-				Cross-	
	all		cross-turn		(all)	
	N	*%*	N	*%*	N	*%*
... || and/but/or ...	748	*39*	306	*42*	116	*36*
... || so/whereas ...	257	*14*	57	*8*	39	*12*
... || because ...	77	*4*	32	*4*	3	*1*
... er/erm || ...	35	*2*	21	*3*	12	*4*
... <pause> || ...	19	*1*	15	*2*	20	*6*
... || which/who/etc ...	26	*1*	11	*2*	4	*1*
... instead of ... || ...	1	*0*	0	*0*	0	*0*
... said/thought/etc ... || ...	12	*1*	5	*1*	0	*0*
... if ... || (then) ...	18	*1*	10	*1*	2	*1*
... when ... || (then) ...	6	*0*	4	*1*	1	*0*
(other)	783	*41*	317	*44*	164	*50*
Total	1902		726		329	

Table 4: Continuation categories

The most common of the CA categories can be seen to be Lerner (1996)'s hesitation-related *opportunistic* cases, which make up 3-5% of same- and 10% of cross-person CCs, meaning cross-person opportunistic cases are more common than same-person ones (same (cross-turn; 36 of 726) vs other (32 of 329) $\chi^2_{(1)} = 8.53, p = 0.003$[10]). Interestingly, the breakdown of cases into those where the antecedent ends with an unfilled pause versus those which end with a filled pause also shows a difference between same- and cross-person cases: an other person is more likely to offer a continuation after an unfilled pause, than after a filled pause (antecedents ending in *'er(m)'* 35 continued by same, 12 by other; ending in '`<pause>`' 19 continued by same, 20 by other $\chi^2_{(1)} = 6.05, p = 0.01$). This finding backs up claims by Clark and Fox Tree (2002), that filled pauses can be used to indicate that the current speaker's turn is not yet finished and thus have the effect of holding the floor.

Lerner's compound TCU cases (*instead of, said/thought* etc, *if-then* and *when-then*) account for 2-3% of same-person and 1% of cross-person CCs, though note that these could be underestimates, as his non-syntactic cues (e.g. contrast stress and prefaced disagreements) could not be extracted. Rühlemann's (2007) *sentence relative* cases come next with over 1%.

9. Note that the categories in Table 4 are not all mutually exclusive (e.g. an example may have both an *'and'*-initial continuation and an antecedent ending in a pause), so column sums will not match totals shown.

10. For completeness, where $p > 0.001$, we report exact probabilities but throughout adopt a criterion probability level of < 0.05 for accepting or rejecting the null hypothesis.

In contrast, by far the most common pattern (for same- and cross-person CCs) is the addition of an extending clause, either a conjunction introduced by *'and/but/or/nor'* (36-42%), or other clause types with *'so/whereas/nevertheless/because'*. There are differences in the proportions of the clause types between same- and cross-person CCs, but further research and annotation is needed to confirm whether this represents systematic differences in pragmatic use (as in Rühlemann's (2007) sentence relative study, where cross-person CCs more often expressed stance (speaker opinion) than same-person CCs).

Split point Other less obviously categorisable cases make up 40-50% of continuations, in both same- and cross- person cases, with the most common first words being *'you'*, *'it'*, *'I'*, *'the'*, *'in'* and *'that'*. In terms of syntactic categories, manual examination of the data suggests that the split point can occur at any point between words,[11] even within what traditional theories of grammar consider to be a single constituent,[12] such as noun phrases and prepositional phrases (11, 12, 13, 14).

(11) **D:** Yeah I mean if you're looking at quantitative things it's really you know how much actual- How much variation happens whereas qualitative is ⟨pause⟩ you know what the actual variations
U: entails
D: entails. you know what the actual quality of the variations are. *[BNC G4V 114-117]*

(12) **M:** We need to put your name down. Even if that wasn't a
P: A proper conversation
M: a grunt. *[BNC KDF 25-27]*

(13) **A:** All the machinery was
G: [[All steam.]]
A: [[operated]] by steam *[BNC H5G 177-179]*

(14) **K:** I've got a scribble behind it, oh annual report I'd get that from.
S: Right.
K: And the total number of [[sixth form students in a division.]]

11. There is anecdotal evidence that CCs can also occur mid-word, as when someone completes a complex multi-syllabic word for another person. Only one of our cross-person CCs occurred mid-word (shown in (i), from a doctor/patient exchange), in which the whole word is also repeated, so we leave such considerations aside for now, though obviously they have implications for e.g. the organisation of the lexicon.

(i) **A:** No it wasn't Marvelon it was that Trin
D: Trin
A: Aye.
D: Trinordiol.
A: Mhm. *[BNC G58 63-68]*

12. Of course, different grammars may have different notions of constituency (such as the *surprising constituents* of CCG (Steedman, 2000)) which these findings may have a bearing on, however, for the purposes of the current discussion, we limit our notion of constituency to that of syntactic elements as in, for example transformational grammars, or HPSG.

S: [[Sixth form students in a division.]] Right.

[BNC H5D 123-127]

To further test the finding that the split point can apparently occur between any types of words, we annotated the *completion* cases for whether the split point occurred within a syntactic constituent, or between constituents.[13] For same-person cross-turn CCs, just over half are between-constituent (111/213; 52%), whilst cross-person CCs appear to be more likely to occur within-constituent although this trend is not significant (52/87; 60%; $\chi^2_{(1)} = 3.49, p = 0.06$). This finding appears to be associated with repair (there seem to be more repairs in the within-constituent cases) but the numbers are too small to be sure.

person:		Same- all		Same- cross-turn		Cross- (all)	
		N	%	N	%	N	%
Antecedent end-complete	Y	1367	*72*	513	*71*	242	*74*
	N	535	*28*	213	*29*	87	*26*
Continuation start-complete	Y	224	*12*	99	*14*	48	*15*
	N	1678	*88*	627	*86*	281	*85*
Repair	Y	77	*4*	34	*5*	32	*10*
	N	1825	*96*	692	*95*	297	*90*
Total		1902		726		329	

Table 5: Completeness and repair

Completeness Examination of the `end-complete` annotations shows that about 8% of s-units in general are incomplete (930/11469), but that (perhaps surprisingly) only 64% (591/930) of these get continued. This compares to 15% of end-complete s-units (1577/10539) that get continued ($\chi^2_{(1)} = 1315.90, p < 0.001$), showing that although incomplete s-units are more likely to be continued, incompleteness does not necessarily prompt the production of a completion.

The majority of both same- and cross-person continuations (71% to 74%) continue an already complete antecedent, with only 26-29% therefore being *completions* in the sense of e.g. Ono and Thompson (1993). Interestingly, though, continuations are no more likely than other s-units to end in a complete way themselves. In fact, continuations are significantly more likely than other s-units to end in an incomplete way (273/2231 (12%) vs. 657/9238 (7%); $\chi^2_{(1)} = 63.34, p < 0.001$).

The frequent clausal categories from Table 4 are all much more likely to continue complete antecedents than incomplete ones.[14] This is not the case for the `(other)` category; again suggesting that split points often occur at random points in a sentence, without regard to particular clausal constructions. The continuations in the `(other)` category are far less likely to continue complete antecedents than the easily classifyable categories from table 4 (220/481; 46% v. 535/574; 93%, $\chi^2_{(1)} = 289.76, p < 0.001$).

13. Here we are concerned with only low-level syntactic constituency; we counted a split point as within-constituent if it fell within a noun phrase (e.g. between a determiner and noun), a prepositional phrase (e.g. between a preposition and a noun phrase) or within a complex noun phrase (e.g. between an auxiliary and a head noun). Other cases (e.g. between a verb and its object, or between clauses) were coded as between-constituent.
14. For the less frequent (e.g. *'if/then'*, *'instead of'*) categories, the counts are too low to be sure.

Looking only at the general `(other)` category, we see that cross-person continuations more often follow antecedents that end in a complete way than same-person continuations (89/164; 54% v. 131/317; 41% $\chi^2_{(1)} = 7.30, p = 0.007$). For both cross-person and same person cases, continuations in the `(other)` category do not often start in a complete way (cross-person: 41/164; 25%, same-person 94/317; 30%).

In general, however, continuations are more than twice as likely to start in a non-complete rather than a complete way, even after complete antecedents.

Repair Explicit repair of the antecedent is not common, only occurring in just under 5% of CCs. As might be expected, incomplete antecedents are more likely to be repaired (cross-turn (same and cross-person); 51/300 17% vs. 15/755 2%, $\chi^2_{(1)} = 82.51, p < 0.001$). Cross-person continuations are also significantly more likely to repair their antecedents than same-person cases (32/329; 10% vs. 34/726; 5%, $\chi^2_{(1)} = 9.82, p = 0.002$).

In those CCs where the split point falls within a syntactic constituent, only 18% (18/102) of same-person cases involve explicit repair at the start of the continuation, compared to 27% (14/52) of cross-person CCs (the equivalent figures for CCs where the split point is between constituents are 12% (13/111) and 18% (6/35)). Although more data are required to see if these are genuine differences, we know that repair in general is not common, so it appears that even when the split point occurs mid-constituent, the participants generally are able to just go on extending the constituent as if they were the original speaker. This might suggest that the parsing and generation mechanisms are not required to back up to the beginning of a constituent in order to process or produce a continuation (i.e. start with a new grammar rule). This seems to favour lexicalised or dependency-based parsing models in that it suggests that the language processing mechanisms directly rely on word-by-word dependencies rather than constituents/grammar rules.

Function of CCs We are concerned in this study primarily with the form, rather than the function, of CCs. However, it is worth noting at this point that they can perform functions beyond merely extending or completing an interlocutor's contribution (see also Szczepek, 2000b); and in some cases are difficult to define functionally, and may even exhibit genuine *multifunctionality* (see e.g. Gregoromichelaki et al., 2009; Bunt, 2009). In (15), for example, J's continuation of M's utterance serves also as a request for confirmation:

(15) **M:** It's generated with a handle and
J: Wound round?
M: Yes *[BNC K69 109-112]*

In many cases, the antecedent explicitly invites the hearer to complete the contribution, so that antecedent and continuation form a question-answer pair, possibly within a single grammatical constituent (16):

(16) **J:** The Holy Spirit is the one who gives us hope.
Mega.
I mean ⟨pause⟩ this generation needs hope.
The Holy Spirit is one who ⟨pause⟩ gives us?
U: Strength.
J: Strength.

Yes, indeed.
⟨pause⟩ The Holy Spirit is one who gives us? ⟨pause⟩
U: Comfort.
Yes. *[BNC HDD 274-283]*

This phenomenon can also happen in cases of clarification-request/clarification-reply pairs (see Purver et al.'s (2003) *gap* category), e.g. (17):

(17) **G:** Cos they ⟨unclear⟩ they used to come in here for water and bunkers you see.
A: Water and?
G: Bunkers, coal, they all coal furnace you see, ⟨clears throat⟩ and we er they'd come in and we used to fill them up with coal, whatever they wanted ⟨cough⟩ lot of that went over the side ⟨unclear⟩ coal, beautiful coal that was. *[BNC H5H 59-61]*

With the range of possibilities regarding where the split point is able to occur, including potentially within a word (see footnote 11) it is hard to see how compound contributions could be characterised as a well-defined syntactic phenomenon, a separate grammatical fragment category, or a sub-class of *non-sentential utterance* (Fernández and Ginzburg, 2002). Moreover, there seems no reason to associate either antecedent or continuation with particular semantic categories or specific pragmatic speech-act information, as they seem to serve a wide range of purposes in dialogue: from assisting a speaker with lexical access, to eliciting a response to a query, to covertly offering a suggestion or asking a clarification.

Summary The results here show that CCs are common in dialogue. Split points may be possible at any syntactic point, but there appear to be (possibly pragmatic) constraints on where they are *likely* to appear: they are far more likely after complete antecedents, although relatively few of them occur in the highly projectable positions studied by e.g. Lerner.

There are interesting differences between same-person CCs and cross-person CCs; firstly, same-person CCs are over twice as common as cross-person. Cross-person continuations are more likely to start with explicit repair/reformulation of the antecedent; this might be considered surprising, as self-repair is preferred in general (Schegloff et al., 1977) although we have no comparable figures for repair at other points in the turn. However, it is interesting to note that a CC, in virtue of being constructed as a continuation of the speakers utterance, may provide a device that enables a less exposed form of other repair.

Outside the frequent clausal or CA categories, cross-person CCs are also more likely than same-person to continue a complete antecedent; and they are more likely where the antecedent ends in an unfilled pause rather than a filled one. This suggests an effect on turn-taking expectations, and that continuations may be systematically invited by a speaker or designed as though they are natural continuations of contributions that could be treated as complete. We will return to these points in the general discussion.

5. Study 2: Experimental Manipulation

While the corpus study of Section 4 provides us with useful information concerning the nature and frequency of CCs and their various sub-categories, it can tell us nothing about the effect of CCs on the dynamics of a conversation.

From a processing point of view, we might intuitively predict that cross-person CCs ought to be more difficult for a third party to process than same-person CCs, as information from potentially conflicting sources must be integrated and interpreted as a single syntactic unit. Conversely, some models (e.g. Dynamic Syntax, Cann et al. (2005)) would predict that there should be no additional processing costs.

The corpus study also suggests pragmatic effects associated with CCs. Are cross-person CCs indicative of particularly close coordination (and thus of Schegloff's 'parties' as Lerner suggests), which might facilitate understanding, or are they viewed as impolite which may add additional implications and disrupt the flow of the conversation?

The experiment reported here is, we believe, the first controlled manipulation of compound contributions during an unfolding interaction. The allows us to directly compare the effects of same-person and cross-person CCs on participants in a dialogue.

The effects of seeing a CC on a dialogue in progress were tested using the Dialogue Experimentation Toolkit (DiET) chat tool, which enables text dialogues to be experimentally manipulated (see Healey et al., 2003).

Of course, text based chat is different to face-to-face dialogue in several ways, and while clearly an interesting field of study in itself (Rosé et al. (2003), for example, compare text and speech based tutoring systems), there are important questions as to whether the results from our corpus study are generalisable to such a different modality. The most obvious differences are attributable to the channel of communication; speech versus text. In text-based chat such as MSN Messenger and the chat tool reported here, participants compose their turns in private before sending them to the other participants. This means that they can revise or even delete their turns without their interlocutors being aware of the revisions, unlike in face-to-face dialogue where overt repairs are necessarily shared. It also means that participants can compose their next turns simultaneously, meaning that the linearity of turn-taking in dialogue is lost. Linked to this is the fact that, unlike in face-to-face dialogue, participants engaged in a text chat are not typically co-present. Although this means that a number of non-linguistic cues are unavailable, this is also true in telephone conversations, for example, so should not be taken as a reason for rejecting the dialogic nature of text chat.

Despite these differences, there are also important similarities between text chat and face-to-face dialogue. Both involve the use of interlocutors' language resources to communicate, and text chat also exhibits many features which are generally seen in spoken dialogue, but not in either spoken monologue or written text. These include the use of non-sentential utterances such as clarification requests (Purver et al., 2003) and acknowledgements (Fernández and Ginzburg, 2002). Importantly for the study reported here, CCs also occur naturally in text-based chat (see, for example, (7) and (18), taken from the DiET chat tool environment).

(18) **U:** i agree tom needs to be there
A: but one of them has to go to save the other 2
R: and what about the cancer research plan ??

According to a preliminary corpus study (Eshghi, 2009) CCs occur as frequently in text chat as they do in face-to-face dialogue. In a total of 2377 text contributions, there were 493 CCs, of which 112 were cross-person and 381 were same-person CCs. Overall this proportion of CCs is not different to that from our BNC corpus study. In the text chat corpus there was a higher proportion of cross-person CCs than expected from our BNC results (112 out of 2377 versus 326 out of 11469 $\chi^2_{(1)} = 22.46, p < 0.001$) which could be related to the task-based nature of the text chat (the dialogues analysed in Eshghi (2009) are three-way tangram task conversations between two directors and a matcher), or possibly due to the way turns are transcribed into consecutive contributions in the BNC, as previously discussed.

5.1 Method

In the DiET chat tool, interventions can be introduced into a dialogue in real time, thus causing a minimum of disruption to the natural ‘flow’ of the conversation. In this experiment, a number of genuine single contributions in a text-based three-way conversation were artificially split into two parts. In some conditions, both parts still appeared to originate from the genuine source (“speaker”), thus appearing as a same-person CC. In other conditions, one or both parts seemed to come from another participant, thus appearing either as an cross-person CC, or as a same-person CC generated by the “wrong” person.

5.1.1 MATERIALS

The Balloon Task The *balloon task* is an ethical dilemma requiring agreement on which of three passengers should be thrown out of a hot air balloon that will crash, killing all the passengers, if one is not sacrificed. The choice is between a scientist, who believes he is on the brink of discovering a cure for cancer, a woman who is 7 months pregnant, and her husband, the pilot. This task was chosen on the basis that it should stimulate discussion, leading to dialogues of a sufficient length to enable an adequate number of interventions.

The DiET Chat Tool The DiET chat tool itself is a custom built Java application consisting of two main components: user interface and server console.

User interface The user interface is designed to look and feel like common instant messaging applications e.g. Microsoft Messenger. It consists of a display split into two windows, separated by a status bar, which indicates whether any other participant(s) are actively typing (see Figure 1). The ongoing dialogue, consisting of both the nickname of the contributor and their transmitted text, is shown in the upper window. In the lower window, participants type and revise their contributions, before sending them to their co-participants. All key presses are time-stamped and stored by the server.

Server Console All text entered is passed to the server, from where it is relayed to the other participants. No turns are transmitted directly between participants. Prior to being relayed, some turns are altered by the server to create fake CCs.

This is carried out automatically. A genuine single-person contribution is split around a space character near the centre of the string. The part of the turn before the space is relayed first, as the antecedent, followed by a short delay during which no other turns may be sent. This is followed by the continuation (the part of the turn after the space), as if they were in fact two quite separate, consecutive contributions. In every case, the server produces two variants of the compound contribution, relaying different information to both recipients. Each time an intervention is triggered, one of the two recipients receives a same-person CC from the *actual* source of the contribution (henceforth referred to as an *AA*-split). The other recipient receives one of three, more substantial, manipulations: a same-person CC that wrongly attributes both antecedent and continuation to the

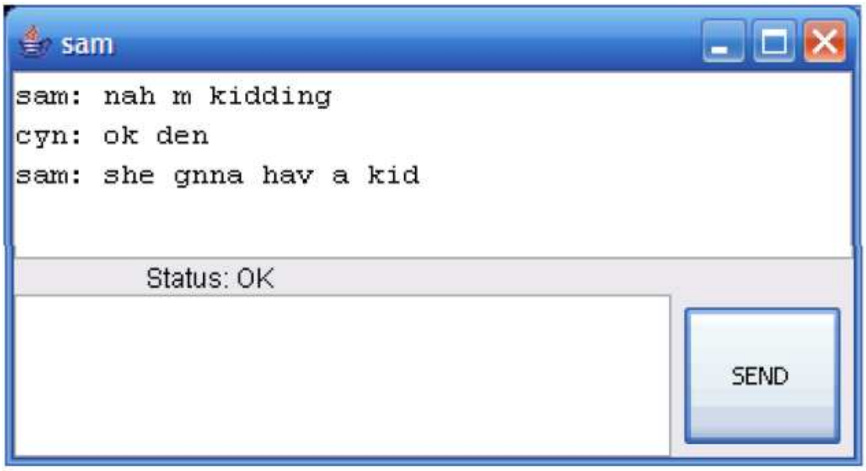

Figure 1: The user interface chat window (as viewed by participant 'sam')

other recipient (a *BB*-split); a cross-person CC whose antecedent comes from from the actual origin and continuation from the other recipient (an *AB*-split), or vice-versa (a *BA*-split).

This allows us to create a 2×2 factorial design which separates potential effects of 'floor change' i.e. whether the original speaker finishes the CC or another participant appears to, from effects of 'same/other' i.e. whether a the two halves of the CC or appear to be produced by the same speaker or by two different speakers. This contrast is shown in table 6.

A types:		
Should we start now		
B sees (AA intervention):		
A: Should we A: start now		
C sees (one of):		
AB intervention:	*BA intervention:*	*BB intervention:*
A: Should we B: start now	B: Should we A: start now	B: Should we B: start now

Table 6: Comparison of split types

The intervention is triggered every 10 turns, and restricted such that the participant who receives the non AA-split is rotated (to ensure that each participant only sees any of the more substantially manipulated interventions every 30 turns). Which of the three non AA-splits they see (AB, BA or BB) is, however, generated randomly.

5.1.2 SUBJECTS

41 male and 19 female native English speaking undergraduate students were recruited for the experiment, in groups of three to ensure that they were familiar with each other. All had previous experience of internet chat software such as Microsoft Messenger and each was paid £7.00 for their participation.

5.1.3 PROCEDURE

Each of the triad of subjects was sat in front of a desktop computer in separate rooms, so that they were unable to see or hear each other. Subjects were asked to follow the on-screen instructions, and input their e-mail address and their username (the nickname that would identify their contributions in the chat window). When they had entered these, a blank chat window appeared, and they were given a sheet of paper with the task description. Participants were instructed to read this carefully, and begin discussing the task with their colleagues via the chat window once they had done so. They were told that the experiment was investigating the differences in communication when conducted using a text-only interface as opposed to face-to-face. Additionally, subjects were informed that the experiment would last approximately 20-30 minutes, and that all turns would be recorded anonymously for later analysis. Once all three participants had been logged on, the experimenter went to sit at the server machine, a fourth desktop PC out of sight of all three subjects, and made no further contact until after at least 20 minutes of dialogue.

5.1.4 ANALYSIS

As production and receipt of contributions sometimes occurs in overlap in text chat, it is not possible to say definitively when one contribution is made in direct response to another.[15] We therefore chose to measure all the contributions produced by both recipients between the most recent intervention and the next intervention, averaged to produce one data point per recipient per intervention. This means that there are two data points for each intervention (one for each of the participants who saw a fake compound contribution).

The data were analysed according to two factors in a 2×2 factorial design; *same/other* – whether both parts of the compound contribution appeared to come from the same-person, or from different sources ([AA and BB] vs [AB and BA]), and *floor change* – whether the continuation part of the CC appeared to come from the genuine source or the other participant ([AA and BA] vs [AB and BB]), with participant as a random factor.

Measures selected for analysis were *typing time of turn* (the time, in milliseconds, between the first key press in a turn and sending the turn to the other participants by hitting the return key) and *length of turn in characters* as measures of production; *deletes per character* (the number of keyed deletes divided by the total number of characters) as a measure of revisions; and *typing time per character* as a measure of speed. Data in tables are displayed in the original scale of measurement. However, as inspection of the data showed that they were not normally distributed, logarithmic

15. In online chat, participants can compose their next contributions simultaneously, and contributions under construction when another is received can be subsequently revised, prior to transmission. This means that a genuine response to a compound contribution might have a negative start time. However, the inclusion of cases where the whole contributions was constructed after receiving the CC (an arbitrary cut-off point, which would catch some contributions that were responses to earlier contributions in the dialogue, and miss some which were begun before the intervention was received and subsequently revised) should impose the same level of noise in all cases.

transformations (using log_e) were applied to the typing time of turn and length of turn in characters measures prior to all inferential statistical analyses, resulting in data distributions that were not significantly different from a normal distribution (using Shapiro-Wilk tests: typing time of turn $W = 0.998, p = 0.882$; length of turn in characters $W = 0.995, p = 0.100$). For the proportional measures of deletes per character and typing time of character, which violate normality assumptions even after transformations, alternative analyses were used.

The Generalized Linear Model (GZLM) extends the General Linear Model (GLM; which includes ANOVAs and linear regression models) to include response variables that follow any exponential probability distribution, including e.g. poisson, binomial and gamma distributions. GZLMs use maximum likelihood estimation to fit the model to the data (and provide parameter estimates). Generalized Estimating Equations (GEE) extend GZLM further by allowing for non-independent data, such as repeated measures and clustered data. Using a GEE analysis (see Liang and Zeger, 1986; Ballinger, 2004) on these variables therefore allows for both the non-normality of the data, and within-subject correlations.

5.2 Results

A post-experimental questionnaire and debriefing showed that, with the exception of one subject, who had taken part in a previous chat tool experiment and was therefore aware that manipulations may occur, none of the participants were aware of any interventions.

Of the 253 interventions to which at least one recipient responded, 89 were AA/AB splits, 99 were AA/BA splits and 65 AA/BB splits. This means there were 506 potential responses, however, in 16 cases, only one of the recipients produced a response, leaving 490 data points. Table 7 shows the actual n values in each case.

Condition	Typing time / turn (ms) Mean (s.d.)		Typing time / char (ms) Mean (s.d.)		N
AA	11122.27	(14413.5)	475.45	(558.92)	246
AB	12500.98	(10944.6)	523.56	(1036.00)	89
BA	9800.77	(8810.3)	357.76	(316.15)	92
BB	11561.67	(10138.4)	479.51	(396.07)	63

Table 7: Typing time of turn and typing time per character by type of intervention

2×2 ANOVAs (with participant as a random effect)[16] show a significant main effect of floor change[17] on the log transformed typing time of turn (see table 7), with participants taking longer over their turns in the AB and BB conditions ($F_{(1,288)} = 6.563$, $p = 0.012$). There was no main effect of same/other ($F_{(1,288)} = 0.001$, $p = 0.980$), and no effect of interaction ($F_{(1,288)} = 1.259$, $p = 0.270$), though there was a main effect of participant ($F_{(59,288)} = 4.565$, $p = 0.008$) showing that there was high individual variation for this measure.

16. We account for between subject variation by including subject as a random factor, meaning that there is more than one datapoint per subject (and, in effect, a $2 \times 2 \times 60$ model). There are 490 datapoints between 60 subjects. As we carried out a full factorial model, the numerator (error) degrees of freedom that resulted from this model was 288.
17. A significant effect is one in which the p-value, the probability of the observed data being sampled if the null hypothesis were true, is below some criterion value. We adopt the criteria of significance as lower than 5% probability.

There were no significant effects on length of turn in characters (same/other, $F_{(1,288)} = 1.709$, $p = 0.194$, floor change $F_{(1,288)} = 0.194$, $p = 0.341$).

A 2×2 GEE model with participant as a subject effect (using the gamma distribution, goodness of fit quasi log-likelihood (QIC) = 149.482)[18] showed a marginally significant main effect of floor change on typing time per character (Model effect; Wald-$\chi^2 = 3.820$, $p = 0.051$. Parameter estimate; $B = -0.281$, Wald-$\chi^2 = 7.192$, $p = 0.007$) and a main effect of participant (Wald-$\chi^2 = 258468$, $p < 0.001$). There was no main effect of same/other, and no interaction effects.

For deletes per character, a 2×2 GEE model with participant as a subject effect (using the negative binomial distribution,[19] goodness of fit quasi log-likelihood (QIC) = 566.574) showed a significant main effect of same/other (Model effect; Wald-$\chi^2 = 9.617$, $p = 0.002$. Parameter estimate; $B = -0.492$, Wald-$\chi^2 = 12.226$, $p < 0.001$) and a main effect of participant (Wald-$\chi^2 = 487986$, $p < 0.001$). There was no main effect of floor change, and no interaction effects.

Condition	Mean (s.d.) (ms/char)	
AA	0.108	(0.16)
AB	0.094	(0.13)
BA	0.071	(0.10)
BB	0.138	(0.17)

Table 8: Deletes per character by type of intervention

As the experiment was looking for generic effects of CCs on the dialogue, the location of the split points was arbitrary. In order to test for effects of split point, post-hoc analyses were carried out to ascertain whether other observed contrasts in the corpus had any effects on processing of apparent CCs. The fake CCs were coded according to three factors; standalone coherence (as judged by the authors) of the antecedent and continuation[20] (see table 9) and whether the split point fell within or between a syntactic constituent. There were no effects of first or second half coherence on any of the variables, and no interaction effects. There were also no main effects of whether the split point fell within or between a constituent; (log transformed) typing time of turn ($F_{(1,204)} = 0.262$, $p = 0.435$); (log transformed) number of characters ($F_{(1,204)} = 1.760$, $p = 0.189$); typing time per character (Wald-$\chi^2 = 0.550$, $p = 0.458$) deletes per character (Wald-$\chi^2 = 0.285$, $p = .594$) and no interaction effects with same/other or floor change. These results are consistent with the finding from the corpus that the split point may be able to occur anywhere syntactically, though the lack of any observed effects could be due to low power caused by the relatively small numbers of some groups.

18. The model distributions were chosen on the basis of being the best fit to the data, as indicated by the lowest quasi log-likelihood score.
19. Each key press can be seen as a delete or not-a-delete.
20. These judgements are simply a yes/no answer to the question 'could this contribution be interpreted as complete in its own right?', i.e. analogous to the *end-complete* and *start-complete* annotation tags in the corpus study, such that an antecedent (first part) judged to be able to stand alone can be considered end-complete and a continuation (second part) judged to be able to standalone can be considered start-complete. The difference in tagging conventions is due to the fact that in the chat tool environment turns can be revised prior to sending, and therefore might be considered to be a unit by its sender, even if the fractured nature of text chat means that it might not constitute a syntactically complete sentence.

Part of CC		Coherent		N
First	Second	1st	2nd	
what the hell	is that	Y	N	131
the woman is pregnant	she should stay	Y	Y	52
these people said	you did something	N	Y	43
I think this is also	the wish of the doctor	N	N	264

Table 9: Examples of standalone coherence judgement examples

5.3 Discussion

Given the novelty of the method and the lack of other experimental studies of CC's to cross-check against, the results of this experiment must be interpreted with caution. Nonetheless, we believe that the results summarised in table 10, below, do bear on the questions raised in section 3.2.

Effect of	Dependent variable and direction of effect
Floor Change	Typing Time (per turn and char) $(AB \wedge BB) > (AA \wedge BA)$
Same/Other	Deletes $(AA \wedge BB) > (AB \wedge BA)$

Table 10: Summary of significant effects

Firstly, it is important to note that introducing fake CCs did have measurable effects on the ongoing dialogue, despite participants being unaware of either the intervention, or their effects. This in itself might be seen as surprising, as if the intervention were highly disruptive, we would presumably expect subjects to notice it.

Though typing time is a fairly crude measure[21] one possible explanation for participants taking longer over the production of a turn (independently of length of turn in characters) could be due to problems arising in the local organisation of turn-taking (Sacks et al., 1974). A participant who has seen a floor change intervention (Participant C) may be taking longer over their turns because there is less pressure on them to take a turn. C will falsely believe that the fake source (Participant B) has just completed a turn, and will therefore not expect them to take the floor. Additionally, the genuine source (Participant A) will not be taking the floor because they have just completed a turn (though C does not know this). However, this effect of floor change could also be due to the confounding fact that when one of the recipients sees a floor change CC, and the other recipient (as always) sees an AA-split, the two are left with different impressions about who made the final contribution (i.e. the continuation part of the fake CC) and thus have potentially conflicting expectations regarding who is entitled to speak next. Whether or not these explanations are correct, the effect does suggest that at some level participants are sensitive to specific interlocutors – note that the difference cannot be

21. For example, the additional typing time may fall at the end of a turn (before pressing enter) suggesting that participants are reviewing their responses more carefully before sending them, or it may be a general effect spread evenly across the turn.

simply attributable to a mismatch between who appears to be speaking and what sort of thing they would say because then we would expect turns following the BA intervention to be equally affected.

Independently of a change of floor, seeing a CC that appears to be shared between speakers also has an impact on the conversation, seen in the amount of revision undertaken in formulating responses (deletes). Perhaps surprisingly, in this case, participants who have seen a CC that was apparently co-constructed by both their interlocutors revise their turns less than after a same-person CC. One reason why participants might worry less about precisely formulating their turns following a cross-person CC is that it could have the effect on the recipient of suggesting that the two other participants are highly coordinated. One possible interpretation of this could be that they have formed a 'party' (Schegloff, 1995) with respect to the decision of who to throw out of the balloon. This might be understood as signalling the formation of a strong coalition between the other two participants, making the recipient behave as though they are resigned to the decision of this coalition. (19), taken from the transcripts shows an example where this appears to be the case (the 'fake' part of the CC is shown in bold).

(19) *AB-Split showing apparent coalition between 'B' and 'D'*

B: and he can tell his formula

D: **to tom and susie**

Note that this is not the same as the effect on the typing time of turn, whereby participants are less rushed when seeing a change of floor. Deletes, in contrast, indicate how carefully participants are constructing their turns.

6. General Discussion and Conclusions

As discussed in the introduction, CCs are of interest for many different groups of researchers. Our corpus study shows that nearly one fifth of all contributions in naturally occurring dialogue continue some previous contribution, indicating the scale of the phenomenon. Just the sub-set of cross-person CCs accounts for 3% of all dialogue contributions, comparable to the frequency of clarification requests (see Purver et al., 2003; Rodríguez and Schlangen, 2004), widely studied by dialogue theorists (e.g. Ginzburg and Cooper, 2004).

Although most of the categories of CC described by conversation analysts appear, these categories do not correspond to the most frequent in the BNC, and do not adequately characterise all the CCs observed in the present analysis. The corpus results show no evidence that syntax places significant constraints on where a split point can occur and the experimental results are consistent with this. Participants were able to process and interpret fake CCs successfully despite their arbitrary split points and were also not explicitly aware of the experimental manipulation. This is consistent with models that advocate highly coordinated resources between interlocutors and, moreover, the need for incremental means of processing that operate on at least a word-by-word basis (Purver et al., 2006; Skantze and Schlangen, 2009).

However, that CCs may be able to occur anywhere in a syntactic sequence is not to say that they necessarily or usually do. Both the corpus study, in which cross-person CCs occur more frequently after an unfilled (rather than a filled) pause and more often follow an end-complete antecedent than same-person CCs (in the not obviously classifiable cases), and the experiment, in which confounding expectations lead to additional response time, suggest that conversational expectations (including for example around turn-taking) play some role above and beyond grammatical/linguistic resources.

	All CCs				Other person CCs			
	BNC		Experiment		BNC		Experiment	
Y	1609	*72%*	183	*37%*	242	*74%*	69	*38%*
N	622	*28%*	307	*63%*	87	*26%*	112	*62%*
Total	2231		490		329		181	

Table 11: Antecedent end-completeness: Comparison of the distribution of actual (corpus) and arbitrary (experiment) split points

These considerations are backed up by the data shown in Table 11. This table shows the distribution of antecedent end-completeness in the annotated corpus, compared to the distribution obtained in the experiment. As can be clearly seen, when strings are artificially split in a arbitrary fashion, the 'antecedent' is far less likely to end in a complete way than actually occurs in the genuine CCs in the corpus. This suggests that continuations are systematically designed (and their split points chosen) as extensions of contributions that could be treated as already complete.

In terms of the effects that CCs have on the ongoing dialogue, the experiment suggests that Lerner's hypothesis that cross-person CCs might demonstrate party membership may well be correct; it also clearly demonstrates that the participants in a dialogue are not interchangeable. These results do not, of course, prejudice the claim that, at a purely mechanistic level, people could anticipate the structures needed to complete a turn (as the interactive alignment model suggests); they do not tell us about the actual production of compound contributions, but rather about the effect they have on the conversation, so do not provide unequivocal evidence in support of one theory over another. They do however indicate that if we wish to treat a jointly produced CC as signalling especially strong alignment, then we need to examine factors other than simply syntax.

They also offer some interesting pointers for further research: if parties are genuine conversational entities, then we might expect dialogue phenomena to have distributional patterns which reflect this. Consistent with our corpus results as regards other-person repair, for example, is the speculative hypothesis that people might be structuring continuations precisely as the preferred 'self'-repairs, where 'self' can be taken to mean within-party.

From a computational modelling point of view, there is some good news: as start-completeness of continuations is rare, a dialogue system may have a chance of detecting continuations from surface characteristics of the input (though note that we did not investigate the general prevalence of *start*-incomplete s-units in the corpus). There is bad news too, though: as there do not seem to be strict syntactic restrictions on where the split point can occur, there may be no grammatical features that can be reliably employed to this end. In addition, antecedents do not end in an incomplete way as commonly as might be expected, and long distances between antecedent and continuation are possible. Detecting continuations and locating their antecedents is therefore unlikely to be a straightforward task for automated systems.

6.1 Further Work

Split point For implementational purposes, additional corpus analysis needs to be carried out regarding the distribution of the different syntactic points at which CCs can and do appear; further experiments are also planned which examine the effect on processing of manipulating the syntactic position of the split point (for example, inserting splits before or after determiners).

Continuation form Further corpus analysis is also required to investigate any systematic differences in the *form* of continuations, and the interaction of this with the properties of the antecedent and conversational genre (including the BNC's "context-governed" or "demographic" face-to-face dialogues and text-based chat).

Ownership A further interesting question regards who can be said to take responsibility for, or 'own' a jointly produced CC. Insights into this might come from Lerner (2004), which discusses CCs that occur in *collaborative turn sequences*: if the original speaker maintains authority over the content of the collaboratively constructed contribution, they may ratify another speaker's continuation (for example by repeating or acknowledging it), but may alternatively strategically offer their own (*delayed completion*, Lerner, 1996). Party-membership, or otherwise, may be influenced by such additional conversational contributions.

Character-by-character experiments Due to the design of the experiment, the floor change effects might, as discussed, be because in floor change cases the two recipients will have been left with the impression that a different person made the final contribution. This means that there may be a an effect of confounded listener expectation (though see Schober and Brennan, 2003, for discussion) – though note that this does not have any bearing on the observed differences on deletes after a cross-person CC. Because of this, and the already noted potential problems of linearity in text-based chat, a follow-up study using a character-by-character chat tool interface is already under way. This more directly enforces turn-taking, as it does not allow participants to formulate their turn before communicating it; each character is transmitted as and when it is entered.

References

G.A. Ballinger. Using generalized estimating equations for longitudinal data analysis. *Organizational Research Methods*, 7(2):127–150, 2004.

Harry Bunt. Multifunctionality and multidimensional dialogue semantics. In *Proceedings of DiaHolmia, 13th SEMDIAL Workshop*, 2009.

Lou Burnard. *Reference Guide for the British National Corpus (World Edition)*. Oxford University Computing Services `http://www.natcorp.ox.ac.uk/docs/userManual/`, 2000. URL `http://www.natcorp.ox.ac.uk/docs/userManual/`.

Okko Buß, Timo Baumann, and David Schlangen. Collaborating on utterances with a spoken dialogue system using an ISU-based approach to incremental dialogue management. In *Proceedings of the SIGDIAL 2010 Conference*, pages 233–236, Tokyo, Japan, September 2010. Association for Computational Linguistics. URL `http://www.sigdial.org/workshops/workshop11/proc/pdf/SIGDIAL42.pdf`.

Ronnie Cann, Ruth Kempson, and Lutz Marten. *The Dynamics of Language*. Elsevier, Oxford, 2005.

Jean Carletta. Assessing agreement on classification tasks: The kappa statistic. *Computational Linguistics*, 22(2):249–255, 1996.

Herbert H. Clark. *Using Language*. Cambridge University Press, 1996.

Herbert H. Clark and Jean E. Fox Tree. Using *uh* and *um* in spontaneous speaking. *Cognition*, 84 (1):73–111, 2002.

Jan Peter de Ruiter and Mattia van Dienst. Completing other people's utterances: evidence for forward modeling in conversation. ms., in preparation.

David DeVault, Kenji Sagae, and David Traum. Can I finish? learning when to respond to incremental interpretation results in interactive dialogue. In *Proceedings of the SIGDIAL 2009 Conference*, pages 11–20, London, UK, September 2009. Association for Computational Linguistics. URL `http://www.aclweb.org/anthology/W/W09/W09-3902`.

Arash Eshghi. *Uncommon ground: The distribution of dialogue contexts*. PhD thesis, School of Electronic Engineering and Computer Science, Queen Mary University of London, 2009.

Raquel Fernández and Jonathan Ginzburg. Non-sentential utterances: A corpus-based study. *Traitement Automatique des Langues*, 43(2), 2002.

Victor Ferreira. Is it better to give than to donate? Syntactic flexibility in language production. *Journal of Memory and Language*, 35:724–755, 1996.

Barbara A. Fox. Principles shaping grammatical practices: an exploration. *Discourse Studies*, 9(3): 299, 2007.

Andrew Gargett, Eleni Gregoromichelaki, Ruth Kempson, Matthew Purver, and Yo Sato. Grammar resources for modelling dialogue dynamically. *Cognitive Neurodynamics*, 3(4):347–363, 2009. ISSN 1871-4080. URL `http://dx.doi.org/10.1007/s11571-009-9088-y`.

Jonathan Ginzburg and Robin Cooper. Clarification, ellipsis, and the nature of contextual updates in dialogue. *Linguistics and Philosophy*, 27(3):297–365, 2004.

Eleni Gregoromichelaki, Yo Sato, Ruth Kempson, Andrew Gargett, and Christine Howes. Dialogue modelling and the remit of core grammar. In *Proceedings of IWCS*, 2009.

Eleni Gregoromichelaki, Ruth Kempson, Matthew Purver, Greg J. Mills, Ronnie Cann, Wilfried Meyer-Viol, and Pat Healey. Incrementality and intention-recognition in utterance processing. *Dialogue and Discourse*, to appear, 2011. to appear.

Markus Guhe. *Incremental Conceptualization for Language Production*. NJ: Lawrence Erlbaum Associates, 2007.

Makoto Hayashi. Where Grammar and Interaction Meet: A Study of Co-Participant Completion in Japanese Conversation. *Human Studies*, 22(2):475–499, 1999.

Patrick Healey, Matthew Purver, James King, Jonathan Ginzburg, and Greg Mills. Experimenting with clarification in dialogue. In *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*, Boston, Massachusetts, August 2003.

Marja-Liisa Helasvuo. Shared syntax: the grammar of co-constructions. *Journal of Pragmatics*, 36 (8):1315–1336, 2004.

Ruth Kempson, Wilfried Meyer-Viol, and Dov Gabbay. *Dynamic Syntax: The Flow of Language Understanding*. Blackwell, 2001.

Gene H. Lerner. On the syntax of sentences-in-progress. *Language in Society*, pages 441–458, 1991.

Gene H. Lerner. On the "semi-permeable" character of grammatical units in conversation: Conditional entry into the turn space of another speaker. In Elinor Ochs, Emanuel A. Schegloff, and Sandra A. Thompson, editors, *Interaction and grammar*, pages 238–276. Cambridge University Press, 1996.

Gene H. Lerner. Collaborative turn sequences. In *Conversation analysis: Studies from the first generation*, pages 225–256. John Benjamins, 2004.

Gene H. Lerner and Tomoyo Takagi. On the place of linguistic resources in the organization of talk-in-interaction: A co-investigation of English and Japanese grammatical practices. *Journal of Pragmatics*, 31(1):49–75, 1999.

Kung-Yee Liang and Scott L. Zeger. Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1):13, 1986. ISSN 0006-3444.

Tsuyoshi Ono and Sandra A. Thompson. What can conversation tell us about syntax. In P.W. Davis, editor, *Alternative Linguistics: Descriptive and Theoretical Modes*. Benjamin, 1993.

Martin Pickering and Simon Garrod. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27:169–226, 2004.

Massimo Poesio and Hannes Rieser. Completions, coordination, and alignment in dialogue. *Dialogue and Discourse*, 1:1–89, 2010. ISSN 2152-9620. URL `http://elanguage.net/journals/index.php/dad/article/view/91/512`.

Matthew Purver, Jonathan Ginzburg, and Patrick Healey. On the means for clarification in dialogue. In R. Smith and J. van Kuppevelt, editors, *Current and New Directions in Discourse & Dialogue*, pages 235–255. Kluwer Academic Publishers, 2003.

Matthew Purver, Ronnie Cann, and Ruth Kempson. Grammars as parsers: Meeting the dialogue challenge. *Research on Language and Computation*, 4(2-3):289–326, 2006.

Kepa Rodríguez and David Schlangen. Form, intonation and function of clarification requests in German task-oriented spoken dialogues. In *Proceedings of the 8th Workshop on the Semantics and Pragmatics of Dialogue (SEMDIAL)*, Barcelona, Spain, July 2004.

Carolyn P. Rosé, Diane Litman, Dumisizwe Bhembe, Kate Forbes, Scott Silliman, Ramesh Srivastava, and Kurt VanLehn. A comparison of tutor and student behavior in speech versus text based tutoring. In *Proceedings of the HLT-NAACL 03 workshop on Building educational applications using natural language processing-Volume 2*, page 37. Association for Computational Linguistics, 2003.

Christoph Rühlemann. *Conversation in context: a corpus-driven approach*. Continuum, 2007.

Harvey Sacks. *Lectures on Conversation*. Blackwell, 1992.

Harvey Sacks, Emanuel A. Schegloff, and Gail Jefferson. A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4):696–735, 1974.

Emanuel A. Schegloff. Parties and talking together: Two ways in which numbers are significant for talk-in-interaction. *Situated order: Studies in the social organization of talk and embodied activities*, pages 31–42, 1995.

Emanuel A. Schegloff. Turn organization: One intersection of grammar and interaction. In Elinor Ochs, Emanuel A. Schegloff, and Sandra A. Thompson, editors, *Interaction and grammar*, pages 52–133. Cambridge University Press, 1996.

Emanuel A. Schegloff, Gail Jefferson, and Harvey Sacks. The preference for self-correction in the organization of repair in conversation. *Language*, 53(2):361–382, 1977.

David Schlangen. From reaction to prediction: Experiments with computational models of turn-taking. In *Proceedings of the 9th International Conference on Spoken Language Processing (INTERSPEECH - ICSLP)*, Pittsburgh, PA, September 2006.

Michael F. Schober and Susan E. Brennan. Processes of interactive spoken discourse: The role of the partner. *Handbook of discourse processes*, pages 123–64, 2003.

Gabriel Skantze and David Schlangen. Incremental dialogue processing in a micro-domain. In *Proceedings of the 12th Conference of the European Chapter of the ACL (EACL 2009)*, pages 745–753, Athens, Greece, March 2009. Association for Computational Linguistics. URL `http://www.aclweb.org/anthology/E09-1085`.

Kristina Skuplik. Satzkooperationen. definition und empirische untersuchung. SFB 360 1999/03, Bielefeld University, 1999.

Mark Steedman. *The Syntactic Process*. MIT Press, Cambridge, MA, 2000.

Beatrice Szczepek. Formal Aspects of Collaborative Productions in English Conversation. *Interaction and Linguistic Structures (InLiSt)*, `http://www.uni-potsdam.de/u/inlist/issues/17/`, 2000a.

Beatrice Szczepek. Functional Aspects of Collaborative Productions in English Conversation. *Interaction and Linguistic Structures (InLiSt)*, `http://www.uni-potsdam.de/u/inlist/issues/21/`, 2000b.

Hongyin Tao and Michael J. McCarthy. Understanding non-restrictive which-clauses in spoken English, which is not an easy thing. *Language Sciences*, 23(6):651–677, 2001.