

**Coordinating in dialogue:  
Using compound contributions to join a party**

Christine Howes

Submitted for the degree of Doctor of Philosophy

Queen Mary University of London

2012

**Declaration**

I hereby declare that the work presented in this thesis is my own work carried out under normal terms of supervision and that the research reported here has been conducted by myself unless otherwise indicated.

**Christine Howes**

**London, 4th January 2011**

# Coordinating in dialogue: Using compound contributions to join a party

Christine Howes

## Abstract

*Compound contributions (CCs)* – dialogue contributions that continue or complete an earlier contribution – are an important and common device conversational participants use to extend their own and each other’s turns. The organisation of these cross-turn structures is one of the defining characteristics of natural dialogue, and *cross-person* CCs provide the paradigm case of coordination in dialogue.

This thesis combines corpus analysis, experiments and theoretical modelling to explore how CCs are used, their effects on coordination and implications for dialogue models. The syntactic and pragmatic distribution of CCs is mapped using corpora of ordinary and task-oriented dialogues. This indicates that the principal factors conditioning the distribution of CCs are pragmatic and that same- and cross-person CCs tend to occur in different contexts.

In order to test the impact of CCs on other conversational participants, two experiments are presented. These systematically manipulate, for the first time, the occurrence of CCs in live dialogue using text-based communication. The results suggest that syntax does not directly constrain the interpretation of CCs, and the primary effect of a cross-person CC on third parties is to suggest to them a strong form of coordination or coalition has formed between the people producing the two parts of the CC.

A third experiment explores the conditions under which people will produce a completion for a truncated turn. Manipulations of the structural and contextual predictability of the truncated turn show that while syntax provides a resource for the construction of a CC it does not place significant constraints on where the split point may occur. It also shows that people are more likely to produce continuations when they share common ground. An analysis using the Dynamic Syntax framework is proposed, which extends previous work to account for these findings, and limitations and further research possibilities are outlined.

Submitted for the degree of Doctor of Philosophy

Queen Mary University of London

2012

# Contents

<b>1</b>	<b>Introduction</b>	<b>15</b>
<b>2</b>	<b>Compound Contributions</b>	<b>20</b>
2.1	Conversation Analysis . . . . .	20
2.1.1	Anticipatory completions . . . . .	21
2.1.2	Opportunistic cases . . . . .	22
2.1.3	Expansions vs. completions . . . . .	22
2.1.4	Functions / effects of CCs . . . . .	24
2.1.5	Ratification / rejection . . . . .	28
2.1.6	Implications of CCs on CA analyses . . . . .	29
2.1.7	Summary . . . . .	30
2.2	Corpus studies . . . . .	30
2.3	Models of communication . . . . .	32
2.3.1	Transmission models of communication . . . . .	33
2.3.2	Inferential communication . . . . .	35
2.3.3	Co-constructing communication . . . . .	39
2.4	Coordination . . . . .	41
2.4.1	Interactive Alignment . . . . .	41
2.4.2	Common Ground . . . . .	47
2.4.3	Interactive Alignment versus Common Ground . . . . .	49
2.5	Linguistic models . . . . .	50
2.5.1	Incrementality . . . . .	51
2.5.2	PTT . . . . .	52
2.5.3	Dynamic Syntax . . . . .	54
2.6	Dialogue models . . . . .	54

<b>3</b>	<b>Hypotheses</b>	<b>56</b>
3.1	Constraints . . . . .	56
3.2	Consequences . . . . .	58
<b>4</b>	<b>CCs in Dialogue</b>	<b>62</b>
4.1	Terminology . . . . .	62
4.1.1	Contributions . . . . .	62
4.1.2	Turns . . . . .	64
4.2	BNC corpus study . . . . .	65
4.2.1	Materials and procedure . . . . .	66
4.2.2	Results and discussion . . . . .	71
4.2.3	Summary . . . . .	82
4.3	Face-to-face versus text dialogue . . . . .	84
4.4	Task-based dialogues comparative corpora . . . . .	86
4.4.1	Materials and procedure . . . . .	87
4.4.2	Results and discussion . . . . .	88
4.5	Summary . . . . .	109
<b>5</b>	<b>Exp 1: Introducing fake CCs</b>	<b>114</b>
5.1	The DiET chat tool . . . . .	115
5.1.1	Line-by-line user interface . . . . .	115
5.2	Method . . . . .	116
5.2.1	Materials . . . . .	116
5.2.2	Subjects . . . . .	118
5.2.3	Procedure . . . . .	118
5.2.4	Analysis . . . . .	118
5.3	Results . . . . .	122
5.3.1	Typing time . . . . .	123
5.3.2	Number of characters . . . . .	124
5.3.3	Deletes per character . . . . .	124
5.3.4	Split point . . . . .	126
5.4	Discussion . . . . .	127

5.4.1	Floor change effects . . . . .	127
5.4.2	Same-person versus cross-person effects . . . . .	128
5.5	Summary . . . . .	129
<b>6</b>	<b>Exp 2: Introducing fake CCs (C-by-C)</b>	<b>130</b>
6.1	Character-by-character interface . . . . .	131
6.2	Method . . . . .	131
6.2.1	Materials . . . . .	132
6.2.2	Subjects . . . . .	135
6.2.3	Procedure . . . . .	135
6.2.4	Analysis . . . . .	135
6.3	Results by dialogue . . . . .	139
6.4	Results by intervention . . . . .	141
6.4.1	Typing time and number of characters . . . . .	141
6.4.2	Onset delay . . . . .	145
6.4.3	Planned post-hoc analyses . . . . .	149
6.5	Discussion . . . . .	149
6.6	Summary . . . . .	151
<b>7</b>	<b>Exp 3: Inducing CCs</b>	<b>154</b>
7.1	Method . . . . .	155
7.1.1	Entropy . . . . .	155
7.1.2	Subjects and materials . . . . .	156
7.2	Analysis . . . . .	157
7.3	Results . . . . .	159
7.3.1	Type of response . . . . .	162
7.3.2	CR responses . . . . .	163
7.3.3	CC responses . . . . .	164
7.3.4	Context . . . . .	166
7.4	Discussion . . . . .	168
7.4.1	Clarification requests . . . . .	169
7.4.2	Compound contributions . . . . .	169

7.4.3	Context . . . . .	172
7.5	Summary . . . . .	174
<b>8</b>	<b>A DS account of CCs</b>	<b>176</b>
8.1	Introduction . . . . .	177
8.2	Parsing . . . . .	178
8.2.1	Lexical actions . . . . .	179
8.2.2	Computational rules . . . . .	179
8.2.3	Underspecification . . . . .	182
8.2.4	LINKed trees . . . . .	185
8.3	Generation . . . . .	188
8.4	Parsing/producing CCs . . . . .	188
8.4.1	Expansion CCs . . . . .	188
8.4.2	Completion CCs . . . . .	189
8.4.3	Type Theory with Records and DS . . . . .	192
8.4.4	Utterance events . . . . .	192
8.4.5	Speech act information . . . . .	193
8.5	Summary . . . . .	197
<b>9</b>	<b>Conclusions</b>	<b>200</b>
9.1	Grammar as a resource . . . . .	201
9.2	Pragmatic constraints . . . . .	202
9.3	Ownership . . . . .	204
9.4	Context . . . . .	206
9.5	Coalitions . . . . .	207
9.6	Future directions . . . . .	209
9.6.1	Experimental . . . . .	209
9.6.2	Formal . . . . .	210
9.6.3	Dialogue Models . . . . .	211
9.7	Summary . . . . .	211
<b>A</b>	<b>Balloon Task</b>	<b>225</b>

<b>B Subarctic Survival Task</b>	<b>226</b>
<b>C Balloon Task Questionnaire</b>	<b>229</b>
<b>D Arctic Survival Task Questionnaire</b>	<b>230</b>
<b>E Questionnaire Results</b>	<b>231</b>
<b>F Exp 1: Additional Stats</b>	<b>232</b>
<b>G Exp 2: Additional Stats</b>	<b>234</b>
<b>H Entropy values</b>	<b>237</b>
<b>I Exp 3: Additional Stats</b>	<b>238</b>
<b>J Dynamic Syntax Rules</b>	<b>240</b>



## List of Figures

2.1	Example of picture card used in syntactic alignment studies . . . . .	42
5.1	Line-by-line DiET chat interface . . . . .	116
5.2	Typing time by floor change $\times$ cont completeness . . . . .	125
5.3	Num chars by same/other $\times$ cont completeness . . . . .	125
6.1	Character-by-character DiET chat interface . . . . .	131
6.2	‘Fake sender’ participant <i>MJBingo</i> ’s view . . . . .	134
6.3	<i>AB-CC</i> , as viewed by <i>Farah</i> . . . . .	134
6.4	<i>BA-CC</i> , as viewed by <i>Del7</i> . . . . .	134
6.5	Typing time by floor change $\times$ cont start-completeness . . . . .	143
6.6	Typing time by ant end- $\times$ cont start-completeness . . . . .	143
6.7	Average typing time by CCs . . . . .	144
6.8	Onset delay by what non-responding recipient saw . . . . .	147
6.9	Onset delay by other FC, other SO and prior speaker . . . . .	149
7.1	All responses by POS entropy $\times$ lexical entropy . . . . .	160
7.2	Onset delay by lexical entropy $\times$ ant end-completeness . . . . .	161
7.3	Onset delay by POS $\times$ lexical entropy . . . . .	162
7.4	CC responses by lexical entropy $\times$ ant end-completeness . . . . .	165
7.5	CC responses by lexical entropy $\times$ POS entropy $\times$ topic . . . . .	167
8.1	Single node tree following INTRODUCTION . . . . .	180
8.2	Tree following PREDICTION . . . . .	180
8.3	Parsing <i>John</i> . . . . .	180
8.4	Parsing <i>John loves</i> . . . . .	181
8.5	Parsing <i>John loves Mary</i> . . . . .	181
8.6	Using *ADJUNCTION . . . . .	182
8.7	Parsing <i>Mary</i> , after *ADJUNCTION . . . . .	182

8.8	Parsing <i>Mary, John loves</i> . . . . .	183
8.9	Parsing <i>It bothers Louise</i> . . . . .	185
8.10	Parsing <i>It bothers Louise that John loves Mary</i> . . . . .	186
8.11	Parsing <i>John</i> (ready for <i>who</i> ), using LINK & *ADJUNCTION . . . . .	187
8.12	Parsing <i>John, who smokes, loves Mary</i> – Completed parse . . . . .	187
8.13	Parsing <i>John loves Mary, which bothers Louise</i> . . . . .	189
8.14	Tree following CC <i>John died</i> . . . . .	194

## List of Tables

4.1	Annotation tags . . . . .	67
4.2	BNC inter-annotator $\kappa$ statistic and absolute agreement (min-max) . . . . .	70
4.3	BNC antecedent/continuation separation . . . . .	72
4.4	BNC completeness and repair . . . . .	73
4.5	BNC continuation categories . . . . .	75
4.6	Comparative overview of CCs by corpus . . . . .	88
4.7	Comparative antecedent/continuation separation . . . . .	90
4.8	Comparative antecedent end-completeness . . . . .	94
4.9	Comparative continuation start-completeness . . . . .	97
4.10	Comparative continuation categories . . . . .	98
4.11	Comparative split point . . . . .	100
4.12	Comparative repair . . . . .	102
4.13	Tuition and tangram corpus within-party CCs . . . . .	105
4.14	Tuition corpus ratification of cross-person CCs . . . . .	107
4.15	Tuition corpus cross-person CCs . . . . .	107
5.1	Exp 1: Comparison of experimental manipulations . . . . .	117
5.2	Exp 1: Standalone completeness . . . . .	119
5.3	Arbitrary split point comparison . . . . .	120
5.4	Exp 1: Typing time of turn and number of chars . . . . .	122
5.5	Exp 1: ANOVA typing time of turn . . . . .	123
5.6	Exp 1: ANOVA number of characters . . . . .	124
5.7	Exp 1: Deletes per character . . . . .	126
5.8	Exp 1: GEE on deletes per character . . . . .	126
6.1	Exp 2: Number of cases by completeness . . . . .	138
6.2	Exp 2: Comparison of experimental and control dialogues . . . . .	140
6.3	Exp 2: First response by type of intervention . . . . .	141

6.4	Exp 2: Mean results of first response . . . . .	141
6.5	Exp 2: GEE typing time . . . . .	142
6.6	Exp 2: GEE number of characters . . . . .	142
6.7	Exp 2: GEE onset delay by completeness . . . . .	145
6.8	Exp 2: GEE onset delay by what other recipient saw . . . . .	146
6.9	Exp 2: Onset delay . . . . .	146
6.10	Exp 2: onset delay following a genuine CC . . . . .	146
7.1	Exp 3: GEE response by predictability and end-completeness . . . . .	159
7.2	Exp 3: GEE onset delay by predictability . . . . .	160
7.3	Exp 3: Response type . . . . .	162
7.4	Exp 3: GEE CRs by predictability . . . . .	163
7.5	Exp 3: GEE CCs by predictability . . . . .	164
7.6	Exp 3: GEE CCs by predictability (inc topic) . . . . .	166
F.1	Exp 1: ANOVA typing time of turn with constituency . . . . .	232
F.2	Exp 1: ANOVA num chars with constituency . . . . .	232
F.3	Exp 1: GEE on deletes per character with constituency . . . . .	233
G.1	Exp 2: Z-tests of proportion of responses . . . . .	234
G.2	Exp 2: GEE onset delay - complete model . . . . .	234
G.3	Exp 2: GEE onset delay by non-responder . . . . .	235
G.4	Exp 2: GEE typing time of response with constituency . . . . .	235
G.5	Exp 2: GEE num chars with constituency . . . . .	236
G.6	Exp 2: GEE onset delay with constituency . . . . .	236
G.7	Exp 2: GEEs onset delay by completeness . . . . .	236
H.1	Exp 3: Entropy values . . . . .	237
I.1	Exp 3: Simple slopes - response by predictability . . . . .	238
I.2	Exp 3: Simple slopes - onset delay by predictability . . . . .	238
I.3	Exp 3: Simple slopes - CC by predictability . . . . .	239
I.4	Exp 3: Simple slopes - CC by predictability (inc topic) . . . . .	239

## Acknowledgements

As is traditional, I will start this section with a big thankyou to my supervisor Pat Healey, who managed to tread the fine line between the Scylla of giving me enough autonomy to make my own mistakes and the Charybdis of steering me away from self-paralysis towards more fruitful endeavour when I hit a wall. His insistence that I stop being so apologetic for anything I think I may have found out has (hopefully...see, I'm at it again!) improved my work immeasurably.

Huge thanks also to Ruth Kempson, without whom I would probably not even have embarked on the adventure of doing a PhD, and who saw the links in my disparate areas of interest (a winding course from philosophy to psychology and artificial intelligence to linguistics and beyond) long before I did, and whose Dynamic Syntax continues to make me think about basics of language in entirely new ways.

Thanks also to the many people involved with the DynDial project that I was lucky enough to be a part of, and my research group Interaction, Media and Communication, both of which have led to fruitful discussions and collaborations. In particular, Matt Purver, who took up the mantle as my second supervisor, Eleni Gregoromichelaki, Arash Eshghi (with special thanks for the hours spent debugging the chat tool) Wilfried Meyer-Viol, Greg Mills and Graham White. Thanks also to Stuart Battersby, who let me use his data for the tuition task corpus study in section 4.4.

Thanks most of all, however, to Martin, who has put up with me finishing his sentences for over a decade.

She's a sentence finisher. It's like dating Mad Libs. – Jerry Seinfeld

## Previously Published Material

The BNC corpus study reported in chapter 4 and the chat tool experiment in chapter 5 are based on work collaboratively produced with Pat Healey, Matt Purver, Eleni Gregoromichelaki, and Greg Mills published in Howes et al. (2011) and presented at SIGDIAL 2009 (Purver et al., 2009; Howes et al., 2009).

The tangram corpus study reported in section 4.4 builds on work reported in Eshghi (2009).

# Chapter 1

## Introduction

---

*Compound contributions (CCs)* – dialogue contributions that continue or complete an earlier contribution,<sup>1</sup> see e.g. (1.1) – are the paradigm case of coordination in dialogue and constitute a critical test case for theories of *natural* language processing. They 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 chapter 2).

(1.1) **Daughter:** Oh here dad, a good way to get those corners out

**Dad:** is to stick yer finger inside.

**Daughter:** well, that's 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

---

<sup>1</sup>These terms will be defined in detail in section 4.1.

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, 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 (e.g. in terms of where the split occurs), 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 (in this thesis these are called *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 hearer’s continuation is not the same as that which the original speaker would have provided, as in (1.1, 1.2).

(1.2) **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 Finnish and Russian



(which have 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; Grenoble, 2006; 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, all things being equal, 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. A similar stance is 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).

As these approaches are mechanistic, they make no claims regarding the necessity of the incoming speaker recognising the original speaker’s intended utterance, though continuation by another speaker can be taken to involve preempting the other interlocutor’s intended content.<sup>2</sup> 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, see section 2.5.2) 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 simpler alignment model such as

---

<sup>2</sup>Note that this says nothing about whether such a continuation successfully matches the initial speaker’s intended continuation. For examples where this cannot be the case see Gregoromichelaki et al. (2011), as well as (1.1), (1.2).

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 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.

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".

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 thesis I present and examine empirical corpus data and carry out the first experimental manipulations of CCs, in order to shed light on some of the questions raised by this phenomenon.

In chapter 2 I will discuss previous work on CCs, and the consequences of the types of CCs people actually produce on our understanding of language in general, and dialogue

in particular. Chapter 4 introduces the technical terminology to be used throughout the empirical studies, and reports several corpus studies mapping the distributions of CCs in a variety of dialogues, comparing face-to-face with text-based and general conversation with task specific dialogues. These indicate that the principal factors conditioning the use of CCs are pragmatic and that same- and cross-person CCs tend to occur in different contexts. Chapters 5 and 6 report the first experiments manipulating CCs in two different text chat environments. The results point towards a pragmatic effect of (fake) cross-person CCs indicating that participants apparently co-constructing an utterance are treated as a coalition. Chapter 7 investigates when people are able or likely to produce continuations, using a further character-by-character text chat experiment which truncates genuine dialogue contributions to prompt continuations. Manipulations of the structural and contextual predictability of the truncated turn show that while syntax provides a resource for the construction of a CC it does not place significant constraints on where the split point may occur. It also shows that people are more likely to produce continuations when they share common ground. In chapter 8, I outline a formal system, using Dynamic Syntax, that can account for both the syntactic and mechanistic properties of CCs, as well as the pragmatic effects observed in the empirical studies. Chapter 9 assesses the viability of this approach and outlines further research possibilities.

## Chapter 2

### Compound Contributions

---

Most previous work on 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), *anticipatory completions* (Fox, 2007) and *split utterances* (Purver et al., 2006) amongst others (with some differences of emphasis in the different terms). As CCs are a phenomena of dialogue, much of the work on them has been carried out in the field of Conversation Analysis, so we begin our discussion there.

#### 2.1 Conversation Analysis

Conversation Analysis (CA) takes the *turn constructional unit* (Sacks et al., 1974, – henceforth SSJ) to be the basic unit in conversation. Identifying turn constructional units (TCUs) in talk-in-interaction is seen to be one of the main issues for turn-taking, to account for the minimal overlap, and lack of long pauses when there is a change of speaker.

First, the existence of organised turn-taking is something that the data of conversation have made increasingly plain. It has become obvious that, overwhelmingly, one party talks at a time, though speakers change, and though

the size of turns and ordering of turns vary; that transitions are finely coordinated; that techniques are used for allocating turns, whose characterisation would be part of any model for describing some turn-taking materials; and that there are techniques for the construction of utterances relevant to their turn status, which bear on the coordination of transfer and on the allocation of speakership (Sacks et al., 1974, p699).

To minimise both gaps in conversations, and overlaps (two or more speakers talking together), thus maintaining the ideal of one-speaker-at-a-time, interlocutors must be able to project when the end of a *turn-constructive unit* (TCU) is approaching, because it is at the ends of TCUs that *transition relevance places* (TRPs) occur, and therefore where speaker change can normally be licensed, though the original speaker may continue, or the conversation may end. Failure to have a coordinated understanding of these issues would lead to a conversation characterised by overlaps and long silences, which, they argue, rarely happens. Instead, “transfer of speakership is coordinated by reference to such transition-relevance places which any unit-type instance will reach” (p703).

TRPs often occur at points of syntactic completion, though of course syntactic units such as sentences, clauses, phrases and lexical constructions (which SSJ identify as potential TCUs in English) can be extended infinitely (e.g. by the addition of adjuncts), and TCUs are not necessarily defined by their syntactic properties (Ford and Thompson, 1996). How such TRPs are anticipated has been the subject of much discussion, with hypothesised cues ranging from syntactic to prosodic and acoustic (Duncan, 1972), with participants better able to predict turn endings when the different cues coincide (Hjalmarsson, 2011). However, de Ruiter et al. (2006) show that TRPs (in Dutch) can be reliably predicted using the lexico-syntactic properties of the string, without any intonational or prosodic information, and cannot be predicted based on prosody alone, suggesting that lexico-syntactic information is crucial to predicting potential TRPs.

### 2.1.1 Anticipatory completions

Lerner (1991) identifies various structures typical of CCs which contain characteristic split points. One group of these are ‘compound’ TCUs, which are structures that include an initial constituent that hearers can identify as introducing some later final compo-

ment. Examples include the IF X-THEN Y, WHEN X-THEN Y and INSTEAD OF X-Y constructions (2.1).

(2.1) **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).

### 2.1.2 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* (2.2).

(2.2) **D:** Well I do know last week that=uh Al was certainly very <pause 0.5s>

**R:** pissed off [Lerner (1996), p260]

As he makes no claims regarding the frequency of such devices for CCs, it is an open question as to 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.

### 2.1.3 Expansions vs. completions

Other classifications of CCs often distinguish between *expansions* and *completions* (Ono and Thompson, 1993). Expansions (or *extensions*; Ferrara, 1992) are continuations which occur at transition relevant places and add, e.g., an adjunct, to an already complete syntactic element or TCU (2.3), (2.4).

(2.3) **T:** It'll be an E sharp.

**G:** Which will of course just be played as an F. [BNC G3V 262-263]

(2.4) **M:** yep dr goes everyones happy

**N:** except the dr [DiET SU1 19 240-241]

Completions involve the addition of syntactic material which is required to make the whole compound contribution (syntactically) complete (2.2), (2.5).

(2.5) **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 I consider both expansions and completions to be CCs, it is possible to distinguish between the two types by considering the completeness or otherwise of the first part of the CC.

Benjamin (2009) argues that expansion type CCs vary along different parameters including who is being addressed and sequential organisation (e.g. whether the expansion projects further action). Because of this range of functional possibilities, he concludes that the strategy of adding syntactic elements to another's talk is a generic strategy in conversation.

Expansion type CCs (which Schegloff (1996) refers to as *increments*), have also been studied in same-person cases (e.g. Goodwin, 1979; Walker, 2004). Like cross-person expansions, same-person ones are viewed as a highly productive way of utilising grammatical constraints for interactional purposes. Walker (2004) notes "it would seem that increments can be added to almost any possibly complete turn at talk, placing the practice alongside other generic conversational practices such as self- and other-initiated repair" (p167).

Ferrara (1992) further subdivides completions into *predictable*, *helpful* and *invited* utterance completions,<sup>1</sup> but it is unclear whether these categories are mutually exclusive,

<sup>1</sup>Note that although Ferrara (1992) names her categories in terms of functional attributes (see section 2.1.4, below), her predictable and helpful categories correspond roughly to Lerner's anticipatory and opportunistic cases as discussed above.

or whether expansions cannot be similarly subdivided. She also analyses only 30 CCs from therapist-client dialogues in detail, so it remains to be seen whether her observations on the characteristics of these devices, e.g. that proffered continuations were rejected in 13% of CCs, apply more generally.

#### 2.1.4 Functions / effects of CCs

CCs can have different properties in terms of how they are reacted to or interpreted. A range of functional properties are outlined (by e.g. Lerner, 1996; Szczepek, 2000b), from collaborative productions and showing understanding (these rely on a continuing speaker accurately guessing the continuation of the utterance), borrowing (as in example 1.2, where the initial speaker's grammatical structure is appropriated to make a joke) and eliciting information (as with Ferrara's (1992) invited utterance completions). Such categorisations along functional lines can be either from the perspective of the initial, or the incoming speaker, but this means that the functional groupings are not mutually exclusive and may not have any explanatory power. Questions also arise regarding whether the categories can be sustained, or whether CCs can exhibit genuine multifunctionality (see e.g. Gregoromichelaki et al., 2009; Bunt, 2009). In (2.6), for example, J's continuation of M's utterance serves both as a completion of the syntactic and semantic content, an indication of understanding and also as a request for confirmation.<sup>2</sup>

(2.6) **M:** It's generated with a handle and

**J:** Wound round?

**M:** Yes

[BNC K69 109-112]

Additionally, analyses according to functional properties do not have anything to say regarding relative frequencies, or whether there are structural differences in the different types.

#### Effects on turn-taking

Despite these concerns, several authors have hypothesised that considerations of turn-taking play a major role in (cross-person) CCs.

---

<sup>2</sup>This multifunctionality is also important in the grounding model proposed in Clark and Schaefer (1989), in which a contribution can act as both an acceptance of a prior contribution and the presentation of further information. See section 2.5.2, which looks at an account of cross-person CCs using the grounding model, and also section 2.4.2.



Firstly, it is noted that in many cases, even where they overlap with material from the original speaker, cross-person CCs are not treated by the participants as interruptive. In Ferrara (1992), 29 out of the 30 cases she studied did not result in the second speaker taking the floor after the CC; in Szczeppek (2000a), this figure is 12 out of 15 cases. Coates (1994) hypothesises that, at least in the all female talk between friends which she analyses, turns are potentially jointly construed, and that one example of this is in the production of cross-person CCs. These observations seem to contradict the classic turn-taking model (Sacks et al., 1974), in which the notion of turn is tied to the individual speaker.

However, Schegloff (1995) claims that although the no-gap no-overlap model of turn-taking is often assumed to apply to individuals, this is not necessarily the case:

...the turn-taking system as described in SSJ organizes the distribution of talk not in the first instance among persons, but among *parties*. Now not uncommonly, of course, parties are composed of persons – *single* persons. But on some occasions, or for some particular phase or topic or sequence *within* some occasion of talk-in-interaction, the aggregate of persons . . . are organized into parties, such that there are fewer parties than there are persons. (p33, italics original)

Eshghi (2009) shows that parties are relevant entities for participants in multiparty conversations. Although these parties may be existing groupings, such as a couple, or an audience (Lerner, 1993), they may be emergent, and only relevant for a stretch of talk within a longer dialogue. In his lecture series in 1967, Sacks (1992) observed that co-constructing a sentence might be one way in which these groupings can be made manifest.

Because a sentence is obviously a prototypical instance of that thing which is done by a unit. Normally, some single person. That then permits it – for those who have the wit to do it – to be a way that some non-apparent unit may be demonstrated to exist.

We get, then, a kind of extraordinary tie between syntactic possibilities and phenomena like social organization. That is, an extremely strong way

that these kids go about demonstrating that, for one, there is a group here, is their getting together to put this sentence together, collaboratively. (p145)

Lerner (1993) agrees that participants can use anticipatory completions to demonstrate their association with an interlocutor, however, he differentiates between these collaborative CCs (which continue both the syntactic form of the prior turn and its action to some other recipient – e.g. (2.1)-(2.5)) and those that are “a distinct turn in response to it but one built off of the prior turn syntactically” (Lerner, 2004, p. 160 – see e.g. (2.6)-(2.11)). Kangasharju (1996) uses this distinction in her discussion of cross-person CCs as one device for aligning as a team in multiparty conversations. However, given the potential multifunctionality of CCs mentioned above, or their possible ambiguity, these distinctions, as well as questions regarding who the intended recipient is (either the initial speaker, or some other participant(s)) are not clear cut. This is a problem noted by Levinson (1988), who extends Goffman’s (1981) participant framework to include additional types of speaker and hearer. He makes the point that jointly authored sentences “. . . raise the fundamental question whether the collaborative nature of verbal interaction does not make inherently problematic the attribution of participant role” (p203).

### **Footing**

Despite this potential problem, attempts to analyse cross-person CCs in terms of *footing* have been made. In Díaz et al. (1996), it is claimed that cross-person CCs can be used to create an association, and that these collectively formulated statements are then assessed to be jointly owned by the contributing participants, adding ‘collective author’ to Levinson’s (1988) list of participant statuses available to speakers. Antaki et al. (1996) claim that to be successful, cross-person CCs must maintain the footing on which the original utterance was made. In the turn after the completion, the original speaker may accept or reject the continuation as conforming to their own footing, as author, relayer or joint author. However, while this may be a way to analyse a subset of CCs, it is unclear how cases such as (1.2) could be assessed in these terms, and Antaki et al. (1996) also reject cases in which there is no response from the original speaker from which to identify the ‘footing-consistency’ (see also section 2.2, below), despite the possibility that a lack of a verbal third part to the sequence may itself be a relevant response to a continuation

of one's own contribution by another.

These analyses of cross-person CCs in terms of turn-taking as it applies to parties and footing as it applies to the individuals who are jointly constructing a sentence are important insofar as they raise questions about what it is that CCs can or do achieve. Are they genuinely used, or taken by conversational participants to be indicative of, group statuses (party membership or collective authorship) or participant roles, or are these just possible functions among many that can only be inferred by an observer after the fact? Additionally, claims in the CA literature are not concerned with issues of cause and effect. Are CCs an indicator of parties which are already usefully oriented to by participants,<sup>3</sup> or do they in some sense create them? Either way, if they are a relevant way to analyse CCs, we would expect different patterns of CCs in dyadic versus multiparty conversations, which raise additional questions about their occurrence in dyadic conversations.<sup>4</sup> Díaz et al. (1996) suggest that, at least in some task-based dyadic conversations, CCs are still indicative of party membership, with reference to some non-present ratified participant (in their case, the experimenter). Whether or not such a claim is sustainable, or whether or not the turn-taking model of SSJ (as applied to parties) can accommodate such a possibility remains to be seen. If so, certain conversations, such as those in Coates (1994), in which the participants all share common ground and common knowledge and all apparently share the floor, might be (counter-intuitively) seen to be more like monologues in nature. This possibility is taken to its logical conclusion by Fais (1994), who sees CCs as evidence towards this view: "If conversation is viewed as simultaneously co-produced "monologue", speakers are merely acknowledging the process that they know their hearers are undertaking; that is, they are acknowledging the fact that their hearers are producing the conversation at the same time that they are." If this were indeed the case then we should not expect any differences between cross-person and same-person CCs.

---

<sup>3</sup>This would be in line with folk psychological notions of 'finishing someone else's sentence' that assume they often occur between people who are already very familiar with each other, such as a married couple, or twins.

<sup>4</sup>The distinction may not be so clear given that one person may speak from more than one position in a conversation, raising the possibility of 'parties' in dyadic dialogues.

### 2.1.5 Ratification / rejection

One of the ways questions regarding parties and footing have previously been assessed is by looking at whether or not the continuation is accepted or rejected by the original speaker in their following contribution. In the *collaborative turn sequences* of Lerner (2004), such responses to a completion show that the original speaker maintains authority over the whole compound contribution, whilst continuations addressed towards a third party (not the original speaker) will not be ratified or rejected, as the authority for the compound contribution is shared by its co-constructors (as with the collective authorship discussed above). In other words, according to Lerner, if a ratification or rejection is appropriate, then the CC cannot be seen as indicating party membership. However, it is important to note that ratification may be non-verbal, so a lack of verbal response cannot be used as an index that the CC is taken as belonging to both its contributors equally.

Ratification can take the form of a repeated element, as in (2.7) – possibly by overlapping material, as in (2.8), or a paraphrase and/or a “yes” (2.9, 2.10).

- (2.7) **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]

- (2.8) **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.]]

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

[BNC H5D 123-127]

- (2.9) **A:** All the machinery was

**G:** [[All steam.]]

**A:** [[operated]] by steam

[BNC H5G 177-179]

(2.10) **S:** Secondly er

**J:** We guarantee P five.

**S:** We we are we're guaranteeing P five plus a noise level.

**J:** Yeah.

[*BNC JP3 167-170*]

### 2.1.6 Implications of CCs on CA analyses

Certain types of cross-person CCs present a challenge to the view that turn-taking proceeds in an orderly fashion with participants waiting until TRPs before they take (or attempt to take) the floor.

Cross-person completion CCs especially require the entering by one person into another's turn space, and it is for this reason that Lerner (1996) talks about the "semi-permeable" nature of grammatical units in conversation. The compound-TCUs (discussed in section 2.1) and other cues (such as mid TCU laughter or silence) present specific points within existing TCUs at which a person may enter another's turn space. Alternative approaches to account for the apparent conflict between CCs and the 'rules' of turn-taking<sup>5</sup> is to assert, as Schegloff (1995) does that turn-taking applies to parties, which may consist of single persons but need not (as discussed in section 2.1.4). The consequences of this move on individual participants in conversation are never explicitly explored, though in the case of CCs, different patterns of overlap might occur within and between parties, with within-party overlap being more common, more acceptable and/or less interruptive (Coates, 1994). If the turn space belongs to a 'party' not an individual, then in what sense can you be said to be interrupting or entering another's turn space if you are also a member of the party? Another way to phrase this question is: if there is a shared entitlement to a turn space, how do the rules of turn-taking apply within that turn space between persons who are equally entitled to it?

Producing a continuation may be a way of actively indicating membership of a party with the initial speaker (i.e. joining a turn), or a means by which to take over someone else's turn. These possibilities have conflicting predictions regarding who would take responsibility for the overall utterance, and it is an open question whether these are equally productive ways of using the grammatical features of dialogue to establish roles

---

<sup>5</sup>Note that these rules may be better conceived of as 'norms', as they are not to be taken as either explicit or compulsory.

or if one type is more common than the other, and whether there are different patterns of turn-taking in each case.

### 2.1.7 Summary

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 coordination, however, it is unclear how representative they are. Additionally, as the emphasis in the CA literature on CCs is in identifying their organisational consequences for the unfolding dialogue, they leave open the question of where a speaker switch may occur.

## 2.2 Corpus studies

In terms of frequency, the only estimates in the CA literature are Szczepek (2000a) and Antaki et al. (1996) (see below). Szczepek (2000a) found approximately 200 cross-person CCs in 40 hours of English conversation, though there is also no mention of the number of sentences or turns this equates to. 75% of her sample were completions, as opposed to expansions, however, this may be influenced by her decision to only include CCs which continue both the original speakers syntactic construction and conversational action (as with Lerner's distinction outlined in section 2.1.4, above), which therefore discounted cases such as appendor questions (2.11) which are specifically designed as syntactic extensions to the utterance they are querying and have been analysed as other-initiated repair.

(2.11) **G:** That's right they had to go on a rota.

**A:** Run by the Dock Commission?

**G:** Run by the Dock Commission. *[BNC H5H 100-102]*

Based on her sample, Szczepek (2000a) claims that, syntactically, people display preferences on how they produce an incoming completion. Cases which she claims are unproblematic (and therefore offer empirical evidence for major constituency boundaries) include a predicate after a copula, a then-clause after an if-clause and an adverbial phrase expansion type CC. Contrarily, if participants continue with an NP after a preposition or a relative clause after a relative pronoun, they are likely to repeat the preposition or

relative pronoun, suggesting weaker syntactic boundaries that are not preserved over a change of speaker. However, as these observations are anecdotal, it is difficult to assess the veracity of these claims, or whether certain cases are preferred at all (for example, is there a ‘cost’ to producing a completion that repeats an element) but, as Szczepek (2000a) asserts, “a thorough investigation of the way syntax is handled by incoming speakers in collaborative productions would certainly offer another source of evidence for the interactional relevance of syntactic constituents” (p22).

Skuplik (1999) has tried to do that for a corpus of German dyadic dialogues, in which one of the interlocutors (the instructor) described how to build a toy plane to their co-participant (the constructor). She annotated for cross-person compound contributions, and found 126 CCs out of 3675 spoken dialogue contributions (3.4%). Syntactically, the most frequent type of completing utterance were prepositional phrases (37%), followed by noun phrases (24%). Participant role also played a part in who provided the completion, with the constructor continuing a contribution offered by the instructor in 79% of cases. Additionally, *expansions* (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). There are several possible reasons for this contrast; for example, there may simply be a difference in the distributions of CCs in different languages,<sup>6</sup> or between experimentally controlled task-oriented dialogue and casual conversational dialogue. Additionally, there may be differences in the classification schemes used, e.g. with appendor questions as outlined above.

Rühlemann (2007) uses corpus analysis on the British National Corpus (BNC; see section 4.2) to examine a subset of *expansion* CCs, *sentence relatives* of one’s own or another’s turn (2.3), (2.12).

---

<sup>6</sup>Note that observers have claimed that different strategies are used in constructing CCs in different languages; for example Couper-Kuhlen and Ono (2007) found different patterns of increments in English, German and Japanese; Sun (1995) found CCs more common in Chinese than English conversations and in French, Chevalier and Clift (2008) assert that incomplete TCUs are not often completed by an interlocutor. These observations will not be elaborated on in this thesis which focuses on English, however, note that the availability of different strategies in differently constrained languages lends weight to the argument that syntactically continuing a possibly incomplete prior contribution is a generally available strategy, which may be taken up in different ways in languages with different typologies.

(2.12) **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 the terminology adopted here (see section 4.1). Rühlemann (2007) also excluded intra-turn cases where the sentence relative was annotated as a separate sentence but there was no intervening material; these would be included under the definitions used in this thesis.

Antaki et al. (1996) collected instances of cross-person CCs (including completion and expansion CCs) from the London-Lund corpus of spoken dialogue. They found 176 instances in 50 dialogues. However, as their focus is on what Lerner (2004) terms “collaborative turn sequences” which are constructed in three-part sequences (the initial contribution, the completion, and a response by the initial speaker, often constituting a ratification or rejection, as discussed in section 2.1.5) they explicitly focus on those cases which are analysable in these terms. This means that there are cases, such as those in which there was no third part to the sequence, or in which the contributor of the second part retained the floor, that they excluded from their analysis. Despite not including all cases of interest, their study raises additional questions. For example, they found a higher number of CCs in dialogues in which the nature of the task required coordination, such as map-task dialogues, which might be because the task is framed in such a way as to require participants to form a party.

### 2.3 Models of communication

One of the main questions raised by the existence of dialogue phenomena such as CCs is regarding what it is that we do when we use language, with a main assumption of most models being that one of the main purposes of language is as a tool for communication.

Though it could be argued that language is neither necessary nor sufficient for com-



munication,<sup>7</sup> even if we accept that an important aspect of language is communicative, there are numerous possibilities as to what this actually means. Some of these will be outlined below with discussions of the consequences for the analysis of CCs and whether they can be unproblematically accounted for in the various frameworks.

### 2.3.1 Transmission models of communication

The term *communication* is suggestive of the folk psychological notion that what happens in a conversation or dialogue is that interlocutors are transferring information to one another. This idea is encapsulated in the Transmission Model of Communication (Shannon and Weaver, 1949) which views communication as the transmission of information from one interlocutor to another. The model states that in communication; i) a sender formulates a message; ii) the sender then encodes the message; iii) they transmit it to a receiver; iv) the receiver decodes the message. While it was not Shannon and Weaver's intention that this would become the pervading model of communication (they had based it on the idea of a telephone exchange, for practical rather than theoretical reasons), it did so, not least because it seems to fit our intuitions that this is really what happens when we communicate. However, the model is flawed in many ways. Not least is that, as pointed out by Reddy (1979), this intuition (which he calls the 'conduit metaphor', as language is seen as a conduit for our thoughts) is mistaken.

Additionally, in the model, communication proceeds in a linear fashion; the sender (speaker) encodes and sends the message (presumed to be their thoughts, in a linguistic form), which is then received and decoded by the receiver (hearer). This does not seem to reflect the fact that communication seems to not just involve, but require feedback, at all points in the process (it is interesting that a 'feedback loop' was added to later versions of the transmission model, but only from the end of the process back to the start, thus requiring the receiver to have decoded the message before offering feedback, which still does not resolve the linearity problem). If this were an accurate template for

---

<sup>7</sup>Communication need not be linguistic (for example, I can communicate my displeasure towards you by kicking you under the table), and whether language is sufficient for communication in itself is subject to a large body of philosophical work (some of which will be discussed in section 2.3.2) which shows that interpretation of any linguistic input relies to a large extent on contextual factors. Additionally, language has often been shown to have extra functions, such as social; for example, what is the 'communicative' aspect of my telling you something you already know (as when we discuss the weather)?

human linguistic communication, then CCs (specifically cross-person completion types where a full message cannot have been received before the hearer takes over as speaker) ought to be impossible.

Perhaps more problematic is the separation of the sender and receiver which this model entails. In a communicative exchange, ‘sending’ and ‘receiving’ are not split between interlocutors; any interlocutor can revise any message they are in the process of sending on the fly (for which they must, in some sense, also be receiving the message as they send it, as with *self-monitoring* in the speech production model of Levelt, 1989), or intervene, or any number of other possibilities which indicate that the roles of sender and receiver are not easily delineated and occur simultaneously. The model thus mistakenly relegates the receiver to a passive role in the communication, waiting for an incoming message to decode.

As Couper-Kuhlen and Selting (2001) point out “...linguistic productions – since they take shape in interaction – can no longer be conceptualized as the product of a single speaker. Instead sentence and clause production, indeed speech production in general, must be thought of as an interactional achievement. In this view, syntax, just like prosody and semantics, is a resource that can be relied on as shared knowledge in the speech community and that can be ‘distributed’ across speakers in collaborative productions” (p5).

Another issue is that accepting the transmission model of communication leads inevitably to the conclusion that the content of a message is equal to its meaning<sup>8</sup>. If this were the case, then decoding a message would be the mirror image of encoding it, and the only possibility for miscommunication would be the presence of extraneous noise, but as Reddy points out, “extraction is a trivial process”. Our intuitions do admit the possibility of miscommunication, and these are not only associated with noise, but also with what we thought the person meant (see e.g. Purver, 2004, for a taxonomy of clarification requests querying different aspects of the communication). This also highlights another presupposition of the code model, namely that everyone has an identical copy of the code<sup>9</sup> (taken to be the syntax and semantics of the language in question). Al-

---

<sup>8</sup>It should be noted that whether this is true or not was irrelevant to Shannon and Weaver’s engineering problem.

<sup>9</sup>Though note that this additional assumption is also not intrinsic to the Shannon and Weaver (1949) model.

though “language is a code linking representations of sound and meaning” (Smith, 1999, p153), it seems that knowledge of this ‘code’ is insufficient for successful communication. Communication seems to involve not just the extraction of meaning, but its construction.

### 2.3.2 Inferential communication

It follows then, that communication cannot be the direct transference of information. How, then, can we characterise it?

One of the greatest contributions to this question came from Grice (1975), who emphasised the inferential nature of communication using language, over and above what can be interpreted directly through the syntax and semantics (the ‘code’).

#### Grice’s conversational implicatures

For Grice (1975), the key lay in the cooperative nature of communicative interaction.

Our talk exchanges do not normally consist of a succession of disconnected remarks, and would not be rational if they did. They are characteristically, to some degree at least, cooperative efforts; and each participant recognizes in them, to some extent, a common purpose or set of purposes, or at least a mutually accepted direction (Grice, 1975, p45).

He accordingly outlines the *Cooperative Principle*; “Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged” (Grice, 1975, p45), which can be paraphrased as “be cooperative”.

The Gricean program differentiates between the truth-conditions of a sentence (taken to be the realm of semantics) and what is actually communicated on any given occasion, via implicatures (taken to be the realm of pragmatics). In other words, it treats *what is said* as quantifiably different from *what is meant*. As Grice puts it:

- (2.13) “*U* meant something by uttering *x*” is true iff, for some audience *A*, *U* uttered *x* intending
- a. *A* to produce a particular response *r*
  - b. *A* to think (recognize) that *U* intends (a)
  - c. *A* to fulfill (a) on the basis of his fulfillment of (b). (Grice, 1969, p151)

What this means is that, over and above what truth conditional meaning we extract directly from the words, we attribute additional meanings which we believe the speaker *intended* to convey. These, as not directly encoded into the words used, must be implied, or, to use Grice's terminology, *implicated*. For Grice, when confronted with any utterance we (perhaps automatically) reason about the speaker's intended inferences.

It is true that, if questioned about how we 'knew' someone meant what we took them to mean, we could come up with such an inferential chain of reasoning (as in (2.13)), but this does not tell us anything about causation: just because we can come up with a chain of reasoning after the event does not imply that this chain of reasoning was how we arrived at our initial interpretation. Breheny (2006) extends this objection to the idea that small children are not capable of such inferential chains of reasoning (using the false belief task), and yet we would be unwilling to accept the conclusion that small children do not actually communicate.<sup>10</sup> It seems therefore, that this cannot be a complete characterisation of communication.

A Gricean account of compound contributions would assert that they carry distinct implicatures regarding mutual knowledge; in example (1.1), repeated below, for example, the father's continuation is taken to imply that he knows a good way to get the corners out himself. A further implicature is that he believes that this is the same method that his daughter was about to tell him, and it is this implicature, not the semantic content, which is then refuted.

(1.1) **Daughter:** Oh here dad, a good way to get those corners out

**Dad:** is to stick yer finger inside.

**Daughter:** well, that's one way. *[from Lerner, 1991]*

This type of analysis (see also section 2.5.1 regarding partial utterances) would only be feasible if the person supplying the continuation had guessed a full proposition prior to offering their completion (which they may or may not assume to be the same). This is because you would need to have a semantic representation which is propositional before being able to calculate any implicatures. Like the CA analyses discussed in section 2.1,

---

<sup>10</sup>Note also that very young children are able to add a syntactically matched continuation to an incomplete antecedent, as in the nursery rhyme 'Old MacDonald'. The adult offers the antecedent "Old MacDonald had a farm (E-I-E-I-O), and on that farm he had a/some..." and the child continues with a singular or plural noun (e.g. "cow/cows"), as appropriate.

this type of analysis would have to be on a case by case basis after the fact and it is not clear that any such supposed implicatures could be generalised, or indeed whether or not they would be (or would need to be for successful communication) calculated on the fly. Once again, just because we can offer a post-hoc explanation with reference to implicatures arising from the different parts of the CC, it is by no means clear that we need them to produce or understand the CC in the first place, and thus, though a Gricean story of CCs can be told, it has little or no explanatory power.

### **Relevance Theory**

Following Grice, Relevance Theorists also see communication as an act of cooperation between speaker and hearer. “The relevance-theoretic account is based on another of Grice’s central claims: that utterances automatically create expectations which guide the hearer towards the speaker’s meaning” (Wilson and Sperber, 2002).

However, Relevance Theorists believe that the fundamental expectation of a hearer is that an utterance will be relevant to the discourse, and not that they will necessarily be told something true. Relevance Theory (RT) tries to show how speakers and hearers can arrive at a shared meaning, via presumptions about what the other takes as relevant in any given context, knowing that the other will be following a path of least effort to arrive at a plausible hypothesis.

Like Grice, Relevance Theorists take an utterance’s literal, or encoded, meaning, as a starting point, but claim that all utterances, not just figurative or loose uses of language, are approached with expectations of relevance. Meaning cannot simply be extracted from the words uttered (the ‘code’) but must be based on additional inferences.

Other philosophers of language (e.g. Cappelen and Lepore, 2005) argue that RT is incoherent, as it states that *any* interpretation of an utterance, literal or otherwise, is dependent on context, and decoding alone is insufficient to recover speaker meanings. But communication *is* possible, so there must be some shared content to what we say.

However, this conclusion is based on an all-or-nothing view of understanding. Cappelen and Lepore (2005) believe that either language allows meanings to be perfectly recreated (which, as we have seen from the discussion of transmission models of communications is a scarcely tenable assumption), or communication is impossible.

According to Cappelen and Lepore’s definition of RT, which presupposes that under-

standing operates on complete propositions<sup>11</sup> compound contributions ought to be impossible, because if we can never fully understand each other, then how can we possibly presume enough to continue anyone else's utterance-in-progress? This is both a misrepresentation of RT, and a denial of the obvious communicational sceptic's<sup>12</sup> response that of course we never *fully* understand each other, because to do so would require intimate knowledge of the conversational situation, your interlocutor's background, prejudices, knowledge etc, which we can never have. This does not mean that we necessarily fail to understand each other suitably for current purposes, because in any given conversational exchange we do not need to have a complete understanding of every possible contributing factor, merely, according to RT, those which we consider relevant (though note that what you or I consider relevant in any exchange may differ, meaning we may indeed draw different conclusions).

For CCs this means that as we may assume that each utterance is relevant to the unfolding discourse we may be able to supply the continuation (or add further, relevant, information as in the case of adjuncts) to another's utterance. This should be especially possible in cases where the subject under discussion is mutually manifest. However, it is unclear how partial utterances would contribute to the cost/benefit model – does interpreting an incomplete utterance use more (because it is predictive) or less (because the interpretation remains vague or underspecified) processing than a complete one? It is reasonable to hypothesise that an RT account, which is based on the processing cost to an individual (as opposed to e.g. Clark and Wilkes-Gibbs, 1986, least collaborative effort, which is concerned with reducing the overall processing effort of all participants, together) would predict that producing a continuation, especially of another's incomplete contribution would be costly, and perhaps only feasible in highly predictable contexts, where the cost of coming up with a probable complete proposition from a partial input is reduced.

Psycholinguistic studies into 'good-enough' language comprehension (Ferreira et al., 2002; Ferreira and Patson, 2007) support the notion that we may not ever exactly un-

---

<sup>11</sup>As discussed in section 2.5.1 this supposition is itself unsustainable in the face of evidence of the pervasive incrementality in dialogue.

<sup>12</sup>Communicational Scepticism, the idea that we can never understand each other fully, is discussed in detail in Taylor (1992).

derstand one another,<sup>13</sup> if to do so is to share every element of our representations, especially in those cases where contextual or world knowledge conflicts with the compositional syntactic meaning: “Our good enough approach to language comprehension holds that language processing is sometimes only partial and that semantic representations are often incomplete” (Ferreira et al., 2002, p11). However, these studies (and RT) leave open the question of how we decide what counts as good enough (or relevant) in any given communicative situation.

### 2.3.3 Co-constructing communication

As Reddy (1979) observed, communication cannot be the direct transmission of thoughts from one head to another. Nor can it be the direct sharing of mental representations, whatever they are taken to be, for exactly the same reason – minds are, of course, closed boxes.

And the Gricean notion, as expounded on by Relevance Theorists seems unable to help in this regard either. In their view, communication is simply the recovery of intended meanings, but even if this is what we think we are doing, it is clear (even in the unlikely event that a conversation proceeds without any problems at all) that you, as speaker, have no way of knowing that I, as hearer, have in mind what you mean, or even, if we choose to treat them as different things, what you have said. Do we always intend things, or mean something explicit that we expect people to try and recover when we use language?

Whilst it is undeniable that in some sense your interpretation of what I say is something intrinsic to you, this is to miss the point of communication. Though it intuitively seems that something must be shared in order for some thought to be “communicated” between two interlocutors, communication is collaborative. It is not that I have something to transfer to you, or you who deciphers what I am trying to say, but the pair of us who build up the interpretations we assign in tandem, via dialogue. Cross-person compound contributions can, under this view, be seen to be just those cases where such joint building up of interpretations is explicit, with, for example, acknowledgements and clarification requests being other indicators of the same processes.

Though we cannot have an objective measure of anything being explicitly shared

---

<sup>13</sup>See also section 2.3.3.

between us as interlocutors, I would argue that if it is enough for me to assume from your behaviour (including linguistic) that we have in some sense converged on a mutual understanding, then, to all intents and purposes, we have done.<sup>14</sup> This is the approach taken by ethnomethodology, as espoused by Garfinkel (1967), and later taken up by conversational analysts, for example Sacks et al. (1974). Taking this position shifts the argument away from questions of whether there is an objective truth or meaning conveyed in our communications from one interlocutor to another. As Taylor (1992) states, there is “. . . no general “in-principle” explanation of HOW communicational order occurs; there are only particular “ad hoc” contingent instances of particular communicators succeeding in understanding each other” (p203).

This also means that listeners have an effect on shaping what we actually mean as speakers, and it often seems to be the case that we may not be clear about what we are trying to say when we start talking, or even, importantly, the underlying thoughts we are trying to communicate.<sup>15</sup> This suggests that such defined thoughts simply don't exist, and we construct not only what other people are saying or communicating, but what we are, as we go along, as evidenced by the numerous studies of disfluencies and repair in dialogue (Schachter et al., 1991; Schegloff et al., 1977, inter alia).

Contra intuitions that communication is usually successful, this may mean that miscommunication is actually the norm (Healey, 2008). As Keysar (2007) points out, it may be that

. . . we do not know how successful communication really is. It seems that miscommunication is relatively rare, but much of it may go unnoticed. You may tell a friend you really liked that movie about the journalist from Kazakhstan who is touring the United States, and the friend may think you were being sarcastic. You proceed to talk about other movies, without ever knowing that he misunderstood you. By definition, we do not know how often miscommunication goes unnoticed. This cluelessness distorts our performance feedback, making it very difficult to make adjustments and know when we are communicating well, and when we are not. (p82)

---

<sup>14</sup>This is again similar to the stance taken by pragmatists studying *good enough* processing (e.g. Ferreira and Patson, 2007).

<sup>15</sup>This is not a new observation in philosophy – see, for example, Kleist and Hamburger (1951)



Further, as interactions typically involve more than one utterance exchanged, as miscommunications can be locally resolved as a conversation progresses, and frequently are (one of the key insights of CA analysis on repair), they should be expected to be ubiquitous as we refine our own and each others interpretations.<sup>16</sup> A mismatch between what you and I think each other is talking about at an isolated point in the conversation should be expected, but not necessarily relevant to whether we understand each other (sufficiently for the task in hand) by the end of the interaction. On this understanding, the use of a syntactic expansion can be seen as an efficient way to perform repair or clarification, with the syntactic tying illustrating exactly where the source of the possible misunderstanding lies, as in (2.11), repeated below.

2.11 **G:** That's right they had to go on a rota.

**A:** Run by the Dock Commission?

**G:** Run by the Dock Commission. *[BNC H5H 100-102]*

## 2.4 Coordination

If we accept that the only meaningful way to consider communication is as co-constructed, with cross-person CCs an explicit example of general processes then how we coordinate at a variety of different levels is a key issue.

Much research has been done in coordination as a communicational strategy, from coordinating attention (e.g. gaze) and gesture, to linguistic features (phonology, morphology, syntax) and meanings (semantic or inferential) to situation models and discourse structure. Whether all or any of these levels are necessary or what level of influence each has on the others, or even if they can be seen as looking at the same type of phenomena are open questions, but for the purposes of this thesis I shall focus on a subset of coordination phenomena, which I believe are relevant to an analysis of compound contributions.

### 2.4.1 Interactive Alignment

Pickering and Garrod (2004) take the evidence of alignment studies, such as those on lexical alignment (Brennan and Clark, 1996; Metzger and Brennan, 2003) and syntactic alignment (Branigan et al., 2000a,b, 2003), coupled with the importance of dialogue, as

---

<sup>16</sup>Note how misinterpretations become more obvious in exchanges by e-mail which are at once chatty and non-interactive (Kruger et al., 2005).

evidence towards their mechanistic account of language – the *Interactive Alignment* model of dialogue processing. They argue that “successful dialogue occurs when interlocutors construct similar situation models to each other” (Pickering and Garrod, 2006, p206), which, for them, means *aligning* on situation models.<sup>17</sup> In other words, conversation is successful when interlocutors come to see the world in the same way.

Importantly, speaker and hearer are assumed to be interchangeable at any point, thus compound contributions are deemed to always be possible (including within words), in contrast to the CA view that they are only licensed in certain circumstances.

Thus, we predict that it should be more-or-less as easy to complete someone else’s sentence as one’s own, and this does appear to be the case. (Pickering and Garrod, 2004, p186)

However, as alignment of situation models is not normally negotiated explicitly, they assert that global alignment (at the level of the situation model) arises automatically from local alignment. They suggest that alignment, or coordination, at local levels, like lexical alignment (repeating or reusing lexical items) and syntactic alignment (repeating or reusing syntactic structures) leads to alignment at other levels.

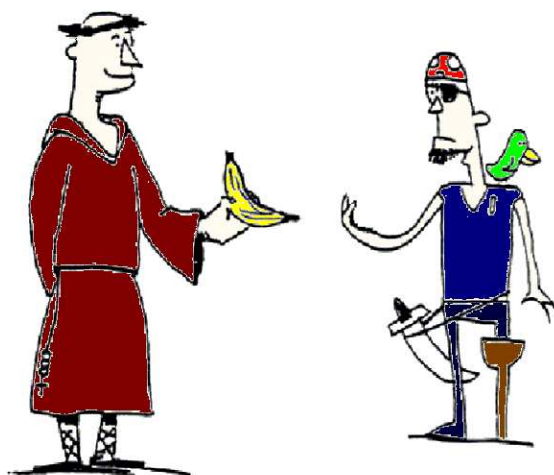


Figure 2.1: Example of picture card used in syntactic alignment studies

Evidence for these claims is taken from various studies (see e.g. Brennan and Clark, 1996; Garrod and Anderson, 1987), including Branigan et al. (2000a), in which syntactic

<sup>17</sup>They do not, however, define such situation models in detail, mentioning only that “key dimensions are space, time, causality, intentionality and reference to main individuals under discussion” and that they are “assumed to capture what people are “thinking about” while they understand a text” (Pickering and Garrod, 2004, p4).

alignment is shown to be stronger when the same verb is used (i.e. when lexical items are aligned). In the basic experimental set-up, there are two participants, one of whom is a confederate of the experimenter. The participants take turns describing picture cards (such as that shown in figure 2.1) to each other, the critical items of which are pictures of actions which require the use of ditransitive verbs in their descriptions. In English, there are two semantically equivalent syntactic structures which can be used to describe the actions; the so-called dative alternation. One of the possibilities uses a double object structure: “The monk handing the pirate the banana”, and the other uses a prepositional object: “The monk handing the banana to the pirate”. By getting the confederate to use a scripted description of the relevant ditransitive prime sentences, they manipulated which type of utterance the subjects had been exposed to. Participants are more likely to use the type of structure that they have just used or been exposed to. This seems to hold across language comprehension and production (Branigan et al., 2000b; Bock et al., 2007; Thothathiri and Snedeker, 2008),<sup>18</sup> from main clauses to relative clauses (Branigan et al., 2006) and even across languages in bilingual speakers (Hartsuiker et al., 2004; Loebell and Bock, 2003). Additional factors found to increase the *strength* of this syntactic alignment include how close the prime and the target are and participant role (Branigan et al., 2007).

These effects have usually been explained as arising from a ‘levels of activation’ model, where different parts of the language system (e.g. lexical items, syntactic representations, ‘lemmas’) are activated as they are encountered meaning they have a higher level of activation than their unexpressed equivalent counterparts and are thus more available for re-activation. This model makes various assumptions about the organisation of language in the mind, which are beyond the scope of the current discussion, but it is interesting to note that the results could also be explained in a memory-based account, such as that proposed by Horton and Gerrig (2005).

A more important question is whether these effects occur in spontaneous dialogue. Whilst Tannen identifies examples of repetition in conversation, these are often rhetorical devices and there is no assessment of their underlying frequency. Additionally, despite the fact that Branigan et al. state that “We can therefore conclude that syntactic alignment

---

<sup>18</sup>This suggests that both rely on the same underlying representations.

is a pervasive phenomenon in dialogue” (2007, p188), it is not clear that they can, in fact, do so. Not only are these results based on a limited number of sentence types (including the dative alternation, passives – *John shot Mary* versus *Mary was shot by John*, and adjectival placement – *the red goat* versus *the goat that is red*) which may or may not be indicative of wider effects, but the experimental set-up does not resemble free-flowing dialogue at all. The confederate is scripted, and the naive participants are also restricted in what they can say. This is clearly an artificial setting, and describing pictures to one another (without saying anything else) is a dialogue only in a limited sense. This raises questions about whether the conclusions can be sustained in more general conversational dialogue (though not about the validity of the results which have been shown repeatedly to hold in experimental settings). Corpus work (Healey et al., 2010; Reitter et al., 2006) shows that cross-person syntactic alignment does not seem to be a genuine effect in more spontaneous dialogue, and suggests that it may be an artifact of the task.

Interestingly, experiments in lexical alignment suggest that high levels of lexical alignment could also be an artifact of the experimental settings used to study them. Hadelich et al. (2004) compared relative lexical overlap in tangram descriptions in conditions allowing verbal feedback or not allowing verbal feedback. They found that in the conditions which were more akin to genuine dialogue (where verbal feedback was permitted), there was in fact less relative lexical overlap. They state:

...there was less overlap in the two verbal-feedback conditions than in the visual or the no-feedback condition. To some extent, this is surprising as the assumptions drawn on the basis of the alignment model pointed into the opposite direction. One way to interpret these results is to consider the overlap showing up in the verbal-feedback conditions as the automatic portion of overlap and the additional overlap in the visual-feedback conditions as stemming from other origins, such as pragmatic or situational influences or an aspect of audience design. (p39)

It is equally possible, however, that both lexical and syntactic alignment actually occur infrequently in ‘natural’ dialogue, especially as compared to highly constrained experimental dialogue (though note that Healey et al. (2010) did find above chance levels of cross-person *lexical* overlap, but not *syntactic* overlap).

## Predictability

Later work by Pickering and Garrod extends the interactive alignment model by emphasising the role of predictability in language comprehension.

Alignment is typically achieved (to an extent that makes people believe that they generally understand each other), because people start off at a very good point. They communicate with other people who are largely similar to themselves, both because they process language in similar ways and because they share much relevant background knowledge. This means that they can, in principle, use knowledge about themselves to understand and, in particular, predict their interlocutor. (Garrod and Pickering, 2009, p294)

This predictability works, as with alignment, at various different levels, and is evidenced by several different psycholinguistic studies, such as those which present subjects with partial sentences and show that processing is faster for highly predictable words (e.g. Schwanenflugel and Shoben, 1985), and those involving the visual world paradigm (e.g. Altmann and Kamide, 1999; Kamide et al., 2003), where participants gaze moves towards semantically predictable pictures before the word is heard (e.g. to the picture of a ‘cake’, the only edible item, after hearing ‘the boy ate...’). Other than semantic predictability, there is also evidence for prediction of grammatical categories, and specific word forms. Wright and Garrett (1984), for example, found that subjects were faster at identifying words from non-words when they syntactically followed from the preceding words, even where they do not semantically follow. These hypotheses are complementary to the observations of CA on cross-party CCs, which, as with Lerner’s compound TCUs, often provide highly predictable contexts for someone to complete another’s utterance (though note that these can be syntactic or not – compare the IF-THEN construction to the more general list environment).

In Pickering and Garrod (2007), they summarise thus: “language comprehension can be highly predictive, so long as linguistic or non-linguistic context supports these predictions. Therefore, comprehenders can get ahead of themselves and have more time to keep up with what they are encountering. They can also use prediction to compensate for problems with noisy or ambiguous input.” They hypothesise that the way this prediction is facilitated in comprehension is by using the production system to emulate what the

speaker might say.<sup>19</sup> The addition of prediction allows the model to account for data such as CCs (see below), which would otherwise be a somewhat mysterious phenomena of participants interchangability, but it remains unclear how the notion of predictibility can complement the previously invoked notion of repetition (as in syntactic alignment studies).

However, the characterisation of the comprehension system facilitating understanding (or being inherent to the notion of understanding) by emulating the production system raises questions. For example, if the comprehension system in some sense relies upon the production system (and notice that in their papers these are distinct notions, though they use the same underlying representations) then how do we explain the intuition that we can understand more of a language (either our first language as children or foreign languages learned later) than we are able to produce. A more parsimonious explanation might be to follow the approach of Fais (1994), which does not have the strict separation between comprehension and production that an emulation model would require. If this is feasible, and how this might be cashed out will be examined in more detail in chapter 8.

### **Syntactic alignment in compound contributions**

In contrast to the usual evidence for syntactic alignment, which relies on repetition of syntactic structures, compound contributions do not. What they assume rather is that parsing and production rely on the same syntax such that a syntactic representation<sup>20</sup> for a string being produced is equivalent to a syntactic representation built up in the parsing process. I can then use the representation I have thus far parsed as a starting point for the production of my completion, and you can switch to hearer, adding my words to the structure you had previously produced.<sup>21</sup> Notice that nothing here is repeated, it is only the underlying structure which is aligned.<sup>22</sup> This makes it both harder to test for experimentally, and also widens the scope of structural alignment; it might be argued that if I have successfully parsed anything you say, then I have aligned my structural representation with yours.

---

<sup>19</sup>Note that more recent work is extended to the action system such that hearers may also emulate what speakers might *do*, though I leave this to one side for now (Pickering and Garrod, 2009).

<sup>20</sup>Questions of what such a representation is will be left to one side for now.

<sup>21</sup>This raises questions regarding the nature of parsing and production, to be discussed later.

<sup>22</sup>This may only mean isomorphic in the sense of similar enough to admit of a mutually acceptable continuation, but for the moment I shall assume equivalence for ease of disposition.

Pickering and Garrod (2009) acknowledge these two different types of evidence for alignment; repetition (overt imitation) and complementary (e.g. continuing another's incomplete utterance). This is seen as analogous to actions; if you raise your arm, I may respond imitatively, by raising mine, or in a complementary fashion, by flinching. However, they assert that both are present in dialogue, whilst corpus evidence from Healey et al. (2010); Howes et al. (2010) suggests that the prevalence of syntactic repetition in dialogue has been overstated, and may not occur at all. This suggests that repetition may not be an indicator of structural alignment at all – or a special case of it. If our underlying representations are aligned then I should not need to repeat what you have said. Nonetheless, if the predictability component of the interactive alignment model is correct, then a complementary response, such as continuing another's (possibly partial) utterance should be easier and/or more common in highly predictable contexts (though as with repetition, predictability may operate on many levels, including lexical, syntactic and semantic, and it is unclear how they might interact). Note that other than CCs, ellipsis could be a measure of this type of complementary alignment, though such an exploration is beyond the scope of this thesis.

### 2.4.2 Common Ground

While Pickering and Garrod do not deny is that there is “common ground” (though this is only one interpretation of what it could mean to have aligned situational models), for them this is *implicit*, which they contrast with Clark (1996), for whom, they say, the common ground is explicit, meaning that it also contains the explicit knowledge that it is shared. Clark (1996) may disagree, and illustrates his usage of common ground using an example of standing on a beach with his son.

It is common ground for my son and me that, among other things, there is a conch shell between us. It is part of our common ground because it is included in a situation that also includes his and my awareness of that very situation. The situation *s* is the *shared basis* for our common ground. In this view, common ground is a form of self-awareness – self-knowledge, self-belief, self-assumption – in which there is at least one other person with the analogous self-awareness. (p94)

For Clark, a crucial point of advancing the joint project of dialogue is in establishing that we are sufficiently coordinated thus far to continue, a process which he calls grounding. “*To ground a thing, in my terminology, is to establish it as part of common ground well enough for current purposes* (Clark, 1996, p221, italics original). Grounding uses, for example, backgrounded responses (such as ‘mm’, or ‘yeah’) or non-linguistic cues (e.g. nods and smiles) to enable interlocutors to indicate that they have understood, and assimilated what has been said into their representation of common ground. Other responses may indicate difficulties in doing so, and signal a need for clarification or repair, but it is only with this feedback that each person’s representation of common ground can be accepted as sufficiently similar to allow continuation. In this way, grounding is an iterative process of demonstrative coordination, which continues throughout any exchange, building on what has already been grounded.

These notions are crucial to the idea of compound contributions if viewed from this perspective. Same-person cases may often be as a response to feedback about who has grounded what (this may of course be non-verbal, and note that even the absence of feedback where it would be expected can change what we consider our interlocutors to have successfully grounded) thus clarifying or making explicit what should be taken to be grounded for participants (Goodwin, 1979). Producing a continuation to another’s (possibly partial) utterance can be seen as a way of demonstrating that you have grounded what they have produced so far, as well as simultaneously being a device to extend the common ground, from a shared starting point. Given the question of whether the continuation is the same (conceptually) as the one that would have been produced, we also need the notion of common ground, because it is precisely our mutual knowledge which determines whether I can accurately complete an utterance that you have started (or deliberately subvert it in the case of “hostile” continuations). In most of the examples shown in chapter 2, the proffered completions are only possible because whatever it is that the dialogue participants are talking about is in their common ground – though note that a loose notion of common ground suffices for these purposes.

How the idea of grounding can be formalised has been explored in Traum (1994), in which *grounding acts* are carried out at the level of individual utterances to build up *discourse units* the level at which core speech acts are realised, through being grounded.



Some of the main grounding acts are *initiate*, to start a new discourse unit; *continue*, which continues a previous act and is syntactically and conceptually part of it; *acknowledge*, to show understanding and *repair*, which changes the content of the discourse unit, by correcting previously uttered material or by adding omitted material. Viewing CCs in these terms (as Poesio and Rieser, 2010, do; see section 2.5.2) same-person cases may be *continues* acts, but cross-person cases cannot, as

... a speaker cannot *continue* an utterance begun by another agent. The speaker could produce an utterance which contains a syntactic continuation of, or conceptually related material to another utterance by another agent, but this would not be a *continue* act. Depending on context it would be interpreted as either an acknowledgement (e.g. if one was just completing the other's thought), a *repair* (if one is correcting what *should* have been said, or an *initiate* of a new DU (if it provides new information) (Traum, 1994, p41).

This contrasts with the view that participants can be seen as grouped into parties, in which case it ought to be possible to perform a *continue* act on another's utterance provided the initial speaker was in a party with the continuing speaker.

If cross-person continuations are treated as a particularly explicit form of acknowledgement, then the initial contribution is grounded, but the speech act carried out by the discourse unit is still taken to be owned by the initial speaker (not the supplier of the continuation), which should mean that it is inappropriate for the next utterance to be an acknowledgement by the initial speaker as one cannot acknowledge one's own immediately prior utterance. Contrarily, if the continuation is taken to be a repair, then the speech act performed by the discourse unit might be considered to be jointly owned, making acknowledgement an appropriate next move.

### 2.4.3 Interactive Alignment versus Common Ground

Although these accounts approach coordination from different angles, it is not clear that they are incompatible. The interactive alignment model does not offer a treatment of the contribution of backchannel responses whilst Clark's grounding model puts them as a central mechanism. Although Pickering and Garrod divorce their situation model from

general knowledge and assume that this constitutes a difference, it clearly maps to Clark's division between personal and cultural common ground, and it is not clear how either general knowledge or cultural common ground gets accessed or updated in either model. It is also not clear that in a grounding model we need be explicitly aware of updates to common ground, and intuitively it seems that we would not be, but that we might be able to access such 'knowledge' only afterwards.

The difference between these models, then, is in the mechanisms used to update your situational model, or common ground and whether these updates are facilitated by automatic low-level coordinative processes, or by (possibly automatic, but often assumed by critics to be accessible and conscious) higher level coordinative processes. Thus the two accounts are not incompatible, as the further possibility exists that both types of coordination operate in tandem.

Both these accounts assume that the reasons we can be confident about communicating successfully are because as people who speak the same language, we have the same linguistic mechanisms, and/or shared knowledge, but both conclude that coordination is vital in dialogue. The difficulties arise in understanding the complex relationships between shared context and shared mechanisms and how coordinated we need to be (or assume we are) in any interaction and at a variety of levels.

## 2.5 Linguistic models

Most linguistic models do not directly address the issues of compound contributions, as they, along with other dialogue phenomena such as disfluencies and repair, are seen to be examples of language-in-use (*performance*) errors which do not accord directly with the 'perfect' (*competence*) grammar people have in their heads (see section 2.5.1). However, even allowing for a grammar which is in some sense independent from parsing and production mechanisms (which translate between the competence grammar and the imperfect performance data) leads to questions about CCs (especially cross-person cases) which would require a switching of roles from speaker to hearer or vice-versa. Given the generally assumed model of language production (Levelt, 1989), in which an intended message is first planned and then encoded into its linguistic form (see section 2.3.1 for

arguments against this view), this switch from parser to producer is a non-trivial task.<sup>23</sup>

### 2.5.1 Incrementality

One consequence of accepting the competence-performance distinction is that competence grammar relies on complete sentences whilst actual dialogue is riddled with incomplete or partial utterances that we nevertheless have no trouble interpreting, such as incremental clarification requests, where a term may be queried (without causing any apparent problems to either speaker) even though the message is not yet complete, as in 2.14, below.

(2.14) **U:** And er they X-rayed me, and took a urine sample, took a blood sample.

**U:** Er, the doctor

**V:** Chorlton?

**U:** Chorlton, mhm, he examined me *[BNC KPY 1005-1008]*

Psycholinguistic evidence (exemplified by Kamide and Mitchell, 1999) shows that interpretations are built up incrementally, from left to right as the words in a string are encountered, and that in languages like Japanese in which the verb comes at the end of the sentence (so-called “head final languages”), representational features can be postulated before ‘head’ words appear (for example, case markers on nouns could indicate structure to be built up prior to the verb being encountered). Arguments can therefore be assigned to an as-yet-unprocessed verb, as they are encountered, with the relationships between them already determined. This type of approach is more in line with speaker intuitions, as Steedman and Baldridge (2003) note, “Dutch, German and Japanese speakers greet with hilarity the suggestion that their languages prohibit any analysis until the verb group . . . has been processed.”

Although incremental parsing algorithms have been developed for head driven grammars, the separation of the grammar from the mechanisms of parsing means that some grammatical rules which could be explained by parsing constraints, appear arbitrary.

---

<sup>23</sup>Note also that while there are many incremental parsing systems implemented (e.g. CCG, Steedman, 2000; Niv, 1994), generation tends to be head-driven (the issue being how to map a complete logical form onto a string of words), which would be problematic for the switch of roles that occurs in cross-person CCs (though see e.g. Kelleher and Kruijff, 2006, for work in this direction).

This means that much of the explanatory power for language-in-use resides not in the grammar, but in the parsing and production mechanisms that use it, but are independent of the grammar itself. In such a set-up (cross-person) compound contributions ought to be difficult to either produce or interpret, as switching roles would be a non-trivial matter. This is especially true in the completion cases in which neither part of the CC is itself complete and therefore ought to be uninterpretable. These points link to observations from interactional linguistics; Fox (2007) outlines several ways in which interaction, including practices such as turn-taking, shapes grammar itself. These claims will be looked at in more detail below.

### 2.5.2 PTT

Based on the corpus data gathered by Skuplik (1999), as discussed in section 2.2, Poesio and Rieser (2010) present a general model for *collaborative completions* (a subclass of cross-person CCs) based in the PTT framework. This model combines 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 essential to dialogue management. They conclude that an intention-based model, based on e.g. Clark’s *common ground* (1996), is therefore more appropriate for modelling the phenomena. Their primary concern is to show how such a model can account for the hearer’s ability to infer a suitable continuation, according to shared plans, 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. In addition, as pointed out in Gregoromichelaki et al. (2011), the use of a string-based syntactic analysis leads to problems in cases where referents switch across the split point, as in (2.15), in which the referent “me” would be inappropriate if C had completed the utterance herself, and (2.16), in which the sentence string produced by the CC is “But have you burned myself”, which is ungrammatical.

(2.15) **C:** Nicola is

**A:** looking for me

[Natural data]

(2.16) (with smoke coming from the kitchen)

**A:** I'm afraid I burnt the kitchen ceiling

**B:** But have you

**A:** burned myself? Fortunately not. *[From Gregoromichelaki et al, 2011]*

Assumptions of intention recognition are, of course, non-trivial, especially given the potential multifunctionality of CCs (as discussed in section 2.1.4). The task-oriented data which Poesio and Rieser (2010) base their model on may be analysable in terms of shared joint plans, due to the specific constraints of the task (in which participants collaborate on a building task so the notion of joint intentions is fixed in advance), but whether this is feasible in generic dialogue, or whether it is necessary for communication in all cases (Gregoromichelaki et al., 2011), is an open question. There are four possible reasons why a cross-person completion is offered in the example examined in Poesio and Rieser (2010) which are, responding to a request, voluntary coordination-level control (acknowledgement), cooperativeness (repair) and blurting out. However, the authors do not identify which reason has prompted the completion in the specific case, and nor is it clear that the incoming speaker had to have in mind a particular intention when producing the completion – it may only be possible to infer such an intention based on the response to the utterance.

Indeed, in cases of “hostile” completions, such as (1.2), or (2.17) it seems clear that there can be no joint intentions, but the grammatical resources used in contributing to the CC are the same.<sup>24</sup> The PTT model is therefore an ambitious one, attempting to explain not just how participants are able to co-construct utterances but also how they choose what the continuation is.

(2.17) *Louis, age 5, is slowly reading out loud to his uncle. His older brother, Miles, age 9, is playing a handheld computer game on the other side of the room, and not obviously paying attention:*

**Louis:** I've ... got ...

**Miles:** a lovely bunch of coconuts.

*[Natural data]*

---

<sup>24</sup>Note that the PTT model discussed does not consider these cases as they are not taken to be collaborative.

### 2.5.3 Dynamic Syntax

Purver et al. (2006) present a grammatical model for cross-person compound contributions (which they call *split utterances*), using an inherently incremental grammar formalism, Dynamic Syntax (Kempson et al., 2001; Cann et al., 2005). In Dynamic Syntax (DS), *interpretations* are built up in a strictly incremental (word-by-word) fashion meaning that the resulting trees are semantic, with no record of the word order of the string. They are constructed using a combination of computational actions (which may apply at any point, subject to restrictions) and lexical actions (which are triggered when the words are encountered). Parsing and generation are tightly coordinated, using exactly the same procedures, with the main difference being that generation requires a goal tree (though note that this may be a partial tree – all that is required is that it be slightly more complete than the existing parse state). These factors mean that CCs are predicted to be strikingly natural, as speaker and hearer are building up the interpretation in parallel and may switch roles at any point. Note also that because the decorations on the trees are semantic, rather than the words themselves, that the issues with referent switching discussed in section 2.5.2, above, do not apply.

DS shows how syntactic and semantic processing can be accounted for no matter where the split point occurs; however, as the focus is on grammatical processing, there is no DS account of any higher-level inferences which may be required (some work towards this has been presented in Purver et al., 2010). Note also that although DS is presented as a grammar formalism that is a model of parsing and generation, there has been little work done on how any particular parse is computed, just the stipulation that there must be a possible parse for a string to count as grammatical (though see Sato, 2011). The formal tools of DS and whether they allow us to account for all aspects of CCs will be discussed in chapter 8 with a comparison to how the PTT model accounts for them.

## 2.6 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 centres mostly on prosodic rather than grammar/theory-based factors.

For cross-person CCs, the only system 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). However, it is to be hoped 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.

Note that these are incremental dialogue management systems which are not currently integrated with the kind of incremental linguistic models discussed in section 2.5, above. Ways in which these approaches might be combined to produce a dialogue model which is incremental at all levels, and thus account for all aspects of CCs, is a fertile area for future research, which the results presented in this thesis offer some pointers towards.

## Chapter 3

### Hypotheses

---

The discussion of CCs in chapter 2, both in terms of previous studies of the phenomena and how they might be incorporated into different models of communication raises a number of questions with potentially conflicting predictions. These will be outlined below, under two broad headings of *constraints* on where and when CCs can and do occur, and the *consequences* of producing or interpreting CCs.

#### 3.1 Constraints

Does syntactic structure place significant constraints on where a split point may occur? Is it possible to produce a continuation at any syntactic point in an utterance, (as Dynamic Syntax – section 2.5.3 – predicts), or do they occur only in highly constrained contexts?

**Hypothesis 1** *There are no syntactic constraints on where a split point may occur*

Relatedly, Szczeppek (2000a) hypothesises that repair is more frequent where split points fall between syntactic constituents, which is evidence for strong syntactic constituency. However, if speaker and hearer really are interchangeable at any point there should be no effects of syntactic constituency on the likelihood of repair.

CA analyses of dialogue phenomena predict that compound contributions should preferably occur at turn-transfer points that are foreseeable by the participants (TRPs). Expansions (section 2.1.3) are CCs with split points at TRPs,<sup>1</sup> and we should therefore

---

<sup>1</sup>Note that some completions may also have a split point at a TRP (as for example in cases of invited completions), but may not, whilst expansions are all necessarily at TRPs.



expect them to be more common than completions.

**Hypothesis 2** *Expansions are more common than completions*

In the case of completions, speaker change should also occur at projectable points (e.g. Lerner’s compound TCUs section 2.1.1) where conditional entry into another’s turn space is permitted. As in the interactive alignment model (section 2.4.1), one of the cues that participants can use for such prediction is syntax.

String-based grammars are sentential, and it is complete sentences that serve as the unit for interpretations in e.g. Relevance Theory (section 2.3.2). Producing a continuation, especially of another’s prior incomplete contribution should be both unlikely and costly to the individual, because in order to do so one would have to have come up with a probable complete proposition from the incomplete input, and then produce only the part of it that had not yet been articulated. In this case, it ought to be only possible to continue another’s utterance if the projected proposition were highly predictable (e.g. in idioms, or in terms of mutual knowledge).

**Hypothesis 3** *Cross-person completions are more likely when they are syntactically and/or pragmatically predictable*

If cross-person CCs are genuinely used to project groups as turn-taking entities (Schegloff, 1995), then we would predict more in task-based dialogues. Antaki et al. (1996), for example, claim that there were more (of their subset of) cross-person CCs in task-based dialogues, which require coordination because such tasks are formulated in such a way as to require participants to form parties (or coalitions). Alternatively, if it is mutual knowledge that influences when people typically produce CCs, we should expect to find different patterns of CCs in task-related dialogues with different patterns of shared goals and mutual knowledge. For example, is it the case that there are more cross-person CCs in tasks where participants share a goal, but do not have the same access to the relevant information, or is it only the case that continuations are produced when participants have a stronger claim to sharing knowledge, not just about what the aim of the task or conversation is, but also the information required to achieve it. More generally, do people continue another’s prior contribution only when they can make assumptions about what

is shared (or more loosely, what is grounded), or are they just more likely to produce something that is a continuation when they do share information?

**Hypothesis 4** *There are more cross-person CCs in collaborative task-based dialogues than in general conversation, and tasks which impose parties have more cross-person CCs than those which do not*

If turn-taking is a key factor in when CCs are produced then the medium of communication should provide additional constraints on their production. In text-based chat, a medium that is non-linear, and in which contributions are constructed in private before being shared (see section 4.3, for a full discussion of these issues) CCs should be more difficult or impossible to produce and there should be different distributional patterns. For example, there will be no overlapping continuations in text chat, and there may be differences in sequential ordering, such that if participants construct a contribution as a continuation to an immediately adjacent contribution, other material may intervene as it can be constructed simultaneously.

**Hypothesis 5** *There are fewer CCs in text-based dialogues than face-to-face conversations*

### 3.2 Consequences

Does a CC function as a single turn that just happens to have been produced in two (or more) parts (potentially by more than one person), or are they quite separate contributions that build parasitically on prior contributions?

This is an important question in regards to CCs as it has been previously used to motivate what is included in the phenomena and what is excluded, as, for example with Szczeppek (2000a), who does not include appendor questions, because though they are syntactically built off a prior contribution they do not continue its actions.

How this question is resolved is crucial to an understanding of why people produce continuations in dialogue, when there is often (if not always) a semantically equivalent alternative formulation available, that does not build syntactically on the prior contribution in such an explicit way. Are CCs used to illustrate one's coordination with one's interlocutor, or are they in fact used as an efficient way to move the conversation forwards

in a variety of ways, whilst explicitly making use of what has already been produced and (syntactically) signposting that this is the case?

If speakers and hearers are interchangeable at any point (as in e.g. DS, section 2.5.3), and parsing and production of language use the same mechanisms then producing a continuation should not be more difficult for an interlocutor than producing an independent syntactic unit. Cross-person CCs should therefore be just as easy as same-person ones, and follow the same pattern of distribution.

**Hypothesis 6** *Same-person and cross-person CCs follow the same distributional patterns*

There should also be no additional processing costs for a third party interpreting a cross-person CC than a same-person CC or single contribution. This is despite the intuition that information from potentially conflicting sources must be integrated and interpreted as a single syntactic, semantic and pragmatic unit, because one is assumed to add incoming information to one's parse in an egocentric manner.

It may even be less costly to start from an existing parse state than starting afresh – the principle of least collaborative effort (Clark and Wilkes-Gibbs, 1986) would predict that cross-person CCs are actually easier to produce and process than non-connected contributions, as the interlocutors have made the link between contributions explicit by utilising the syntactic structure of the contributions.

**Hypothesis 7** *Cross-person CCs are at least as easy to interpret as same-person CCs*

If a CC is treated as a single turn, there are different possibilities regarding who has responsibility for it. According to Lerner (2004), ratifications are appropriate only when the original speaker maintains authority over the complete CC. However, if, as claimed in a grounding model, “completions themselves can be viewed as a particularly explicit form of acknowledgement” (Poesio and Rieser, 2010, p20), then the preferred next move ought to be to initiate a new discourse unit (as grounding acts do not themselves need to be grounded, preventing an infinite regress). Contrarily, expansions are often analysed as repair acts, adding omitted material, and the preferred next move ought therefore to be an acknowledgement.

**Hypothesis 8** *Acknowledgements are more likely following expansions than completions*

As discussed in section 2.1.6, if parties really are turn-taking entities, upon which the no-gap, no-overlap model of turn-taking applies, then there ought to be more overlapping contributions in within-party turns, where all members of the party are equally entitled to the turn space. Such parties should also be treated, for the purposes of the conversation at the point at which they are relevant, in the same way as an individual. This has consequences for CCs, in terms of ratifications, repair, and who is entitled to speak next.

Lerner (1993) hypothesises that extending a prior contribution may be a way to perform other-repair in a way that does not intrude upon the turn space of the TCU-in-progress. As self-repair is generally preferred, it may be that such cases are better analysed as ‘within-party’ repair analogously to self-repair.<sup>2</sup> If cross-person CCs are often used as face-preserving ways to perform other-repair, then ratifications should be uncommon (as people do not ratify their own repairs). Even if not taken to be repair, as both contributors to the CC have just completed a TCU jointly, a different person (not a member of the party) should be more entitled to take the subsequent turn.

**Hypothesis 9** *If cross-person CCs are taken to be jointly owned, then the co-constructors of the CC should be less entitled to provide the subsequent contribution than their interlocutors, and ratification should be rare*

For Lerner, ratifications following cross-person CCs are appropriate when the continuation is addressed to the original speaker, but not when addressed towards a third party. This must always be the case in dyadic dialogue,<sup>3</sup> but not in multiparty dialogues, meaning that ratification following CCs ought to be more common in dyadic than multiparty dialogues.

Similarly, in multiparty dialogues, if there are *cross-party* CCs in which the speaker of the antecedent is in a party with people other than the supplier of the continuation, then it should be possible for any member of their party to provide ratification, not just the actual speaker. It should not be generally possible, however, for a third party to provide ratification.

**Hypothesis 10** *Cross-party CCs can be ratified by any member of the antecedent owner’s party*

---

<sup>2</sup>This possibility would also map to Ginzburg et al. (2007), which provides a model for unifying self- and other-repair, but such considerations will be left aside for future research.

<sup>3</sup>Discounting the possibility that there is some non-present ratified participant.

However, even if cross-person CCs are taken to be indicative of parties, does their use demonstrate party membership or are they just more likely if there is a party? In other words am I more likely to produce a CC if I believe I am in a party with you or does producing a CC make it useful to treat it as if we are a party?

**Hypothesis 11** *The presence of cross-person CCs will lead interlocutors to act as if parties have been formed*

The answers to these questions have implications for how we understand the nature of CCs and will be assessed with regards to the empirical analyses of CCs, to which we now turn.

## Chapter 4

# Compound Contributions in Dialogue

---

This chapter will first introduce the terminology of CCs to be adopted throughout this thesis, and then report on the findings from a corpus study of a spoken dialogue sample, followed by a comparative study using two text-based and one spoken task-based corpora. These corpus studies demonstrate the frequency of CCs, and their distributional properties, addressing some of the questions raised in chapter 3.

### 4.1 Terminology

The earlier definition of compound contributions – dialogue contributions that continue or complete an earlier contribution – begs several questions; most importantly, what is a ‘dialogue contribution’? The use of the term in this thesis can be best explained by reference to a short extract of dialogue taken from the British National Corpus (4.1).

#### 4.1.1 Contributions

Each of the transcribed lines (1-13) is a *contribution*. This usage is intended to correspond to Clark’s (1996): “a *contribution* to discourse – [is] a signal successfully understood” (p227).<sup>1</sup> With transcribed corpus text, of course, it is not always possible to determine whether contributions have been successfully understood, as we do not have access to

---

<sup>1</sup>Note that Clark uses contribution to refer to both “the joint act of . . . completing the signal and its joint construal” and for the interlocutor’s “participatory act, his *part* of that joint act, as when we speak of Roger’s contribution to the discourse.” Contribution is used in this thesis in the second sense only, and of course transcription protocols may mean that this is a loose mapping at best.

non-verbal signals (such as nodding). Contributions are therefore taken to be stretches of talk bounded by a change in speaker, a significant pause, or the end of the sentence, and I further 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.

- (4.1)
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]

*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. 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 (4.1) 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.<sup>2</sup>

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 double 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 the definition above, 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 a single contribution. To avoid possibly misleading conclusions in the corpus studies CC figures are reported both including and excluding such cases (see section 4.2). Note, however, that these concerns only apply to same-person CCs and not to cross-person CCs.

#### 4.1.2 Turns

I 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*. This notion allows CCs which span multiple turns to be distinguished 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.

In this thesis, as my interest is general, I use the term **compound contributions**

---

<sup>2</sup>Though of course it may turn out that CCs and other tying techniques are not, in the end, reliably different in kind, such considerations are beyond the scope of this thesis.



(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. I 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 they are termed **same-person** CCs.

As not all cases will lead to complete propositions, and not all will be split over exactly two contributions, terms like *first-half*, *second-half* and *completion* are avoided: instead the contributions on either side of a split point will be referred to as the **antecedent** and the **continuation**. In cases where a 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. It is then possible to talk about the **completeness** of each portion independently, with the traditional completion/expansion distinction corresponding to the completeness (or otherwise) of the antecedent. See the section on the annotation scheme (in 4.2.1) for details of how completeness is assessed.

Given the above definitions I will now look at a general corpus of spoken English, and three task-based corpora (two of text-based dialogues, and one spoken) to see how the evidence bears on a number of questions surrounding the phenomena of CCs.

## 4.2 BNC corpus study

To answer some of the questions raised in chapter 3, a corpus analysis of CCs in the spoken portion of the British National Corpus (Burnard, 2000) was carried out. This corpus was chosen because it contains a vast number of genuine spoken dialogues across a wide range of people, thus allowing us to examine the prevalence of CCs in a variety of situations not restricted to the task-based dialogues which previous corpus studies (section 2.2) tend to have analysed. The BNC consists of 100 million words of British English collected in the early 1990s. The spoken portion includes over 10 million words in approximately 1 million contributions, split between what the BNC defines as context-governed dialogue (tutorials, meetings, doctor’s appointments etc.) and demographic dialogue (casual unplanned conversations).

Specific questions to be addressed in this section are concerned primarily with the dis-

tributional properties of CCs – how frequently they occur and whether there are distinct types which are preferred. In terms of the hypotheses outlined in chapter 3, we expect there to be no syntactic constraints on where CCs occur (hypothesis 1), but with a preference for expansion CCs over completions (hypothesis 2), and continuations more likely to occur at if what follows the split point is predictable such as Lerner’s compound TCUs (hypothesis 3). If speakers and hearers are interchangeable (as the interactive alignment model predicts), we also expect there to be no differences in distributional patterns between same and cross-person CCs (hypothesis 6). Looking specifically at cross-person CCs, if a CC operates as a single turn that just happens to have been jointly produced, then we expect acknowledgements to be more likely following expansions rather than completions (hypothesis 8), and the next turn following the CC to be likely to be taken by other interlocutors not involved in the joint construction of the CC (hypothesis 9).

#### 4.2.1 Materials and procedure

For this exercise, the portion of the BNC annotated by Fernández and Ginzburg (2002), chosen to maintain a balance between context-governed dialogue and demographic dialogue, was used. This portion comprises 11,469 *s-units* – roughly equivalent to sentences<sup>3</sup> – 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, annotators were allowed to use punctuation 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 3), where utterances may contain several *s-units* from the same speaker. Annotation was at the level of individual *s-units*, to allow self-continuations within a turn to be examined; the BNC’s *s-unit* therefore corresponds to my notion of dialogue *contribution*, and the BNC’s *utterance* to my notion of *turn*.

The BNC forces speaker turns to be presented in linear order, which is vital if we are

---

<sup>3</sup>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.

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. I will discuss this further below.

### Annotation Scheme

Tag	Value	Explanation
<b>end-complete</b>	y/n	For all s-units: does this s-unit end in such a way as to yield a complete proposition or speech act?
<b>continues</b>	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?
<b>repairs</b>	number of words	For continuations: does the start of this continuation explicitly repair words from the end of the antecedent? If so, how many?
<b>start-complete</b>	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 4.1: Annotation tags

The initial stage of manual annotation involved four tags: **end-complete**, **continues**, **repairs** and **start-complete** – these are explained in table 4.1, above.

For antecedents, the important question is 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 the interesting question is whether or not they would be complete if they did get finished (**start-complete**). 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.” These notions are by no means entirely clear cut (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, in the corpus study reported in section 4.2, in most cases there was a high level of agreement between annotators on what constitutes completeness.

S-units marked **end-complete=n** are those contributions which somehow *require* a

continuation – whether or not they receive one, whilst those marked **end-complete=y** could be taken by conversational participants to have reached a TRP. Similarly, continuations which could begin new TCUs are those which are annotated **start-complete=y**.

S-units which act as continuations are those marked with non-empty **continues** tags; and their antecedents are the values of those **continues** tags. Annotating continuations with the value of the antecedent, in a backwards looking fashion, allows us to have more than one continuation associated with a single antecedent (as for example in overlapping cases), and also to look up the end-completeness or otherwise of the antecedent, thus enabling us to distinguish expansion type CCs from completion ones. 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.

As for **repair**, as the focus is on CCs, my 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 (4.2), where the last word of the antecedent is repeated in the continuation. Repairs at other points in the contribution or turn are not taken into consideration.<sup>4</sup>

(4.2) **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]

Returning to the extract in (4.1), repeated below, 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

---

<sup>4</sup>Consequently, the 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. As such this is a necessarily impoverished notion of repair and this should be borne in mind in the following results and discussion.

**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*’.<sup>5</sup> Finally, the **start-complete** tag (also only applied to continuations) indicates whether the contribution starts in such 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.

(4.1)		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

---

<sup>5</sup> ‘*I’ve*’ is counted as two words as a contraction of ‘*I have*’.

### 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 – e.g. (4.3), where G’s second contribution could be seen as continuing either his own prior utterance, or A’s intervening contribution.

- (4.3) **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  
**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 annotators. First, all three annotated one dialogue independently, then compared results and discussed differences. They then annotated three further dialogues independently and agreement was measured; kappa statistics (Carletta, 1996) are shown in Table 4.2 below. This doesn’t, of course, give us definitive answers in specific cases, rather it assesses whether the annotators were able to make sufficiently similar judgements over a number of cases to support statistical analysis.

Tag	BNC Dialogue Code					
	Cohens’ Kappa			Absolute agreement		
	KND	KBG	KB0	KND	KBG	KB0
<b>end-complete</b>	.86-.92	.80-1.0	.73-.90	.97-.99	.96-1.0	0.96-0.99
<b>continues (y/n)</b>	.81-.89	.76-.85	.77-.89	.93-.96	.92-.95	0.91-0.96
<b>continues (ant)</b>	.82-.90	.74-.85	.76-.86	.93-.96	.91-.94	0.91-0.95
<b>repairs</b>	1.0-1.0	.55-.81	1.0-1.0	1.0-1.0	.98-.99	1.0-1.0
<b>start-complete</b>	.59	.68	.62	.87	.92	0.90

Table 4.2: BNC inter-annotator  $\kappa$  statistic and absolute agreement (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. These differences in reliability should be borne in mind in the interpretations of the results reported below.

#### 4.2.2 Results and discussion

The 11,469 annotated s-units yielded 2,231 CCs, of which 1,902 were same-person and 329 cross-person cases; 112 examples involved an explicit repair of the antecedent at the start of the continuation. The data come from the full range of dialogues; all dialogues had at least three same-person cases, though 4 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 Across-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 (4.1)). As explained in chapter 3, however, same-person cases are interesting in their own right. From a processing/psycholinguistic point of view, it would be useful 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 across-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 (4.1) lines 1-3). In the tables and results below, same-person CC figures are presented both including all cases, and excluding those cases which are either within-turn or separated only by non-verbal or overlapping material. These figures are labelled as *all* and *across-turn* respectively.

**General observations**

person:	Same-				Cross-	
	all		across-		(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 4.3: BNC antecedent/continuation separation

Looking at across-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 4.3). 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 (4.1), 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). This difference in distributions following a backchannel between same- and cross-person CCs, contrary to the predictions of hypothesis 6, provides evidence that people are not interchangeable at any point. Same-person cases are more often produced as a response to a backchannel, suggesting that shaping one's next turn as a response to feedback is a common strategy in dialogue. Note also that 13% of all s-units in the corpus sample were backchannels<sup>6</sup> so there are actually a greater proportion of same-person cases following a backchannel

<sup>6</sup>Note that this figure is based on the BNC part of speech tags, and as such may incorrectly include some answers to yes/no questions.



than would be expected by chance, suggesting that backchannels may be used as a cue for participants to perform a *continue* grounding act (Traum, 1994, as discussed in section 2.4.2).

### Completeness

In order to assess whether expansions are more common than completions (hypothesis 2) we looked at the end-completeness of each contribution. 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% of these get continued. This compares to 15% of end-complete s-units that get continued (591/930, 64% vs. 1577/10539, 15%;  $\chi^2_{(1)} = 1315.90, p < 0.001$ ),<sup>7</sup> showing that although incomplete s-units are more likely to be continued, incompleteness does not necessarily prompt the production of a completion. This implies that partial sentences can be interpreted, and the need for complete propositions does not drive the production of a continuation.

	person:	Same-				Cross-	
		all		across-		(all)	
		N	%	N	%	N	%
Antecedent end-complete	Y	1367	72	513	71	236	72
	N	535	28	213	29	93	28
Continuation start-complete	Y	224	12	99	14	45	14
	N	1678	88	627	86	284	86
Repair	Y	77	4	34	5	32	10
	N	1825	96	692	95	297	90
Total		1902		726		329	

Table 4.4: BNC completeness and repair

Confirming hypothesis 2, the majority of both same- and cross-person continuations (71% to 72%) continue an already complete antecedent, with only 28-29% therefore being *completions* in the sense of e.g. Ono and Thompson (1993). These figures contrast with those found by both Skuplik (1999) and Szczepek (2000a). As discussed in section 2.2, the large discrepancy between this BNC study and Szczepek (2000a), who found 75% completions and only 25% expansions is probably to do with her criteria, which only included cross-person CCs that also continued the action of the previous contribution, whilst we include all those that syntactically build on it, regardless of whether they continue the action of the antecedent or not. Many of those which do not may be

<sup>7</sup>Where  $p > 0.001$ , I report exact probabilities but throughout adopt a criterion probability level of  $< 0.05$  for accepting or rejecting the null hypothesis.

expansions, as with the extending clauses which may be a convenient way to initiate a new discourse unit, whilst indicating that it is linked (e.g. topically) to some prior discourse unit. In contrast, the difference between this BNC study and Skuplik's Bielefeld Toy Plane Corpus (BTPC) data (BNC 93/329, 28% vs. BTPC 54/126, 43%;  $\chi^2_{(1)} = 8.867, p = 0.003$ ) cannot be accounted for in these terms. It may be that different languages have different patterns of CCs, or that cross-person completions are more common in task-based dialogue (in line with hypothesis 4). This possibility will be revisited in section 4.4.

It is interesting to note from table 4.4 that the proportions of end-complete antecedents and start-complete continuations is not different for same-person CCs and cross-person CCs. This supports the hypothesis that they follow the same distributions (hypothesis 6), though as we have already seen this was not the case for the proportion of CCs following a backchannel response, suggesting a more complicated relationship between interlocutors in dialogue than simply being able to take over another's utterance at any point.

### CA categories

In order to test hypothesis 3, that CCs usually occur at predictable points, 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 performed 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 overestimates). 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 4.5 (where the || token represents the split point).<sup>8</sup>

The most common of the CA categories are Lerner's (1996) 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, across-turn 36/726, 5% vs. other 32/329, 10%;  $\chi^2_{(1)} = 8.53, p = 0.003$ ). Interestingly, the breakdown

---

<sup>8</sup>Note that the categories in Table 4.5 are not 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.

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, 13 by other; ending in ‘<pause>’ 19 continued by same, 19 by other;  $\chi^2_{(1)} = 4.77, p = 0.03$ ). 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.

person:	Same-		across-		Cross-	
	all				(all)	
	N	%	N	%	N	%
...    and/but/or ...	748	39	306	42	118	36
...    so/whereas ...	257	14	57	8	42	1
...    because ...	77	4	32	4	3	1
... er/erm    ...	35	2	21	3	13	4
... <pause>    ...	19	1	15	2	19	6
...    which/who/etc ...	26	1	11	2	5	2
... instead of ...    ...	1	0	0	0	0	0
... said/thought/etc ...    ...	12	1	5	1	0	0
... if ...    (then) ...	18	1	10	1	3	1
... when ...    (then) ...	6	0	4	1	1	0
(other)	783	41	317	44	161	49
Total	1902		726		329	

Table 4.5: BNC continuation categories

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 1-2%.

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*’, and the (other) category (see below).

These seem to be cases in which the content is *less* predictable, contra hypothesis 3, suggesting that people often utilise the grammar to tie a contribution to something prior, using this as a device to move the conversation forwards, without constraining themselves to the predictability of what their interlocutor might have gone on to say.

### Split point

The less obviously categorisable cases of the (other) category make up 41-49% of continuations, suggesting that though there are preferred split points (e.g. the extending clauses, as per hypothesis 2), there are no strict syntactic constraints on where they can occur (providing evidence for hypothesis 1). The most common first words were ‘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,<sup>9</sup> even within what traditional theories of grammar consider to be a single constituent,<sup>10</sup> such as noun phrases and prepositional phrases (e.g. (2.7), (2.8), (2.9), (4.2), repeated here).

(2.7) **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]

(2.8) **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.]]

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

[BNC H5D 123-127]

---

<sup>9</sup>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 such considerations are left 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]

<sup>10</sup>Of course, different grammars 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, the notion of constituency is limited to that of syntactic elements as in e.g. transformational grammars, or HPSG.

- (2.9) **A:** All the machinery was  
**G:** [[All steam.]]  
**A:** [[operated]] by steam [BNC H5G 177-179]
- (4.2) **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]

To further test hypothesis 1, that the split point can apparently occur between any types of words, the *completion* cases were annotated for whether the split point occurred within a syntactic constituent, or between constituents.<sup>11</sup> For same-person across-turn CCs, just over half are between-constituent (52%), whilst cross-person CCs appear to be more likely to occur within-constituent although this trend is not significant (102/213, 48% vs. 55/93, 60%;  $\chi^2_{(1)} = 3.28, p = 0.07$ ). This finding appears to be associated with repair (there seem to be more repairs in the within-constituent cases), which would be in line with the anecdotal evidence from Szczepek (2000a) about major constituency boundaries, but the numbers are too small to be sure.

The frequent clausal categories from Table 4.5 are all much more likely to continue complete antecedents than incomplete ones.<sup>12</sup> This is not the case for the (**other**) category; suggesting that split points often occur at random points in a sentence, without regard to particular clausal constructions, as predicted by hypothesis 1.

The continuations in the (**other**) category are far less likely to continue complete antecedents than the easily classifiable categories from table 4.5 (211/478, 44% vs. 535/574, 93%;  $\chi^2_{(1)} = 304.38, p < 0.001$ ). This shows that most of the easily classifiable categories are expansions, which are those cases where people use something that is grammatically a continuation even though they need not necessarily do so.

## Repair

If repair, as formulated, is to be taken as an index of the difficulty of integrating the continuation to the syntactic material offered in the antecedent, then under a model in

<sup>11</sup>Here I am concerned with only low-level syntactic constituency; a split point was counted 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.

<sup>12</sup>For the less frequent (e.g. *'if/then'*, *'instead of'*) categories, the counts are too low to be sure.

which speakers and hearers are interchangeable (such as the interactive alignment model) the proportion of repairs should be the same in same-person CCs as cross-person CCs. There should also not be any effects of where the split point occurs on the prevalence of repair.

However, cross-person continuations are significantly more likely to repair their antecedents than same-person cases (32/329, 10% vs. 34/726, 5%;  $\chi^2_{(1)} = 9.82, p = 0.002$ ), showing that there are differences between distributions of same- and cross-person CCs. In other words, although the distributions between same- and cross-person cases were equivalent in terms of completeness thus supporting hypothesis 6, it isn't this simple, and there appear to be additional constraints associated with continuing another's prior contribution that do not apply when continuing one's own.

Incomplete antecedents are also more likely to be repaired (across-turn (same- and cross-person) 51/300, 17% vs. 15/755, 2%;  $\chi^2_{(1)} = 82.51, p < 0.001$ ). Of the *completion* 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)). This pattern might be one that we would expect if cross-person continuations are often formulated as less exposed forms of other-repair.

Although more data are required to see if these are genuine differences, we know that repair of the start of a continuation is not common, so it appears that even when the split point occurs mid-constituent, people are able to just go on extending the constituent as if they were the original speaker. This suggests 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.

### **Next speaker**

In order to see if there are any effects on turn-taking or apparent party-membership (as per hypothesis 9, if cross-person CCs are treated as jointly owned then we should expect

a third person to be more likely to take the floor subsequently), the 329 cross-person CCs were further annotated according to who spoke next (after the continuation) and whether the conversation was dyadic or multiparty. Of the 53 dialogues, 34 were dyadic and 19 multiparty (though as observed in Eshghi (2009), many segments of multiparty dialogue are also dyadic in nature, I leave this complication to one side). This equates to 4919 turns in dyadic dialogues, in which there were 204 cross-person CCs (4.15%) and 2961 turns in multiparty dialogues in which there were 125 cross-person CCs (4.22%). These proportions are not different (204/4919, 4% vs. 125/2961, 4%;  $\chi^2_{(1)} = 0.03, p = 0.87$ ), which is surprising – if cross-person CCs are typically used to indicate party-membership we might expect a greater proportion in the multiparty dialogues. This could be taken to suggest that parties are not common in these annotated dialogues.

There is no difference in the proportion of occasions in which the participant who contributed the continuation also provides the next contribution, thus holding the floor (50/204, 25% vs. 26/125, 21%;  $\chi^2_{(1)} = 0.600, p = 0.44$ , in line with the figure of 3/15, 20% reported in Szczeppek, 2000a). However, in all dialogues the proportion in which the supplier of the continuation retains the floor is lower than in the general case. For all annotated s-units in the dialogues there is no change of speaker following an s-unit in 41% of cases, compared to 23% of cases following a cross-person CC (4791/11469, 41% vs. 76/329, 23%;  $\chi^2_{(1)} = 44.424, p < 0.001$ ), suggesting that the continuation is treated as a separate turn, and not as part of a single unit and that interlocutors supplying a continuation do not assume they have a right to retain the floor.

In the multiparty case, contrary to the predictions of hypothesis 9, the contributor of the antecedent is also twice as likely to take the next turn rather than a third person (70/125, 56% vs. 29/125, 23%;  $\chi^2_{(1)} = 28.112, p < 0.001$ ). This shows that people are more likely to re-take the floor after supplying an antecedent, meaning that the continuation is taken to be addressed at the antecedent owner, but without baseline figures for how often a third party should normally take the floor (given that many stretches of multiparty dialogue are dyadic in nature) is not clear. This point will be returned to in the comparative corpus studies.

### Ratification

Supporting the idea that ratifications ought to be more common in dyadic dialogues, if only appropriate when addressed to the original speaker, cross-person CCs are ratified or rejected by the initial speaker in (marginally) more cases in dyadic than multiparty cases (82/204, 40% vs. 37/125, 30%;  $\chi^2_{(1)} = 3.769, p = 0.052$ ). This does suggest that in dyadic dialogues cross-person CCs are more often interpreted by the antecedent owner as addressed towards them, potentially as a form of repair which requires acknowledgement or ratification, and not interpreted as simply the mechanistic articulation of predictable material by another (Fais, 1994; Pickering and Garrod, 2007).

Rejections are much less common than ratifications, occurring in just 13 of the 329 CCs (e.g. (4.4), where the rejection leads to a repair sequence, or (4.5) where the initial speaker continues as they had intended to do despite the interruption – compare this to (1.1), where the rejection takes the form of a rebuttal) compared to 103 clear ratifications and 22 cases where the CC overlapped the initial speakers continuation. These cases are interesting in their own right; in 17 of them the overlapping material was either the same as the original speakers continuation (2.8), or semantically equivalent (2.9) whilst in 5 cases the continuation offered was different (4.6), (4.7).

- (4.4) **A:** Erm because as Moira said that Kraft is erm ⟨pause⟩ now what was she saying, what was she saying Kraft is the same as ⟨pause⟩
- M:** Craft?
- Traidcraft?
- A:** No.
- C:** Maxwell.
- Maxwell House.
- A:** Maxwell House is Kraft *[BNC G3U 412-418]*
- (4.5) **R:** Whereas some er normal recruiting is one person one job and
- L:** That's it.
- R:** hard luck to the rest. *[BNC JA1 14-16]*



- (4.6) **J:** People don't mind [[waiting if they know]]  
**S:** [[the frustration.]]  
**J:** how long they're waiting for *[BNC H61 29-31]*
- (4.7) **J:** And I couldn't remember whether she said at the end of the  
three months or  
**A:** [[End of the month.]]  
**J:** [[just now?]] *[BNC H4P 17-19]*

Contrary to the predictions of hypothesis 8, cross-person CCs are more likely to be ratified or rejected in completions than expansions (59/93, 63% vs. 79/236, 34%;  $\chi^2_{(1)} = 24.600, p < 0.001$ ). This is surprising if completions are merely the vocalisation of already predicted material (as in the interactive alignment model, for example), or if they are taken to be explicit acknowledgements (in a grounding model) as they shouldn't need explicit evaluation or additional completion by the contributor of the antecedent.

In total, 138/329 (42%) of cross-person CCs are ratified; which, contrary to hypothesis 9 is not rare,<sup>13</sup> suggesting that cross-person continuations are often treated as not part of the same single unit as the antecedent. In a grounding model this suggests that these cases are those which are taken to be repairs, or new discourse units though note that we cannot distinguish between these possibilities. However, if they are treated as repairs then they are not treated as within-party repairs analogously to self-repairs (as discussed in chapter 3), because these should also not require ratification.

There are very few cases that we can identify as obviously demonstrating that participants are organised into parties, although there are a couple of instances in which a different person to the original speaker ratifies the CC, which are suggestive of party-formation (4.8), (4.9). Questions specifically regarding parties will be further explored in the comparative corpora in section 4.4.

- (4.8) **T:** You mean so w  
**A:** Shorter  
**J:** Yeah, not physical length of words. *[BNC J40 184-186]*

---

<sup>13</sup>Though see section 4.4 for comparative figures from other corpora.

(4.9) **M:** <unclear> what we need more is <unclear>

**G:** To keep the rain out.

**J:** Yeah. [BNC JNX 132-134]

### 4.2.3 Summary

With the range of possibilities regarding where the split point is able to occur, including potentially within a word (see footnote 9) 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.

Of course, it could be argued that this is an inevitable consequence of the way in which we have defined a CC, but note that if there is no way to distinguish the possibilities from surface features (not including prosodic factors, here, though these may prove vital) then this is a complex problem for dialogue models. Note also that the multifunctionality holds at all levels, including for example within the sub-class of cross-person completions that has previously been focused on.

The results show that CCs are common in dialogue. In line with hypothesis 1, there is no evidence that there are any specifically syntactic constraints on where CCs may occur with split points potentially possible at any syntactic point. However, contrary to hypothesis 3, relatively few of them occur in the highly projectable positions studied by e.g. Lerner, which may have consequences for accounts such as the interactive alignment model of which predictability is a key feature. How predictability at different levels affects the likely production of a continuation will be explored further in chapter 7.

There do appear to be pragmatic constraints on where CCs are *likely* to appear, with about three quarters of CCs occurring after complete antecedents, showing that interlocutors are sensitive to possible transition relevance places, as per hypothesis 2, with continuations systematically invited by a speaker or designed as though they are natural continuations of contributions that could be treated as complete.

Contrary to the predictions of hypothesis 6, although there are similarities between same- and cross-person CCs, e.g. in the distributions of expansions versus completions, there are also interesting differences. 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. This pattern might be one that we would expect if cross-person CCs, in virtue of being constructed as a continuation of the speakers utterance, provide a device that enables a less exposed form of other repair (as Lerner (1993) hypothesised).

Opportunistic CCs (following a '<pause>' or an '*erm*') are in general more likely to be cross-person cases; however there are again pragmatic constraints – cross-person CCs are more likely where the antecedent ends in an unfilled pause rather than a filled one. This also suggests participants are aware of turn-taking or sequential expectations, and that speaker and hearer roles carry different responsibilities.

There also seem to be different places when same- and cross-person continuations are offered; the majority of cross-person continuations are adjacent to their antecedents, whilst even considering within-turn cases this is not so for same-person continuations. Same-person continuations are far more likely to follow a backchannel or single other s-unit than cross-person cases, suggesting that it may be the feedback from one's interlocutor(s) that leads to producing something syntactically tied to one's own prior contribution.

Cross-person CCs do not seem to typically operate as a single turn that just happens to have been produced by more than one interlocutor, but might be better characterised as separate contributions that build parasitically on prior contributions. This is apparent in the ratifications offered following a cross-person CC, which should not occur if the speaker of the antecedent were treating the continuation as if they themselves had just finished their turn. That ratifications were also more likely to be offered following a completion rather than an expansion (contrary to hypothesis 8), suggests that completions cannot be taken to be solely grounding devices, but must also be being treated by the antecedent owner as at least potentially repairing the incomplete antecedent (in which case an acknowledgement is appropriate).

This means that although one can unproblematically finish or continue another's utterance, this does not give it the same status as if they had completed or continued it themselves.

The evidence regarding who speaks next following a CC does not clearly support the hypothesis that cross-person CCs are jointly owned turns (hypothesis 9), though there are some differences between who is entitled to speak next and ratifications between dyadic and multiparty dialogues, as well as a few apparently within-party ratification cases. Though there is no clear evidence in this corpus that cross-person CCs are often used or indeed treated as devices to demonstrate party-membership, this is at least suggestive of the idea that some CCs can be treated as if they project parties which will be explored further in section 4.4.

This corpus study remains neutral regarding whether cross-person CCs are more likely if participants are in parties, or make participants act as if they are in parties, and cannot answer questions regarding whether collaborative task-based dialogues encourage cross-person CCs (hypothesis 4) and text-based communication inhibits them (hypothesis 5). I therefore now turn to a comparative study of some task based corpora, using different modes of communication.

### 4.3 Face-to-face versus text dialogue

The evidence presented so far suggests that the grammar used in producing CCs has no specific syntactic constraints. There is also evidence that conversants are sensitive to turn-taking considerations, but whether these observations are generally true or hold only in face-to-face spoken dialogue is a question that will be addressed in this section, with a view to motivating the use of experiments using text-based dialogues (chapters 5, 6 and 7).

Of course, face-to-face dialogue and text-based chat are different in several respects. However, text chat is clearly an interesting field of study in itself, especially given the increasing prevalence of computer mediated communication. In 2009, for example, there were over 1 billion users of instant messengers, with 47 billion instant messages sent per day (9 billion of which were sent using the Windows Live Messenger application).<sup>14</sup>

---

<sup>14</sup>Source: <http://royal.pingdom.com/2010/04/23/amazing-facts-and-figures-about-instant-messaging-infographic/>

There are also important applications; Rosé et al. (2003), for example, compare text and speech based tutoring systems. However, there remain questions as to whether findings from face-to-face conversations (such as the BNC corpus study reported in section 4.2) apply in such a different modality.

The most obvious differences between face-to-face spoken dialogue and text-based chat are attributable to the channel of communication; speech versus text. In text-based chat such as MSN Messenger, 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 spoken dialogue is lost. Turns in text chat also persist physically in a way that they do not in spoken dialogue, as participants are able to scroll back through preceding dialogue.

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 to a lesser extent 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 empirical 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).

With a comparative study of map task dialogues using either text chat or speech, Newlands et al. (2003) show that while computer mediated communication has been thought to provide limitations on establishing mutual understanding, users very quickly adapt to the medium. Although novice users in the text condition took longer to complete a map route, by the third iteration, their accuracy on the task was as good as those in the spoken context, and used fewer words (which they explain using Clark and Wilkes-Gibbs's 1986 principle of least collaborative effort coupled with the non-linearity and ability to revise turns prior to sending them, as discussed above).

Importantly, CCs also occur naturally in text-based chat, as reported in Eshghi (2009). See, for example, (2.4) repeated here and (4.10), taken from the DiET chat tool environment.

(2.4) **M:** yep dr goes everyones happy

**N:** except the dr *[DiET SU1 19 240-241]*

(4.10) **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 ?? *[DiET SU1 13 35-37]*

#### 4.4 Task-based dialogues comparative corpora

In this section I present three corpus studies of task-based dialogues (two text-based and one face-to-face) with comparisons to the data from the BNC discussed in section 4.2.2.

The specific questions from chapter 3 that this section seeks to address involve questions related to task and medium, which the BNC corpus study provides a baseline for. Firstly, we have predicted that CCs will be less likely in a text based medium, due to the turn-taking constraints imposed (hypothesis 5). Conversely, cross-person CCs ought to be more likely in collaborative task-based dialogues, with a further increase if the task imposes parties from the outset (hypothesis 4). More specifically, if parties are genuine conversational entities then there should be more within-party overlap than across-party (as the turn space is taken to be jointly owned) and cross-party CCs should be ratifiable by any member of the antecedent owner's party, not just the supplier of the antecedent themselves (hypothesis 10).

This section also seeks to provide additional evidence for the findings from the BNC study regarding the apparent lack of syntactic constraints (hypothesis 1), the preference for expansions over completions (hypothesis 2) and the fact that predictability does not seem to have a significant influence on whether participants produce a continuation (hypothesis 3). The impact of the medium and task on who is entitled or expected to speak next (hypothesis 9), or whether acknowledgements are offered (hypothesis 8) as in the BNC will also be explored, as will whether or not the similarities and differences found between same- and cross-person CCs (hypothesis 6) hold across such a variety of communicational situations.

#### 4.4.1 Materials and procedure

Three corpora of task-based dialogues were annotated according to the scheme outlined in section 4.2.1. As can be seen from the descriptions below, these corpora vary along a number of different parameters, including task, interactional medium and number of interlocutors. The possible contributions of these different factors will be returned to section 4.5.

##### **Tangram task corpus**

This corpus was produced by Eshghi (2009), and consists of 2377 text contributions from triadic Tangram task dialogues (as used in Clark and Wilkes-Gibbs, 1986, *inter alia*) conducted using the line-by-line DiET interface. This is an MSN Messenger style chat interface which will be described in detail in section 5.1.1.

##### **Character-by-character corpus**

The character-by-character corpus consists of 3056 text contributions from 14 task-based dialogues using the character-by-character interface of the DiET chat tool. This interface does not allow participants to type simultaneously, and text appears as it is typed (in a character-by-character fashion) on all interlocutors' screens, in the same window in which contributions are typed. It will be described in detail in section 6.1. The corpus consists of 681 contributions from dyadic conversations of the balloon task (see appendix A), and 2375 contributions from 4-way discussions of the arctic survival task (see appendix B).

##### **Tuition task corpus**

For the third task-based corpus, I used the data from Battersby and Healey (2010), which includes 6600 contributions from conversations between 11 triads of students. Participants were seated close together in triangular formation and the conversations were both filmed and motion captured to produce a multi-modal corpus. I shall only be concerned with the speech data, though of course there may be implications for CCs in e.g. gesture (Bolden, 2003; Furuyama, 2002).

Each group completed three 'tuition task' conversations, in which one of the three participants was randomly assigned the role of 'learner' whilst the other two were 'instructors'. Each participant played the role of 'learner' in one of the three conversations. The task was for the two instructors to collaborate in describing hierarchical systems – either short Java programs, or the organisation of government in different countries,

to the learner, after which the learner was tested on their comprehension. Instructors therefore started the dialogue with the same information as each other, which was not available to the learner. This manipulation, and the nature of the task (collaborating to impart this information to a third party; the learner) has the effect of creating a ‘party’ between the two instructors.

#### 4.4.2 Results and discussion

With the exception of the same-person cases in the character-by-character corpus,<sup>15</sup> in all three corpora, CCs occur at least as frequently as they do in the BNC. These figures are shown in table 4.6.

Corpus	Person: Contributions	Same-		Cross-	
		all N %	across- N %	(all) N %	
Tangram	2377	396 17	154 6	120 5	
Character-by-character	3056	315 10	82 3	167 5	
Tuition	6600	1567 24	450 7	692 10	
BNC	11469	1902 17	726 6	329 3	
Total	23502	4180 18	1412 6	1308 6	

Table 4.6: Comparative overview of CCs by corpus

#### Within- and across-turn cases

As with the BNC data, we consider the results both including and excluding those cases which are within a single speakers turn. However, there are differences in these definitions between the corpora. In the tangram corpus though it is initially easier to delineate these cases, as non-verbal material is not recorded, the fact that participants can construct their turns simultaneously may have the effect of adding intervening material even when the participant meant their consecutive contributions to be considered as a single turn. Conversely, in the character-by-character text corpus participants cannot construct their turns simultaneously – only one participant may type at a time. For the face-to-face dialogues in the tuition corpus, there is also the added possibility that within-turn cases are affected by different transcription protocols. In the following discussions, I therefore focus on across-turn cases.

<sup>15</sup>Possible reasons for this will be discussed below.



### Backchannels

Contrary to hypothesis 5, which predicted fewer CCs in text-based dialogues, the proportion of same-person CCs in the text-based tangram corpus is not different to that from our BNC corpus study (tangram 154/2377, 6% vs. BNC 726/11469, 6%;  $\chi^2_{(1)} = 0.07, p = 0.787$ ); likewise for the tuition task spoken corpus compared to the BNC (tuition 450/6600, 7% vs. BNC 726/11469, 6%;  $\chi^2_{(1)} = 1.64, p = 0.200$ ). However, at only 3% when within-turn cases are excluded, the proportion of same-person CCs in the character-by-character corpus is lower, in line with predictions from hypothesis 5 (character-by-character 82/3056, 3% vs. BNC 726/11469, 6%;  $\chi^2_{(1)} = 61.09, p < 0.001$ ). However, because this difference applies only to the character-by-character text corpus and does not apply in the tangram corpus which uses a line-by-line text interface, it cannot be attributable to the use of text per se, but is likely to be due to the specific constraints of the character-by-character interface. Specifically, only one participant is able to type at a time – this means that e.g. mid-turn backchannel responses, which in the BNC (as previously discussed) lead to turns being split over several contributions, are simply not possible. In the BNC, a high proportion of same person CCs were only separated by overlapping material or a backchannel (across-turn 456/726, 63%), which, as can be seen from table 4.7 is not the case for the character-by-character corpus. If we remove these cases, then there is no difference in the distributions of same-person CCs (BNC 270/11469, 2% vs. 82/3056, 3%;  $\chi^2_{(1)} = 1.11, p = 0.293$ ). This suggests that, as people are constructing contributions as single stretches of talk in text chat that our treating same-person cases as CCs in face-to-face dialogue is not the correct analysis. Conversely, as previously discussed, it may be that feedback such as backchannels actually leads conversants to revise their contributions on the fly (as Goodwin, 1979, suggests), with CCs being a particularly good mechanism to use to do so. The design of the current chat tool does not allow us to distinguish between these possibilities, but there are also fewer backchannels overall in the character-by-character corpus (113/3056, 4% vs. BNC 1501/11469, 13%;  $\chi^2_{(1)} = 215.40, p < 0.001$ )<sup>16</sup> suggesting that this type of feedback is simply more difficult in this interface, possibly because turns may not overlap.

There are also differences in the distributions of same-person CCs separated only by a

---

<sup>16</sup>This is also a significant difference when we include text specific backchannels, such as ‘lol’, or smiles (202/3056, 7%).

backchannel between the BNC and the tangram corpus (BNC 456/726, 63% vs. tangram 29/154, 19%;  $\chi^2_{(1)} = 99.33, p < 0.001$ ) and a single (other) contribution (BNC 229/726, 32% vs. tangram 78/154, 51%;  $\chi^2_{(1)} = 20.42, p < 0.001$ ), with the BNC having more same-person CCs separated by a backchannel, and the tangram corpus more separated by a single other contribution.

person:		Same-				Cross-	
		all		across-		(all)	
corpus		N	%	N	%	N	%
Tangram	adjacent	242	61	0	0	67	56
	sep. by backchnl	29	7	29	19	14	12
	sep. by 1 s-unit	78	20	78	51	22	18
	sep. by 2+ s-units	47	12	47	31	17	14
	Total	396		154		120	
C-by-C	adjacent	228	72	0	0	143	86
	sep. by backchnl	11	3	10	12	7	4
	sep. by 1 s-unit	69	22	65	79	12	7
	sep. by 2+ s-units	7	2	7	9	5	3
	Total	315		82		167	
Tuition	overlapping	0	0	0	0	11	2
	adjacent	1048	67	0	0	487	70
	sep. by overlap	16	1	0	0	14	2
	sep. by backchnl	138	9	113	25	41	6
	sep. by 1 s-unit	258	16	232	52	94	14
	sep. by 2+ s-units	107	7	105	23	45	7
	Total	1567		450		692	
BNC	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	43	2	41	6	6	2
	Total	1902		726		329	

Table 4.7: Comparative antecedent/continuation separation

However, this difference cannot be accounted for by the medium, as the pattern was the same for the tuition corpus, with fewer same-person CCs following a backchannel and more following a single (other) contribution than in the BNC (113/450, 25% vs. 456/726, 63%;  $\chi^2_{(1)} = 158.09, p < 0.001$ ; 232/450, 52% vs. 229/726, 32%;  $\chi^2_{(1)} = 46.68, p < 0.001$ ). There is also an equivalent proportion of backchannels overall in the tangram and tuition corpora (263/2377, 11% vs. 771/6600, 12%;  $\chi^2_{(1)} = 0.654, p = 0.419$ ),<sup>17</sup> contrary

<sup>17</sup>Including text specific backchannels actually resulted in a greater proportion of backchannels in the tangram corpus (315/2377, 13%)

to Newlands et al. (2003) who found fewer overt grounding cues (i.e. backchannels) in text based dialogues.

This suggests that the observed differences in same-person CCs following a backchannel in task-based dialogues compared to the BNC are associated with the nature of the task and not the medium. In both the tangram and tuition corpora two participants must impart information to a third (either which tangram to choose next or the hierarchy being taught). Of course there may be different reasons for the differences between both these corpora and the BNC, including interactional constraints associated with the mode of communication. For example, the setup of the experiment in the tuition task means that participants are specifically seated so as to be able to provide non-verbal feedback (e.g. nodding). Non-verbal feedback is, of course, impossible in the tangram corpus, but might also be rarer in the annotated BNC dialogues, where it is not necessarily the case that interlocutors are even looking at each other. Additionally, in the tangram task the increased proportion of same-person CCs with one (non-backchannel) intervening contribution could just be due to those cases in which participants are constructing contributions simultaneously, with the continuation *planned* as being adjacent to the antecedent (4.11).

(4.11) **Match:** one leg is down

**Dir1:** just do it man

**Match:** & other one is up

[DiET Tang16 26-28]

### Multiple continuations

Antecedents receive more than one competing continuation in 2% of cases in the character-by-character corpus, in line with the proportion in the BNC (8/473, 2% vs. 53/2177, 2%;  $\chi^2_{(1)} = 0.95, p = 0.329$ ). However, there is a greater proportion in both the tangram corpus (39/477, 8% vs. 53/2177, 2%;  $\chi^2_{(1)} = 152.56, p < 0.001$ ), and the tuition corpus (244/1979, 12% vs. 53/2177, 2%;  $\chi^2_{(1)} = 152.56, p < 0.001$ ).

This difference is also likely to be related to the nature of the tasks, which, as discussed above, are similar in the tuition and tangram corpora. Contributions from the instructors (directors) presenting information, and from the learner (matcher) acknowledging it (or vice versa) may be both constructed and/or construed as continuations of the same antecedent (see e.g. (4.12) and (4.13), where multiple continuations overlap). In the

case of the tangram corpus, it is likely that the two continuations were constructed simultaneously (4.14), (4.15), which is impossible in the character-by-character corpus.

- (4.12) **Learn:** chief minister is in charge of the  
**Inst1:** [[executive]]  
**Learn:** [[executive part]] *[AHI-2Jersey 55-57]*
- (4.13) **Inst1:** so there's these are the responsibilities of the people's  
**Inst2:** consul  
**Inst1:** consultative  
**Learn:** [[assembly]] yup  
**Inst1:** [[assembly]] *[AHI-3Indo 53-57]*
- (4.14) **Dir1:** big triangle  
**Dir1:** flat side left  
**Match:** like a sitting dog  
**Dir1:** with a smaller triangle on its right *[DiET Tang25 731-734]*
- (4.15) **Dir1:** camel  
**Dir1:** facing left  
**Match:** FACING LEFT *[DiET Tang38 299-301]*

### Cross-person CCs

In all three task-based corpora there was a greater proportion of cross-person CCs than in the BNC, as predicted by hypothesis 4, following Antaki et al. (1996), who found a higher proportion of cross-person CCs in cooperative task-based (spoken) dialogues (tangram 120/2377, 5% vs. BNC 326/11469, 3%;  $\chi^2_{(1)} = 30.73, p < 0.001$ ; character-by-character 167/3096, 5% vs. BNC 3%;  $\chi^2_{(1)} = 50.60, p < 0.001$ ; tuition 692/6600, 10% vs BNC 3%;  $\chi^2_{(1)} = 460.22, p < 0.001$ ).

However, though there was an equivalent proportion in the two text-based corpora (120/2377, 5% vs. 167/3096, 5%;  $\chi^2_{(1)} = 0.46, p = 0.496$ ) there was a higher proportion of cross-person CCs in the spoken tuition task than in the text-based dialogues (692/6600, 10% vs. 287/5433, 5%;  $\chi^2_{(1)} = 107.91, p < 0.001$ ). There are several possible explanations

for this. Firstly, the tuition task is, as discussed, one which imposes parties, which it is hypothesised leads to an increase in cross-person (within-party) CCs (hypothesis 4). This possibility will be pursued in more detail in subsequent sections. Alternatively, it could be the case that though there is an increase in cross-person CCs in task-based dialogues it is not as marked in text mediated communication, which we might expect if CCs in general are less likely in text-based communication (as per hypothesis 5). There is also the possibility that both factors are at play. As we do not have a corpus of non-task-based dialogues in text chat it is difficult to distinguish between these possibilities, but note that the party explanation is the one that seems to best fit the data, as discussed below.

Though in all four corpora cross-person continuations tend to be adjacent to (or overlapping)<sup>18</sup> their antecedents, there are differences between them. As can be seen from table 4.7, this figure is higher in the BNC than the tangram corpus (BNC 290/329, 88% vs. tangram 67/120, 56%;  $\chi^2_{(1)} = 56.35, p < 0.001$ ), and the tuition corpus (290/329, 88% vs. 512/692, 74%;  $\chi^2_{(1)} = 26.53, p < 0.001$ ), but no different to the character-by-character corpus (290/329, 88% vs. 150/167, 90%;  $\chi^2_{(1)} = 0.31, p = 0.578$ ).

For the tangram corpus, this is likely to be because of the linearity issues discussed in section 4.3 (CCs which appear non-adjacent in the transcripts may have been started prior to intervening material being present), or due to the fact that the tangram corpus dialogues are all triadic, whilst many of the BNC dialogues were dyadic, as discussed previously. The pattern of results suggests that it is in fact a combination of these factors which lead to the decreased likelihood of adjacency in the tangram corpus. This is because the tuition dialogues are also all triadic and it too has a reduced proportion of adjacent cross-person continuations compared to the BNC, but not as few as the tuition corpus, suggesting that the linearity issue adds to the effect.

#### **Antecedent end-completeness**

Exploring the effects of task and medium on the finding that expansions are more common than completions (confirming hypothesis 2) we find that although in the BNC 8% of all contributions did not end in a complete way (e.g. because the person tailed off, or was interrupted), in the tangram corpus only 2% of all contributions did not end in a complete

<sup>18</sup>Note that overlap does not apply in the text-based dialogues.

way (930/11469, 8% vs. 54/2377, 2%;  $\chi^2_{(1)} = 101.62, p < 0.001$ ). This means that almost all of the CCs (96%) in this corpus were of the *expansion* type (see table 4.8), adding e.g. an adjunct to an already complete sentence. This is likely to be due to the constraints of line-by-line text-based communication, which, as discussed, means that participants do not transmit their turns until they are satisfied with them and if interlocutors interrupt whilst a contribution is under construction it can be revised prior to transmission (or even deleted without being sent). It is likely that this occurs fairly frequently, but the result is still that what is actually transmitted and therefore shared is end-complete. This private construction of contributions means that it is also impossible to project another's upcoming TRP.

	person:	Same-				Cross-	
		all		across-		(all)	
		N	%	N	%	N	%
Tangram	Y	366	92	147	95	118	98
	N	30	8	7	5	2	2
(Total)		396		154		120	
Character-by-character	Y	230	73	26	31	140	84
	N	85	27	56	68	27	16
(Total)		315		82		167	
Tuition	Y	916	58	302	67	422	61
	N	651	42	148	33	270	39
(Total)		1567		450		692	
BNC	Y	1367	72	513	71	236	72
	N	535	28	213	29	93	28
(Total)		1902		726		329	

Table 4.8: Comparative antecedent end-completeness

In the character-by-character chat tool dialogues, by contrast, 15% of all contributions did not end in a complete way. This figure seems to be inflated by cases where two participants attempted to start typing almost at the same time with the result that one of them appeared to type a single letter. If these cases are removed then there is no difference in the proportion of contributions which did not end in a complete way to the BNC (260/3056, 9% vs. 930/11469, 8%;  $\chi^2_{(1)} = 0.60, p = 0.439$ ). This is not the same as was seen in the tangram corpus, in which only 2% of contributions did not end in a complete way, demonstrating that it is not the fact that it is text-based communication per se but the specific constraints of the interface that accounts for the difference between the tangram corpus and the BNC.

In the tuition corpus, about 17% of s-units end in an incomplete way (1096/6600), and 70% of these get continued. This compares to 22% of end-complete s-units that get continued (764/1096, 70% vs. 1215/5504, 22%;  $\chi^2_{(1)} = 987.80, p < 0.001$ ), showing again that although incomplete s-units are more likely to be continued, incompleteness does not necessarily prompt the production of a completion.

That there are approximately twice as many incomplete contributions in the tuition corpus compared to the BNC may be due to differing transcription conventions, but is likely to be influenced by the nature of the task which encourages invited utterance completions (Ferrara, 1992), which are “questions masquerading as statements, with the addressee intended to supply the missing Wh-information” (p221). This stems from the fact that the task is in effect a shared memory task – first the two instructors must remember the hierarchical information, which they then impart to the learner, who has to make sure that they, in turn, remember it for the testing phase after the conversations have taken place. This does seem to be the case in a number of examples, as in (4.16), (4.17) and especially (4.18), in which the learner makes clear that they expect one of the instructors to continue their unfinished contribution, as they had forgotten who the highest member was and thus how the sentence should appropriately end.

- (4.16) **Inst1:** which includes  
**Inst2:** natural resources and the environment [AHI-2Indo 162-163]
- (4.17) **Inst1:** the executive ummm  
 make the decisions about what’s best for Jersey  
**Learn:** ok  
**Inst1:** and the scrutiny  
 erm  
**Learn:** scrutinises them [AHI-3Jersey 80-85]
- (4.18) **Learn:** the highest member is  
 totally forgotten what you said  
**Inst1:** the states assembly [AHI-3Jersey 52-54]

There are differences in the proportions of *completion* type CCs in the different corpora. For example, in the character-by-character corpus there are large differences in the

distributions by same- and cross-person CCs, as can be seen in table 4.8. In this corpus cross-person CCs are less likely to be completions than in the BNC (27/167, 16% vs. 93/329, 28%;  $\chi^2_{(1)} = 8.84, p = 0.003$ ), though they are far more likely to be completions than in the tangram corpus (27/167, 16% vs. 2/120, 2%;  $\chi^2_{(1)} = 16.16, p < 0.001$ ). Despite this, the overall proportion of cross-person completions by all contributions is the same as that in the BNC (27/3096, 1% vs. 93/11469, 1%;  $\chi^2_{(1)} = 0.11, p = 0.738$ ), meaning that the higher proportion of cross-person CCs overall in this corpus are of the expansion type. This could be because of the constraints of the task or the text environment (or a combination of the two). Newlands et al. (2003) note that different strategies are used in text chat, with typically shorter turns – one of the manifestations of this may be in the use of expansions which serve to tie the contribution to a prior one and thus maintain the topic under discussion, without having to formulate a complete syntactic sentence, as in e.g. (4.19) and (4.20).

(4.19) **X:** fine newspaper for kindling makes sense

**A:** with ax for cutting wood *[DiET CbyCcont24 156-157]*

(4.20) **P:** so what do u actualy have LOL

**D:** on ur list?! *[DiET CbyCcont17 117-118]*

Same-person cases are more likely to continue an incomplete antecedent than in the BNC (56/82, 68% vs. 213/726, 29%;  $\chi^2_{(1)} = 50.34, p < 0.001$ ). This may be due to cases which are not necessarily intended as CCs but where the speaker leaves the floor open for long enough that one of their interlocutors makes a contribution before they have completed their own, as in examples (4.21) and (4.22).

(4.21) **D:** but we need him as he

**S:** y

**D:** ibelieves he has cure for he most common types of cancer so he would save  
many lives in the futre *[DiET CbyCcont31 7-9]*

(4.22) **L:** a small

**Q:** huh#

**L:** a ball\* of stell wool *[DiET CbyCcont12 25-27]*



The majority of both same- and cross-person continuations in the tuition corpus (67% and 61% respectively) continue an already complete antecedent, with 33-39% therefore being *completions*. The proportion of completions for same-person cases is not different to the BNC (148/450, 33% vs. 213/726, 29%;  $\chi^2_{(1)} = 1.65, p = 0.20$ ), but the proportion of cross-person completions is significantly higher (270/692, 39% vs. 93/329, 28%;  $\chi^2_{(1)} = 11.25, p = 0.001$ ). The proportion of cross-person completions is also higher than same-person completions (270/692, 39% vs. 148/450, 33%;  $\chi^2_{(1)} = 4.41, p = 0.04$ ). That this is likely to be due to the nature of the task is evidenced by the fact that the numbers are in line with the proportion of completion CCs that Skuplik (1999) found in the Bielefeld Toy Plane Corpus (270/692, 39% vs. BTPC 54/126, 43%;  $\chi^2_{(1)} = 0.66, p = 0.42$ ).

### Continuation start-completeness

As can be seen in table 4.9, continuations were also less likely to be start-complete in the tangram corpus than in the BNC (22/516, 4% vs. 269/2231, 12%;  $\chi^2_{(1)} = 26.65, p < 0.001$ ) which may be related to the already discussed differences in proportions of expansions and completions. Note however, that in both text-based corpora, the proportion of same-person CCs is in line with the proportion of cross-person CCs (as was the case for the BNC and in line with hypothesis 6), though this is not the case for the tuition corpus.

	person:	Same-				Cross-	
		all	across-	(all)			
		N	%	N	%	N	%
Tangram	Y	19	5	6	4	3	3
	N	377	95	148	96	117	98
	(Total)	396		154		120	
Character-by-character	Y	46	15	19	23	30	18
	N	269	85	63	77	137	82
	(Total)	315		82		167	
Tuition	Y	126	8	23	5	71	10
	N	1141	92	427	95	621	90
	(Total)	1567		450		692	
BNC	Y	224	12	99	14	45	14
	N	1678	88	627	86	284	86
	(Total)	1902		726		329	

Table 4.9: Comparative continuation start-completeness

### CA categories

As with the BNC, to assess hypotheses 1 and 3 regarding syntactic constraints and predictability, the data from the character-by-character corpus and the tuition corpus

were analysed according to particular lexical items on either side of the split point, however due to differences in end-completeness (as discussed), the split point in the tangram corpus could not be analysed in these terms.

As can be seen from table 4.10 (where the || token again represents the split point), the type of task and/or interface does constrain which types of CCs are common.

As with the BNC, the most common of the CA categories are Lerner’s (1996) hesitation-related *opportunistic* cases, which make up 7% of same- and 8% of cross-person CCs in the tuition corpus, meaning that for this corpus, cross-person opportunistic cases are not more common than same-person ones. Contrarily, in the character-by-character corpus, these are 4% of same- and 29% of cross-person CCs.

person:		Same-				Cross-	
		all		across-		(all)	
		N	%	N	%	N	%
C-by-C	...    extending clause ...	90	29	14	17	95	57
	...er/...    ...	4	1	0	0	13	8
	... <pause>    ...	13	4	3	4	35	21
	...    which/who/etc ...	3	1	0	0	0	0
	... Compound TCUs ...	98	31	7	9	26	16
	(other)	126	40	62	76	47	28
Total		315		82		167	
Tuition	...    extending clause ...	486	31	152	34	212	31
	...er/erm    ...	84	5	11	2	24	3
	... <pause>    ...	90	6	19	6	33	8
	...    which/who/etc ...	87	6	35	8	25	4
	... Compound TCUs ...	13	1	2	0	2	0
	(other)	940	60	260	58	442	64
Total		1567		450		692	
BNC	...    extending clause ...	1082	57	395	54	163	50
	...er/erm    ...	35	2	21	3	13	4
	... <pause>    ...	19	1	15	2	19	6
	...    which/who/etc ...	26	1	11	2	5	2
	... Compound TCUs ...	37	2	19	3	4	1
	(other)	783	41	317	44	161	49
Total		1902		726		329	

Table 4.10: Comparative continuation categories

This could be because what was defined as a ‘<pause>’ for these corpora was defined in numerical terms, calculated based on the top 10% of positive onset delays for all contributions, whilst for the BNC it was a more subjective measure. However, it does suggest that (contrary to the results of de Ruiter et al., 2006), it may be harder for participants to anticipate potential TRPs, and where it is therefore appropriate for them

to take the floor with only lexicosyntactic cues.<sup>19</sup>

The observed difference in the likelihood of same- or cross-person CCs between filled and unfilled pauses seen in the BNC (that an other person is more likely to offer a continuation after an unfilled pause) does not hold for the tuition corpus, though again this could be due to the definitions used for a ‘<pause>’.

Note also that although the figures for split points following a filled pause in the tuition corpus were in line with the BNC, for the character-by-character corpus there was only one CC in which the antecedent finished in an explicit filled pause (4.23); the figures shown in the table include those where participants added a text specific filler (“...” (4.24), (4.25)), or a question mark after an incomplete antecedent (4.26), which serves the purpose of explicit turn handover, indicating that the contribution is inviting completion, which may be achieved by e.g. intonational means in face-to-face dialogue, and in contrast to filled pauses in spoken dialogue which may, as we’ve seen, be used to retain the floor.

- (4.23) **S:** s yes exactly..hmm  
**D:** so back to tom i thionk [DiET CbyCcont31 12-13]
- (4.24) **R:** maybe it was too hih in the air ...  
**O:** and what are they doing with sandbags [DiET CbyCcont18 65-66]
- (4.25) **M:** has anyone watched the program Lost? What would they do? There might  
be polar bears ...  
**C:** in canada?! crazy! [DiET CbyCcont22 138-139]
- (4.26) **Pi:** osooo other food, choco and?  
**Po:** and canned goods?? [DiET CbyCcont17 86-87]

It also appears that different types of task facilitate different types of CCs; in the tuition corpus the next most common of the CA categories is Rühlemann’s (2007) *sentence relative* cases with 4% of cross-person cases and 8% of same-person cases, whilst in the character-by-character corpus many of the CCs (cross- 16%, same- 9%) are Lerner’s Compound TCUs. In fact, for this corpus, the majority of these are listing environments

<sup>19</sup>Though of course this may be peculiar to the character-by-character text chat environment.

– 6 of the same person cases, and 25 of the cross-person cases – as the arctic survival task required participants to put objects in order. See (4.27) for an example where all four participants contribute to the list.

(4.27) **N:** 1.pistol.

2.food

**S:** 3. fishing rod

**A:** 44 small axe

**S:** 5. tent

**N:** 6 board skate wooden board

**A:** 7 ball of steen wstye stell wool

**M:** 8.compass.

9map

[DiET CbyCcont13 239-247]

As with the BNC, however, the most common groups are the extending clauses and the (other) category suggesting a generally available strategy to syntactically tie a contribution to a potentially incomplete prior contribution; whether one’s own or someone else’s.

### Split point

Due to the differences in end-completeness already discussed, the split point in the tangram corpus is not directly comparable to that in the BNC. However, to further test the finding that observed syntactic constraints from previous studies may actually be related to the nature of the task under discussion we can break down the cross-person data from all corpora by the first word of the continuation, as Skuplik (1999) did – see table 4.11.

corpus:	Tangram		C-by-C		Tuition		BNC		BTPC
	N	%	N	%	N	%	N	%	%
...    and/or/but/so ...	27	23	95	57	214	31	170	52	
...    noun/noun phrase ...	28	23	43	26	365	53	64	20	30
...    adjunct/preposition ...	51	43	16	10	31	4	22	7	44
...    sentential modifier ...	1	1	8	5	36	5	8	2	
...    other ...	13	11	5	3	46	7	64	19	
Total	120		167		692		329		(126)

Table 4.11: Comparative split point

The proportions of different types of split point in the tangram corpus are in line with those observed by Skuplik (1999) in the Bielefeld Toy Plane Corpus of German. In her data, 44% were prepositional or adverbial phrases, whilst 43% of cross-person CCs in the Tangram corpus were adjuncts or prepositions. Additionally, she found 30% were nouns or noun phrases; the equivalent for this data is 23%. Reasons for these similarities could be to do with the nature of the tasks; although the tasks are different in each case, they have very similar qualities; in both corpora, the key is in identifying the appropriate next piece (be it a tangram or construction element). In both cases, adding e.g. prepositional elements to an already complete antecedent may serve to narrow down the frame of reference and serve as a check that participants are talking about the same element. Given the large differences between the other corpora and these two (especially in terms of extending a contribution with an adjunct or prepositional phrase), this suggests that the *type* of task is influential in determining the type of CCs produced, and not the communication medium used.

### Constituency

To further investigate the apparent lack of any strict syntactic constraints (hypothesis 1) *completion* cases in the tuition and character-by-character corpus<sup>20</sup> were annotated for whether the split point occurred within a syntactic constituent, or between constituents. For both corpora, there was no difference in the proportion of same- and cross-person CCs that fell within or between a constituent (tuition 86/148, 58% vs. 152/270, 56%;  $\chi^2_{(1)} = 0.13, p = 0.72$ ; character-by-character 32/56, 57% vs. 12/27, 44%;  $\chi^2_{(1)} = 1.18, p = 0.28$ ).

For the cross-person CCs in the tuition corpus, there was more likely to be repair in the within constituent cases, in line with the trend in the BNC and the anecdotal observations of Szczepek (2000a) (within 27/118, 23% vs. between 20/152, 13%;  $\chi^2_{(1)} = 4.37, p = 0.037$ ).<sup>21</sup> This may be taken as evidence for the interactional relevance of syntactic constituents, though note that it is impossible to tell from this study whether participants are backing up to the start of the constituent or some other point, or whether this is dictated by their own processing needs or awareness of their interlocutor (audience design). It is also important to note that in the majority of cases, even where the split

<sup>20</sup>As discussed, there were too few completion CCs in the tangram corpus to analyse.

<sup>21</sup>In the character-by-character corpus there were too few cases to analyse.

point falls within a syntactic constituent, participants do not perform any kind of repair, so this is at best a weak effect.

### Repair

For same-person cases, there was more repair in all the three task-based corpora than in the BNC suggesting that both task and medium provide additional costs in integrating turns to prior material. (tangram 14/154, 9% vs. BNC 34/726, 5%;  $\chi^2_{(1)} = 4.79, p = 0.029$ ; tuition 44/450, 10% vs. BNC 5%;  $\chi^2_{(1)} = 11.64, p = 0.001$ ; character-by-character 19/82, 23% vs. BNC 5%;  $\chi^2_{(1)} = 41.09, p < 0.001$ ).

	person:	Same-				Cross-	
		all		across-		(all)	
		N	%	N	%	N	%
Tangram	Y	42	11	14	9	12	10
	N	354	89	140	91	108	90
(Total)		396		154		120	
Character-by-character	Y	46	15	19	23	6	4
	N	269	85	63	77	161	96
(Total)		315		82		167	
Tuition	Y	97	6	44	10	135	20
	N	1470	94	406	90	557	80
(Total)		1567		450		692	
BNC	Y	77	4	34	5	32	10
	N	1825	96	692	95	297	90
(Total)		1902		726		329	

Table 4.12: Comparative repair

This is perhaps surprising, as in text-based chat, there is an explicit record of what has been said, and in the line-by-line interface of the tangram corpus participants also have the opportunity to revise their turns prior to sharing them. In many cases, however, it appears that the repair is explicitly designed to indicate that the contribution functions as a continuation to some prior contribution (4.28), with different reasons for this in the two different text corpora. In the line-by-line interface of the tangram corpus this needs to be made clearer because of the non-linearity issues and the possibility of interleaved turns which might otherwise break up the coherence of interlinked contributions, as in (4.29). In the character-by-character corpus participants often use new contributions as a way to correct their own typos (4.30), (4.31), and is therefore an explicit form of repair that participants have developed specifically due to the constraints of this text environment (where deletes are unavailable).

- (4.28) **R:** its a boat like shape with chimney  
**R:** chimney arrow shape *[DiET Tang16 219-220]*
- (4.29) **J:** WELL THERE IS A SQUARE  
**R:** not at all  
**J:** AT THE LEFT  
**J:** SQUARE AT THE LEFT TOP *[DiET Tang16 188-191]*
- (4.30) **F:** i tink it would be easuer to make less typos if i typed shower  
**F:** slower\* *[DiET CbyCcont23 100-101]*
- (4.31) **Q:** we're msii  
**Q:** no  
**Q:** missing our lectures *[DiET CbyCcont12 11-13]*

For cross-person continuations, the patterns of repair are different in the different corpora. The character-by-character corpus has significantly fewer cross-person repairs than the BNC (6/167, 4% vs. 32/329, 10%;  $\chi^2_{(1)} = 5.89, p = 0.015$ ), the tangram corpus has the same proportion (12/120, 10% vs. 32/329, 10%;  $\chi^2_{(1)} = 0.007, p = 0.931$ ), and the tuition task has significantly more (135/692, 20% vs. 32/329, 10%;  $\chi^2_{(1)} = 15.60, p < 0.001$ ). This appears to be a consequence of making sure the imparted information is correct, and related to the nature of the task as a collective memory exercise with unfamiliar terms, but once again there may be a combinatorial effect of task and medium.

### Next speaker

As with the BNC, questions of who is expected or entitled to speak next (hypothesis 9) are assessed by examining who actually provides the turn following the CC. In the character-by-character corpus, the proportion of cases in which the participant who supplied the continuation retains the floor is not different between multiparty and dyadic conditions (12/118, 10% vs. 8/49, 16%;  $\chi^2_{(1)} = 1.25, p = 0.264$ ). In both conditions this likelihood is lower than the baseline proportion of occasions in which a participant generally keeps the floor in this interface (dyadic 8/49, 16% vs. 312/675, 46%;  $\chi^2_{(1)} = 16.555, p < 0.001$ ; multiparty 12/118, 10% vs. 524/2367, 22%;  $\chi^2_{(1)} = 9.517, p = 0.002$ ). This is also true for the tuition corpus, in which the supplier of the continuation retained the floor in just

17% of cases compared to the same person continuing in 36% of cases in the baseline case (116/692, 17% vs. 2396/6567, 36%;  $\chi^2_{(1)} = 107.605, p < 0.001$ ). This means that participants are sensitive to turn-taking considerations, and do not assume that providing a continuation to another's contribution entitles them to retain the floor in turns following cross-person CCs.

Contrary to predictions made by hypothesis 9, in the multiparty condition of the character-by-character corpus, the floor is taken by a third person (not involved in supplying either the antecedent or the continuation of the CC) in the same proportion of cases as the baseline (66/118, 56% vs. 1215/2367, 51%;  $\chi^2_{(1)} = 0.953, p = 0.329$ ). This means that the antecedent owner is providing the contribution following a cross-person CC in a disproportionately high number of cases, which is also true of the tangram corpus, issues of linearity notwithstanding (71/120, 59% vs. 607/2373, 26%;  $\chi^2_{(1)} = 65.079, p < 0.001$ ) and the tuition corpus (408/692, 59% vs. 1958/6567, 30%;  $\chi^2_{(1)} = 242.020, p < 0.001$ ), in which the proportions are not different to the multiparty cases from the BNC.

### Parties

As the tuition task was a task that imposed party membership from the outset, and such a setup did lead to more cross-person than same-person CCs (692/6600, 10% vs. 450/6600, 7%;  $\chi^2_{(1)} = 56.14, p < 0.001$ ), in line with hypothesis 4, I analysed the cross-person CCs according to whether they were within-party (the antecedent came from one of the instructors and the continuation came from the other instructor) or across-party (the antecedent or the continuation came from the learner and the other part came from an instructor), which could be learner led, or instructor led. All other things being equal, we would expect two thirds of cross-person CCs to be between an instructor and a learner because there are twice as many possible combinations of antecedent and continuation contributor in these cases (especially given that of the 6600 contributions, almost exactly one third (2310; 35%) were from the learner).

However, as can be seen from table 4.13, 312 of the 692 cross-person CCs were between the two instructors, which at 45% is statistically significantly higher than the expected 33% ( $Z = 6.55, p < 0.001$ ). Similarly, in the tangram corpus, there were differential matching conditions which changed as the dialogues unfolded. In the *dyadic* condition, only one director had access to the tangram to be described to the matcher, and in the *col-*



*lective* condition, both directors had access to the tangram to be described (this case is therefore analogous to the task in tuition corpus, as there were then two directors who had equal access to the information to be imparted to the matcher). As can be seen from table 4.13, in the dyadic cases there were only 3 cross-person CCs started by one director and continued by the other. In the collective case, however, 49% of cross-person CCs were of the within party type. This proportion is not different to that of within-party CCs in the tuition corpus (26/53, 49% vs. 312/692, 45%;  $\chi^2_{(1)} = 0.31, p = 0.58$ ).

Within or across party:	Across-				Within-	
	learn-inst		inst-learn		inst-inst	
Corpus:	N	%	N	%	N	%
Tuition	267	39	113	16	312	45
Tangram Dyadic	18	58	46	77	3	10
Collective	13	42	14	23	26	90
Total	31		60		29	

Table 4.13: Tuition and tangram corpus within-party CCs

There are no differences in the distributions of within- and across-party CCs in terms of CA categories or repair, however, across-party CCs are more likely to continue an antecedent that does not end in a complete way than within-party (163/380, 43% vs. 107/312, 34%;  $\chi^2_{(1)} = 5.33, p = 0.021$ ), which could be a consequence of the high proportion of invited completions. This means that the within-party proportion of CCs with an incomplete antecedent is in line with the same-person proportion, and the earlier observed difference between same-person and cross-person CCs is therefore a difference between within-party (same-person CCs are necessarily within-party) and across-party CCs.

### Within-party overlap

If we look at all turns in the tuition corpus, we can see that, as predicted, the number of turns which overlap the previous turn is higher when the incoming speaker is in a party with the current speaker (479/1123, 43% vs. 1015/3048, 33%;  $\chi^2_{(1)} = 51.26, p < 0.001$ ). This suggests that the turn space really is taken to belong, not to the individual, but to the party, especially given the lack of explicit turns discussing interruptions. Note that the seemingly high measures of overlap are in line with those reported in Shriberg et al. (2001), who found between 31% and 54% of *spurts* of dialogue<sup>22</sup> exhibited some overlap, with differences accounted for by types of dialogue.

<sup>22</sup>Defined as “speech regions uninterrupted by pauses longer than 500 ms.”

### Ratification

In the character-by-character corpus dyadic cases were more likely to be ratified or rejected than multiparty cases (39/49, 80% vs. 70/118, 59%;  $\chi^2_{(1)} = 6.27, p = 0.012$ ), as expected if all cross-person CCs in dyadic dialogues are taken to be directed at the speaker of the antecedent, with some in the multiparty dialogues genuinely taken to project parties and be addressed to the other interlocutors.

In the tuition corpus, cross-person CCs are ratified or rejected (possibly by overlapping material, as seen in (4.32), below) in 467 of the 692 cases, which is higher than in the BNC (467/692, 67% vs. 138/329, 42%;  $\chi^2_{(1)} = 60.25, p < 0.001$ ). Interestingly, CCs after which the third participant takes the floor are as likely to be ratified or rejected as those continued by the antecedent owner (130/168, 77% vs. 337/408, 83%;  $\chi^2_{(1)} = 2.111, p = 0.146$ ). However, confirming hypothesis 10, which predicted that any member of a party ought to be able to ratify a CC which another member of their party supplied the antecedent for, there is a difference in third party ratifications when we take role into account. Of the 113 CCs in which an instructor supplied the antecedent and the learner continued, 29 were ratified by the other instructor, either by adding a ‘yes’ response (4.33), or repeating (4.32) or paraphrasing elements (4.34). These could be considered to be within-party ratifications, and they are more frequent than ratifications from the third participant when they are not in a party with the antecedent owner (29/113, 26% vs. 101/579, 17%;  $\chi^2_{(1)} = 4.19, p = 0.041$ ).

- (4.32) **Inst1:** chief decision making is the assembly  
**Inst2:** which has the [[executive]]  
**Inst1:** [[executive]] and [[the scrutiny]]  
**Learn:** [[and the scrutiny]]  
**Inst2:** and the scrutiny [AHI-1Jersey 129-133]
- (4.33) **Inst1:** the umm chief min um minister  
**Inst1:** is in charge of  
**Learn:** everything  
**Inst2:** yeah [AHI-2Jersey 51-54]

(4.34) **Inst1:** so the king yeah

**Learn:** is god

**Inst2:** is the god

**Inst1:** yeah

[AHI-2Saudi 7-10]

### Mutual knowledge

As discussed above, there were more within-party than across-party CCs in the tuition corpus. But the picture is not quite as clear as this, as we can see from table 4.14 below.

Firstly, note that there is an imbalance in the across-party cases. CCs in which an instructor continues an antecedent offered by the learner are twice as common as those in which the learner continues an instructor's antecedent.

Within or Across party:	Across-				Within-	
	learn-inst		inst-learn		inst-inst	
	N	%	N	%	N	%
Ratified by: Antecedent owner	151	57	42	37	144	46
Third person	47	18	29	26	54	17
No-one	69	26	42	37	114	37
Total	267		113		312	

Table 4.14: Tuition corpus ratification of cross-person CCs

There is also a difference in the patterns of ratifications depending on who supplied the antecedent and who supplied the continuation. Overall, there is more likely to be a ratification if the learner supplied the antecedent and an instructor contributed the continuation (198/267, 74% vs. 269/425, 63%;  $\chi^2_{(1)} = 8.819, p = 0.003$ ), so if they are learner led then they are more likely to be treated as belonging to the learner.

Within or Across party:	Across-				Within-	
	learn-inst		inst-learn		inst-inst	
	N	%	N	%	N	%
Phase of dialogue: teaching	99	37	42	37	196	63
recall	168	63	71	63	116	37
Total	267		113		312	

Table 4.15: Tuition corpus cross-person CCs

Additionally, as per table 4.15, in the first half of the dialogues (*teaching phase*; very crudely determined by the number of contributions in each dialogue divided by two to prevent possible interpretational effects), there are more within-party (cross-person) CCs

than across-party, whilst in the second half (the *recall phase*), the reverse is true (196/312, 63% vs. 141/380, 37%;  $\chi^2_{(1)} = 45.35, p < 0.001$ ).

These observations could explain the differences in the pattern of results to those found in the Bielefeld Toy Plane corpus (Skuplik, 1999), in which 79% of cross-person CCs had the constructor (analogous to our learner, here) continuing or completing a prior contribution by the instructor, whilst in this corpus, even only considering the across-party CCs (i.e. the ones where the learner and an instructor co-construct the utterance), only 29% had an antecedent by the instructor and continuation by the learner. This is because of the anticipatory nature of her task – the participants had the same collection of components and could therefore predict what sort of action is likely to come next. Contrarily, in the tuition corpus, the instructors have all the information and the learner does not until towards the end of the dialogues – in other words, the distribution of shared knowledge (or context) is different in the two tasks.

In the case of the tangram (line-by-line) corpus, as discussed above, in cases where the distribution of knowledge is more akin to that in the tuition corpus (in the collective contexts where both directors have access to the tangram to convey to the matcher) then so is the distribution of cross-person CCs. Conversely, where the distribution of knowledge is equivalent to Skuplik’s dyadic task (in the dyadic contexts where only one of the directors has access to the tangram to be described) then the distribution of cross-person CCs is in line with her figures (with the matcher continuing an offering by the director in 46/67 cases or 69% vs. 99/126, 79%;  $\chi^2_{(1)} = 2.30, p = 0.13$ ).

These results suggest that cross-person CCs are more common where interlocutors share context, and it might therefore be that the use of CCs demonstrates shared knowledge or common ground and not party membership per se. However, one of the key reasons for believing yourself to be in a party with someone else and behaving accordingly by e.g. continuing their prior contributions, may just be the common ground or shared knowledge.<sup>23</sup>

---

<sup>23</sup>Note that this interpretation matches the folk psychological notions of cross-person CCs being more likely between people who are particularly close (because they should have a lot of accumulated shared knowledge).

## 4.5 Summary

Although there are differences between the corpora in terms of task and communicational medium, it is clear from these studies is that CCs do occur in text based chat, contrary to hypothesis 5, and there are more cross-person CCs than in the BNC in the task based corpora studied here, regardless of medium, as hypothesis 4 predicted.

Given the wide range of contexts in which CCs occur, their production can be seen as a generally available dialogue strategy, however, the different corpora also provide evidence for different patterns of effects which are relevant to the production of CCs. Though there are similarities between the corpora, there are also differences in the distributions of CCs according to a number of interacting factors (which varied for practical rather than methodological reasons), most notably task-type and communicational interface.

Firstly, in the two text-based corpora, non-verbal feedback is not available, and this places constraints on strategies for managing turn handover, which are relevant for the production of CCs. In the line-by-line interface, as turns are constructed in private and may be constructed simultaneously, it is very rare for participants to transmit something that is noticeably not end-complete, meaning that although there are more CCs than in the face-to-face dialogues, these are almost exclusively expansions not completions. In the character-by-character interface, participants cannot type simultaneously so that backchannels, for example, are rarer, which has the knock-on effect of making same-person CCs rarer, as many in the face-to-face dialogues are presented in response to feedback that is unavailable in this interface. Interestingly, participants develop text specific strategies to deal with such issues, for example, in the use of ellipsis in the character-by-character interface to explicitly indicate ceding the floor. In the text-based dialogues repair of the antecedent at the start of the continuation is also used to explicitly indicate that a turn is tied to a prior one, which might be achieved by other means (e.g. intonational, non-verbal) in the face-to-face dialogues.

Although the communicational interface does have a bearing on CCs, as discussed, the constraints of the task are more influential. All the task-based corpora required participants to cooperate and all of them resulted in a greater proportion of cross-person CCs, regardless of the medium. However, the different tasks had different distributions of prior knowledge and the different goals in each case did lead to different patterns of

CCs. Where the task required some participants imparting knowledge to another, as in the tuition task and the tangram task, there were a greater number of competing continuations, where more than one continuation was offered to the same antecedent. Where the distribution of shared knowledge was skewed between different participants, cross-person CCs were more likely between those who shared knowledge. This is especially clear in the tangram corpus, where different matching conditions mean that whether the knowledge was shared or not differs systematically, and the distribution of cross-person CCs changes accordingly.

Although there were differences in the types of CCs offered, depending on the communicational medium, such as the increased proportion of expansions in the line-by-line interface, there were also effects of task on the types of continuation offered, with, for example, Lerner's listing environment providing a specific locus for CCs in the arctic survival task. Similarly, the types of CCs seen in the tangram task were equivalent to those from the Bielefeld toy plane corpus, with expansion type CCs, especially e.g. preposition phrases being used to narrow down the reference for the required component, regardless of the different communicational medium. Likewise, although repair is influenced by the medium, there are also effects of the task, with the increased number of repairs in the tuition task attributable to the need to ensure that the imparted information is correct and the use of unfamiliar terms, which was not the case in the other corpora. Supporting this interpretation, ratifications following cross-person CCs are also more common in the task-oriented than generic face-to-face dialogues.

Despite these differences between the communicational medium and type of task in these corpora and their effects on the distributions of CCs, there are also similarities between them. For example, regardless of the interface or task (or interaction between them), the next speaker is disproportionately likely to be the supplier of the antecedent, showing that participants are sensitive to turn-taking considerations and do not assume that supplying a continuation entitles them to retain the floor. Likewise, ratifications following a cross-person CC are more likely in dyadic dialogues in both text-based (character-by-character) and spoken (BNC) corpora.

Generally, the corpus studies presented in this chapter strengthen the claim of hypothesis 1, that there are no specifically syntactic constraints on where a split point may

occur. Despite the different tasks and interfaces, at least a quarter of all CCs could not be easily categorised. However, there is some weak evidence for syntactic constituency, in the higher proportion of repair in the within constituent cross-person completions in the tuition corpus.

Over 60% of cross-person CCs in all corpora also continued a complete antecedent, consistent with hypothesis 2, showing that participants are sensitive to TRPs, and tend to wait for a transition relevance place even when constructing their contribution as a continuation of another's. This was even more marked in the line-by-line text interface of the tangram corpus, where design factors mean that participants cannot project another's TRP, as turns are constructed in private. This means that completions, such as terminal item completions, are not used. This reduced flexibility in turn-taking leads to different patterns of CCs, but these are caused not by the fact that the communication is text-based per se, but the particular interactional constraints of the interface, as can be seen by the differences in distributions between CCs in the line-by-line text dialogues and the character-by-character ones.

Aspects of predictability do seem to encourage the use of cross-person CCs, as per hypothesis 3. For example, patterns of CCs are influenced by constraints of the task, including making specific types of CC more likely (e.g. lists in the arctic survival task) and making CCs more likely between participants who share knowledge (as seen in the different patterns of CCs in the learning and recall phase of the tuition task). These findings also appear to be independent of medium, though note that strictly syntactic predictability is harder to assess, and will be returned to in chapter 7.

These task-based corpora also strongly confirm hypothesis 4. All three have more cross-person CCs than the BNC, and the tuition task, which explicitly imposes parties at the outset, has twice as many cross-person CCs as the others. The evidence suggests that hypothesis 9 is incorrect, as cross-person CCs do not seem to be treated as single turns with a joint owner in terms of who is entitled to speak next and patterns of ratifications. However, there is other evidence that suggests that despite this, parties are relevant entities in the dialogues oriented to by interlocutors in the course of ongoing dialogue. In both text and spoken dialogues, CCs are one way in which such groupings can be made manifest, as seen by the increased likelihood of within-party CCs rather than across-party

CCs. In addition, in the spoken tuition corpus, there was more within-party overlap, and within-party ratifications were more common than across-party ratifications (confirming hypothesis 10). It remains an open question, however, as to whether cross-person CCs demonstrate party membership or are just more likely if there is a party. In other words whether I am more likely to produce a CC if I believe I am in a party with you or if producing a CC makes it useful to treat it as if we are a party is not answered by these frequency counts, and will be explored in the next two chapters.

The apparently contradictory evidence for both parties and continuations being treated as independent contributions (in terms of turn-taking and ratifications) could be explained by there being a disconnect between who is taken to be responsible for the *grounding act* (the continuation) which may require acknowledgement, and who holds authority over the *speech act* that is co-constructed (the compound contribution) which could be taken to be jointly owned. However, this suggests that these are not parties in the sense of Schegloff (1995), as the turn-taking behaviour is not taken to be jointly applicable, but the content is (suggesting that the *collective author* of Díaz et al. (1996) may apply here, but does not extend to the level of turn-taking). The evidence of differential patterns of CCs at different stages in the tuition corpus backs this up, suggesting that it may be presumed common ground which drives the production of a continuation and not presumed party-membership.

It is possible that what determines who is held to be responsible for the speech act is actually dependent on the response (or lack of response) it receives, such that an acknowledgement serves precisely to assert ownership of the CC which would otherwise be treated as jointly owned. Contrarily, ratifying a cross-person CC may act as an acceptance of the jointly constructed material, such that it validates the ‘party’ and is a way of indicating that one has accepted another into one’s turn space.

In any case, that CCs do occur in text chat means that the chat tool methodology is validated as a means to study compound contributions, and this will be pursued in subsequent chapters. Chapters 5 and 6 will address how cross-person CCs are interpreted by other interlocutors (hypothesis 7 predicts that there should be no additional difficulty in interpreting a cross-person CC as opposed to a same person one), and whether they have the effect of making it appear that participants have formed parties (hypothesis 11,



which predicts that people will respond as if there are parties formed of others who appear to have co-constructed an utterance), and chapter 7 will look at the notion of predictability more closely.

## Chapter 5

# DiET chat tool experiment 1: Introducing fake CCs

---

This chapter reports an experiment which is the first controlled manipulation of compound contributions during an unfolding interaction. It uses the DiET chat tool (section 5.1) to directly compare the effects of same-person and cross-person CCs on participants in a text-based dialogue. In this experiment, single contributions from one of the dialogue participants are altered to appear as either a same- or cross-person CC.

While the corpus studies of sections 4.2 and 4.4 provide us with useful information concerning the nature and frequency of CCs and their various sub-categories, they cannot tell us about the causal effects of CCs on the dynamics of a conversation.

Despite their different distributional properties, we expect cross-person CCs to be at least as easy to interpret as same-person CCs, as per hypothesis 7. However, the evidence from the corpus studies suggests that, although CCs can apparently occur at any point (between words) in a string, there may be preferential points at which they are more likely to, suggesting that these cases – e.g. at the end of a potentially complete TCU – should be less problematic to interpret.

The evidence regarding parties additionally suggests that using a cooperative task means that cross-person CCs are not unusual, and we have hypothesised that their presence will lead interlocutors to act as if parties have been formed (hypothesis 11).

## 5.1 The DiET chat tool

The Dialogue Experimental Toolkit (DiET) chat tool is a text-based chat interface into which interventions can be introduced into a dialogue in real time. These interventions can take a number of forms; turns which have been entered by a participant may not be relayed to their conversational partners, additional turns may be added (which can appear to come from active participants, or an outside source) and transmitted to one or more participants (as in Healey et al. (2003), in which spoof clarification requests are added to the dialogue), or turns may be altered prior to transmission (by altering the apparent sender, or by changing, adding or omitting some of the words in the turn). Any combination of these manipulations is also possible, however as they all occur as the dialogue progresses, they cause a minimum of disruption to the ‘flow’ of the conversation.

There are two different interfaces available in the DiET chat tool environment, the line-by-line interface, and the character-by-character interface. Both are based on the DiET custom built Java application, and consist of two main components: the server console and the user interface. The server console is the same for both interfaces. All key presses are time-stamped and stored by the server, which acts as an intermediary between what participants type and what they see. 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, real turns can therefore be altered by the server or not relayed, or fake turns can be introduced.

The line-by-line user interface will be briefly outlined below, with the character-by-character user interface described in section 6.1.

### 5.1.1 Line-by-line user interface

In this version of the DiET chat tool, the user interface is designed to look and feel like common instant messaging applications such as MSN Messenger. The display is split into two windows, separated by a status bar, which indicates whether any participants are actively typing (see figure 5.1). 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 by either clicking on the send button or pressing the return key.

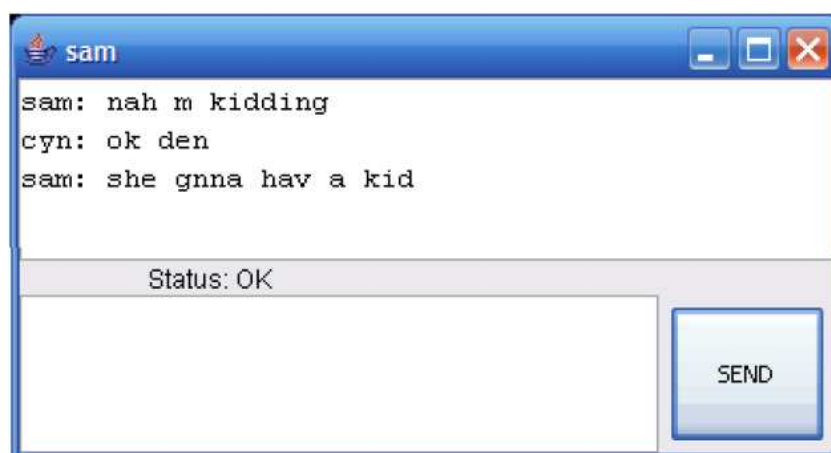


Figure 5.1: The user interface chat window (as viewed by participant ‘sam’)

## 5.2 Method

In this experiment, a number of genuine single contributions in text-based triadic conversations 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 a cross-person CC, or as a same-person CC generated by the “wrong” person. We can then compare the effects of seeing a cross-person CC or a same-person CC independently of the turn-taking expectations arising from who apparently made the last contribution (see below for full details). As the BNC corpus study did not yield any obvious syntactic constraints on the position of the split point, in this experiment the position of the split point was arbitrary, but each CC was categorised according to where the split point occurred for the analysis.

### 5.2.1 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 (see appendix A). This task was chosen on the basis that it is known to stimulate discussion, leading to dialogues of a sufficient length to enable an adequate number of interventions.

### The Intervention

In this experiment, some turns are automatically altered by the server to create fake CCs. 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 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  $2 \times 2$  factorial design which separates effects of ‘floor change’ i.e. whether the original speaker finishes the CC or another participant appears to (which, as we have seen in the corpus studies of chapter 4 has an effect on who is normally entitled or expected to speak next) from effects of ‘same/other’ i.e. whether a the two halves of the CC appear to be produced by the same speaker or by two different speakers. This contrast is shown in table 5.1.

<i>A types:</i>		
Should we start now		
<i>B sees (AA intervention):</i>		
A: Should we		
A: start now		
<i>C sees (one of):</i>		
<i>AB intervention:</i>	<i>BA intervention:</i>	<i>BB intervention:</i>
A: Should we	B: Should we	B: Should we
B: start now	A: start now	B: start now

Table 5.1: Comparison of experimental manipulations

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 generated randomly.

### 5.2.2 Subjects

60 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 MSN Messenger and each was paid £7.00 for their participation.

### 5.2.3 Procedure

Each subject 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 using a text interface as opposed to face-to-face.<sup>1</sup> Additionally, subjects were informed that the experiment would last approximately 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 with them until at least 25 minutes of dialogue had been carried out.

### 5.2.4 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

---

<sup>1</sup>Ethics approval was obtained from the Queen Mary Research Ethics Committee, reference QMREC2008/26, and subjects were fully debriefed as to the actual nature of the intervention following their participation.

to another.<sup>2</sup> We therefore measured 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 the participant who saw an AA-split, and one for the participant who saw an AB-, BA- or BB-split).

### Design

The data were analysed according to four factors in a  $2 \times 2 \times 2 \times 2$  factorial design with participant as a random factor. The two major factors were *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 appeared to come from the genuine source or the other participant ([AA and BA] vs [AB and BB]). If the misattribution of (parts of) utterances has an effect per se, this will show up as an interaction between floor change and same/other (as misattributions occur in AB, BA and BB interventions). Following the corpus studies in which antecedent and continuation completeness are important factors in determining when people do produce continuations, the two additional factors were standalone completeness of the antecedent and continuation<sup>3</sup> (see table 5.2 for examples).

Antecedent	Text of CC		Complete		
		Continuation	Ant	Cont	N
what the hell		is that	Y	N	68
the woman is pregnant		she should stay	Y	Y	27
these people said		you did something	N	Y	23
I think this is also		the wish of the doctor	N	N	135

Table 5.2: Examples of standalone completeness judgements

<sup>2</sup>As discussed in section 4.3, in online chat participants can compose contributions simultaneously, and contributions under construction when another is received can be revised prior to transmission. Genuine responses to a compound contribution might have a negative start time. However, the inclusion of cases where the whole contribution was constructed after receiving the CC (an arbitrary cut-off point, which would catch some contributions that were responses to earlier contributions, and miss some which were begun before the intervention was received) should impose the same level of noise in all cases.

<sup>3</sup>These judgements are 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 studies, such that an antecedent judged to be able to stand alone is end-complete and a continuation judged to be able to stand alone is start-complete. The difference in tagging conventions is due to the differences in completeness found in the tangram corpus (discussed in section 4.4.2) and the fact that in this 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.

It is interesting to note at this point that because the position of the split point was arbitrary we can compare how likely expansions and completions are when there are no interactional constraints. As can be seen from table 5.3, in all four corpora the split point was much more likely to occur after a complete antecedent than in the two experiments where the split point was arbitrary. This shows that in genuine dialogues, people clearly wait for TRPs a lot of the time even when constructing a turn as a continuation, rather than butting in at random.

	Antecedent completeness				Total
	Y	%	N	%	
Experiment 1	94	37	159	63	253
Experiment 2 (chapter 6)	75	45	90	55	165
Tangram	484	94	32	6	516
Character-by-character	370	77	112	23	482
Tuition	1338	59	921	41	2259
BNC	1603	72	628	28	2231

Table 5.3: Arbitrary split point comparison

### Dependent variables

The DiET chat tool environment allows us to analyse variables which measure different things. Contribution production is measured in both length and time – *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*. The amount of revision participants make to a contribution prior to transmitting it (which could be seen as indexing how carefully participants construct their turns) is measured by *deletes per character* (the number of keyed deletes divided by the total number of characters).

As this is the first experimental manipulation of compound contributions, what follows is necessarily exploratory, however, the potential interpretations of these dependent variables – and consequent pattern of effects we would expect given the hypotheses will be briefly explored here.

In general, we would expect that if participants are having trouble integrating the content of what is said in a dialogue, they need to do more work by producing more words in order to clarify or repair their understanding. Taking a longer turn can mean that you are typing more (indexed by the number of characters measure) or that you



are taking longer producing a turn (regardless of its length). All things being equal, we would expect these to be highly correlated, but differences may indicate a) how rushed participants feel, and independently b) how much information participants are adding.

Using deletes in a text based dialogue is taken to be analogous to repair, where repair is seen as a way to locally manage mutual misunderstandings that arise at different levels in the course of a dialogue. Because turns in the line-by-line chat tool are constructed privately, it can be seen as an index for how much participants revise what they communicate to their interlocutors before they do so, such that they attempt to ensure that their contribution is understandable and relevant to the ongoing dialogue. If there are few apparent misunderstandings, then there should not be the necessity to revise one's contribution before transmitting it, meaning that this can be seen as a proxy measure for how coherent the participant feels the dialogue is up to this point. Although there is the added complication that participants use deletes for correcting typos, there should be the same proportion of typos in turns of the same length regardless of the preceding context (all else being equal).

If speakers are interchangeable from a parsing perspective, because one interprets utterances from an egocentric point of view, then cross-person CCs should be as easy to interpret as same-person ones, as per hypothesis 7. We would not expect, therefore, to see any differences (on any of the measures) between turns following cross-person or same-person CCs. If, however, the presence of cross-person CCs leads interlocutors to behave as if parties have been formed (as per hypothesis 11) then we should expect participants to type shorter turns and use fewer deletes than otherwise, because they feel that more of the preceding context is shared, and there ought to be less scope for misunderstandings.

Additionally, the distributional evidence from the corpus studies suggests that CCs which have an antecedent that ends in an end-complete way should be more natural, so that lack of end-complete antecedents should result in more disruption to participants, which could manifest itself in longer turns or greater proportions of deletes.

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 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  $W = 0.995, p = 0.100$ ), meaning that an ANOVA analysis is appropriate for these (transformed) measures. For the proportional measure of deletes per character, which violates 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 deletes per character therefore allows for both the non-normality of the data, and within-subject correlations.

### 5.3 Results

A post-experimental questionnaire<sup>4</sup> 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 explicitly aware of any interventions (see appendix E).

Condition	Typing time (ms)		Num chars		N
	Mean	(s.d.)	Mean	(s.d.)	
AA	11122.27	(14413.5)	26.41	(20.4)	246
AB	12500.98	(10944.6)	32.12	(23.9)	89
BA	9800.77	(8810.3)	28.27	(18.4)	92
BB	11561.67	(10138.4)	25.78	(13.6)	63

Table 5.4: Typing time of turn and number of characters by type of intervention

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 5.4 shows the actual n values in each case.

<sup>4</sup>See appendix C.

### 5.3.1 Typing time

The results of a  $2 \times 2 \times 2 \times 2$  ANOVAs (with participant as a random effect)<sup>5</sup> can be seen in table 5.5.

IV	F	p
Same/Other (SO)	0.342	0.559
Floor Change (FC)	10.117	0.002**
Antecedent Completeness (Ant)	1.940	0.164
Continuation Completeness (Cont)	0.031	0.859
Participant	4.797	<0.001**
SO $\times$ FC	0.911	0.341
FC $\times$ Ant	0.199	0.656
FC $\times$ Cont	8.108	0.005**
SO $\times$ Ant	0.181	0.671
SO $\times$ Cont	2.639	0.105
Ant $\times$ Cont	1.344	0.247
SO $\times$ FC $\times$ Ant	0.546	0.460
SO $\times$ FC $\times$ Cont	0.010	0.921
FC $\times$ Ant $\times$ Cont	0.292	0.589
SO $\times$ Ant $\times$ Cont	0.399	0.528
SO $\times$ FC $\times$ Ant $\times$ Cont	0.141	0.707

\* indicates significance at the  $p < 0.05$  level

\*\* indicates significance at the  $p < 0.01$  level

Table 5.5: ANOVA on log transformed typing time of turn

There was a main effect of participant ( $F_{(59,415)} = 4.797$ ,  $p < 0.001$ ) showing that there was high individual variation for this measure. There was also an interaction effect of floor change  $\times$  continuation completeness ( $F_{(1,415)} = 8.108$ ,  $p = 0.005$ ). Post hoc pairwise comparisons showed that where the continuation was complete on its own, respondents typed more and for longer in the AB and BB cases (floor change effects where continuation completeness = y;  $F_{(1,415)} = 11.126$ ,  $p = 0.001$ , floor change effects where continuation completeness = n;  $F_{(1,415)} = 0.126$ ,  $p = 0.722$ ). In other words, recipients only type more when there is a change of floor and the continuation looks like an independent new turn.

This interaction is shown in figure 5.2.

<sup>5</sup>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  $2 \times 2 \times 2 \times 2 \times 60$  model). There are 490 datapoints between 60 subjects. As we carried out a full factorial model, but not all participants saw all four types of manipulation the numerator (error) degrees of freedom that resulted from this model was 415.

### 5.3.2 Number of characters

As can be seen from table 5.6, there is a similar pattern of results on the number of characters. There was a main effect of participant, and an interaction effect of floor change  $\times$  continuation completeness.

IV	F	p
Same/Other (SO)	0.106	0.745
Floor Change (FC)	5.510	0.019*
Antecedent Completeness (Ant)	2.694	0.102
Continuation Completeness (Cont)	0.010	0.922
Participant	4.358	<0.001**
SO $\times$ FC	0.080	0.777
FC $\times$ Ant	0.046	0.831
FC $\times$ Cont	7.446	0.007**
SO $\times$ Ant	0.002	0.967
SO $\times$ Cont	5.537	0.019*
Ant $\times$ Cont	0.314	0.575
SO $\times$ FC $\times$ Ant	0.467	0.495
SO $\times$ FC $\times$ Cont	0.055	0.814
FC $\times$ Ant $\times$ Cont	0.528	0.468
SO $\times$ Ant $\times$ Cont	0.793	0.374
SO $\times$ FC $\times$ Ant $\times$ Cont	0.246	0.620

Table 5.6: ANOVA on log transformed number of characters

There was also an interaction effect of same/other  $\times$  continuation completeness ( $F_{(1,415)} = 5.537$ ,  $p = 0.019$ ). Pairwise comparisons show that if the continuation did not appear complete then participants typed more characters in subsequent contributions if they had seen a cross-person CC (cont complete = n;  $F_{(1,415)} = 5.353$ ,  $p = 0.021$ ).

These are cases which must be interpreted as cross-person CCs, as the continuation cannot be treated as an independent turn, and is shown by the dashed line in figure 5.3.

### 5.3.3 Deletes per character

For deletes per character, a GEE model with participant as a subject effect (using a Tweedie distribution ( $p=0.0001$ ) with log link and independent correlation matrix, QIC = 107.582)<sup>6</sup> showed a significant main effect of same/other (model effect; Wald- $\chi^2 = 4.067$ ,  $p = 0.044$ ) with subjects seeing a cross-person CC (AB or BA) using *fewer* deletes per character than those seeing a same-person CC (see table 5.7). There were no other main effects or interaction effects.

<sup>6</sup>The model distributions and correlation matrices were chosen on the basis of being the best fit to the data, as indicated by the lowest quasi log-likelihood (QIC) score.

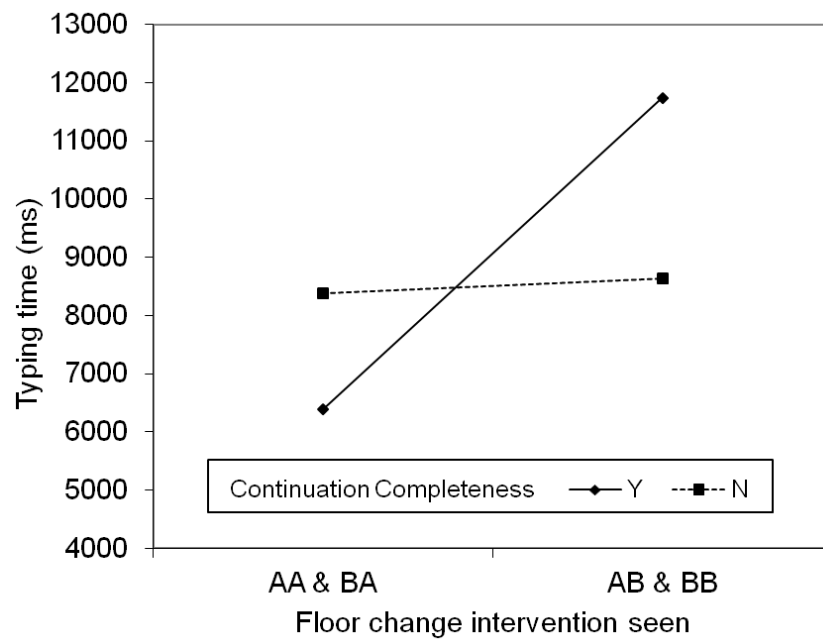


Figure 5.2: Marginal means of typing time of turn by floor change  $\times$  continuation completeness

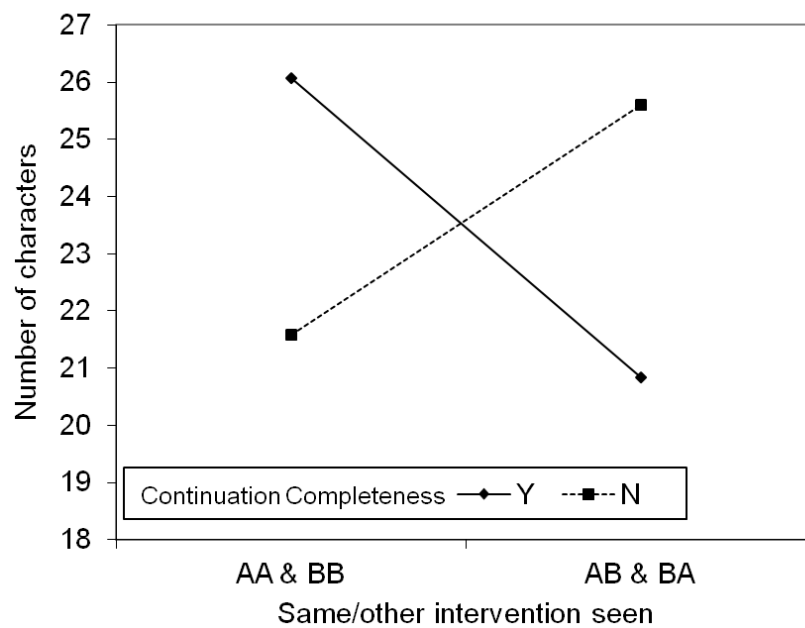


Figure 5.3: Marginal means of number of characters by same/other  $\times$  continuation completeness

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 5.7: Deletes per character by type of intervention

The model in table 5.8 was not as good a fit to the data as a simpler model including only same/other and floor change and their interaction and not including antecedent and continuation completeness (QICC<sup>7</sup> = 107.582, simpler model = 85.044), but the significant main effect shown in the more complex model is also present in the simpler model.

IV	Model effects		Parameter estimates		
	Wald $\chi^2$	p	B	Wald $\chi^2$	p
Same/Other (SO)	4.067	0.044*	-0.509	5.093	0.024*
Floor Change (FC)	0.391	0.532	0.362	2.350	0.125
Antecedent (Ant)	3.317	0.069	-0.030	0.019	0.891
Continuation (Cont)	0.474	0.491	0.063	0.031	0.861
SO $\times$ FC	0.264	0.607	0.173	0.247	0.619
FC $\times$ Ant	0.563	0.453	-0.198	0.216	0.642
FC $\times$ Cont	0.139	0.709	-0.065	0.012	0.911
SO $\times$ Ant	0.374	0.541	0.298	1.009	0.315
SO $\times$ Cont	0.426	0.514	0.322	0.330	0.566
Ant $\times$ Cont	0.515	0.573	-0.083	0.025	0.875
SO $\times$ FC $\times$ Ant	1.503	0.220	-0.717	1.319	0.251
SO $\times$ FC $\times$ Cont	1.222	0.269	-0.757	0.633	0.426
FC $\times$ Ant $\times$ Cont	0.891	0.345	-0.830	1.022	0.312
SO $\times$ Ant $\times$ Cont	0.653	0.419	-0.946	1.624	0.203

Table 5.8: GEE on deletes per character

### 5.3.4 Split point

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, the fake CCs were coded according to whether the split point fell within or between a syntactic constituent, and post-hoc analyses were carried out.

There were no additional main effects of whether the split point fell within or between

<sup>7</sup>Corrected Quasi Likelihood under Independence Model Criterion; a lower number indicates a better fit to the data.

a constituent and no interaction effects with same/other or floor change. These results are consistent with the findings from the corpora that the split point can 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 (see appendix F for details).

## 5.4 Discussion

Given the novelty of the method and the lack of other experimental studies of CCs to cross-check against, the results of this experiment must be interpreted with caution. Nonetheless, the results do bear on the questions raised in chapter 3.

Firstly, it is important to note that the introduction of fake CCs did have an effect on the ongoing dialogue, despite participants being unaware of the interventions. This in itself might be seen as surprising – if the intervention were highly disruptive, we would expect subjects to notice it.

### 5.4.1 Floor change effects

Though typing time is a fairly crude measure<sup>8</sup> one possible explanation for participants taking longer over the production of a turn (including in length of turn in characters) is that it 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 take 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).

That the effects of floor change were only significant in cases where the continuation could be considered as an independent contribution, or as a continuation to a prior contribution, further supports this interpretation. In these cases, participants take longer over their turns than when the continuation must unambiguously be integrated with the preceding material.

---

<sup>8</sup>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.

However, the effects 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. Either way, the effect does suggest that participants are sensitive to specific interlocutors, and the pattern of integration to prior turns, though the difference cannot be 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.

The lack of significant effects of antecedent completeness is especially surprising given the results from the tangram corpus, which showed that contributions almost always ended in a complete way, meaning that antecedents which do not end in a complete way ought to be more marked in text chat. Interestingly it seems that participants do not orient themselves to this difference caused purely by the mode of communication, indicating that the language resources used are the same, and available regardless of the medium. This is despite the fact that evidence (e.g. Newlands et al., 2003) suggests different patterns of interaction in text and spoken dialogue.

#### 5.4.2 Same-person versus cross-person effects

Independently of these effects, seeing a CC that appears to be shared between speakers has an impact on the conversation, seen in the amount of revision undertaken in formulating responses (deletes). Perhaps surprisingly, 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.

Participants may worry less about precisely formulating their turns following a cross-person CC because it could have the effect on the recipient of suggesting that the two other participants are highly coordinated and 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. Excerpts (5.1), (5.2) and (5.3), taken from the transcripts show examples where this appears to be the case (the ‘fake’ part of the CC is shown in bold).



(5.1) *AB-Split showing apparent coalition between ‘B’ and ‘D’*

**B:** and he can tell his formula

**D:** to tom and susie

(5.2) *AB-Split showing apparent coalition between ‘M’ and ‘B’*

**M:** i dont know i’m confused between

**B:** the doctor and the husband

(5.3) *BA-Split showing apparent coalition between ‘S’ and ‘B’*

**S:** these people said

**B:** you did something

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, may indicate how carefully participants are constructing their turns.

The interaction between same/other  $\times$  continuation completeness on number of characters such that participants type more in contributions following a cross-person CC if it had to be treated as a continuation suggests that it is harder to integrate the information from different interlocutors than from a single interlocutor, contrary to the predictions of hypothesis 7. However, as with the floor change effects reported above, this difference may also be due to the mismatch between what the two recipients saw, showing that they need to do more work (by using more characters) to retain a mutual understanding of the dialogue in progress.

## 5.5 Summary

The experiment reported in this chapter backs up the findings from the corpus studies that interlocutors are sensitive to turn-taking considerations but syntactic constituency does not seem to play a major role in interpreting CCs.

Further, as per hypothesis 11 these studies provide tentative evidence that, at least in task related dialogues with an explicitly shared goal, cross-person CCs are interpreted as evidence of collaboration between the contributors of the antecedent and continuation. This possibility will be pursued further in the next chapter.

## Chapter 6

### DiET chat tool experiment 2: Introducing fake CCs in the character-by-character interface

---

As there were differences in the distributions of CCs in the corpora according to the interactional properties of the interface, a further experiment into the effects of same- versus cross-person CCs was carried out using the character-by-character interface of the DiET chat tool. This enables us to examine the validity of the results presented in chapter 5 with regards to the interpretational effects (hypothesis 7) and parties (hypothesis 11) whilst controlling for the possible confound of conflicting expectations regarding who is entitled or expected to take the floor.

The experiment reported in chapter 5 offered some suggestive results as to the effects of introducing fake CCs into an ongoing dialogue in real-time, in line with proposed effects of party-formation. However, although we can look at examples where it seems to be the case that a fake split indicates a coalition (as in (5.1)-(5.3)) it is not possible to directly analyse the responses of participants who had been exposed to an apparently cross-person CC to see if they were indeed interpreting them in this way. Due to the linearity issues in line-by-line text chat (outlined in section 4.3), it is a matter of interpretation as to whether a contribution is a response to the immediately preceding contribution or something that occurred earlier in the dialogue. Additionally there were potentially confounding effects of conflicting expectations which may have led to the observed floor change effects (as discussed in section 5.4.1). The second experiment, using the character-by-character

interface, addresses these issues.

As discussed in section 5.1, the server console in the character-by-character interface of the DiET chat tool is the same as for the line-by-line version. Differences to the user interface will be outlined below.

## 6.1 Character-by-character interface

In the character-by-character version of the DiET chat tool, the user interface consists of a single chat window. Below this, there is a status bar, which indicates if any participants are actively typing (see figure 6.1).

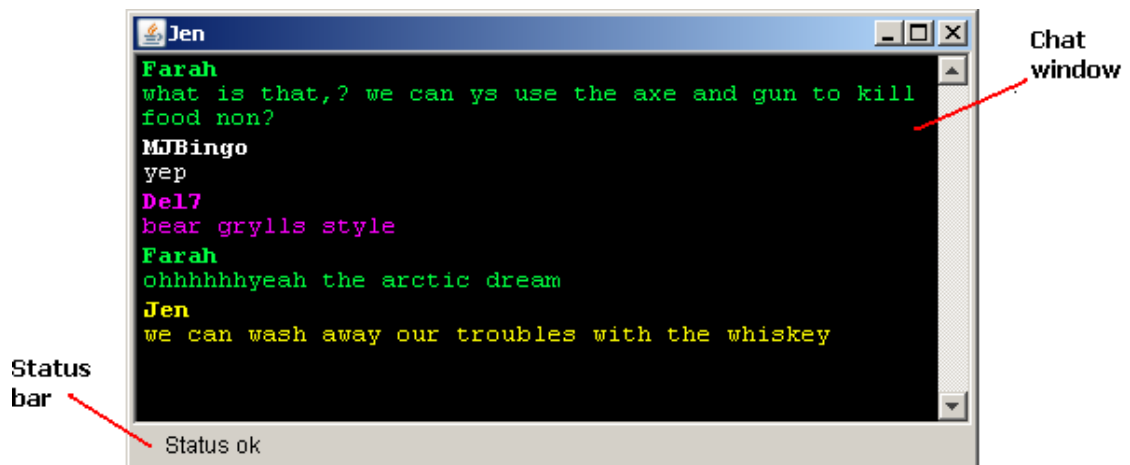


Figure 6.1: The user interface chat window (as viewed by *Jen*)

Unlike traditional chat interfaces (such as MSN Messenger), users type directly into the same window in which they see their interlocutors' contributions. This means that each character that any of the participants type is displayed in the window at the time it is entered – i.e. users see both their own and their interlocutors' contributions unfold in a character-by-character fashion. Consequently, only one participant may type at a time.

## 6.2 Method

This experiment used the same intervention as experiment 1, in that genuine turns from conversants were artificially split into two parts round an arbitrary split point in the string (always a space character to prevent mid-word CCs). The experiment differs from that reported in chapter 5 in several respects. First, and most importantly, it used

the character-by-character interface of the DiET chat tool (as described in section 6.1). Secondly, the experiment used four person conversations (instead of triads) in order to control for the possibility that the floor change effects seen in experiment 1 were as a result of differing expectations about who will speak next. The design of this experiment means that for each intervention, two of the four participants saw a manipulated fake CC, but these could be any one of the four possibilities as shown in table 5.1 (and reiterated in section 6.2.1 below). This differs from experiment 1, in which one of the participants always saw an AA type intervention. This meant that in cases where one recipient had seen either an AB- or BB-split (though not when they had seen an BA-split), the two recipients would have believed a different interlocutor to have produced the continuation.

### 6.2.1 Materials

#### The arctic survival task

The *arctic survival task* is a decision making task requiring agreement on which objects, salvaged from a plane crash, will aid survival in arctic conditions, and which are red herrings. 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, and, as has already been seen in the comparative corpus study, cross-person CCs occur naturally in text discussions of this task. The task also requires the participants to come to an agreement about the order of importance of the objects (see appendix B).

#### The intervention

As in experiment 1, certain single contributions are automatically altered so as to appear as if they are two part CCs. When the intervention is ready to be triggered, an incoming turn is assessed for its suitability, based on the length of the contribution and the participant typing it (consecutive interventions are never carried out on the same person's contributions). The text is then relayed to the other participants in the same character-by-character fashion as genuine contributions, with a split point inserted arbitrarily at a space character. There is a short delay between the two parts to make it appear as if they are in fact separate contributions. During the intervention, participants are unable to type, which, depending on their status in the intervention is either because there appears to be someone else typing (even though who this is may be incorrect) or

due to network errors. Inevitably there is a slight buffer whilst a turn's suitability to be manipulated is assessed before it is relayed to the other participants; random network errors are introduced so that this buffer does not systematically cue an intervention. Note also that for the purposes of this experiment (due to screen synchronisation issues), the delete function was disabled.

For each intervention, the server produces two compound contributions, relaying different information to all three of the recipients. The first recipient is the 'fake sender'; this person receives the turn as a single turn from the genuine source (i.e. they do not see a manipulation, though there is a slight delay as discussed above). The two other recipients each see one of four different CC manipulations. The split types were a within-subject manipulation, meaning that each participant may see any (or all) of the different manipulations during the dialogue.

**AA-split** – Both antecedent and continuation parts of the CC appear to come from the actual source of the utterance.

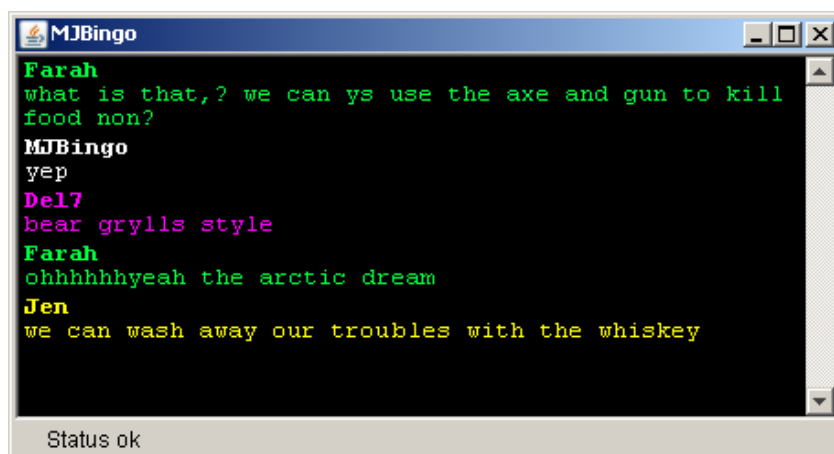
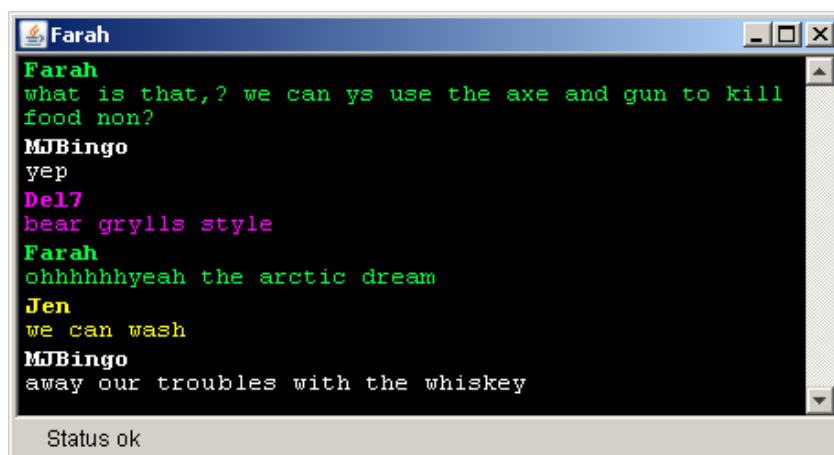
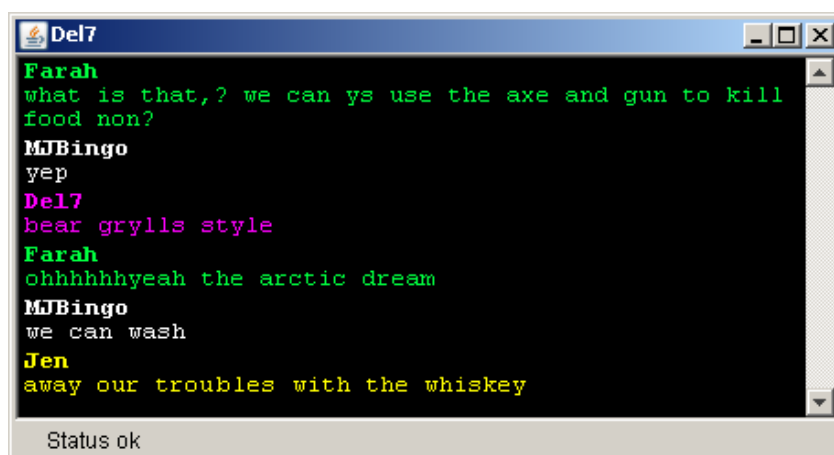
**BB-split** – Both parts of the CC appear to come from the fake sender (i.e. not the actual source of the utterance).

**AB-split** – The antecedent part of the CC appears to come from the actual source, and the continuation appears to come from the fake sender.

**BA-split** – The antecedent part of the CC appears to come from the fake sender, and the continuation appears to come from the actual source.

This factorial design allows us to separate the effects of a change in conversational momentum (floor change) from the effects of a same- versus cross-person CC. An example of what each participant sees in a genuine intervention taken from the dialogues is shown in figures 6.1 to 6.4. Figure 6.1 is the actual source of the intervened turn and 6.2 shows what the fake sender sees whilst figures 6.3 and 6.4 show what the two recipients see (an AB- and BA-split, respectively).

The intervention is triggered every 8 turns, restricted so that the same person does not contribute consecutive intervention turns. Which type of intervention is seen by each of the participants is generated pseudo-randomly, such that if an intervention has been

Figure 6.2: 'Fake sender' participant *MJBingo*'s viewFigure 6.3: *AB-CC*, as viewed by *Farah*Figure 6.4: *BA-CC*, as viewed by *Del7*

seen by a participant then they are less likely to see the same intervention subsequently. Whether the two participants saw the same or different interventions, and which types were seen at the same time by the two different participants was random.

Additional analyses were performed in a between-subjects design comparing the responses of participants in the manipulated dialogues to those from the 4-way arctic survival dialogues from the character-by-character corpus. These *control* dialogues were collected in an identical manner in all respects (e.g. task, number of subjects per conversation etc), except that there were no interventions.

### 6.2.2 Subjects

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

### 6.2.3 Procedure

As with experiment 1, each subject was sat in front of a desktop computer in a separate cubicle, so they were unable to see or hear each other. After inputting their e-mail address and username, a blank chat window appeared, and they were given the task description. They were informed that the experiment was investigating the differences in communication when conducted using text as opposed to speech, that the experiment would last approximately 45 minutes, and that all text would be recorded anonymously for later analysis. Once all four participants had been logged on, the experimenter made no further contact with them until at least 40 minutes of dialogue had been carried out.

### 6.2.4 Analysis

The character-by-character version of the chat tool enforces strict notions of turn-taking – there can be no overlapping or simultaneous contributions as only one participant may hold the floor at any given time. If, however, a participant pauses whilst constructing their contribution, one of their interlocutors may take the floor. For the purposes of this experiment the floor timeout value was set at 200ms. The contribution immediately following the intervention was therefore taken to be a response to it.

### Qualitative

Given the results from the first experiment and the tuition and tangram corpora, we would expect that in cases where one of the recipients sees what appears to be a cross-person CC, they would react as if the other two recipients have formed a party (as per hypothesis 11). If the hypothesised explanation for using fewer deletes after an apparent cross-person CC is correct then we might also expect them to respond in a way which shows that they believe there to be a coalition. As can be seen from examples (6.1)–(6.9) below (misattributed parts are shown in bold), this does sometimes appear to be the case. Notice that in all these examples, although different types of response are formulated, they respond to the whole CC (antecedent and continuation), regardless of the fact that it appears to be split between two different interlocutors.

(6.1) *AB-Split showing apparent coalition between ‘Y’ and ‘B’*

**Y:** clothes don’t

**B:** burn that well. wood would be better.

**L:** agreed [DiET SU2 18 211-212]

(6.2) *BA-Split showing apparent coalition between ‘H’ and ‘F’*

**H:** but u can eat

**F:** the fat if needs be contain lots of energy if u low on anything

**L:** yeh so then maybe the fat atfet the newspapers [DiET SU2 12 164-165]

(6.3) *BA-Split showing apparent coalition between ‘H’ and ‘S’*

**H:** msn is way

**S:** better

**T:** i agree [DiET SU2 2 187-188]

(6.4) *AB-Split showing apparent coalition between ‘C’ and ‘J’*

**C:** what are youyr

**J:** items agin again?

**M:** yep

lets list them again [DiET SU2 4 75-77]





and compared global factors from the dialogues in which fake CCs had been introduced to the control dialogues. There were 8 control dialogues, with an average length of 288 contributions. The manipulated dialogues were also on average 288 contributions long, though participants saw, on average, 4.3 interventions during the dialogue.

The second type of analysis investigated differential effects on the responses to the different types of intervention, as with experiment 1. As with the corpus studies, reported in chapter 4, the fake CCs were coded according to whether or not the antecedent was end-complete and whether or not the continuation was start-complete.<sup>1</sup> The number of each type is shown in table 6.1, below.

Antecedent	Complete				
	n		y		
Continuation	n	y	n	y	
AA	15	3	4	7	29
AB	6	7	4	6	23
BA	12	5	7	5	29
BB	6	2	5	7	20

Table 6.1: Number of cases by antecedent end and continuation start-completeness

The data were once again analysed according to several factors in a factorial design; *same/other* – whether both parts of the utterance appeared to come from the same person, or from different sources ([AA and BB] vs [AB and BA]), *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]), antecedent end-completeness and continuation start-completeness.

As the two interventions were assigned randomly, but only one of the recipients could possibly respond first, two additional factors were available to control for possible effects of differing expectations about who should speak next. These were whether the *other* recipient (i.e. the one who did not respond first) saw a *same/other* CC and whether or not they saw a *floor change* CC.

<sup>1</sup>The difference between these codings and those for experiment 1 are because the enforced linearity of the character-by-character interface means that participants do not expect contributions to be necessarily end-complete, as with spoken dialogue.

### Dependent variables

Measures selected for analysis were *number of characters* and *typing time* as measures of response length and *onset delay* as a measure of processing time. As with experiment 1, these dependent variables are open to interpretation, however, we again expect turn length (in both typing time and number of characters) to be an indication of how much work participants need to do in order to maintain mutual understanding.

Contrarily, onset delay can be viewed as a more direct measure of how much trouble participants are having in integrating the preceding turn. As we do not expect any additional difficulty in parsing the syntax of what is offered regardless of who has apparently supplied the contribution(s), longer onset delays could therefore indicate either a processing difficulty at the discourse level, or that a participant does not feel that they are entitled to the floor.

Once again, if speakers are interchangeable then we would not expect to see any differences (on any of the measures) between turns following cross-person or same-person CCs. However, if, as is indicated by experiment 1, the presence of cross-person CCs leads interlocutors to behave as if parties have been formed (as per hypothesis 11) then we should expect participants to type shorter turns than otherwise, because mutual understanding is presupposed. Additionally, the evidence from the corpus studies suggests that participants would expect the turn following a cross-person CC to be taken by the provider of the antecedent so we ought to see longer onset delays by participants who have seen an apparently cross-person CC as they do not feel entitled to take the floor.

We would again expect that end-complete antecedents are more natural, which should manifest itself in shorter onset delays as there should be less integrational difficulty.

As the data were not normally distributed, even after transformations were applied, they were analysed using a Generalised Estimating Equation (GEE) model, using a gamma distribution with same/other, floor change, antecedent end-completeness and continuation start-completeness as fixed factors and participant as subject variable.

### 6.3 Results by dialogue

A post-experimental questionnaire (see appendix D) and debriefing showed that although participants did not feel that the conversations went as smoothly as face-to-face dialogue,

they found it easy to understand and neither overly easy or difficult to come to a decision. Importantly there were no differences in their questionnaire answers compared to the control group who had not experienced any interventions (appendix E). None of the participants reported awareness of any interventions.

Contributions	Conversation Type		
	CC	Control	All
Number of contributions	5471	2302	7773
Mean onset delay	2725.03	2079.94	2533.98
Mean typing time of contribution	6948.40	7222.25	7029.50
Mean num of characters per contribution	30.88	28.42	30.15
Mean number of words per contribution	6.17	5.74	6.04
Mean num of contributions per dialogue	287.95	287.75	287.89

Table 6.2: Comparison of experimental and control dialogues

As can be seen from table 6.2, there are a number of global differences between the experimental and control dialogues. T-tests<sup>2</sup> show that control conversations have lower onset delays ( $t_{(7336.387)} = 6.204, p < 0.001$ ), and participants use on average fewer characters ( $t_{(5153.864)} = 2.821, p = 0.005$ ) and words ( $t_{(5176.586)} = 2.528, p = 0.011$ ) than in the experimental dialogues, suggesting that the introduction of CCs leads to additional processing costs (though these differences could simply be an artefact of seeing apparent network errors in the experimental dialogues, and not in the control dialogues).

To explore whether the presence of fake CCs had the effect of making participants behave as if parties or coalitions have been formed (hypothesis 11), we looked at the characteristics of specific contributions that participants make during the conversations. Although uncommon, in the experimental dialogues, participants used explicit agreement tokens (“I agree”, “I concur” etc) twice as often as in the control dialogues (104/5471, 2% vs 25/2302, 1%;  $\chi^2_{(1)} = 6.593, p = 0.010$ ), suggesting that they are indeed behaving as if there are more coalitions and adjusting their own linguistic behaviour accordingly.

Similarly, participants were more likely to explicitly express propositional attitudes (using verbs like “think”, “believe” etc) in the experimental dialogues (263/5471, 5% vs 75/2302, 3%;  $\chi^2_{(1)} = 9.348, p = 0.002$ ), which suggests that they are explicitly outlining their own beliefs, and/or checking whether these are shared. This fits with the idea of cross-person CCs being more likely if participants share knowledge or beliefs (as suggested

<sup>2</sup>Equal variances not assumed as Levene’s test for equality of variance was significant for all measures.

by the corpus studies), and suggests that if there are CCs, then participants make the inverse inference (perhaps automatically) that knowledge is shared.

## 6.4 Results by intervention

There were 165 interventions in the data. In each case there were three possible first respondees; either of the two participants who had seen an intervention, and the fake sender. Table 6.3 shows who responded first in each case. Z-tests (see appendix G) show there that there are no significant differences between the proportions of first responses according to which type of intervention was seen. As this holds true for what each recipient saw, including the fake sender, this adds weight to the notion that the interventions were not disruptive per se.

	possible	actual	percent
AA	82	29	35.37
AB	68	23	33.82
BA	72	29	40.28
BB	61	20	32.79
fake sender	165	64	38.79

Table 6.3: First response by type of intervention

### 6.4.1 Typing time and number of characters

The results from GEE models on typing time and number of characters (using the gamma distribution, with an unstructured correlation matrix and an identity link function),<sup>3</sup> with participant as subject effect show that the first response after a fake CC is shorter (Wald- $\chi^2 = 9.951$ ,  $p = 0.002$ , table 6.6) and takes less time (Wald- $\chi^2 = 11.700$ ,  $p = 0.001$ , table 6.5) if it follows a cross-person CC rather than a same-person CC (see also table 6.4).

	n	typing time	number of characters	Onset delay
AA	29	15294.79	60.14	8880.79
AB	23	5964.00	28.17	11416.35
BA	29	8578.14	35.00	8947.21
BB	20	11620.45	56.55	10722.50
fake sender	64	12427.02	51.22	10898.63

Table 6.4: Mean results of first response

<sup>3</sup>Goodness of fit (QIC) - typing time 116.866; number of characters 116.715.

For both typing time and number of characters there is an interaction effect of floor change by continuation start-completeness (typing time, Wald- $\chi^2 = 6.784$ ,  $p = 0.009$ ; number of characters, Wald- $\chi^2 = 9.265$ ,  $p = 0.002$ ). There was also an interaction effect of antecedent end-completeness  $\times$  continuation start-completeness (typing time, Wald- $\chi^2 = 6.612$ ,  $p = 0.010$ ; number of characters, Wald- $\chi^2 = 6.432$ ,  $p = 0.011$ ).

IV	Model effects		Parameter Estimates		
	Wald $\chi^2$	p	B	Wald $\chi^2$	p
Same/Other (SO)	11.700	0.001**	13114.383	10.334	0.001**
Floor Change (FC)	2.707	0.100	965.919	0.122	0.727
Antecedent (Ant)	0.312	0.576	1251.551	0.174	0.676
Continuation (Cont)	1.217	0.270	-0.077	0.624	0.430
SO $\times$ FC	0.095	0.758	2061.736	0.149	0.700
SO $\times$ Ant	0.008	0.927	2241.602	0.046	0.830
SO $\times$ Cont	0.178	0.673	1012.758	0.009	0.922
FC $\times$ Ant	<0.001	0.993	6120.700	0.732	0.392
FC $\times$ Cont	6.784	0.009**	-2063.968	0.253	0.615
Ant $\times$ Cont	6.612	0.010**	-29013.547	6.356	0.012*
SO $\times$ FC $\times$ Ant	1.632	0.201	2737.528	0.033	0.856
SO $\times$ FC $\times$ Cont	2.770	0.083	-13303.225	0.746	0.388
SO $\times$ Ant $\times$ Cont	0.908	0.341	22964.244	2.645	0.104
FC $\times$ Ant $\times$ Cont	2.904	0.088	-122.029	<0.001	0.987
SO $\times$ FC $\times$ Ant $\times$ Cont	2.745	0.098	-29884.769	2.745	0.098

Table 6.5: GEE of typing time

IV	Model effects		Parameter Estimates		
	Wald $\chi^2$	p	B	Wald $\chi^2$	p
Same/Other (SO)	9.951	0.002**	64.421	11.800	0.001**
Floor Change (FC)	0.016	0.899	1.696	0.022	0.882
Antecedent (Ant)	0.977	0.323	7.575	0.363	0.547
Continuation (Cont)	2.119	0.145	22.949	1.432	0.231
SO $\times$ FC	0.282	0.595	2.996	0.013	0.911
SO $\times$ Ant	0.390	0.532	32.169	0.298	0.585
SO $\times$ Cont	1.357	0.244	16.168	0.187	0.665
FC $\times$ Ant	0.007	0.931	21.041	0.678	0.410
FC $\times$ Cont	9.265	0.002**	94.974	2.129	0.145
Ant $\times$ Cont	6.432	0.011*	-88.247	6.680	0.010*
SO $\times$ FC $\times$ Ant	2.246	0.134	-15.997	0.054	0.817
SO $\times$ FC $\times$ Cont	1.395	0.103	-92.640	1.703	0.192
SO $\times$ Ant $\times$ Cont	0.427	0.514	23.738	0.176	0.675
FC $\times$ Ant $\times$ Cont	0.340	0.560	24.351	0.641	0.423
SO $\times$ FC $\times$ Ant $\times$ Cont	1.611	0.204	-93.390	1.611	0.204

Table 6.6: GEE of number of characters

Pairwise comparisons on the marginal means show that if the recipient saw a floor change CC (AB or BB), then they spend less time typing and type less if the continuation is not start-complete. Similarly, if participants saw a continuation which did not appear to start in a complete way then they typed less if they had seen a floor change CC. These

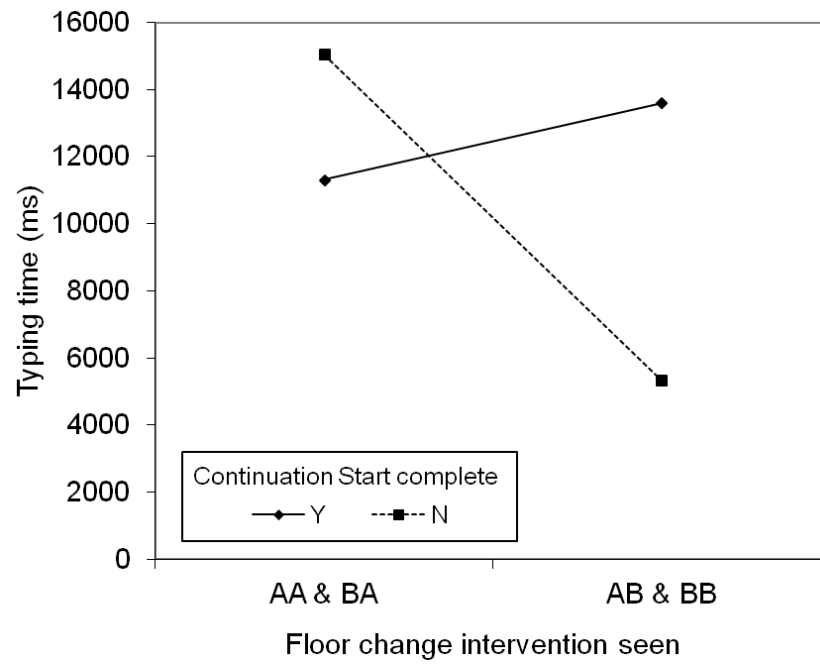


Figure 6.5: Marginal means of contribution typing time by floor change  $\times$  continuation start-completeness

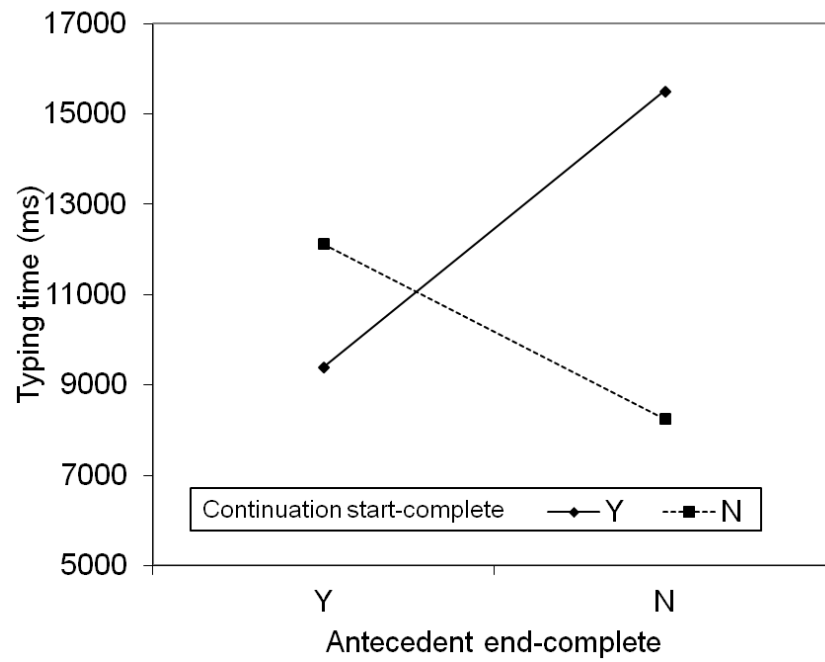


Figure 6.6: Marginal means of contribution typing time by antecedent end-completeness  $\times$  continuation start-completeness

are cases in which there is both a change of floor and the continuation is clearly not an independent turn. These differences can be seen in figure 6.5.

For antecedent end-completeness  $\times$  continuation start-completeness (see figure 6.6), pairwise comparisons show that if participants had seen a CC with an antecedent that was not end-complete but a continuation that was start complete, then they typed more than other cases where either both appeared to be complete or both did not. Note however that the results of same/other are unaffected by these factors.

### By participant

Averaged over all contributions by participant, subjects who saw a greater number of apparently cross-person CCs typed shorter turns (both in typing time and number of characters<sup>4</sup>) than those who saw fewer cross-person CCs (see figure 6.7), independently of how many floor change CCs they saw or fake CCs in total.

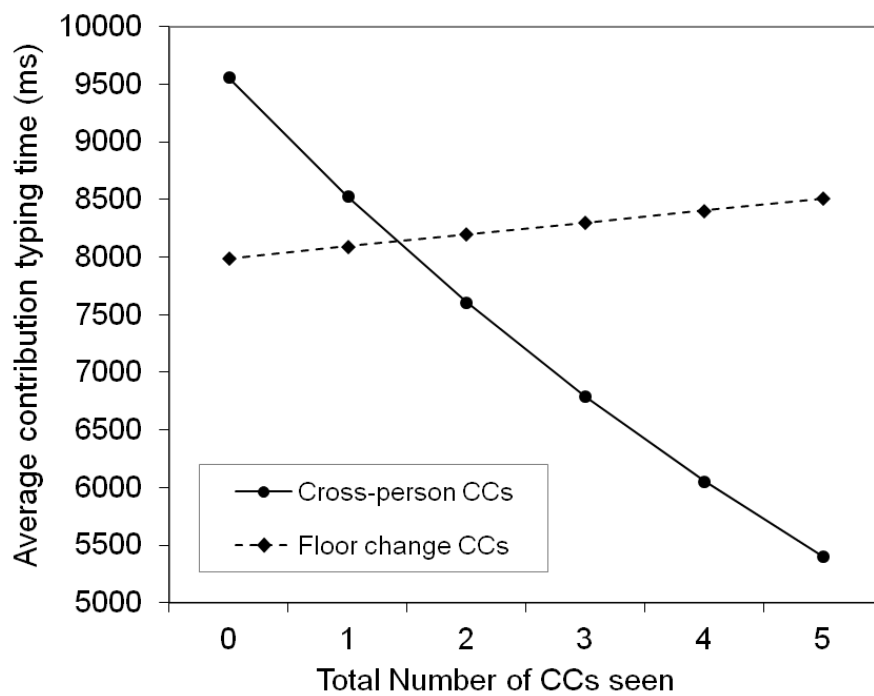


Figure 6.7: Marginal means of average contribution typing time by number of CCs seen

These results are again strongly suggestive of a belief that the other participants are

<sup>4</sup>GEEs using a gamma distribution and independent correlation structure with conversation ID as subject variable and total number of interventions seen, total number of floor change interventions seen and total number of cross-person interventions seen as covariates; Average typing time of turn, QIC = 20.454, Wald- $\chi^2 = 5.169$ ,  $p = 0.023$ ; number of characters, QIC = 15.379, Wald- $\chi^2 = 6.098$ ,  $p = 0.014$ .



forming (or are already members of) parties, as, if there are more apparent coalitions (indexed by the number of cross-person CCs seen) then participants do less, in terms of typing shorter turns. Note that there is no difference in the number of contributions by each participant based on these measures, so it is not the case that participants are making fewer contributions, rather that they may believe there is less up for discussion so are making shorter contributions.

#### 6.4.2 Onset delay

There are no effects on onset delay (the time between seeing the end of a fake CC and the response, see table 6.7),<sup>5</sup> although there appears to be a trend towards longer onset delays if the respondent saw a floor change CC (see table 6.4).

IV	Model effects		Parameter Estimates		
	Wald $\chi^2$	p	B	Wald $\chi^2$	p
Same/Other (SO)	0.122	0.727	-0.232	0.338	0.561
Floor Change (FC)	1.225	0.268	-0.125	0.089	0.765
Antecedent	0.270	0.603	-0.565	2.311	0.128
Continuation	0.039	0.844	0.351	0.624	0.429
SO $\times$ FC	0.086	0.769	0.092	0.086	0.769
SO $\times$ Ant	0.757	0.384	0.308	0.757	0.384
SO $\times$ Cont	0.031	0.859	-0.082	0.031	0.859
FC $\times$ Ant	1.630	0.202	0.435	1.630	0.202
FC $\times$ Cont	3.183	0.074	-0.707	3.183	0.074
Ant $\times$ Cont	0.150	0.699	0.166	0.150	0.699

Table 6.7: GEE of onset delay by same/other, floor change, antecedent end-completeness and continuation start-completeness

However, when the two additional factors (included to control for differences in speaker expectations) were included in the analysis, there were effects on onset delay depending on differences in what the other person, who did not respond first, saw.<sup>6</sup>

There was an interaction effect between non-responding recipient same/other and non-responding recipient floor change (see table 6.8).<sup>7</sup> Pairwise comparisons show that if the non-responder saw a cross-person CC that did not involve a floor change (i.e. a BA-intervention) responses from the actual responder took longer than all of the other

<sup>5</sup>The model used a gamma distribution with a log link function and exchangeable correlation matrix, goodness of fit (QIC) = 70.896. Note that though there were higher order effects in a complete model (see appendix G), removing the 3- and 4-way interactions as shown in the model here resulted in an improved QICC of 81.112 (from 84.569), indicating that the simpler model is in fact a better fit to the data.

<sup>6</sup>Recall that, unlike in experiment 1, any pair of interventions could be seen.

<sup>7</sup>3-way and 4-way interactions were not significant and have been removed from the model shown, which improves the QICC from 77.725 to 69.002; full model shown in appendix G.

possibilities – again independently of what type of intervention the actual responder saw. See table 6.9, and figure 6.8.

IV	Model effects		Parameter Estimates		
	Wald $\chi^2$	p	B	Wald $\chi^2$	p
Same/Other (SO)	0.390	0.532	-0.346	1.831	0.176
Floor Change (FC)	0.003	0.955	-0.033	0.016	0.900
Other Same/Other (Oth SO)	15.232	<0.001**	-0.462	3.757	0.053
Other Floor Change (Oth FC)	14.189	<0.001**	0.884	9.191	0.002**
SO $\times$ FC	0.018	0.892	.028	0.018	0.892
SO $\times$ Oth SO	1.811	0.178	0.332	1.811	0.178
SO $\times$ Oth FC	0.434	0.510	0.170	0.434	0.510
FC $\times$ Oth SO	1.784	0.182	0.293	1.784	0.182
FC $\times$ Oth FC	1.126	0.289	-0.273	1.126	0.289
Oth SO $\times$ Oth FC	6.164	0.013*	-0.677	6.164	0.013*
QIC = 55.558; QICC = 69.002					

Table 6.8: GEE of onset delay of response including what non-responding recipient saw

	same-		other-		Total	
	N	onset	N	onset	N	onset
no floor change	35	7906.17	22	17739.91	57	11701.65
yes floor change	21	6839.14	23	7974.87	44	7432.82
Total	56	7506.04	45	12748.89	101	9841.96

Table 6.9: Mean onset delay of first response by what the non-responding recipient saw

If the non-responding recipient saw a cross-person CC, then the onset delay for the actual responder was slower. Conversely, if the non-responding recipient saw a floor change CC, then the onset delay for the actual response was quicker.

In other words, there seem to be cases where at least one of the participants is leaving the floor open – i.e. they may not think it appropriate for themselves to take the floor.

### Onset delay following genuine CCs

As we have precise timing data from conversations using the character-by-character interface without interventions, we can look at the onset delay following genuine CCs in order to explore how disruptive they were on participants.

person:	Same-	Cross-
	all	cross-turn
		(all)
Onset Delay	2521.91	2137.44
Standard Deviation	3302.45	2067.88
Total	315	82
		167

Table 6.10: Onset delay following a genuine CC

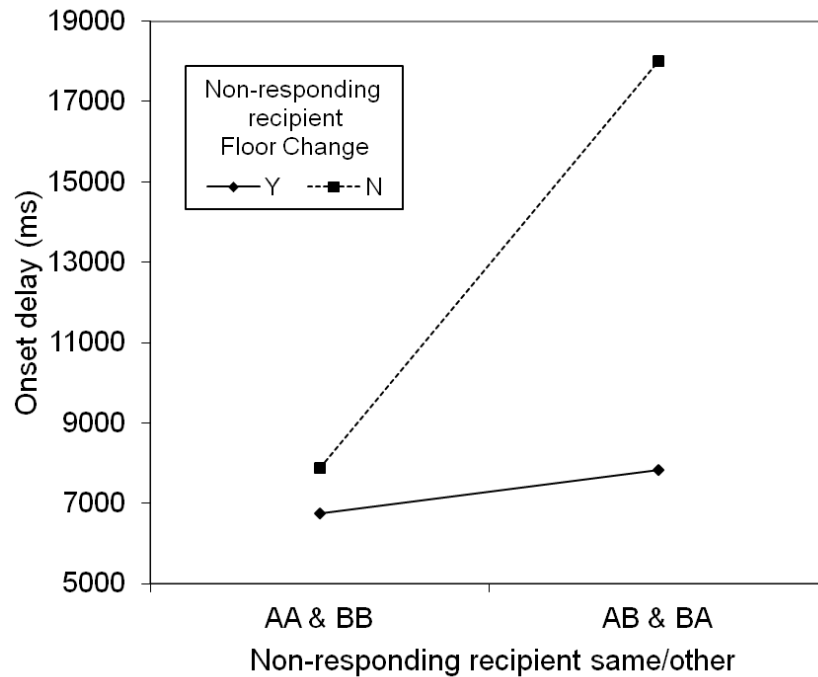


Figure 6.8: Marginal means of onset delay by what the non-responding recipient saw

As can be seen from table 6.10, the onset delay following a genuine cross-person CC is higher than that following a same-person CC ( $t_{171.426} = -2.349, p = 0.020$ ). This suggests that there are differences between who is entitled to or expected to speak next following a cross-person CC, as with the experimental results.

Additionally, a one-way ANOVA shows that for the cross-person CCs, there is a difference in the onset delay depending on who subsequently takes the floor (A (81), 7445.57; B (20), 1462.70; C (66), 2577.17;  $F_{(2,166)} = 3.796, p = 0.024$ ). Post-hoc comparisons show that if the next speaker is the supplier of the antecedent (A), then the onset delay is greater than if it is either the supplier of the continuation (B) or a third person (C).

This result, showing that interlocutors leave the floor open for longer in specifically those cases in which they expect the antecedent owner to supply the next contribution (and note that there are no cues to the contrary as there might be in face-to-face dialogue) could help to explain the apparently counterintuitive results of the influence of what the other person saw on onset delay.

Onset delay following *all* contributions in the control dialogues was also found to

be correlated with the contribution length ( $r_{(2208)} = 0.109$ ,  $p < 0.001$ ), such that longer contributions typically resulted in a longer delay. Reasons for this are unclear, but it may be that if participants are typing longer turns, due to the lack of visual feedback, their interlocutors may pay less attention, or e.g. remove their hands from the keyboard etc. The length of the manipulated turn (the fake CC) was therefore included as a covariate in the models reported below.

### Other recipient effects on onset delay

Returning to the onset delay data, a number of post hoc tests were performed to try to identify the source of the interaction. Looking only at the 48 cases where a non-responding recipient had seen a *BA*-split (which were the ones with a large onset delay, as discussed, and including those where the fake sender produced the first response), we find that there is an effect of who supplied the contribution *prior* to the fake CC. For those cases where the fake sender ('B') had made the contribution immediately prior to the CC, participants were quicker in responding than if one of the recipients had made the immediately preceding contribution ( $t_{(43.918)} = -2.487$ ,  $p = 0.017$ ).

Looking again at all the first responses to the CC, we find that participants are disproportionately more likely to respond first if they saw a cross-person CC and the speaker immediately prior to the intervention was not the fake sender (43/73, 58% vs. 9/28, 32%;  $\chi^2_{(1)} = 5.803$ ,  $p = 0.016$ ). This shows that recipients of a cross-person CC may be more likely to leave the floor open if the sequence of contributions appears as a dyadic exchange between the actual and fake senders ('BAB' or 'BBA') rather than a more collaborative portion of the dialogue ('CAB' or 'CBA'). It seems therefore that the effects of not just who appeared to supply the antecedent and continuation but also who supplied the previous contribution, may be relevant in the organisation of turn-taking.

That this is the case can be seen illustrated graphically in figure 6.9. If prior speaker (*fake sender* or *other*) is added to the model, then the previously seen interaction effects are subsumed by two new interactions; other floor change  $\times$  previous speaker (Wald- $\chi^2 = 3.926$ ,  $p = 0.048$ ) and other same/other  $\times$  previous speaker (Wald- $\chi^2 = 3.847$ ,  $p = 0.050$ ) though note that there is no 3-way interaction between all these three factors.<sup>8</sup>

More complex models indicate that the other same/other  $\times$  previous speaker are not

---

<sup>8</sup>Model shown in appendix G.

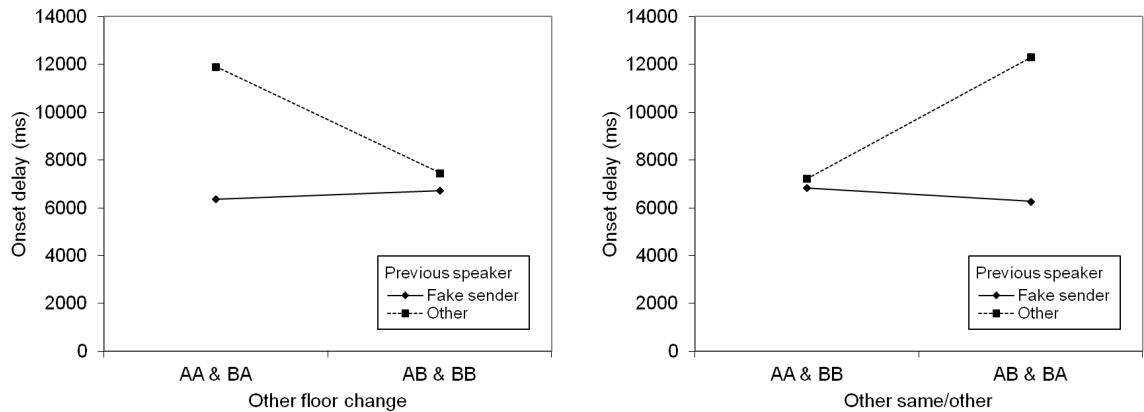


Figure 6.9: Marginal means of onset delay by what the non-responding recipient saw and prior speaker

influenced by higher order effects, but that the other floor change  $\times$  previous speaker effect may be modified by the completeness or otherwise of the antecedent. Lack of data and the complexity of the model means that these results are suggestive and not definitive, but it at least appears as though there are complex issues involved in tracking who may be entitled to speak next which may be more difficult to resolve in these 4-way text dialogues.

#### 6.4.3 Planned post-hoc analyses

Once again, as the experiment was looking for generic effects of CCs on the dialogue, the location of the split points was random. The data were therefore annotated as to whether the split point occurred within or between constituents, as before, but there were no additional effects of this factor on any of the measures (see appendix G).

## 6.5 Discussion

As with the line-by-line version of the experiment, introducing fake CCs did have measurable effects on the ongoing dialogue, despite participants being unaware of the interventions.

Participants who saw more cross-person CCs typed shorter contributions over the entire dialogue, and immediate responses to cross-person CCs were shorter than responses

to same-person CCs, which backs up the previously reported corpus and experimental results. As in the line-by-line experiment, participants are doing less in those cases which would be indicative of strong coordination between two of their interlocutors, with respect to the task in hand,<sup>9</sup> even though this apparent coordination is a construct of the manipulation. Thus we can see that subjects are orienting to potential parties, which may change as the conversation progresses, providing evidence for hypothesis 11.

The differences in onset delay are harder to explain. There is no obvious reason why, on some occasions, participants who receive a BA intervention respond first but on those occasions where they do not the floor is left open for longer, with the other participant also failing to respond as quickly. The results suggest, however, that there are some subtle effects associated with both sequential ordering and who is expected or entitled to take the floor. This is harder to negotiate in this text interface than in face-to-face dialogue, where participants may not overlap (even with e.g. grounding cues, which might serve to break up longer contributions into more easily integrated ‘chunks’) or interrupt, and may be exacerbated by the fact these are four-way dialogues. Following a CC this manifests itself in a number of different but interacting ways with participants leaving the floor open because they believe that another participant is ‘more’ entitled to the next turn. This can be for a variety of reasons, for example, where they have seen an apparently cross-person CC and the fake sender took the immediately preceding turn.

The interaction effects of floor change  $\times$  continuation start-completeness, and antecedent end-completeness  $\times$  continuation start-completeness show that, as with experiment 1, conversational momentum can affect production of subsequent contributions. Although the corpus studies found that a split point could apparently occur between any words in a string, some forms were more common than others. Those cases in which the antecedent is not end-complete and the continuation is not start-complete correspond to more unusual CCs, and may therefore lead to interpretation difficulties when combined with misattribution of the continuation. This raises questions regarding how interpretations are built up, online and incrementally. Although speaker switching at arbitrary points does not obviously cause processing difficulties for a third party, and is unaffected by the split point, there are effects associated with floor changes, in just those cases

---

<sup>9</sup>Note that, as with the previous experiment, and the tuition and tangram corpora, the task is designed such that the joint goal of the participants is to achieve coordination.

where the interpretations must be unambiguously integrated.

For antecedent end-completeness  $\times$  continuation start-completeness participants typed less in cases where either both appeared to be complete or both did not. In other words, they are doing more work in the cases which are ambiguous as to whether the continuation is explicitly tied to the antecedent or not. This appears to be another integration effect, when potentially competing interpretations are available.

## 6.6 Summary

The studies reported in the preceding chapters show that there are more cross-person CCs where the nature of the task requires coordination (either in imparting knowledge as in the hierarchy task corpus, or in coming to an agreement as in the character-by-character corpus and experiment, as per hypothesis 4). Furthermore, where the task explicitly defines those parties by means of what knowledge is shared, a greater proportion of cross-person CCs are between the members of the party than across party lines. This extends to Lerner's collaborative turn sequences whereby if the continuation is offered by someone external to the party, the third part to the sequence (a ratification or rejection) may be supplied by a different party member than the one who supplied the antecedent (confirming hypothesis 10).

In both the line-by-line DiET experiment (reported in chapter 5) and the character-by-character experiment (chapter 6), seeing alleged coalitions or parties between interlocutors led to recipients subsequently doing less work; in the line-by-line experiment this was manifest in participants using fewer deletes per character, and in the character-by-character experiment, subsequent turns were shorter, both immediately following a cross-person CC and over the duration of entire conversations with participants who had seen more apparent parties producing shorter turns overall. This is taken as evidence that the presence of cross-person CCs does lead interlocutors to act as if parties have been formed (hypothesis 11).

In both experiments there were also effects of floor change related to a change in conversational momentum caused by the misattribution of the most recent contribution seen. In both experiments there were differences related to whether the continuation could be treated as an independent contribution or not. Responses were shorter when

there was a floor change and the continuation did not start in a complete way (and thus had to be unambiguously integrated with some prior contribution), independently of the interface (which, as previously discussed imposes additional constraints).

In experiment 2, using the character-by-character interface to prevent turns being constructed and repaired in private (and potentially simultaneously), issues regarding who is entitled or expected to speak next also have effects, for example in cases where participants were leaving the floor open for their colleagues; particularly after seeing an apparent BA-split. The reasons for this are unclear, but the data from the genuine CCs in the control dialogues demonstrate that onset delays are also extended after a real cross-person CC, especially where the antecedent owner subsequently takes the floor, suggesting that they tend to be the expected next speaker following a cross-person CC (a tendency borne out by the corpus results in which the antecedent owner takes the floor in a disproportionately high number of cases, contra hypothesis 9). This fits with the idea, discussed in section 4.5, that low-level expectations of who should provide the next (possibly grounding) contribution are independent of who is taken to be responsible for the complete speech act. However, note that these expectations do not fit with the idea that CCs indicate parties in Schegloff's (1995) sense, as the expectations of turn-taking are applied to the contributing individuals and not to the apparent party. It might therefore be more parsimonious to speak of the type of party which is demonstrated by a CC as a *coalition*, which acts like a party at the level of shared content, but not necessarily at the level of turn-taking.<sup>10</sup>

It may also be the case that, as the dialogues are between 4 people, participants are simply unsure of the appropriateness or otherwise of themselves taking the floor. This again may be due to the lack of feedback cues in text chat (either backchannels or non-verbal) with which to orient themselves as a group.

In sum, in line with previous evidence, the studies in this chapter offer little or no evidence for syntactic constraints on either production or processing of CCs (as per

---

<sup>10</sup>Note that this is similar to the notion of *collective author* introduced in Díaz et al. (1996), though I use the term *coalition* in a broader way. The apparent cross-person CCs are treated as producing coalitions despite the lack of explicit feedback from the antecedent owner, which conflicts with the notion as used by Antaki et al. (1996), in which the crucial determinant is footing consistency. Note also that by *coalition* I do not necessarily mean to suggest that the members of the coalition are necessarily in agreement, merely that they are treated as being coordinated at some level, but not as much as would be expected if they were a full party in Schegloff's terms.



hypotheses 1 and 7), but plenty of evidence for strong pragmatic effects – specifically, the formation of, or appearance of, parties (hypothesis 11), or coalitions and different patterns of results for same- and cross-person CCs (contra hypothesis 6).

## Chapter 7

### DiET chat tool experiment 3: Inducing CCs

---

This chapter examines in more detail the notion of predictability (hypothesis 3) in relation to when CCs can be or usually are produced. The results of experiments 1 and 2 suggest that CCs create the perception, amongst other participants, that a coalition has formed. These coalitions alter the patterns of grounding and turn-taking in ways that are not compatible with the notion of party as a conversational entity essentially equivalent to an individual participant.

The analyses of constituent structure in experiments 1 and 2 also suggest that these effects on the pragmatics of conversational organisation are unconstrained by the low level syntactic organisation of the CC itself. However, these experiments have two important limitations. First, they focus on the impact of apparent CCs on third parties rather than their production. Second, they rely on post-hoc comparisons of the effects of constituent structure instead of manipulating it directly.

As noted in chapter 2, for linguistic and psycholinguistic accounts of CCs the integration of low-level structure, especially syntactic structure, is a key issue. However, the results from the corpus studies in chapter 4 showed that some aspects of predictability had more influence on the production of CCs than others, and that the predictions of hypothesis 3, that syntactic and pragmatic predictability lead to more cross-person CCs, are not so simple. It is therefore important to try to separate the different possible effects of lexical, syntactic, semantic and pragmatic structure to the construction of continuations to some prior contribution.

Different sources of predictability, such as lexical and syntactic predictability, shared knowledge and antecedent completeness and how they might contribute to the conditions for the production of CCs are therefore examined using a further DiET chat tool experiment. The evidence points towards shared knowledge being a key factor with other sources of predictability also contributing. How this fits with the analysis presented so far in this thesis is then discussed.

## 7.1 Method

In this experiment, to see what factors influence people’s likelihood of producing a continuation, a number of genuine single contributions in dyadic text-based conversations were artificially split into two parts. The experiment again used the character-by-character interface outlined in section 6.1.

The first part was transmitted to the other participant as it was typed, with the turn truncated according to various factors as discussed below. Following a pilot study, which showed that people were more likely to supply a response after a filler “...” or “...?” than if there were no filler (after a filler: 18/26, 69%, no filler: 12/45, 27%;  $\chi^2_{(1)} = 12.24, p < 0.001$ ), the truncated first part of the genuine turn was followed by a text filler.<sup>1</sup> Subsequently, there was a large delay (of 12 seconds), during which the other person could respond if they wished. Any response was trapped by the server and not relayed to the original sender, before the rest of the original (interrupted) contribution was transmitted.

Split points are manipulated according to measures of a) syntactic and b) lexical predictability calculated as each turn is produced (see below).

### 7.1.1 Entropy

Entropy is a measure of uncertainty; the higher the entropy, the higher the uncertainty, and the lower the entropy value the higher the predictability. In this case the two types we looked at were part-of-speech entropy (how likely is one part-of-speech (POS) to follow another, e.g. nouns often follow determiners) and lexical entropy (how likely a particular lexical item is to follow a specific POS, e.g. although nouns often follow determiners,

---

<sup>1</sup>Of course there may be additional pragmatic effects associated with these fillers, however I leave this complication to one side for now.

there are lots of different nouns so determiners have a relatively low POS entropy, and a relatively high lexical entropy).

Since predictability depends to a significant extent on dialogue context the entropy values were calculated by taking a corpus of prior balloon task chat tool dialogues (with a total of 53663 tokens) and running the Stanford part-of-speech tagger on them, with a misspellings map for common text abbreviations and typos. For each POS, we then calculated how many possible continuations there were in the corpus for that POS by both the following lexical item and its POS, and computed the entropy based on these values ( $\sum_{i=1}^n p(x_i) \log_2 p(x_i)$ ). This gave us two entropy values for each POS in our list – one for the predictability of the following word, and one for the predictability of the POS of the following word. We discounted those which were e.g. punctuation, or comprised less than 1% of the total corpus (as if they were that rare we couldn't be sure the entropy values were valid) and then assigned each POS into groups based on total means and standard deviations for entropy across the whole corpus. This manipulation produced a range of POS tags with high, medium and low POS entropy, and, independently, high, medium and low lexical entropy (see appendix H).

During the experiment, a POS-tagger analysed the strings in real time and triggered an intervention if there were more than 9 words (this was an arbitrary value based on the mean length of all contributions) and there was a balance between the entropy level types that had already been seen per pair, according to the lexical and syntactic entropy such that truncations occurred at a range of combinations of the two experimental factors.

### 7.1.2 Subjects and materials

The experiment was carried out on 18 dyads. The 36 participants were students at Queen Mary University of London and were paid £7.00 or given course credit (for psychology undergraduates) for providing an hour of their time. The task was once again the balloon task, described in section 5.2 and shown in appendix A.

Two dialogues had to be discarded from the analysis – one due to a software update error which meant that the dyad were able to use the delete key and thus had different conditions to the other experimental subjects, and one because the dialogue was conducted in a language other than English (and thus no interventions were possible).

As with experiments 1 and 2, the subjects were seated at desktop computers in sepa-

rate rooms, asked to input their e-mail address and username and given a sheet of paper with the task description. They were again told that the experiment was investigating the differences in communication when conducted using a text-only interface as opposed to face-to-face, that the experiment would last approximately 45 minutes, and that all turns would be recorded anonymously for later analysis. Once both participants had been logged on, the experimenter went to the server machine and made no further contact with them until at least 40 minutes of dialogue had been carried out.

## 7.2 Analysis

Each intervention was annotated according to a number of factors. Firstly, whether or not there was a response to the intervention during the timeout period (before the second part of the original contribution was relayed). If there had been a response, the type of response was coded according to whether it was constructed as a compound contribution, a clarification request (CR) or a yes/no response.<sup>2</sup> Note that these are not mutually exclusive – as can be seen from examples (7.1)<sup>3</sup> and (7.2) which are CRs constructed as CCs, and example (7.3), which is both a CC and a yes/no answer. The lexical and POS entropy values have been standardised for the purposes of the analyses, such that 0 is the mean, and +/-1 is one standard deviation from the mean, though note that all lexical entropies are higher than POS entropies as might be expected. The minimum POS entropy was 1.44, maximum 4.16, mean 3.27 (standard deviation 0.87); for lexical entropy those values are 5.59–8.14, mean 7.03 (s.d. 0.60) – see appendix H for values associated with each POS tag.

(7.1)            *POS: 0.837; Lexical: 0.106*

**D:** I suppose you need to ask a ...

**H:** ask a what

**D:** nother question. Is it who do we think should be thrown our or

who would get thrown out?

*[DiET CCInd4 559-62]*

---

<sup>2</sup>These response types were chosen on the basis of an examination of the response data.

<sup>3</sup>Responses to the intervention are shown in bold.

(7.2) POS: -1.929 Lexical: -0.366

**B:** also surely the guy who knows how to ...

**N:** fly?

**B:** fly the baloon should know how to inscrease its height?

[DiET CCIInd9 1277-80]

(7.3) POS: 0.523; Lexical: 0.700

**J:** do you assess their value to society ...

**Q:** in milliseconds yes =

**J:** firstim with nick qne wuwi and susie - tom can explain how

toise use the hot air balloon before he jumps [DiET CCIInd13 2048-51]

The intervened turn was also annotated for whether it was potentially end-complete and could therefore be responded to as if it were a complete contribution, as this was found to be a major factor in the corpus studies in which CCs were predominantly expansions. Antecedent end-completeness can therefore be used as a proxy measure for pragmatic completeness, with 40 of the 241 truncated contributions appearing to end in a complete way.

The other major factor found to increase production of CCs in the corpus studies was whether the subject under discussion was known to be shared. Although lexical entropy gives us a measure of the predictability of the local context, a more general measure of contextual salience was also required. Each intervened contribution was therefore classified as either contributing to an ongoing topic of discussion, or introducing a new topic, as a loose measure of common ground. 170 of the contributions were found to be about an existing topic under discussion, with 71 introducing some other topic.

An additional measure for analysis was onset delay – the time between receiving the last character of the antecedent and typing the first character of the response (if there was one). As with experiment 2, we expect the onset delay to be higher if participants have difficulties interpreting the preceding contribution – in this experiment these ought to be cases in which the continuation is unpredictable. Participants should also be more likely to respond quickly if they believe their interlocutor has reached a turn relevance place, i.e. if the antecedent appears end-complete.

Given the predictions of hypothesis 3, we should see a greater proportion of continuation responses where the antecedent is more predictable, lexically, syntactically and pragmatically. The data from the corpus study also suggest that one factor influencing the production of CCs is shared mutual knowledge, so we would expect even more CCs where the topic under discussion was also shared.

### 7.3 Results

Of the 241 interventions, 171 elicited a response (71%). A GEE analysis with whether there was a response to the intervention as dependent variable (using a binary model with a logit link function)<sup>4</sup> with POS and lexical entropy values as covariates, antecedent end-completeness as a fixed factor and participant as subject effect (goodness of fit QIC = 294.562; see table 7.1) showed a main effect of antecedent end-completeness such that responses were more likely in cases that could be considered complete ( $B = -1.078$ , Wald- $\chi^2 = 4.286$ ,  $p = 0.038$ ), showing that people really are sensitive to TRPs.

IV	Model effects		Parameter Estimates		
	Wald $\chi^2$	p	B	Wald $\chi^2$	p
Antecedent end-complete (Ant)	4.286	0.038*	1.078	24.286	0.038*
Lexical entropy (Lex)	0.148	0.700	0.233	4.566	0.033*
POS entropy (POS)	0.593	0.441	0.205	2.549	0.110
Ant $\times$ Lex	3.251	0.071	-0.590	3.251	0.071
Ant $\times$ POS	2.546	0.111	-0.816	2.546	0.111
Lex $\times$ POS	6.460	0.011*	0.273	6.460	0.011*
Lex $\times$ POS $\times$ Ant	0.287	0.592	-0.260	0.287	0.592

Table 7.1: GEE of response or not by lexical entropy, POS entropy and antecedent end-completeness

There was also an interaction effect of POS entropy by lexical entropy ( $B = 0.237$ , Wald- $\chi^2 = 5.893$ ,  $p = 0.015$ ). This effect is illustrated in figure 7.1.

To assess where the significant effects lay, simple slopes analysis – following procedures by Aiken et al. (1991) – was conducted on a simpler model, discounting antecedent end-completeness (see appendix I for details). This shows that for high levels of POS entropy (here set at 1 standard deviation above the mean, and shown by the dashed line in figure 7.1) there is a main effect of lexical entropy ( $B = 0.390$ , Wald- $\chi^2 = 5.840$ ,  $p = 0.016$ ), but there is no such effect at low POS entropy (1 standard deviation below the mean). Likewise, there is an effect of POS entropy at high lexical entropy ( $B = 0.443$ ,

<sup>4</sup>All models in this chapter use an independent correlation structure unless otherwise noted.

Wald- $\chi^2 = 7.446$ ,  $p = 0.006$ ), but not at low lexical entropy. Responses are more likely in cases where both POS and lexical entropy were high (the highly unpredictable cases) than in cases where one or both levels of entropy were low.

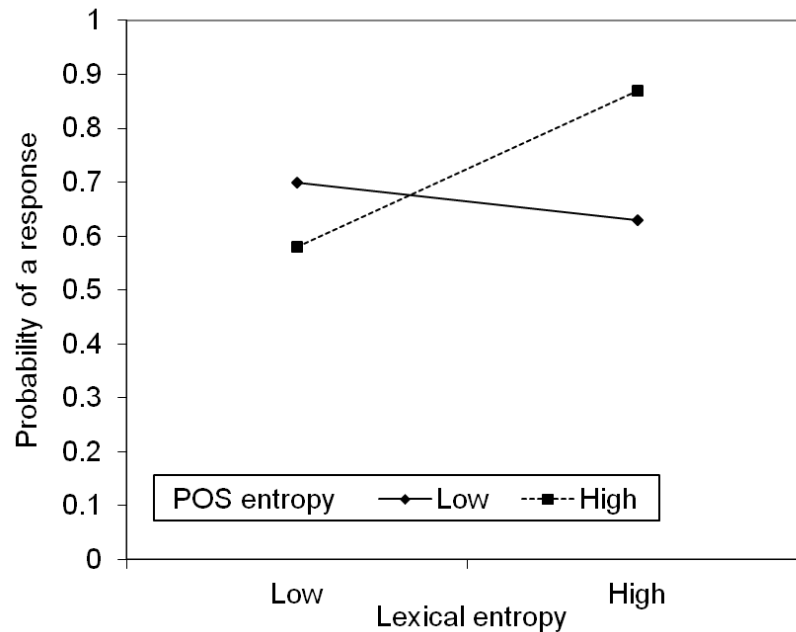


Figure 7.1: Marginal means of probability of a response by POS entropy  $\times$  lexical entropy

### Onset delay

With onset delay as the dependent variable, a GEE model with the same covariates and factors (using the gamma distribution with a log link function, goodness of fit QIC = 26.003) had a number of significant effects (table 7.2).

IV	Model effects		Parameter Estimates		
	Wald $\chi^2$	p	B	Wald $\chi^2$	p
Antecedent end-complete (Ant)	4.900	0.027*	0.132	4.900	0.027*
Lexical entropy (Lex)	0.492	0.483	0.134	6.445	0.011*
POS entropy (POS)	1.137	0.286	-0.120	1.669	0.196
Ant $\times$ Lex	6.302	0.012*	0.138	1.878	0.012*
Ant $\times$ POS	1.878	0.165	0.141	2.726	0.165
Lex $\times$ POS	4.751	0.029*	-0.087	0.548	0.111
Ant $\times$ Lex $\times$ POS	0.356	0.551	0.037	0.356	0.551

Table 7.2: GEEs of onset delay by lexical entropy, POS entropy and antecedent end-completeness

There is an interaction effect of antecedent end-completeness  $\times$  lexical entropy (Wald- $\chi^2 = 6.302$ ,  $p = 0.012$ ). Simple slopes analysis (see appendix I) shows that lexical entropy has an effect where the antecedent could be end complete ( $B = 0.075$ , Wald- $\chi^2 = 6.445$ ,



$p = 0.011$ ). As shown by the solid line in figure 7.2, if the antecedent could be considered end-complete then participants take longer to respond in the lexically uncertain cases. Whether the antecedent is end-complete or not has an effect on onset delay in the low lexical entropy conditions ( $B = 0.480$ ,  $\text{Wald-}\chi^2 = 9.189$ ,  $p = 0.002$ ), i.e. when the next word is highly predictable participants take longer to respond if they have not reached a potential TRP.

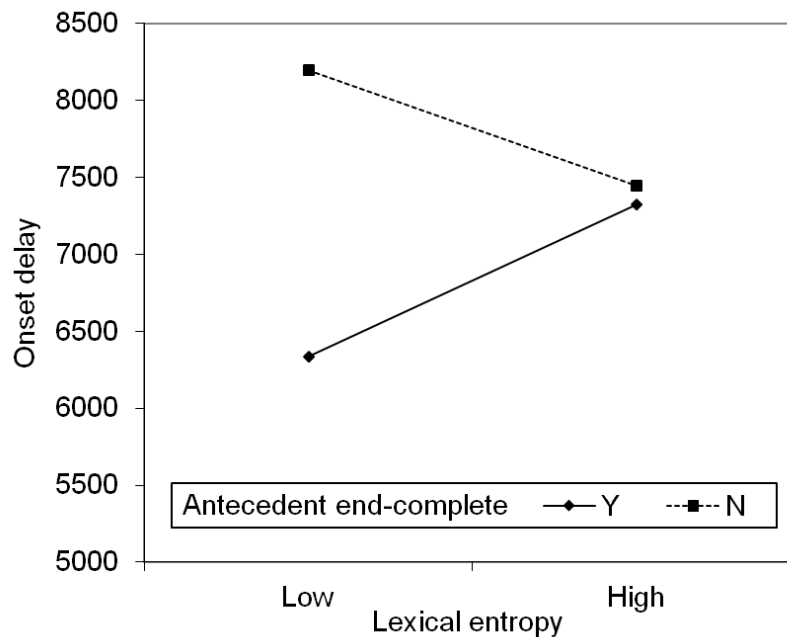
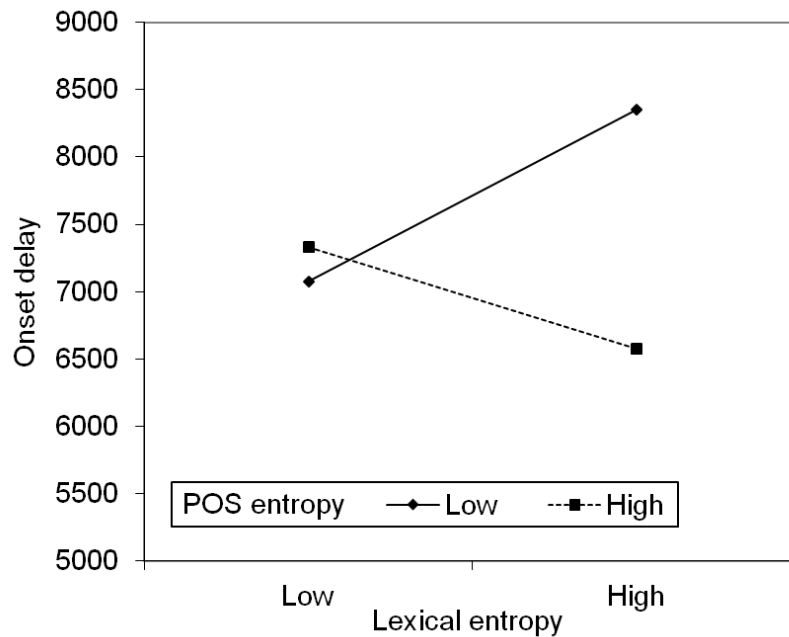


Figure 7.2: Marginal means of onset delay by lexical entropy × antecedent end-completeness

There is also an interaction of POS entropy × lexical entropy ( $\text{Wald-}\chi^2 = 4.751$ ,  $p = 0.029$ )<sup>5</sup> as shown in figure 7.3, such that lexical entropy is significant at low levels of POS entropy ( $B = 0.163$ ,  $\text{Wald-}\chi^2 = 4.720$ ,  $p = 0.030$ ) shown by the solid line in figure 7.3. If the POS of the next word is highly predictable participants respond faster if the lexical item is also highly predictable. POS entropy is also marginally significant at high levels of lexical entropy ( $B = -0.207$ ,  $\text{Wald-}\chi^2 = 3.626$ ,  $p = 0.057$ ); if the next word is unpredictable then participants respond faster if the POS is also unpredictable.

<sup>5</sup>Significance levels taken from model effects; note that the difference in significance between the model effect and the parameter estimates is because they test different things; “a variable effect may be significant while a corresponding parameter coefficient may be non-significant. If there is a difference, hypothesis-testing whether the effect of a variable is significantly different from 0 should use the significance levels reported in the “Test of Model Effects” table.” Source: [http://faculty.chass.ncsu.edu/garson/PA765/gz1m\\_gee.htm](http://faculty.chass.ncsu.edu/garson/PA765/gz1m_gee.htm)

Figure 7.3: Marginal means of onset delay by POS entropy  $\times$  lexical entropy

### 7.3.1 Type of response

The complex results of onset delay and whether there is any response or not outlined above may conflate different effects which are specifically associated with different kinds of response. Analyses were therefore carried out separately on the types of responses, with special attention paid to those which were formulated as a CC.

		Antecedent end-complete					
		N	%	Y	%	Total	
Yes/No	Y	20	15	12	36	32	19
	N	118	85	21	64	139	81
CR	Y	39	28	2	6	41	24
	N	99	72	31	94	130	76
CC	Y	62	45	10	30	72	42
	N	76	55	23	70	99	58
Total		138	69	33	83	171	71

Table 7.3: Response type

The breakdown of the 171 responses depending on whether the antecedent of the intervened contribution appeared to be end-complete at the truncation point by the type of response is shown in table 7.3.

As can be seen, participants were more likely to produce a Yes/No response if the antecedent is potentially end-complete ( $\chi^2_{(1)} = 8.374, p = 0.004$ ), and they are also less likely to respond with a clarification request ( $\chi^2_{(1)} = 7.201, p = 0.007$ ), though there is no difference in the proportion of responses constructed as CCs based on whether the antecedent was end-complete or not, which is unexpected given the preference for expansions over completions in the corpus studies.

### 7.3.2 CR responses

With the data filtered to responses only, GEE analyses<sup>6</sup> on whether or not the response was formulated as a CR, with the POS and lexical entropy values as covariates and participant as subject effect (goodness of fit = 186.828) showed a main effect of POS entropy ( $B = -0.442$ , Wald- $\chi^2 = 5.135$ ,  $p = 0.023$ , see table 7.4 – note that end-completeness could not be included in this model due to the scarcity of the data; only 2 cases where the antecedent appeared end-complete resulted in requests for clarification). Lower POS entropy (greater syntactic predictability) increased the probability of the response being a clarification request. There was no effect of lexical entropy, or interaction between POS entropy and lexical entropy. Additionally, onset delays are higher for CRs than other responses (8561.05 vs. 7211.01;  $t_{169} = 2.712, p = 0.007$ ), reflecting the difficulty in interpretation that precedes the necessity for requesting clarification or the preference for self-repair (in allowing the original speaker more time to continue their own contribution).

IV	Model effects		Parameter Estimates		
	Wald $\chi^2$	p	B	Wald $\chi^2$	p
Lexical entropy	2.207	0.137	-0.198	2.207	0.137
POS entropy	5.135	0.023*	-0.442	5.135	0.023*
Lex $\times$ POS entropy	0.176	0.674	-0.054	0.176	0.674

Table 7.4: GEEs CRs by lexical entropy, POS entropy and antecedent end-completeness

CRs are often formulated as CCs, as in (7.4), (7.5) and (7.6) which is particularly true in the low POS entropy condition (independently of lexical entropy) where the syntactic category of the next word was highly predictable. Of the 72 CCs, 21 occurred in low POS entropy conditions with 12 of these also being CRs. Of the other 51 CCs, only 13 were also CRs (57% vs. 25%;  $\chi^2_{(1)} = 6.575, p = 0.010$ ).

<sup>6</sup>All binary responses used a binary model with a logit link function unless stated otherwise.

(7.4) POS: -0.994; Lexical: 1.856

**Y:** ithink we should go through each ...

**A f each of the ppl?**

**Y:** one [DiET CCInd11 1639-42]

(7.5) POS: 0.834; Lexical: 0.477

**N:** i think susie because she is t ...

**B: a woman?**

**N:** ehe least important out of the three if you think about it ... dr

nick is a doctor and could be really useful in the world

[DiET CCInd9 1214-7]

(7.6) POS: 0.683; Lexical: -0.303

**J:** mine is about the three people in t ...

**W: in... the balloon?**

**J:** he hot air ballono . balloon. doctor, pregnant mom and the air

balloon pilot . which do we save?!

[DiET CCInd16 2736-9]

### 7.3.3 CC responses

GEE analyses on whether or not the response was formulated as a CC, with the POS and lexical entropy values as covariates, participant as subject effect and antecedent end-completeness as a fixed effect (goodness of fit = 234.351) showed an interaction between antecedent end-completeness  $\times$  lexical entropy (parameter estimate;  $B = -1.247$ , Wald- $\chi^2 = 15.835$ ,  $p < 0.001$ , table 7.5).

IV	Model effects		Parameter Estimates		
	Wald $\chi^2$	p	B	Wald $\chi^2$	p
Antecedent end-complete (Ant)	1.951	0.162	-0.767	1.951	0.162
Lexical entropy (Lex)	3.586	0.058	0.293	3.075	0.080
POS entropy (POS)	0.235	0.627	0.073	0.196	0.658
Ant $\times$ Lex	15.835	<0.001**	-1.247	15.835	<0.001**
Ant $\times$ POS	0.018	0.894	0.067	0.018	0.894
Lex $\times$ POS	0.344	0.558	0.129	0.344	0.558
Ant $\times$ Lex $\times$ POS	0.005	0.945	-0.025	0.005	0.945

Table 7.5: GEE of CCs by lexical entropy, POS entropy and antecedent end-completeness

Simple slopes analysis shows that antecedent end-completeness is significant at high lexical entropy ( $B = -2.015$ , Wald- $\chi^2 = 8.728$ ,  $p = 0.003$ ) – if the next lexical item is

unpredictable then you are more likely to produce a CC if the antecedent is not end-complete. There is also an effect of lexical entropy where the antecedent is end-complete (Wald- $\chi^2 = 11.073$ ,  $p = 0.001$ ; the solid line in figure 7.4) – when the antecedent is end-complete, responses are more likely to be continuations in more predictable contexts (e.g. (7.7)), but when it is not end-complete CCs are more likely in the unpredictable cases (e.g. (7.8)).

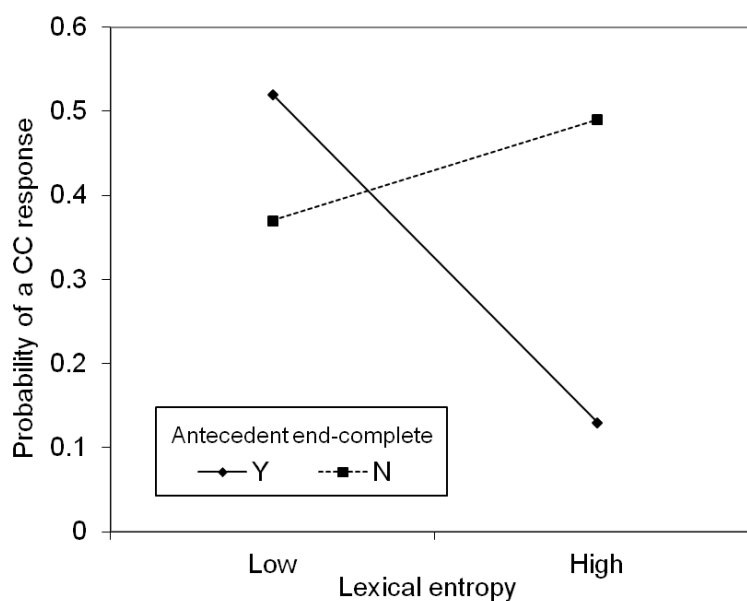


Figure 7.4: Marginal means of probability of a CC response by lexical entropy × antecedent end-completeness

(7.7) POS: 0.641; Lexical: -1.158

**W:** I feel like we should be talking ...?

**J:** about the prompt?

**W:** about something important. [DiET CCInd16 2846-9]

(7.8) POS: 0.882 High; Lexical: 0.394

**W:** nope we are not god we are ...?

**M:** [M] and [W] ini lol we are [M] and [W] u fool lol so s

**just shut up npw please ad thank u for ur c kindness**

**W:** not making dis di decision i knw we got bre spellintg werrorz

man i r we even aloowed to talk type in slang? [DiET CCInd6 929-32]

### 7.3.4 Context

As the corpus results suggested that CCs are more common where participants share information or common ground about the subject under discussion, and local context (indexed by lexical entropy) plays an important role, planned post hoc analyses were carried out using the topic under discussion (as a proxy measure of common ground). Of the 241 intervened contributions, 170 were about an existing topic under discussion, whilst 71 introduced some new topic.

Participants were no more likely to respond if the turn was about the current topic or not; nor were they more likely to respond with a yes/no answer, or a clarification request. However, they were more likely to construct their response as a CC if it was about the current topic than if it was about something else (topic 59/121, 49% vs. Off-topic 13/50, 26%;  $\chi^2_{(1)} = 7.519, p = 0.006$ ).

Adding topic to the GEE model with CC response as dependent variable (QIC = 227.895, table 7.6)<sup>7</sup> resulted in an additional three-way interaction effect of lexical entropy  $\times$  POS entropy  $\times$  topic (Wald- $\chi^2 = 8.63, p = 0.003$ ), with the original effect of lexical entropy  $\times$  end-completeness now marginal (Wald- $\chi^2 = 3.435, p = 0.064$ ).

IV	Model effects		Parameter Estimates		
	Wald $\chi^2$	p	B	Wald $\chi^2$	p
CC response; QIC = 227.895; QICC = 236.728					
Antecedent end-complete (Ant)	0.046	0.830	0.874	1.008	0.315
Topic	0.276	0.600	0.101	0.002	0.965
Lexical entropy (Lex)	2.545	0.111	-0.693	1.915	0.166
POS entropy (POS)	0.018	0.892	-0.637	0.247	0.619
Line number	2.361	0.124	0.004	2.361	0.124
Ant $\times$ Topic	0.381	0.537	-1.422	0.381	0.537
Ant $\times$ Lex	3.435	0.064	1.077	4.890	0.027*
Ant $\times$ POS	0.183	0.669	0.517	0.151	0.697
Topic $\times$ Lex	2.103	0.147	-0.537	0.582	0.446
Topic $\times$ POS	0.281	0.596	0.980	0.083	0.773
Lex $\times$ POS	0.034	0.853	-0.466	0.190	0.663
Ant $\times$ Topic $\times$ Lex	0.091	0.763	-0.218	0.091	0.763
Ant $\times$ Topic $\times$ POS	0.005	0.946	-0.244	0.005	0.946
Ant $\times$ Lex $\times$ POS	0.133	0.716	0.382	0.133	0.716
Topic $\times$ Lex $\times$ POS	8.635	0.003**	0.751	8.635	0.003**

Table 7.6: GEE of type of CC responses by lexical entropy, POS entropy, antecedent end-completeness and topic

<sup>7</sup>The model also included line number as an additional covariate as it was found that participants were more likely to introduce a new topic later on in the conversation; 110.8 vs 86.2,  $t_{239} = 2.75, p = 0.006$ , but this did not affect the results as reported. Note also that a model with the dependent variable as only those CCs which were not CRs gave the same pattern of results as shown in table 7.6.

Exploring the interaction effect of lexical entropy  $\times$  POS entropy  $\times$  topic using a simplified model (discounting antecedent end-completeness, see appendix I) shows a significant interaction of topic  $\times$  POS entropy in the high lexical entropy cases (Wald- $\chi^2 = 6.462$ ,  $p = 0.011$ ). There was a significant interaction of topic  $\times$  lexical entropy in the low POS entropy conditions (Wald- $\chi^2 = 10.794$ ,  $p < 0.001$ ). Drilling down further, there is a significant effect of topic at high lexical and low POS entropy (Wald- $\chi^2 = 8.838$ ,  $p = 0.003$ ; Wald- $\chi^2 = 7.456$ ,  $p = 0.006$ ) and of lexical entropy at low levels of POS entropy (Wald- $\chi^2 = 4.054$ ,  $p = 0.044$ ). These effects can be seen in figure 7.5.

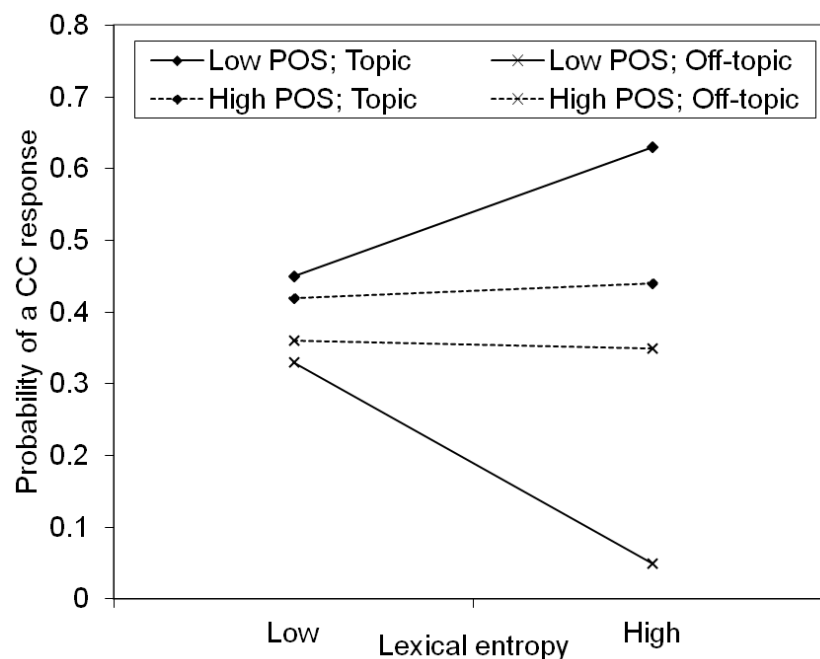


Figure 7.5: Marginal means of probability of a CC response by lexical entropy  $\times$  POS entropy  $\times$  topic

In other words, if the next lexical item were highly predictable, participants were equally likely to construct their response as a CC regardless of the topic of the conversation and POS entropy. Additionally, if both the next lexical item and POS were unpredictable, they were also equally likely to construct their response as a CC regardless of the topic.

However, in lexically unpredictable cases, which were syntactically predictable, then they were more likely to construct their response as a CC if they were talking about some topic which they had already been discussing, and which was therefore contextually salient.

## 7.4 Discussion

These results offer some insights regarding the conditions influencing whether and how conversational partners respond to an incomplete utterance, and when they can and do construct those responses as continuations.

There is a response to 71% of the interventions, with this proportion affected by the predictability of the upcoming material. Perhaps counterintuitively, people are more likely to respond to unfinished contributions and also responded faster when they did so (in text chat)<sup>8</sup> if both syntactic and lexical items were unpredictable. This is not what we would expect if a simple model of levels of predictability were correct, as intuitively the most predictable cases ought to elicit the most responses. However, it is what we would expect if one of the drivers of human communication is in locally managing and resolving potential sources of misunderstanding (as in the interactive misalignment of Healey, 2008).

The main effect of potential completeness also demonstrates that people are more comfortable responding at all, and quicker doing so if the other person has reached a potential TRP – backing up the findings from the corpus studies and Schegloff’s assertion that people are sensitive to possible endings.

The interaction between lexical entropy and antecedent end-completeness on onset delay show that if the local context is predictable (low lexical entropy) and the other person appears to have finished a turn then participants simply take the floor. Conversely, if the local topic is predictable but the other participant does not seem to have finished, they allow longer for possible completion. This is equivalent to the larger gaps allowed for possible ‘transition space’ repairs (Schegloff et al., 1977), and may be indicative of the preference for self-repair in general, showing a reluctance to intrude on another’s incomplete turn space.

However, while interesting, these general effects do not tell us about the particular conditions under which participants construct their responses as continuations as opposed to any other type of response. Indeed, they may serve to disguise the fact that different types of responses are produced in different circumstances, which does appear to be the case.

---

<sup>8</sup>This could be a genuine difference because there may be other cues in spoken dialogue, but I leave a discussion of this to one side.



### 7.4.1 Clarification requests

Participants are extremely unlikely to produce a CR response in the cases where the truncated contribution appears to be a complete turn. Where it does not (i.e. there is something obviously missing), they are more likely to produce a CR if the type of word that is coming next is highly predictable. These are also likely to be syntactically formulated as a CC (by e.g. repeating fragments of the incomplete antecedent to localise the trouble source). Having a predictable POS means that what follows is more syntactically constrained – in other words, these are precisely the cases where syntax may be exploited to formulate a clarification request. Examples can be seen in (7.2), (7.4), (7.9), (7.10).

(7.9) *POS: -0.994; Lexical: 1.856*

**M:** the doctor nd susie r having an ... ?

**I:** wat

**having wat**

**M:** affair, so tom shud jump coz

[*DiET CCInd1 43-47*]

(7.10) *POS: -2.096; Lexical: -1.249*

**Lo:** i kno can u imagin if they locked us in eer well the thing is i

can ...

**La:** i can wot?

**Lo:** see a pack of buiscits on the flood someone left behind that

wud eb my meal i guess loooooool

[*DiET CCInd3 523-6*]

That onset delays are greater prior to CR responses than other responses also illustrates the preference people have for allowing others more time to make their own self-repairs prior to initiating a repair sequence.

### 7.4.2 Compound contributions

What is critical to whether people construct a response as a continuation part of a CC, thus tying it directly to the incomplete contribution they have thus far seen, seems to be the actual and presumed accessibility of common ground. If the local content of what comes next is salient from the (presumed shared) context then people will produce com-

pletions. They do this by taking advantage of the syntactic structure of the antecedent, but syntactic predictability alone is not sufficient to prompt a completion.

A continuation response is more likely if the antecedent is complete but the local content is predictable (in terms of lexical predictability) or if the antecedent is incomplete, suggesting that people complete where they can.

Examples (7.7), (7.11) show that end-complete, lexically predictable cases may be ones with looser ties to the preceding material such as the common expansion type CCs seen in the various corpora with e.g. an adjunct or extending clause.

(7.11) *POS: 0.154; Lexical: -2.404*

**M:** you can't throw a pregnant woman out ...?

**K:** no but people wont like the dr dying either

but then again he has his research people to cure

cancer too

hmmm

yeah pregnant ladies need \*need t be protected

**M:** of a balloon like

[DiET CCInd14 2326-32]

For the cases in which the antecedent is not end-complete, responses were, perhaps surprisingly, more likely to be constructed as CCs in lexically unpredictable cases (i.e. where local content is not taken to be shared). However, if the next lexical item is highly predictable, then it can be interpreted as if it had actually been produced, and responded to on that basis, as in (7.12), (7.13), (7.14), (7.15).

(7.12) *POS: 0.642; Lexical: -1.158*

**O:** im not typing slowly, i think the screen has massive lag, yours is taking ...?

**T:** yer to be fair it is quite slow going accross the screen.

to be fair we probably both type fast ... all thast time

spemnt chatting on msn rather than doing rrevision for

that exam shiz

**O:** ages to come up on mine

[DiET CCInd8 1174-7]

(7.13) POS: 0.642; Lexical: -1.158

**A:** I cant belive they did not have any saftley kik like\* flying ...?

**P:** they prob threw it out

**A:** things - that last sentance just shows how much “flying esp

exp\* i ahve! haha [DiET CCInd15 2697-700]

(7.14) POS: 0.206; Lexical: -0.645

**W:** but hes gna find a cure for ...

**M:** but without tom no one cannnnnnnnn fly the ballon and

as tom and susie are husband n wife and she is pregant

and about to start a family it makes more snse if nick

dies cos hes just thier freind u geh wa me saying

**W:** cancer which is gna hek help millions of pplo [DiET CCInd6 856-9]

(7.15) POS: 0.642; Lexical: -1.158

**T:** its not that fair on the girl doing th ...

**H:** exactly, you need to think of others and not be so

selfish :P

**T:** study we should do lots of chatting although i doubt she'll read

past the exercise what with it not being standardised etc

[DiET CCInd4 685-8]

This result is not perhaps as surprising as it first appears when we recall that in the BNC corpus study, only 64% of end-incomplete contributions get continued, meaning that 36% never do. These are cases in which the local context is so predictable that it can be taken to be shared without the words themselves being produced.

Additionally, the unpredictable cases might be considered to be more open ended, meaning that interlocutors may use the proffered syntax to move the conversation along, but without considering how their conversational partner might have himself completed the contribution. Examples can be seen in (7.16), (7.17), (7.18) – notice that in (7.16) the proffered continuation is semantically equivalent to that of the actual continuation, whilst in (7.17) it is quite the opposite.

(7.16) *POS: 1.032; Lexical: 1.394*

**F:** if this cure is so great surely ...?

**S:** **there musy muis lol must be other people working with  
him**

**who know the work**

**F:** he must of dicsciussed it with otehr scientists so they can nst  
still use the cure to treat people? *[DiET CCInd17 2929-33]*

(7.17) *POS: 1.032; Lexical: 1.394*

**J:** aslo susies a primary tra teacher so ...?

**Q:** **she might go on to molest slash abduct slash....**

**J:** presumably that makes her a good 'un m not that there was  
any real doubt about our decision over her on i wonder if  
anyone else chucked her in order to garuantree the cure for  
cancer arrivin safely on the ground *[DiET CCInd13 2158-61]*

(7.18) *POS: 0.834; Lexical: 0.477*

**J:** suise can many marry nickthis typo thing is ...?

**Q:** **very**

**J:** embraassing - i cant even type embarrassing *[DiET CCInd13 2055-8]*

### 7.4.3 Context

The three-way interaction of POS entropy by lexical entropy by topic adds weight to the notion that what is critical is the actual and presumed accessibility of common ground.

If the local context (next lexical item) is predictable, then you are equally likely to produce a continuation whether or not you are talking about something which you have already been discussing, and the predictability or otherwise of the syntax doesn't help or hinder production of a CC. However, if the local context is unpredictable then syntactic predictability aids production of CCs in cases where the topic of the truncated contribution is shared, thus acting as a resource which helps frame the offered continuation as such. Syntax does not however help at all in cases where the topic is new so the gist of

the contribution cannot be predicted and the local context also offers no clues as to a plausible continuation.

The interesting cases are therefore those in which the word is unpredictable but the syntax is predictable, as it is then that whether or not you are talking about something that is contextually salient makes a difference. See e.g. (7.4), (7.19), (7.20).

(7.19) *POS:-1.929; Lexical: -0.366*

**D:** well obviously he might be able to ...?

**H:** **help Susie if she goes into labour or something like that. or if one of them gets hurt he will know what to do he is a good friend so she will trust him and feel more comfortable so the situation wil seem a bit less?**

**D:** save people from cancer. then again I believe a lot of things but often turn out to be mistaken, so it could be a waste of time if a loving marriage is cut short just because dr nick is over confident!

[DiET CCInd3 580-5]

(7.20) *POS:-1.929; Lexical: -0.366*

**Lo:** og my goodness and the prize goes to ...

**La:** meeeeeeeeeeeeeeeeeeeee...loooooool

**Lo:** [La] lol

[DiET CCInd3 504-7]

This pattern of predictability corresponds to cases in which the high lexical entropy equates to lots of different words of a single type, as in the determiner case, rather than the high lexical entropy being associated with lots of different words of many different types (as with e.g. adverbs). This means that the syntactic category is highly constrained and the additional information associated with contextual salience can significantly narrow down an appropriate continuation.

However, it is important to note that although CC responses are more or less likely in these specific circumstances, they occur at every level of syntactic and lexical predictability, as the examples throughout this chapter attest.

## 7.5 Summary

This experiment, to the best of my knowledge the first to ever systematically attempt to induce continuations in an ongoing dialogue, shows that different types of predictability have different effects on what type of response participants produce to incomplete contributions, if any.

As with the earlier studies, it shows that although syntax can be mobilised in constructing a response, it is not the crucial determinant of whether people respond or not, or even whether they construct their responses as continuations to the immediately preceding contribution. Participants are able to produce a continuation in a variety of differentially predictable cases from different syntactic starting points, and make use of syntactic predictability only if the context is sufficiently constrained. Though people respect the constraints of the syntax (so they don't produce any old word), different points in the sentence do not cause greater difficulty in producing something that syntactically builds of what has been offered than others (as predicted in hypothesis 1). However, that the grammar is a mutually available resource does not mean that it is used in the same way by all interlocutors, and indeed the different distributions between same- and cross-person CCs (contra hypothesis 6) have made it clear that it is not. Further evidence for this notion of grammar as a resource is the finding that clarification requests are more likely, and more likely to be formulated as continuations, when the syntactic category of the upcoming material is more constrained, as these are cases where the syntax may be exploited to localise the source of a potential misunderstanding.

Another of the main findings, as in the earlier corpus studies, is that people are sensitive to potential turn endings. These may be syntactic (in the antecedent end-complete cases) but they are not necessarily so. Some cases which appear to be syntactically incomplete can be responded to as if they are complete, provided that the continuation is predictable (either lexically or from the context). If there are indeed cases which are interpreted as complete when they are not – as if the hearer is supplying the missing material internally, but does not necessarily produce it,<sup>9</sup> this has implications for any grammatical model. Incomplete syntactic strings must be not only successfully analysed, but also assigned potentially complete semantic representations.

---

<sup>9</sup>Note that this matches observations of French incomplete utterances (Chevalier and Clift, 2008).

The evidence from this experiment, coupled with the results from the tuition corpus reported in section 4.4.2 show that when people are likely to produce CCs (or produce more CCs) is principally driven by common ground. They are possible (or more likely) when it is shared, and other participants recognise and orient to this fact, as shown in experiments 1 and 2. Of course, whether or not the subject of an utterance is already being talked about or is newly introduced is a loose measure of context, and might be better thought of as a measure of something like the ‘question under discussion’ (QUD) of Ginzburg (in press). However this relies entirely on linguistic context, whilst the notion of shared knowledge (as in the tuition corpus) is a broader one which can include shared perceptual experience (perhaps making it more akin to both common ground and the situation model of Pickering and Garrod (2004)). It is not clear how these notions might be (or whether they should be) separated. Nonetheless, it is apparent that some formal notion of context is crucial for a thorough understanding of CCs, especially if we are to ever hope to model them appropriately in a dialogue system (though of course the restricted domain for most dialogue systems means this is not an insurmountable task).

With respect to parties, if shared knowledge is key to when you are more likely to produce a continuation, then interpreting CCs as indicating a coalition (though, as previously discussed, not necessarily a party for turn-taking purposes) makes sense. Although it is not possible to ascertain cause and effect, if there is a statistical correlation between sharing knowledge with someone and syntactically tying your contribution to a prior one of theirs, then the inference from this to assuming that other people do likewise and are therefore coordinated at a higher level when they continue each other’s utterances is a natural one, and could be a useful heuristic in helping us delimit the world into social groups. Notice that although we cannot say anything about the directionality of this here, there are implications for how we learn associations in language, and the added implication that common inferences (though defeasible, as when the speaker of the antecedent refutes the proffered continuation thus negating the ‘coalition’ reading) can also be learned. How both the syntactic facts about CCs, and such inferences that arise from them may be formalised will be discussed in the following chapter.

## Chapter 8

# A Dynamic Syntax Account of Compound

## Contributions

---

All the data reported in this thesis show that CCs can not be explained by reference to only syntactic or only pragmatic factors, but that an interplay between interactional factors and the constraints provided by the grammar is necessary to explain and model the practice/phenomenon. This chapter attempts to formalise the empirical findings reported earlier in the thesis, using the framework of Dynamic Syntax. Extensions to and limitations of the framework for capturing all the facets of the phenomena will be discussed, with a comparison to the treatment of CCs in PTT described in Poesio and Rieser (2010).

Dynamic Syntax has already been shown to be an appropriate grammar formalism for the syntactic analysis of cross-person completion CCs, due to its inherent incrementality, and tight coordination between parsing and generation (Purver et al., 2006; Gregoromichelaki et al., 2011), able to cope with the switch from speaker to hearer between any words in the string, consistent with our lack of empirical evidence for strictly syntactic constraints on CCs. It is also a formalism in which interpretations may rely on prior context (either linguistic or non-linguistic), and this does appear to be the case with CCs, which are more likely when interlocutors share knowledge or common ground. I will outline DS and show how expansions can also be easily modelled. Recent extensions to DS (Purver et al., 2010) introduce the notion of optional inference rules, which can



model pragmatic effects. In this chapter I will explore how this machinery can be used to model the different interpretations available to interlocutors following receipt of a CC, including coalitions, or, more strongly, parties, which this thesis emphasises. My aim however is more modest than that of Poesio and Rieser (2010), as I seek only to outline the mechanisms which give rise to the ability of participants to produce and interpret CCs, and to try to see how much of the phenomena can be accounted for without reference to planning, or specific dialogue situations. Of course these may be necessary in order to establish why a particular continuation is offered on any particular occasion, and as such DS may need to have a suitable interface with an appropriate dialogue model (e.g. KoS, Ginzburg, in press), but that is not my concern here.

## 8.1 Introduction

Dynamic Syntax (DS, Cann et al., 2005) is a grammar formalism which is based on the idea of monotonic tree growth. The theoretical notions upon which it depends are outlined below, with the formal tools laid out in more detail in section 8.2.

Briefly, the foundations of DS are based in the recognition of the fact that what are usually considered independent features of language; syntax, semantics and pragmatics, are in fact mutually dependent features of human communication. Parsing and processing are taken to be two sides of the same coin, meaning it is well placed to analyse dialogue phenomena such as CCs.

Importantly, words are analysed in the order in which they occur in a string, thus taking *how* an interpretation is built up to have a central role. Complete trees in DS have no representation of word order, and are thus semantic trees. DS rejects the notion that a separate descriptive level is required for syntax, postulating instead that phenomena usually described as syntactic can be explained and described by the dynamics of the growth of a (potentially partial) parse tree. DS also seeks to incorporate notions of context into the theory, thus formalising ideas that are usually consigned to the “pragmatic wastebasket” (Bar-Hillel, 1971).

## 8.2 Parsing

In DS, trees are built up word-by-word incrementally via a combination of generally applicable computational rules and specific lexical actions. Any string encountered is grammatical just in case there is a sequence of steps which leads to a completed (requirement free) tree when all the words have been parsed.

Nodes are terms in the typed lambda calculus and are annotated with decorations giving information about the current tree node, for example, what type of node it is (e.g. propositions  $Ty(t)$  and entities  $Ty(e)$ ). Complex types are built up from the basic types, for example, intransitive verb phrases are functions from entities to truth values, i.e.  $Ty(e \rightarrow t)$  and so on. Other decorations are a formula value (the semantic content) of the form  $Fo(John')$ , which, by the rules of the grammar can be combined to form complex expressions such as  $Fo(Love'(Mary')(John'))$ , and the tree node address. This is stated, using the logic of finite trees (LOFT Blackburn and Meyer-Viol, 1994), either in relative terms, based on the root node of the tree under construction being  $Tn(0)$ , and each daughter node being assigned an additional one for a functor daughter or zero for an argument daughter, or by its relation to any other tree node. This uses two modal operators to signal daughters or mothers of a node ( $\langle \downarrow \rangle$ ,  $\langle \uparrow \rangle$ ). Additional subscripts 0 or 1 indicate whether this refers to an argument or functor daughter/mother ( $\langle \downarrow_0 \rangle$ ,  $\langle \downarrow_1 \rangle$ ,  $\langle \uparrow_0 \rangle$  and  $\langle \uparrow_1 \rangle$  respectively), and can be underspecified using the Kleene star as a subscript.  $\langle \uparrow_* \rangle Tn(0)$ , for example, means that the root node is somewhere above (or at) the current node.

All these labels can show type and formula information about what has already been parsed, or be *requirements*, indicating what else is required to complete the current partial tree. Unlike complete descriptive decorations (which are facts, hence persistent), requirements are preceded by a question mark;  $?Ty(t)$  is a requirement for a proposition,  $? \exists \mathbf{x}. Tn(\mathbf{x})$  is a requirement for a fixed tree node address. In addition, there is a pointer in the tree, depicted by the  $\diamond$  symbol, which indicates the node currently under construction. This is an important feature of DS, as it immediately explains the ungrammaticality of certain strings by setting a restriction on what rules or lexical actions can occur at any given point in the parse.

### 8.2.1 Lexical actions

Like Head-driven Phrase Structure Grammar (HPSG, see Sag et al., 2003) and Combinatory Categorical Grammar (CCG, see e.g. Steedman, 2000), DS is a lexicalised grammar, acknowledging the fact that complexity in language relies to a large extent on information that is stored in the lexicon. In the case of DS, what is stored in the lexicon is a set of procedures, known as lexical actions. This contrasts with HPSG and CCG which store collections of feature structures and category specifications respectively.

The lexical action for ‘John’, shown in 8.1, will be accessed when the word *John* is encountered in a string, and states “if there is a requirement for  $Ty(e)$  at the current node (determined by the position of the pointer  $\diamond$ ), then put type  $Ty(e)$ , and formula  $Fo(John')$  at the node, otherwise abort the parsing process.”<sup>1</sup>

(8.1)	<i>John</i>	<table style="border-collapse: collapse; margin-left: 1em;"> <tr> <td style="padding-right: 1em;">IF</td> <td><math>?Ty(e)</math></td> </tr> <tr> <td style="padding-right: 1em;">THEN</td> <td><math>put(Ty(e), Fo(John'), [\downarrow]\perp)</math></td> </tr> <tr> <td style="padding-right: 1em;">ELSE</td> <td>ABORT</td> </tr> </table>	IF	$?Ty(e)$	THEN	$put(Ty(e), Fo(John'), [\downarrow]\perp)$	ELSE	ABORT
IF	$?Ty(e)$							
THEN	$put(Ty(e), Fo(John'), [\downarrow]\perp)$							
ELSE	ABORT							

Lexical actions can also account for some aspects of syntactic predictability. For example, in the lexical entry for the transitive verb *loves*, nodes for the verb and its object are created with the pointer left at the object node with the expectation that the next lexical item will be of  $Ty(e)$ .

(8.2)	<i>loves</i>	<table style="border-collapse: collapse; margin-left: 1em;"> <tr> <td style="padding-right: 1em;">IF</td> <td><math>?Ty(e \rightarrow t)</math></td> </tr> <tr> <td style="padding-right: 1em;">THEN</td> <td><math>go(\langle \uparrow_1 \rangle); go(\langle \downarrow_1 \rangle); make(\langle \downarrow_1 \rangle); go(\langle \downarrow_1 \rangle);</math>  <math>put(Ty(e \rightarrow (e \rightarrow t)), Fo(Love'), [\downarrow]\perp);</math>  <math>go(\langle \uparrow_1 \rangle); make(\langle \downarrow_0 \rangle); go(\langle \downarrow_0 \rangle);</math>  <math>put(?Ty(e))</math></td> </tr> <tr> <td style="padding-right: 1em;">ELSE</td> <td>ABORT</td> </tr> </table>	IF	$?Ty(e \rightarrow t)$	THEN	$go(\langle \uparrow_1 \rangle); go(\langle \downarrow_1 \rangle); make(\langle \downarrow_1 \rangle); go(\langle \downarrow_1 \rangle);$ $put(Ty(e \rightarrow (e \rightarrow t)), Fo(Love'), [\downarrow]\perp);$ $go(\langle \uparrow_1 \rangle); make(\langle \downarrow_0 \rangle); go(\langle \downarrow_0 \rangle);$ $put(?Ty(e))$	ELSE	ABORT
IF	$?Ty(e \rightarrow t)$							
THEN	$go(\langle \uparrow_1 \rangle); go(\langle \downarrow_1 \rangle); make(\langle \downarrow_1 \rangle); go(\langle \downarrow_1 \rangle);$ $put(Ty(e \rightarrow (e \rightarrow t)), Fo(Love'), [\downarrow]\perp);$ $go(\langle \uparrow_1 \rangle); make(\langle \downarrow_0 \rangle); go(\langle \downarrow_0 \rangle);$ $put(?Ty(e))$							
ELSE	ABORT							

### 8.2.2 Computational rules

Computational rules can apply at any point in the parse (DS does not give us a strategy for choosing a rule at any given point, although of course there are restrictions on when they can apply) and can be expressed in the same way as lexical actions.<sup>2</sup>

<sup>1</sup>The last decoration in this lexical entry  $[\downarrow]\perp$ , is the bottom restriction, which means that the node cannot be developed further (i.e. it may not have daughter nodes).

<sup>2</sup>The computational rules used in this chapter are shown in appendix J.

The AXIOM states that we begin a parse with a single node consisting of a requirement for a proposition, and with the pointer at that node. Using INTRODUCTION and PREDICTION effectively creates a blank tree waiting for a subject and predicate,<sup>3</sup> as shown in figures 8.1 and 8.2.

$$?Ty(t), ?\langle \downarrow_0 \rangle Ty(e), ?\langle \downarrow_1 \rangle Ty(e \rightarrow t)$$

Figure 8.1: Single node tree following INTRODUCTION

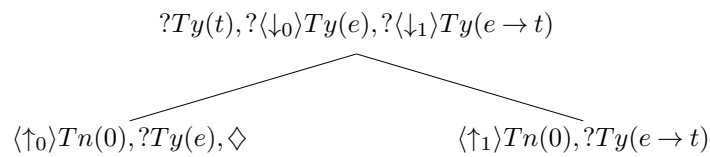
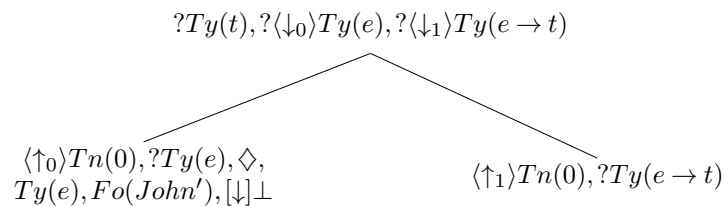


Figure 8.2: Tree following PREDICTION

Given a simple example *John loves Mary* using the lexical entries shown in (8.1) and (8.2), the next step would be to parse the word *John*, as per the lexical actions discussed above, leaving the resulting tree, as in figure 8.3.

Figure 8.3: Parsing *John*

Further rules allow us to update the tree such that the pointer is at the node with a requirement  $?Ty(e \rightarrow t)$ , where we can parse the word *loves*. They are THINNING, which removes a requirement from a node if the completed form is also present, COMPLETION, which moves the pointer to a mother node if no requirements are outstanding at the current node, and ANTICIPATION, which moves the pointer to any daughter node with outstanding requirements.

The application of the lexical actions in 8.2 would result in the tree shown in figure 8.4.

This leaves the pointer at the object node, where the trigger condition for *Mary* ( $?Ty(e)$ ) is met, and the tree can then be completed using THINNING, COMPLETION and

<sup>3</sup>Note that these computational rules are language specific, and are appropriate for English as it has subject-verb-object (SVO) word order. As discussed in Cann et al. (2005), these rules may not generalise to free word order and prodrop languages (e.g. Japanese, Spanish), though this will not be a concern for the current treatment.



### 8.2.3 Underspecification

Central to DS is the notion of underspecification. At every non-final point in the parse, the partial tree may be underspecified, with each type of tree decoration ( $Fo()$ ,  $Ty()$ ,  $Tn()$ ) a potential source of underspecification.

#### Unfixed nodes

Unfixed nodes are nodes whose position ( $Tn()$ ) in a tree is initially underspecified, with a requirement to be fixed at a later point in the parse. This means that a parse tree can unfold with certain elements not yet in the positions they will occupy in the final tree, without having to resort to the notion of movement. A canonical example of this is left topic dislocation, as in *Mary, John loves*, in which, although it is the first item encountered in the string, *Mary* is the object of *loves*, not the subject. In a transformational account, it is assumed that it is moved from its usual object position, as a focus effect, but in DS a parse may proceed using the weak structural relation rule of \*ADJUNCTION (figure 8.6) from where the trigger requirement for parsing *Mary* is met, as in figure 8.7.

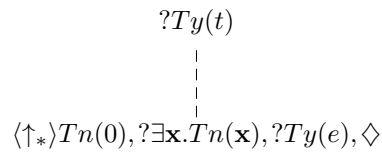


Figure 8.6: Using \*ADJUNCTION

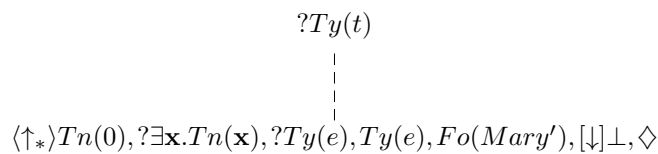
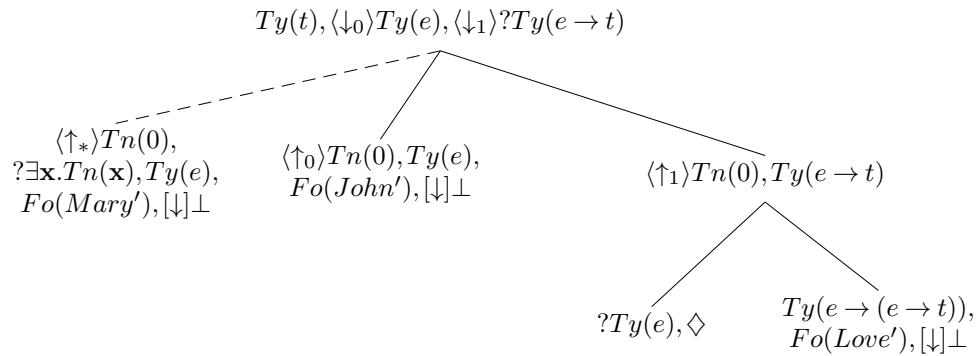


Figure 8.7: Parsing *Mary, . . .*, after \*ADJUNCTION

THINNING and COMPLETION leave the pointer back at the  $?Ty(t)$  node, from where *John* and *loves* can be parsed as before, with the results shown in figure 8.8.

However, the parse is not yet complete, as the pointer is at a node with an outstanding requirement ( $?Ty(e)$ ) and there is a requirement for a fixed position in the tree on the unfixed node. These can both be satisfactorily resolved by merging the two nodes (using MERGE), resulting in the same tree shown in figure 8.5, with the only difference being in the steps used to get there.

Figure 8.8: Parsing *Mary, John loves ...*

### Metavariables

Underspecification of content ( $Fo()$ ) is achieved in DS through metavariables, which formalise one way in which we use context to interpret strings in, for example, pronouns, anaphora and ellipsis.

In DS, a sentence such as *He loves cakes* would be ungrammatical if there were no contextual indication of who to interpret *he* as. This underspecification is important in that it assumes that, as processors, we constantly update our interpretations of utterances based on what we know about the world, previous discourse, or other perceptual indicators (e.g. pointing). Parsing a string with a pronoun in it involves a pragmatic process of SUBSTITUTION; for example, if the string *He loves cakes* follows *John ate all the meringues*, we would be able to substitute the formula value  $Fo(John')$  for *he* in the second string, resulting in the parse leading to the complete formula  $Fo(Love'(Cake')(John'))$ .

The lexical entry for *he* (see 8.4) contains a formula value that is underspecified; this is in the form of a *metavariable*,  $Fo(\mathbf{U}_{Male'})$ , with a requirement for a fixed formula value ( $?\exists \mathbf{x}. Fo(\mathbf{x})$ ), which must be filled for the parse to be complete.<sup>4</sup>

$$(8.4) \quad he \left| \begin{array}{ll} \text{IF} & ?Ty(e) \\ \text{THEN} & \text{put}(Ty(e), Fo(\mathbf{U}_{Male'})), \\ & ?\exists \mathbf{x}. Fo(\mathbf{x}), [\downarrow] \perp \\ \text{ELSE} & \text{ABORT} \end{array} \right.$$

Pronouns can thus be seen as place-holders for some other information to be assigned from context. First (and second) person pronouns would have similar lexical entries,

<sup>4</sup>Note that this is a simplified version of the lexical entry for *he*, which would also require a case condition to prevent strings such as *Mary liked he* being licensed.

except that the metavariable is further restricted as to what value it can take according to who is currently speaker (hearer), i.e.  $Fo(\mathbf{U}_{Spkr'})$ , which enables DS to account easily for speaker switches across CCs which involve a switch of pronoun.

### Expletive *it*

In English, even where the semantic subject of a sentence has been extraposed to the end of the sentence, there needs to be a syntactic subject, as in example (8.5), below, which can be paraphrased by (8.6). Because of this, we can see that *it* in this example is not referential in the same way that the pronoun *he* is, taking its value from context (as seen above in section 8.2.3), rather it takes its value from what follows the verb.

(8.5) It bothers Louise that John loves Mary.

(8.6) That John loves Mary bothers Louise.

This is known as *expletive it*, and leads some theories to accept that there is an unavoidable disjunct between syntax and semantics (as in HPSG, see Sag et al. (2003)). In Dynamic Syntax, *it*, in its expletive use, is a pronoun which has lost its bottom restriction (meaning that there is, in principle, no limit to the complexity of the structure with which it can be updated), and its function is to move the pointer from the subject node, in order to allow the parse to continue. The lexical entry for expletive *it* (taken from Cann et al. (2005), p195) is shown in 8.7, below.<sup>5</sup>

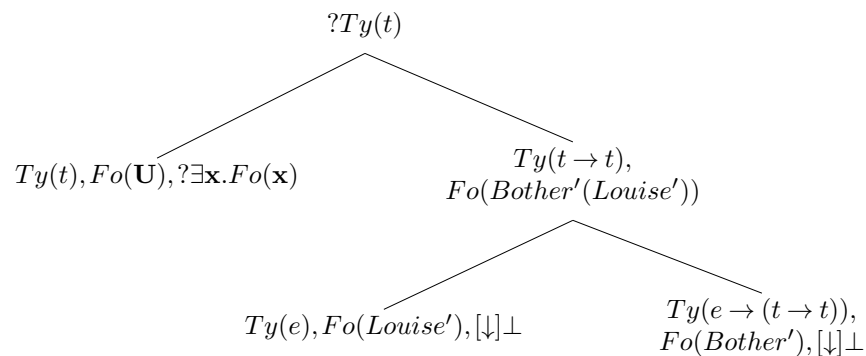
(8.7)	$it_{expl}$	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="padding-right: 10px;">IF</td> <td><math>?Ty(t)</math></td> </tr> <tr> <td style="padding-right: 10px;">THEN</td> <td>IF <math>[\uparrow]\perp</math></td> </tr> <tr> <td style="padding-right: 10px;"></td> <td>THEN ABORT</td> </tr> <tr> <td style="padding-right: 10px;">ELSE</td> <td><math>put(Ty(t), Fo(\mathbf{U}), ?\exists \mathbf{x}. Fo(\mathbf{x}));</math></td> </tr> <tr> <td style="padding-right: 10px;"></td> <td><math>go(\langle \uparrow_0 \rangle \langle \downarrow_1 \rangle)</math></td> </tr> <tr> <td style="padding-right: 10px;">ELSE</td> <td>ABORT</td> </tr> </table>	IF	$?Ty(t)$	THEN	IF $[\uparrow]\perp$		THEN ABORT	ELSE	$put(Ty(t), Fo(\mathbf{U}), ?\exists \mathbf{x}. Fo(\mathbf{x}));$		$go(\langle \uparrow_0 \rangle \langle \downarrow_1 \rangle)$	ELSE	ABORT
IF	$?Ty(t)$													
THEN	IF $[\uparrow]\perp$													
	THEN ABORT													
ELSE	$put(Ty(t), Fo(\mathbf{U}), ?\exists \mathbf{x}. Fo(\mathbf{x}));$													
	$go(\langle \uparrow_0 \rangle \langle \downarrow_1 \rangle)$													
ELSE	ABORT													

Once *bothers Louise* has been parsed, the tree is in the state shown in figure 8.9. Although all type requirements have been fulfilled, there is still a requirement for a

<sup>5</sup>Note that there is an assumption that a subject node can be of  $Ty(t)$ , and that INTRODUCTION and PREDICTION can apply to create a blank tree with a  $Ty(t)$  and a  $Ty(t \rightarrow t)$  node. Once these rules have been applied the lexical actions in 8.7 can be processed. In a fuller account such as that of Gregoromichelaki (2006), which introduces *event* terms, these typing issues can be resolved, but this need not concern us for the purposes of this thesis.



formula value at the subject node, so we cannot yet evaluate the overall formula value and complete the parse. Instead, we can move the pointer through the tree to the node with the outstanding requirement. Once at this  $Ty(t)$  node, because there is a metavariable at the node, we can either complete the parse from context (as would be the case if there were a dialogue along the lines of that shown in 8.8), or from following lexical material. To do this, we need the rule of LATE \*ADJUNCTION. Like \*ADJUNCTION, discussed earlier, this introduces an unfixed node. Unlike \*ADJUNCTION, however, the unfixed node is of the same type as the one from which it is projected.

Figure 8.9: Parsing *It bothers Louise ...*

(8.8) **A:** Did you know that John loves Mary?

**B:** Yes. It bothers Louise.

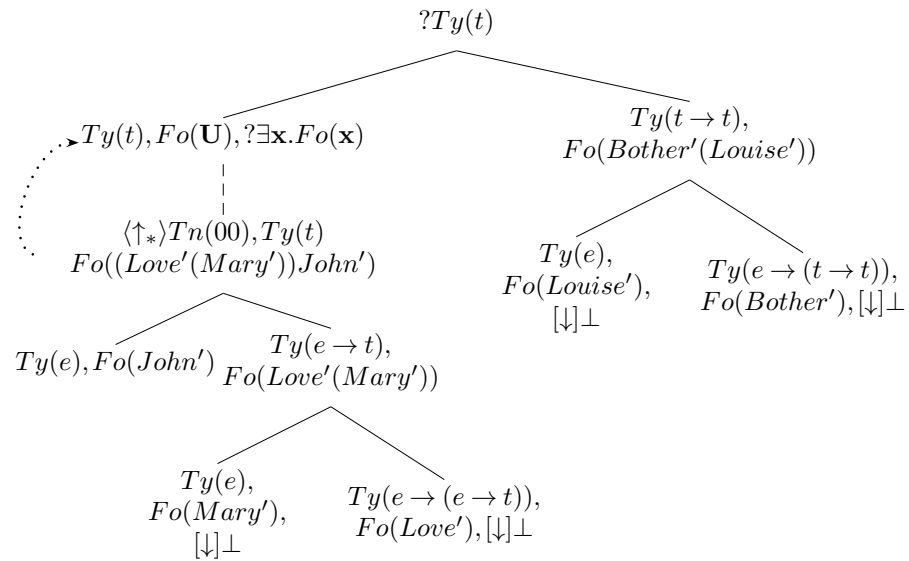
Following the application of LATE \*ADJUNCTION, we can parse the proposition *that John loves Mary*. Subsequently, we can MERGE the two completed  $Ty(t)$  nodes, as shown in figure 8.10, after which the usual processes of THINNING, COMPLETION and ELIMINATION will allow us to complete the parse.

The resulting formula value will be:

$Fo((Bother'(Louise'))((Love'(Mary'))John'))$ , which will be the same for both (8.5) and (8.6), as it should be.

#### 8.2.4 LINKED trees

As can be seen in earlier examples in which a formula value from one tree can be pragmatically substituted into another, trees are not constructed in isolation. This has implications for many different types of construction, including coordination and relative

Figure 8.10: Parsing *It bothers Louise that John loves Mary*

clauses and adjunction in general.<sup>6</sup> The way these are dealt with in the DS framework is in the building of separate, but linked informational trees, in tandem. The rule of LINK ADJUNCTION allows us to construct a new tree, linked to the tree currently under construction, and carrying the requirement for a copy of the formula value from the node at which LINK ADJUNCTION is applied.

### Relative clauses

In combination, the rules of LINK ADJUNCTION and \*ADJUNCTION allow us to project a linked tree (using the modal operators  $\langle L \rangle$  and  $\langle L^{-1} \rangle$ ) which must contain a copy of the node it is linked to, so that sentences such as *John, who smokes, loves Mary* can be parsed as shown in figure 8.11. Later in the parse, shown in figure 8.12, the unfixed node is merged to a fixed node position, carrying the copy of the node the linked tree is from. This can then be evaluated using LINK EVALUATION, giving the correct interpretation that it is John who both smokes and loves Mary. Notice that, as with the pronoun *he*, above, there is no trace of the word *who* on the final tree; the lexical actions merely provide a metavariable which is then updated by the copy of  $Fo(John')$ .

<sup>6</sup>For an account of GENERALISED ADJUNCTION as an available strategy whereby a node can be introduced from any node to another of the same type across any arbitrary relation, see Cann et al. (2005), p206.

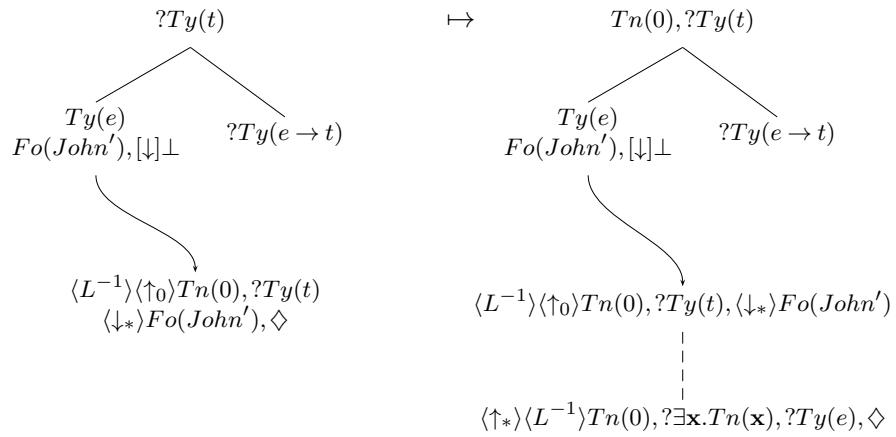


Figure 8.11: Parsing *John ...* (ready to parse *who*), using LINK ADJUNCTION and \*ADJUNCTION

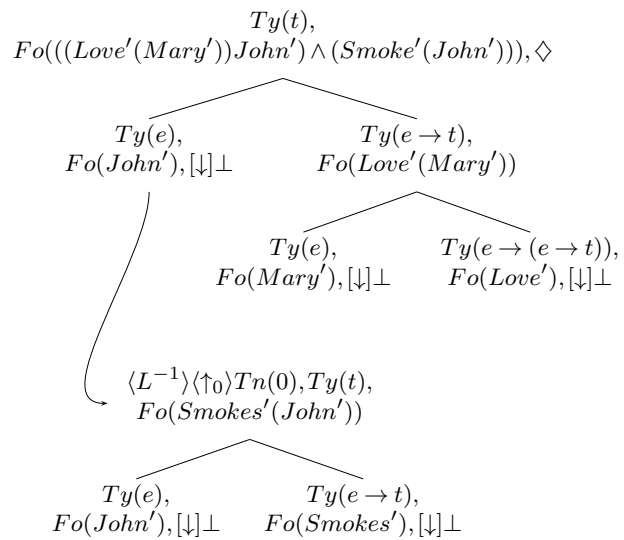


Figure 8.12: Parsing *John, who smokes, loves Mary* – Completed parse

### 8.3 Generation

In DS, generation uses the exact same tree representations and actions as parsing, with the addition of a goal tree which is subject to a subsumption check of the partial tree under construction at every step. Note that the goal tree can itself be partial; the only stipulation is that it is more advanced than the current parse state.

### 8.4 Parsing/producing CCs

With this apparatus in place DS can deal with same- and cross-person CCs of both expansion and completion types in terms of syntactic licensing, as the (potentially partial) antecedent parse tree can be used as input to the parsing and generation of the continuation, such that the complete CC will be well-formed provided it results in a complete tree with no outstanding requirements (Purver et al., 2006; Gregoromichelaki et al., 2011). Cross-person CCs are just as easy in this system as same-person ones, provided that the person who supplies the continuation has in mind a goal tree which both extends and is subsumed by the potentially partial tree used as the antecedent.

#### 8.4.1 Expansion CCs

Cross-person completion CCs have been previously outlined in DS (see section 8.4.2, below), and we can now see how the apparatus of DS also allows us to account for expansions in both same- and cross- person cases. Expansions can be seen in DS to be extremely natural, as adding a LINKed tree, e.g. by adding an adjunct or the sentence relatives studied by Rühlemann (2007) is a generally available option – it may even be less costly to start from an existing parse state than starting afresh.<sup>7</sup>

To illustrate one way in which a parse might proceed using a simple sentence relative example, consider the expansion *which bothers Louise* uttered (by either the same or a different person) following a complete sentence such as *John loves Mary*. Following parsing *John loves Mary* which produces the tree shown in 8.5. LINK ADJUNCTION may then apply<sup>8</sup> with the requirement that a copy of the formula value  $Fo((Love'(Mary'))John')$  exists in the LINKed tree. The parse may then proceed exactly as in the expletive *it* case,

<sup>7</sup>This is, of course, an empirical question.

<sup>8</sup>Note that the availability of the rule of LINK ADJUNCTION from a  $Ty(t)$  node is independently motivated by the possibility of sentential subjects as discussed in section 8.2.3.

with the relative pronoun *which* contributing a metavariable that can be updated with the copy of the formula to resolve all outstanding requirements. The resulting tree (prior to LINK EVALUATION) is shown in 8.13, below.<sup>9</sup>

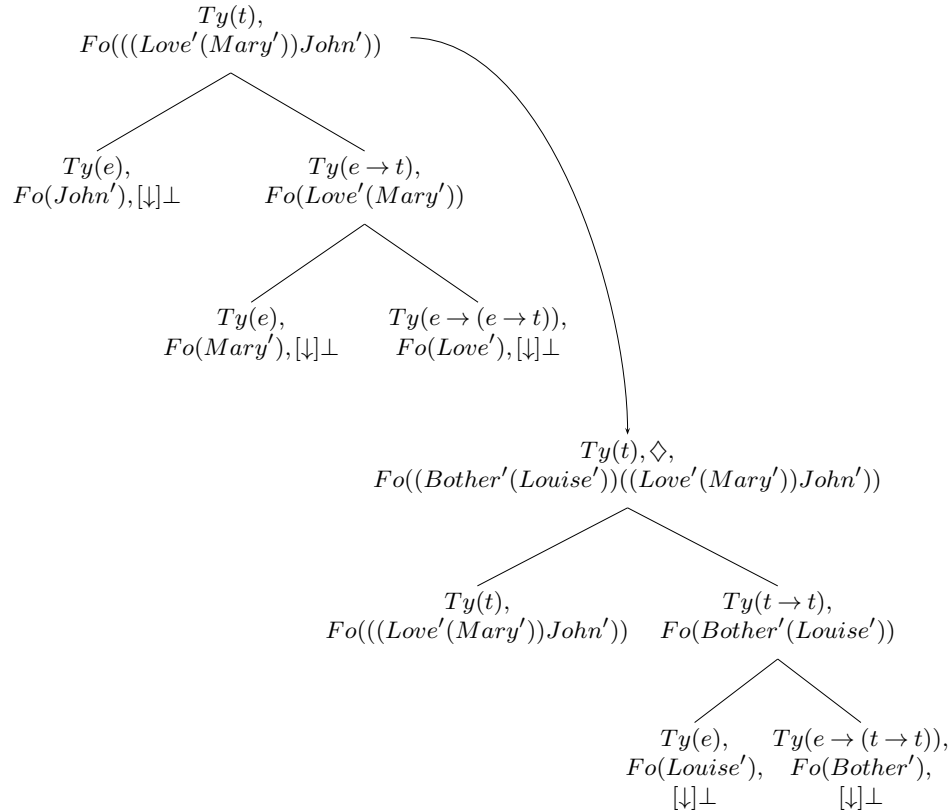


Figure 8.13: Parsing *John loves Mary, which bothers Louise*

### 8.4.2 Completion CCs

Cross-person completions (outlined in DS in Purver et al., 2006; Gregoromichelaki et al., 2011) are also easily explicated – even those cases that are problematic for string based accounts, such as (2.16), repeated below, because the representation is an interpretational one, not a string based one. This means that grammaticality judgements are based on the semantic interpretation of the pronouns (*I*, *you*, *myself*) and not their syntax in isolation (for details, see Gregoromichelaki et al., 2011).

<sup>9</sup>Applying LINK EVALUATION as it is currently formulated would result in the formula at the top node of the original tree being:

$Fo((Love'(Mary'))John') \wedge ((Bother'(Louise'))((Love'(Mary'))John'))$ . I do not consider here whether this is the correct interpretation, with the repeated information as a direct result of the explicit semantic tying of the second contribution or whether the LINK EVALUATION rule needs to be either reformulated or optional. Note also the typing issues mentioned in footnote 5.

(2.16) (with smoke coming from the kitchen)

**A:** I'm afraid I burnt the kitchen ceiling

**B:** But have you

**A:** burned myself? Fortunately not. [From Gregoromichelaki et al. (2011)]

Extending these ideas with specific reference to the data in this thesis, we can see that whilst providing the means for speaker switch at any point in the parse/production of an utterance, the grammar also provides the means to account for effects of syntactic predictability. The highly predictable parts of speech are those in which the lexical actions for a word introduce additional nodes with type requirements to be fulfilled, as seen in the lexical entry for the transitive verb *'loves'* (8.2), which introduces an object node with a requirement for a  $Ty(e)$ . These cases elicited more CR responses in experiment 3, often formulated as continuations, showing that participants were able to predict the syntax, and use it in forming their response, but were unable to adequately resolve the requirement, leading to the request for clarification.

Additionally, without having to postulate additional machinery, DS may also be able to account for situations in which no continuation is offered to an incomplete contribution, but it is responded to as if it were complete, with the respondent 'filling in the blanks' in the interpretation, as with (7.12)–(7.15) from experiment 3, and (8.9), from the BNC.

(8.9) **C:** Yes I know.

**S:** It makes you feel very

**C:** No, I don't think about it, just let it run. [BNC KBG 25-27]

Hypothetically, such interpretations could be underspecified, with the proposed conceptual or semantic completion supplied directly by the hearer from a potential goal tree without requiring articulation. Of course, there is no guarantee that the hearer's goal tree matches the original speaker's, as this is an unconstrained process, but this is also the case in all situations where underspecification and update from context plays a role (e.g. anaphora and ellipsis). When a continuation is offered, there may also be a mismatch between interlocutors' goal trees, but this may not be apparent in the resulting dialogue because different continuations may be accepted as appropriate even if not what

the original speaker had in mind. Cases where there is an obvious match or mismatch in goal tree are those where competing continuations overlap, as for example in (2.8) and (4.6), repeated here.

(2.8) **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.]]

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

Right.

[BNC H5D 123-127]

(4.6) **J:** People don't mind [[waiting if they know]]

**S:** [[the frustration.]]

**J:** how long they're waiting for

[BNC H61 29-31]

However, given that many of the factors found to empirically affect both CC production and processing are interactional, this account can only be a partial one. Despite one of the core elements of DS being that interpretations occur in context, these analyses have little to say about exactly when transitions are likely to occur, or how the hearer decides what the completing goal tree should be. As Poesio and Rieser (2010) point out, by "... joining parsing and production DS provides a necessary condition for Cnst's production of some object NP but this will not necessarily be *eine Schraube*,<sup>10</sup> it could be anything else, say *a car, a black hole etc*" (p69).

This is true, because the story of generation as told so far in DS is tactical; not strategic. DS has nothing to say about how it is that you actually decide what you are going to say before you begin speaking, but note that even so, in cross-person CCs such as (1.1), repeated below, in which the continuation is not the same as the original antecedent owner's continuation would have been, the *grammaticality* of the continuation is not in question.

1.1 **Daughter:** Oh here dad, a good way to get those corners out

**Dad:** is to stick yer finger inside.

**Daughter:** well, that's one way.

[from Lerner (1991)]

<sup>10</sup>The specific example to which this relates will be discussed further below.

In DS terms, the mismatch is not at the level of content, but between the different goal trees which the interlocutors had in mind (notice the rejection clearly indicates that daughter has a different goal tree in mind but that Dad’s continuation is a semantically and syntactically valid one). How either goal tree is arrived at is not answered by DS.

### 8.4.3 Type Theory with Records and DS

To address some of these issues, Purver et al. (2010), suggest some extensions to DS which allow the semantic content to access information beyond the combinatorial semantics of the string, using a Type Theory with Records (TTR) representation. These additions will be outlined briefly before discussing how or whether they can account for the pragmatic effects of CCs explored in this thesis.

TTR (Cooper, 2005) is a structured notational representation in which *record types* can be inhabited by specific *records* with both using the same representations of fields, as sequences of *label : type* ordered pairs. In a record, these are pairs of a label and an object, whilst in a record type they are pairs of a label and a type, such that a record  $\left[ x = john \right]$  is of type  $\left[ x : Ind \right]$ . Even if the record has additional fields in it, it is still of this type, meaning that a subtype relationship can be defined. The value of a field may also itself be a record (i.e. its type may be a record type), and each successive field may depend on previous fields within the record (or record type), giving a one-way notion of dependency.

Purver et al. (2010) replace the unstructured *Fo()* DS node decorations with TTR record types and *Ty()* simple type labels (and requirements) are then interpreted as referring to final TTR field types. The lexical entry for ‘John’ is thus modified as shown in (8.10).

$$(8.10) \quad john \left| \begin{array}{ll} \text{IF} & ?Ty(e) \\ \text{THEN} & \text{put}(Ty(e), \left[ x : john \right], [\downarrow] \perp) \\ \text{ELSE} & \text{ABORT} \end{array} \right.$$

### 8.4.4 Utterance events

Additional information about the utterance event itself must be available to interlocutors because pronouns (e.g. ‘I’, ‘you’) are interpreted with respect to who is speaking, and such utterance events are available to be referred to anaphorically (Ginzburg, in press;



Poesio and Rieser, 2010). Purver et al. (2010) therefore propose a structured *context* field on which the *content* field may depend, containing, minimally, information about utterance event and speaker/addressee. As this information is available at each tree node, the change of referents across a split point can be trivially accounted for.<sup>11</sup>

In this revised system, ELIMINATION both performs beta reduction on the *content* (as before) and concatenation ( $\oplus$ ) on the *context* (8.11).

(8.11) ELIMINATION

$$\begin{array}{l}
 \text{IF} \quad ?Ty(X), \\
 \quad \langle \downarrow_0 \rangle (Ty(Y), \left[ \begin{array}{l} ctxt : c_1 \\ cont : \alpha \end{array} \right]) \\
 \quad \langle \downarrow_1 \rangle (Ty(Y \rightarrow X), \left[ \begin{array}{l} ctxt : c_2 \\ cont : \beta \end{array} \right]) \\
 \text{THEN} \quad \text{put}(Ty(X), \left[ \begin{array}{l} ctxt : c_1 \oplus c_2 \\ cont : \beta(\alpha) \end{array} \right]) \\
 \text{ELSE} \quad \text{ABORT}
 \end{array}$$

#### 8.4.5 Speech act information

Similar considerations lead Purver et al. (2010) to include speech act information, which can also be reasoned over (e.g. an appropriate response to “are you asking me a question” doesn’t rely on the semantic content of the prior contribution, but its speech act). However, unlike in PTT, such information is optional and can, like other levels of information in DS, be highly underspecified. The mechanism for computing such speech act information, which may only be constructed if subsequent discourse makes it necessary (as in the Turn Taking Puzzle outlined in Ginzburg (1997)), uses the device of LINKed trees, and in general is of the form sketched in 8.12, “where **A** is a metavariable ranging over speech act specifications, **V** the agent responsible for the speech act, **U** an utterance event (or sequence of events), and **F** some function over the semantic content of the utterance (**p** and **x** are rule-level variables binding terms on the nodes where the rules apply)” (p47).

<sup>11</sup>The availability of this type of information at such a fine-grained level corresponds loosely to the *micro conversational events* in Poesio and Rieser (2010).

(8.12) IF  $Ty(\mathbf{x}), Fo(\mathbf{p})$   
 THEN  $make(\langle L \rangle), go(\langle L \rangle)$   
 $put(\mathbf{A}(\mathbf{U}, \mathbf{V}, \mathbf{F}(\mathbf{p})))$   
 ELSE ABORT

Using this apparatus, we can now model both co-participants' and a third person's possible inferences regarding the functions of a CC and how such inferences might differ between participants. Note, however, that in all cases these inferences do not need to be computed to parse the string and may only be computed if required to formulate an appropriate response.

### Inferences over CCs

Even for a simple assertion co-constructed by two interlocutors, there are several possible inferences derivable, relating to who is responsible for the complete CC. This can be illustrated by the simple constructed example “John died” with *John* uttered by participant A, and *died* uttered by participant B, as shown in the tree in 8.14.<sup>12</sup>

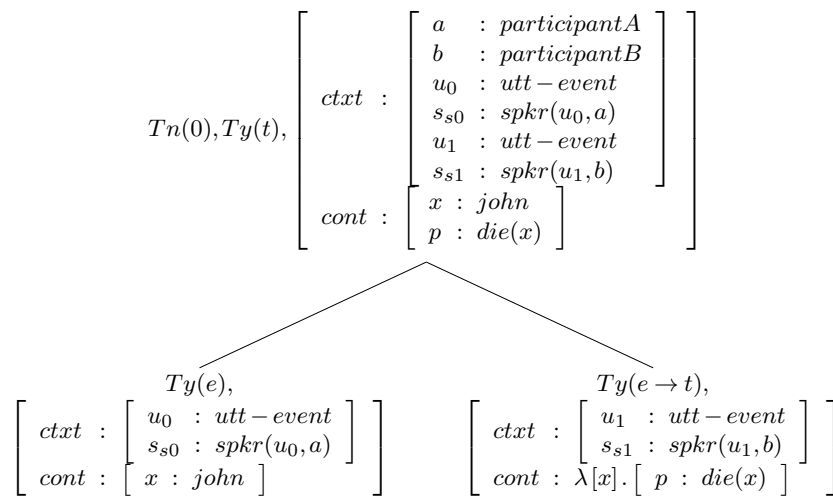


Figure 8.14: Tree following CC *John died*

If a third-person were asked “who said what” about this CC, there are (at least) four potentially appropriate responses, listed in (8.13)–(8.16), and corresponding to different potential inferences from the tree shown in 8.14.<sup>13</sup>

<sup>12</sup>Note that *participantA* and *participantB* should be understood as schemata standing in for the appropriate record or record type, but the details are not important for my purposes.

<sup>13</sup>Of course there are potentially many more possible interpretations, for example, if it is assumed that B is guessing how A was going to continue, then we might paraphrase the inference as ‘B said A said John died’. I leave such complications to one side, but see no practical reason why these more complex inferences could not also be modelled in the system outlined.

- (8.13) A said John and B said died [No inference]
- (8.14) A said John died [Inference: A is responsible agent]
- (8.15) B said John died [Inference: B is responsible agent]
- (8.16) A and B said John died [Inference: A & B jointly responsible]

For (8.14), the LINKed tree including the inferred information (in the *inf* field, which depends on the information in both the context and content fields) is shown in (8.17).<sup>14</sup>

$$(8.17) \quad \langle L \rangle Tn(0), \quad \left[ \begin{array}{l} \text{ctxt} : \left[ \begin{array}{l} a : \text{participant}A \\ b : \text{participant}B \\ u_0 : \text{utt} - \text{event} \\ s_{s0} : \text{spkr}(u_0, a) \\ u_1 : \text{utt} - \text{event} \\ s_{s1} : \text{spkr}(u_1, b) \end{array} \right] \\ \text{cont} : \left[ \begin{array}{l} x : \text{john} \\ p : \text{die}(x) \end{array} \right] \\ \text{inf} : \left[ \begin{array}{l} p' : \text{assert}(a, u_0, p) \end{array} \right] \end{array} \right]$$

For (8.15), the *inf* field would contain  $\left[ \text{inf} : \left[ p' : \text{assert}(b, u_1, p) \right] \right]$ , whilst for (8.16) both would potentially be available, with an additional combinatory inference<sup>15</sup>  $\left[ \text{inf} : \left[ \begin{array}{l} p' : \text{assert}(a, u_0, p) \\ p'' : \text{assert}(b, u_1, p) \\ p''' : \text{assert}(c, u_2, p) \end{array} \right] \right]$ , where *c* is an entity which consists of *a* and *b*, and *u*<sub>2</sub> is an utterance event composed of *u*<sub>0</sub> and *u*<sub>1</sub> (i.e. the full record type might look like 8.18).<sup>16</sup>

<sup>14</sup>In all subsequent records, the arguments in the *inf* field are shortened from their full path specifications for readability. The full *inf* field for (8.17) is actually  $\left[ \text{inf} : \left[ p' : \text{assert}(\text{ctxt}.a, \text{ctxt}.u_0, \text{cont}.p) \right] \right]$

<sup>15</sup>Of course it could be that only the combined inference is available. but potential inference rules and how they are made (including considerations of the added computational complexity that such optional rules bring in) are beyond the scope of this thesis. In principle, however, certain default inferences could be lexicalised.

<sup>16</sup>Note that one way in which this might be consistently instantiated is to have speech act predicates taking arguments that are lists of entities rather than single entities. Certainly this would be a natural way to conceive of utterance events, with single word utterances then being lists with only one member, and longer fragments being lists of more than one utterance event. Whether this would also be appropriate for the speaker field would need more careful consideration. I do not pursue this here.



## 8.5 Summary

In sum, DS offers a suitable underlying grammar formalism for modelling CCs, which allows us to account for the apparent lack of syntactic constraints on where a CC may occur, by presenting the grammar itself as a resource from which interpretations are built. It does not suffer from some of the issues other formalisms face by virtue of its strict incrementality and the tight coupling of parsing and generation. It is also able to account for some aspects of syntactic and lexical predictability, given the underspecified nodes and/or requirements that lexical and computational actions create on the tree under construction. However, although DS claims to utilise context it is not entirely clear what is available as context at any given time – and only prior linguistic context has been formalised in DS (Purver et al., 2006). From the corpus and experimental results in this thesis we can see that semantic or contextual predictability (in the form of shared knowledge or common ground) are crucial in determining whether participants are able to continue another’s (possibly incomplete) contribution. The evidence from CCs suggests, in line with previous studies (e.g. Eshghi, 2009), that contexts may vary between different interlocutors in the same conversation. Such considerations have a bearing on the differential distributions of CCs as seen in the various corpora, such that the coalitions that can be seen by examining CCs are diagnostic of uneven levels of shared context amongst participants in a conversation.

And, as discussed, DS as currently formulated has nothing to say about strategic generation. As Poesio and Rieser (2010) state, “[t]he targeted production of language material is clearly plan-based. So, in order to make DS do that it would have to be tied to some theory of dialogue *via* a suitable interface” (p69). Whether such a theory would need to be strictly intentional as theirs is is an open question, but it at least seems plausible that it is not in the general case; notice that even in their example (shown here as (8.19)), there is no need for *cnst* to make inferences over intentions before producing the continuation. As *cnst* has several screws available and cannot know *which* one she needs to use, *eine Schraube* is underinformative, and could just be based on her own knowledge of her own context, as she can predict that she needs to use one of the screws available to her to attach the two pieces together, though not which one. This corresponds to Poesio and Rieser’s ‘blurring out’ explanation, which they list as one possible reason

for producing the continuation. As we have no way in practise of knowing whether *cnst* relied on intentional inferences or not before generating the continuation it is hard to judge whether the addition of complicated interpretations which assume she did (as one plausible interpretation) are either necessary or desirable.

(8.19) **inst:** So, jetzt nimmst du *<pause>*

*Well, now you take*

**cnst:** eine Schraube

*a screw*

**inst:** eine *<->* orangene mit einem Schlitz

*an <-> orange one with a slit*

**cnst:** Ja

*Yes*

*[Poesio and Rieser (2010), 1.1-1.4]*

The evidence suggests that as long as you can come up with a suitable continuation, based on your own context supplying a subsumed goal tree, you can do so. Why you would do so in any given situation (which Poesio and Rieser, 2010, motivate in terms of shared plans) is not clear cut, though it appears that presumptions about what is shared, with whom at a given point in a dialogue influences the likelihood of producing a contribution as a continuation. It is also true that why or how you come to formulate any utterance at all isn't clear either, so this is a general point, and not specifically related to whether or not the contribution in question continues something prior.

The extensions to DS offered by Purver et al. (2010) do offer ways in which DS could be extended to account for the differing interpretations available to all utterances, and I have shown how different interpretations of CCs as e.g. being indicative of coalitions can be modelled. Indeed, such additions are necessary to transform DS from a theory of grammar to a theory of dialogue more generally, including, for example, adding an account of grounding. However, it does give us a framework in which it is possible to model the dissociation seen in the corpora and the experiments between who is expected to contribute next (possibly with a grounding act) which is related to the utterance event and thus part of the *ctxt*, and who holds authority for the overall speech act, which may be jointly owned, and is stored in the *inf* field. As with everything in DS, these

fields are a potential source of underspecification which may mean that certain words (such as discourse markers, or Wh-words) introduce underspecified inferences, but do not themselves leave any explicit representation on the tree (as with e.g. pronouns).

## Chapter 9

### Conclusions

---

This chapter assesses what the studies reported in this thesis tell us about compound contributions in particular and the nature of dialogue and communication in general. Future directions and research to follow up and capitalise on these results are then discussed.

The evidence presented points towards the general conclusion that CCs are principally driven by common ground. They are possible when it is shared and other participants recognise and orient to this fact. How this is made manifest and affects the various aspects of CCs in terms of the hypotheses from chapter 3, is discussed below under five broad headings.

- (9.1) There are no major syntactic constraints on where CCs *can* occur; grammar can thus be seen as a resource for constructing common ground.
- (9.2) There are pragmatic constraints regarding under what circumstances CCs *are likely* to occur, including turn-taking considerations.
- (9.3) CCs do not function as a single turn that just happens to have been produced in two or more parts, potentially by more than one person.
- (9.4) CCs are possible when context is shared between participants in a dialogue. Moreover, they are diagnostic of uneven levels of common ground between different participants at different stages in the conversation.



- (9.5) Cross-person CCs are often interpreted as indicative of coalitions, particularly in collaborative task-based dialogues.

## 9.1 Grammar as a resource

Firstly, the evidence in this thesis shows that CCs, as I have defined them (see below), occur frequently in dialogue, in both different media (text-based and speech) and different contexts (task-based dialogue and general conversation), with 3-10% of all contributions in dialogue being continuations by one person of another's prior contribution and 10-24% being continuations of one's own prior contribution. Continuations are spontaneously produced and interpreted by conversational interlocutors with no apparent problems, despite having to switch role from speaker to hearer (or vice versa) or integrate syntactic and semantic information from potentially disparate sources.

In support of hypothesis 1, there was no evidence for strong syntactic constraints on where CCs can occur, demonstrating that the unfolding grammatical structure is a generally available resource which any interlocutor may continue (provided, of course that they can formulate a suitable continuation). In the corpora, split points were found between all different word types, both within and between syntactic constituents. In experiments 1 and 2 there were no effects to distinguish responses to those where the (arbitrary) split point fell within or between a constituent. In experiment 3, although there were effects of syntactic structure on overall dialogue structure (specifically whether or not there was a response at all and whether responses were clarification questions) participants were able to formulate responses to an incomplete antecedent as a continuation after words in all syntactic categories.

Despite this, it is clear that what I have been calling compound contributions for the purposes of this thesis are not a single unified category. The same continuation (in syntactic terms) can be used for a variety of different functions, from showing agreement and understanding to indicating quite the opposite by requesting clarification. This shows that language provides the structure for building interpretations but does not determine them. The obvious question that arises from this is whether my original definition of CCs – dialogue contributions that continue or complete an earlier contribution – can, or should be sustained. This is a complex question. While it is true that CCs can and do have

different interpretations and effects on dialogue, it is also clear that if we are looking only at the CC itself (and not the preceding or following dialogue) then we cannot distinguish between these. But we are able to identify the phenomenon as defined in this thesis from just the antecedent and continuation parts of the CC, as seen in the corpus studies of chapter 4. For this reason, I believe that my definition is a useful one for a group of phenomena that utilise the same grammatical resources for their expression. Further subcategorisation, however, might be helpful for those interested in the interpretations of CCs in dialogue, and could be motivated by the data in this thesis. For example, the completion/expansion distinction is based purely on syntactic factors, but it might be useful to separate those expansions that really are extensions of the existing sentence from those which use discourse markers such as *and* to start a new discourse unit, that is in some sense tied to a prior one (see section 9.2). The response to a cross-person CC might also help to categorise them, with acknowledgements indicating that the continuation was treated as a repair, for example. However, it should be remembered, especially in terms of creating dialogue systems that can respond appropriately to CCs, that at the time of its production, whether a continuation is intended as a repair or an acknowledgement, for example, can be ambiguous or undecided, so many different responses might be equally appropriate.

Additionally, explicit repair of antecedent material in the continuation was not common, though in the tuition corpus there was more likely to be repair in the within-constituent cross-person completions, showing a weak effect of syntactic constituency.

In view of this, Dynamic Syntax is a good grammatical model for the syntactic analysis of CCs, because the syntax is a resource which interlocutors use to build up *interpretations* of language input. Its inherent incrementality, and tight coordination between parsing and generation enable the switch from speaker to hearer between any words in the string, consistent with our lack of empirical evidence for strictly syntactic constraints on CCs.

## 9.2 Pragmatic constraints

In all corpora studied, the majority of continuations were of the expansion type, adding syntactic material to a contribution that could be perceived to be already complete (as

predicted by hypothesis 2). Such continuations typically use weaker syntactic ties, such as adding an extending clause (by starting a contribution with ‘*and*’ or ‘*but*’, for example), which serves to confirm a contribution’s link to a prior one and thus add to the overall coherence of the dialogue. In a grounding model, expansions, which are in some sense not as strongly tied to their antecedents, could be taken as potentially initiating a new discourse unit, whilst maintaining the coherence of the dialogue at the discourse level. This especially appears to be the case with the common category of extending clauses (beginning with e.g. ‘*and*’), which then serves the dual purpose of both presenting new information and tying a contribution to a prior one (syntactically and e.g. topically).

In text-based corpora, an even greater proportion of naturally occurring CCs were expansions, especially in dialogues conducted using the line-by-line interface, in which participants cannot project upcoming TRPs. Despite this, there were no differences in responses to CCs which appeared to be completions rather than expansions in experiments 1 and 2, in which the position of the split point was arbitrary. This is especially notable when we consider that the arbitrary splits were expansions only 37-45% of the time, whilst in the genuine text dialogues it was over 77%. This therefore appears to be a pragmatic consideration of participants generally waiting for a TRP before taking the floor rather than an issue of processing or interpretation of the string from apparently disparate sources.

The majority of all turns in the corpora end in a potentially complete way, showing that participants are sensitive to each other’s turn space, though even this can be manipulated via the unfolding grammar, as when people produce hearably incomplete contributions which somehow invite completion, as in the tuition corpus examples. Importantly, having not finished a contribution with a recognised ending does not itself prompt production of a completion, as evidenced by the proportion of incomplete contributions in the corpora which never get completed, and also by the disjunction of syntactic and pragmatic completeness seen in experiment 3. That participants may be able to interpret incomplete strings in highly predictable cases – without necessarily articulating the missing lexical material – has implications for both dialogue models and the relationship between syntax and semantics more generally.

In the BNC there was a greater proportion of same-person than cross-person con-

tinuations, but this was not necessarily the case for the other corpora, which looked at cooperative task-based dialogues and showed higher proportions of cross-person CCs than in the BNC. Of the specific CA categories, there were more cross-person opportunistic cases (after a '<pause>' or an '*er/erm*') suggesting that this may be taken as a cue for another person to enter one's turn space. In the BNC, the cross-person continuations were also more likely after an unfilled pause than a filled one, with '*er*' or '*erm*' acting as a pragmatic marker indicating the desire to retain the floor. Contrarily, it seems that filled pauses in the text-based corpora and experiment 3 are taken as explicitly relinquishing the floor despite the incompleteness of what preceded them.

### 9.3 Ownership

CCs do not seem to function as a single unit that happens to have been produced in more than one part by potentially more than one person. Despite the lack of strong syntactic constraints there is evidence that participant role affects both the purposes of producing a contribution as a continuation, and how a continuation is treated in the dialogue. There is clear evidence that the distributions between same- and cross-person cases are different (contra hypothesis 6), and that there are differences in responses to CCs that appear to be same- or cross-person.

For example, same-person continuations are more likely to follow a backchannel or single other s-unit than cross-person cases, suggesting that it may be the feedback from one's interlocutor(s) that leads to producing something syntactically tied to one's own prior contribution.

Finishing or continuing another's utterance does not give it the same status as if they had completed or continued it themselves – in all corpora the antecedent owner takes the next turn in a disproportionate number of cases following a CC, contrary to the predictions of hypothesis 9. In the character-by-character text-based corpus, the expectation that this should be the case leads to an increase in onset delay for turns following a cross-person CC, especially where the next speaker is the antecedent owner. This also leads to the pattern of results seen in experiment 2, whereby onset delays are affected by who is entitled or expected to speak next.

Patterns of ratifications also demonstrate that CCs are not treated as a single unit.

If the speaker of the antecedent treated the continuation as if they themselves had just finished their own contribution (either in terms of a cashing out of predictable material, or as a within-party repair), then acknowledgement or ratification should be inappropriate. That ratifications were also more likely to be offered following a completion rather than an expansion (contrary to hypothesis 8), suggests that completions cannot be taken to be solely explicit acknowledgements.

There are several ways in which this could be accounted for in a grounding model. Either continuations by another are generally treated as repairs (and not exclusively as particularly strong forms of acknowledgement) or they are not taken to be acknowledgements at all. Given that continuations tend to be offered when common ground is presumed to be shared (see section 9.4, below) it could be the case that it is this presumption of shared common ground which requires acknowledging, or rejecting. Other possibilities are available – for example, it might be the fact that the incoming participant is aligning themselves with the initial speaker that requires acknowledgement, and not specifically the content itself, but such questions remain open for future research.

We have seen how compound contributions can reflect ownership in many different ways. However, what determines who is held to be responsible for the speech act is independent of who is expected or entitled to speak next, or the form of the CC, and may be dependent on the response (or lack thereof) that the CC receives. An acknowledgement can therefore either assert ownership of the CC which would otherwise be treated as jointly owned or act as an acceptance of the jointly constructed material, such that it validates the ‘coalition’. Note that a lack of response to an apparent CC from the antecedent owner in experiments 1 and 2 leads to this coalition reading. It seems plausible that this interpretation becomes appropriate not because there was a CC *per se*, but because there was a CC that the antecedent owner implicitly accepted, and perhaps took to be addressed at a third party. This disjunction can be modelled in DS, with underspecification allowing participants to remain neutral about exactly who can be claimed to have said what.

## 9.4 Context

Hypothesis 3 can be seen as partly correct insofar as predictability at some levels has an influence on when a contribution is constructed as a continuation. The results of experiment 3 highlight the role of context or shared knowledge in the production of CCs - specifically in the highly restricted cases whereby the next word is highly unpredictable and its syntactic category is predictable, where if subjects were able to project a suitable continuation from context they were more likely to produce one. There was not, however, any evidence that predictability of e.g. syntactic category has an independent effect on when continuations are produced.

Data from the task-based dialogues also shows that having shared knowledge is a crucial factor in when participants produce CCs, as there are more cross-person CCs in task-based dialogues regardless of the communicational medium (and contrary to hypothesis 5). Additionally, as predicted by hypothesis 4, there are more cross-person CCs in tasks which impose parties. If participants have the same information available (which they are talking about), they produce more CCs than if they do not, and crucially this distribution of knowledge and the associated likelihood of producing a CC can shift over the course of a conversation. Just having a shared goal does not account for the differences in proportions of CCs at different stages in the tuition corpus and in different conditions in the tangram corpus, as the joint goals were the same in each case.<sup>1</sup> These findings mirror those of Eshghi (2009), who found contexts between conversants in the same dialogue were not necessarily the same between all participants at all times (“uncommon ground”). The differential distributions at different stages in the tuition and tangram corpora especially show that CCs can be seen as diagnostic of these uneven levels of (presumed and actual) shared context amongst participants in a conversation.

As the effects in experiments 1 and 2 (that coalitions were formed) were clearly interpretational, the apparent intentions to produce an utterance together of the producers of the fake CC are implied by a third person. It may be that questions about intentionality (above having the intention to produce an utterance) and speech act information is only ever worked out after the event, if at all (see Gregoromichelaki et al., 2011).

---

<sup>1</sup>It is of course possible to invoke a complex analysis of differing subgoals between participants at different points in the dialogues to explain this, but this would need careful working out and is an explanation at a higher order level than the ones proposed in this thesis.

## 9.5 Coalitions

CCs are more common in task-based dialogues, and even more so where the task itself designates participants into parties (in line with hypothesis 4). There was also a greater proportion of within-party CCs where the task imposes party-membership, as well as a greater proportion of within-party ratifications than cross-party (hypothesis 10).

However, as the turn-taking and grounding behaviour of participants around CCs shows (as discussed in section 9.3), they cannot be seen as indicative of parties in the sense used by Schegloff (1995), in which parties are single entities for the purposes of turn-taking. The data in this thesis show that co-constructing a sentence can be used to demonstrate groups as Sacks (1992) speculated, but that this does not in and of itself give the joint producers of the utterance the same rights and responsibilities towards the turn space as a single person or a party. For this reason we have referred to the groupings apparently formed by CCs as coalitions, not parties. The interactional behaviour produced by and in response to a coalition is clearly distinct from that produced by a single individual (which would not be the case if they were parties in the full sense). One person can be a party (and is necessarily so) but cannot be a coalition. In effect, the analysis I have proposed in this thesis assumes that there is an intermediate level between parties composed of more than one person for the purposes of turn-taking and single persons, which I have called coalitions. In a grounding model, one can illustrate this distinction as being whether participants are coordinated at the level of discourse unit but not at the level of utterances (coalitions) or if they are coordinated at both levels (parties). This means that, contrary to Traum (1994), which focused on dyadic interactions, a participant could be seen to perform a *continue* act on another's utterance, but only if they were taken to be in a party with them, and not if they were merely a coalition.

Of course, people may produce continuations of another's utterance to indicate that they believe themselves to be in agreement or a coalition with another, but the status of partyhood or otherwise is still negotiable. As discussed in section 9.1, producing a continuation can have different interpretations, and the different possibilities with respect to coalitions or parties are also dependent on the subsequent response (or lack of response) to the CC. In practise, the interpretation of co-constructors of a CC as coalitions (rather than, for example, the incoming speaker taking over the turn, or acknowledging

understanding of the original speaker’s turn) may be dependent on how cooperative participants are seen to be. Certainly this reading is more natural in the cooperative task dialogues in the corpus studies and experiments presented in this thesis. Note also that both notions of coalition and party are fluid, and who is coordinating with whom and at what level can change throughout the course of a dialogue.

In terms of effects on the dialogues, participants in the chat tool experiments behaved differently after seeing a cross-person CC compared to a same-person one. This manifests itself in different ways in the different interfaces, but in each case resulted in participants doing less after a cross-person CC (in terms of deletes or length of contributions). This is evidence that, in multi-party collaborative task-based dialogues, cross-party CCs are often interpreted as indicative of coalitions between participants, as suggested by hypothesis 11, but these aren’t necessarily parties in the turn-taking sense. Of course, nothing precludes participants in coalitions also being a party – in fact a party with more than one participant might be a particularly strong coalition, but this does not seem to be generally the case, and it is not clear how these possibilities would be separated.

And this should not be too surprising a finding, given what we have said about context, above – as you are more likely to produce a continuation if you share context or information, CCs can be reliably treated, by a third party, as an indicator of where participants are coordinated in terms of common ground, though need not be taken as evidence for the stronger claim that participants are parties in terms of turn-taking.

In a sense, due to the tight coupling of parsing and generation, this is what DS gives us. It supplies the mechanisms for aligning on situation models<sup>2</sup> or common ground, though of course it does not resolve the issue of how much we need to share in any given situation to be said to have communicated successfully. In terms of sharing common ground or aligning on situation models, the purpose of communication could then be reframed as being to form a coalition with one’s interlocutors in terms of the subject at hand (with different metrics applying in different scenarios in line with the evidence for ‘good enough’ processing). Certainly that is the explicit aim of the task-based dialogues reported in this thesis (though whether it happens is, of course, a moot point) and the production of continuations to another’s contributions can be viewed either as a strategy

---

<sup>2</sup>Though note that explicit repetition should not be either needed or expected under this view.



to enable such coalitions, or as an indicator of success.

## 9.6 Future directions

The work outlined here shows that continuing or extending one's own or another's prior contribution in dialogue is a generally available productive linguistic strategy that has specific effects on the unfolding interaction. However, as this thesis contains the first ever experimental manipulations of CCs in real time ongoing dialogues, there are at least as many questions raised as answered, and many avenues for further research present themselves. Some of the possible extensions to and applications of this work will be outlined below.

### 9.6.1 Experimental

Firstly, given the results from the three chat tool experiments there are a number of different experiments which could be carried out to help untangle the effects reported in this thesis. A few possible directions would continue to explore the possible effects of CCs (both on those supplying the antecedent and continuation, and third parties) and the conditions under which they are preferentially produced, using the chat tool methodology. For example, to disentangle cause and effect of context, we might design different tasks, which systematically vary the amount of shared knowledge interlocutors do share or need to share at different stages in the conversation, by introducing new information to one or more participants as the task progressed. This would also allow us to ascertain whether simply having shared knowledge is sufficient for producing CCs or whether the shared knowledge needs to be known to be shared (or grounded) in some sense. The type of task also seems to play a role in the production and interpretation of CCs – does a competitive task, for example, influence the interpretation of cross-person CCs as indicating coalitions, or is it a general heuristic, which still leads to the doing of less work that we saw in the cooperative tasks? Other experiments could investigate the effects of apparent ratification or rejection of a fake cross-person CC to see how defeasible the inferences of coalitions are (is it something that is automatic and may need subsequent revision or only available if there is no other evidence to the contrary).

More ambitiously, a future direction to see if the results from our experiment are directly transferable to the spoken modality, which some commentators have queried,

would be using a speech analogue of the experiments reported here. The misattribution of spoken dialogue would have to be accomplished through a system using avatars, or talking heads, and is a non-trivial engineering task, but it is not beyond the realms of possibility. Experiment 3 would present less technical challenges in translating to a spoken domain, and it would be informative to see how the availability or not of verbal backchannels might affect the results.

### 9.6.2 Formal

We have seen how the grammatical framework of Dynamic Syntax lends itself to the analysis of the specifically dialogic phenomena of CCs and how it may be extended to account for pragmatic inferences that may arise. However, it remains to be seen whether this initial sketch of an enhanced DS can scale up to account for all the data, and it is unclear how it might be translated into a parsing strategy which would operate in the same way that people do – e.g. when the wrong continuation is offered, people don't have to restart the whole sentence (4.4).<sup>3</sup>

Another shortcoming of DS as it currently stands is that the model of context (which is essential for resolving anaphora and ellipsis, as well as implicated in when participants can produce a continuation) is not well worked out. Though there is a basic model of linguistic context (based on retaining access to previous parse trees and the words and actions used to produce them), it is unclear how non-linguistic context could be modelled, or how such linguistic context might decay, and this is essential if DS is ever to realistically account for all facets of dialogue. This is also related to the question raised by Poesio and Rieser (2010) regarding how one decides what to say in any given situation. Whilst this is clearly non-deterministic (as one can say something unrelated to anything that has gone before at any point), a full model of context would at least be able to constrain the range of appropriate contributions in any given situation (as e.g. questions under discussion (QUD) attempt to do, Ginzburg, *in press*).

Dynamic Syntax, then, is an appropriate underlying grammar formalism for dialogue, which avoids a number of common issues arising from the usual disjunction of syntax and semantics, parsing and processing, and competence and performance. Much needs

---

<sup>3</sup>See Sato (2011) for some initial ideas in this direction, and note that this issue extends to a wide range of dialogue phenomena, such as repair in general.

to be done to scale it up, however, including the incorporation of dialogue markers and grounding cues, with appropriate underspecification.

### 9.6.3 Dialogue Models

From a dialogue modelling perspective, we would want to be able to tell when a human agent's contribution continues some prior contribution – either their own or the systems – in order to correctly analyse the semantics of the discourse, which is a non-trivial matter given that antecedents do not have to be (and often are not) incomplete, or adjacent to the continuation.

We would also want to be able to allow the system to produce naturalistic continuations, though as cross-person completions are not that common DeVault et al.'s (2009) approach (with the system finishing syntactically incomplete user contributions) might not be the most appropriate. Given the common types of CC seen in the corpus studies, an appropriate type of continuation for the system to offer might instead be an expansion, i.e. the system would not need to compute a complete sentence, but use previously parsed input as a starting point. As dialogue models are very often in highly constrained contexts in which the system seeks information from the user, appropriate strategies involving CCs could be using incomplete antecedents to invite a user completion (for example, the travel agent system might ask “You want to go to...?”) and appendor questions (“...by bus?” see Hough, 2011, for a preliminary outline of such a system). How this would work in practice and whether it would create more natural human-computer interaction is yet another open research question.

## 9.7 Summary

*Compound contributions* have been shown to be an important and common feature of dialogue which constitute a critical test case for theories of natural language processing. Unsurprisingly, this thesis merely provides the tantalising tip of the iceberg into understanding the phenomena and how it illuminates aspects of dialogue, and much, as ever, remains to be done.

## Bibliography

- Aiken, L. S., West, S. G., and Reno, R. R. *Multiple Regression: Testing and Interpreting Interactions*. Sage Publications, 1991.
- Altmann, G. and Kamide, Y. Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3):247–264, 1999.
- Antaki, C., Díaz, F., and Collins, A. Keeping your footing: Conversational completion in three-part sequences. *Journal of Pragmatics*, 25(2):151–171, 1996.
- Ballinger, G. A. Using generalized estimating equations for longitudinal data analysis. *Organizational Research Methods*, 7(2):127–150, 2004.
- Bar-Hillel, Y. Out of the pragmatic wastebasket. *Linguistic Inquiry*, 2:401–407, 1971.
- Battersby, S. A. and Healey, P. G. T. Head and hand movements in the orchestration of dialogue. In *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*. Portland, OR, 2010.
- Benjamin, T. M. *Grammatical expansion as a generic resource for the construction of responses*. Master’s thesis, University of York, 2009.
- Blackburn, P. and Meyer-Viol, W. Linguistics, logic and finite trees. *Logic Journal of the Interest Group of Pure and Applied Logics*, 2(1):3–29, 1994.
- Bock, K., Dell, G., Chang, F., and Onishi, K. Persistent structural priming from language comprehension to language production. *Cognition*, 104(3):437–458, 2007.
- Bolden, G. Multiple modalities in collaborative turn sequences. *Gesture*, 3(2):187–212, 2003.
- Branigan, H., Pickering, M., and Cleland, A. Syntactic co-ordination in dialogue. *Cognition*, 75:13–25, 2000a.

- Branigan, H., Pickering, M., Stewart, A., and McLean, J. Syntactic priming in spoken production: Linguistic and temporal interference. *Memory and Cognition*, 28(8):1297–1302, 2000b.
- Branigan, H., Pickering, M., Pearson, J., McLean, J., and Nass, C. Syntactic alignment between computers and people: The role of belief about mental states. In *Proceedings of the Twenty-fifth Annual Conference of the Cognitive Science Society*. 2003.
- Branigan, H., Pickering, M., McLean, J., and Stewart, A. The role of local and global syntactic structure in language production: Evidence from syntactic priming. *Language and Cognitive Processes*, 21(7-8):974–1010, 2006.
- Branigan, H., Pickering, M., McLean, J., and Cleland, A. Syntactic alignment and participant role in dialogue. *Cognition*, 104(2):163–197, 2007.
- Breheeny, R. Communication and folk psychology. *Mind & Language*, 21(1):74–107, 2006.
- Brennan, S. E. and Clark, H. H. Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22:482–1493, 1996.
- Bunt, H. Multifunctionality and multidimensional dialogue semantics. In *Proceedings of DiaHolmia, 13th SEMDIAL Workshop*. 2009.
- Burnard, L. *Reference Guide for the British National Corpus (World Edition)*. Oxford University Computing Services <http://www.natcorp.ox.ac.uk/docs/userManual/>, 2000.
- Buß, O., Baumann, T., and Schlangen, D. 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. Association for Computational Linguistics, Tokyo, Japan, 2010.
- Cann, R., Kempson, R., and Marten, L. *The Dynamics of Language*. Elsevier, Oxford, 2005.
- Cappelen, H. and Lepore, E. *Insensitive Semantics: A Defence of Semantic Minimalism and Speech Act Pluralism*. Blackwell, Oxford, 2005.

- Carletta, J. Assessing agreement on classification tasks: The kappa statistic. *Computational Linguistics*, 22(2):249–255, 1996.
- Chevalier, F. H. and Clift, R. Unfinished turns in French conversation: Projectability, syntax and action. *Journal of Pragmatics*, 40(10):1731–1752, 2008.
- Clark, H. H. *Using Language*. Cambridge University Press, 1996.
- Clark, H. H. and Fox Tree, J. E. Using *uh* and *um* in spontaneous speaking. *Cognition*, 84(1):73–111, 2002.
- Clark, H. H. and Schaefer, E. A. Contributing to discourse. *Cognitive Science*, 13:259–294, 1989.
- Clark, H. H. and Wilkes-Gibbs, D. Referring as a collaborative process. *Cognition*, 22:1–39, 1986.
- Coates, J. No gap, lots of overlap: Turn-taking patterns in the talk of women friends. In Graddol, D., Maybin, J., and Stierer, B., editors, *Researching language and literacy in social context*, pages 177–192. Multilingual Matters Ltd, 1994.
- Cooper, R. Records and record types in semantic theory. *Journal of Logic and Computation*, 15(2):99–112, 2005.
- Couper-Kuhlen, E. and Ono, T. ‘Incrementing’ in conversation. A comparison of practices in English, German and Japanese. *Pragmatics*, 17(4):513–552, 2007.
- Couper-Kuhlen, E. and Selting, M. Introducing interactional linguistics. In Selting, M. and Couper-Kuhlen, E., editors, *Studies in Interactional Linguistics*, pages 1–21. John Benjamins, Amsterdam, 2001.
- DeVault, D., Sagae, K., and Traum, D. Can I finish? Learning when to respond to incremental interpretation results in interactive dialogue. In *Proceedings of the SIGDIAL 2009 Conference*, pages 11–20. Association for Computational Linguistics, London, UK, 2009.
- Díaz, F., Antaki, C., and Collins, A. Using completion to formulate a statement collectively. *Journal of Pragmatics*, 26(4):525–542, 1996.

- Duncan, S. Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, 23(2):283 – 292, 1972.
- Eshghi, A. *Uncommon Ground: The Distribution of Dialogue Contexts*. Ph.D. thesis, Queen Mary University of London, 2009.
- Fais, L. Conversation as collaboration: Some syntactic evidence. *Speech Communication*, 15(3-4):231–242, 1994.
- Fernández, R. and Ginzburg, J. Non-sentential utterances: A corpus-based study. *Traitement Automatique des Langues*, 43(2), 2002.
- Ferrara, K. The interactive achievement of a sentence: Joint productions in therapeutic discourse. *Discourse Processes*, 15(2):207–228, 1992.
- Ferreira, F. and Patson, N. The ‘good enough’ approach to language comprehension. *Language and Linguistics Compass*, 1(1-2):71–83, 2007.
- Ferreira, F., Bailey, K., and Ferraro, V. Good-enough representations in language comprehension. *Current Directions in Psychological Science*, 11(1):11, 2002.
- Ferreira, V. Is it better to give than to donate? Syntactic flexibility in language production. *Journal of Memory and Language*, 35:724–755, 1996.
- Ford, C. E. and Thompson, S. A. Interactional units in conversation: Syntactic, intonational, and pragmatic resources for the management of turns. In Ochs, E., Schegloff, E. A., and Thompson, S. A., editors, *Interaction and Grammar*, pages 134–184. Cambridge University Press, 1996.
- Fox, B. A. Principles shaping grammatical practices: an exploration. *Discourse Studies*, 9(3):299, 2007.
- Furuyama, N. Prolegomena of a theory of between-person coordination of speech and gesture. *International Journal of Human-Computer Studies*, 57(4):347–374, 2002.
- Garfinkel, H. *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice Hall, 1967.
- Gargett, A., Gregoromichelaki, E., Kempson, R., Purver, M., and Sato, Y. Grammar resources for modelling dialogue dynamically. *Cognitive Neurodynamics*, 3(4):347–363, 2009.

- Garrod, S. and Anderson, A. Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, 27:181–218, 1987.
- Garrod, S. and Pickering, M. Joint action, interactive alignment, and dialog. *Topics in Cognitive Science*, 1:292–304, 2009.
- Ginzburg, J. On some semantic consequences of turn taking. In Dekker, P., Stokhof, M., and Vennema, Y., editors, *Proceedings of the 11th Amsterdam Colloquium on Formal Semantics and Logic*, pages 145–150. 1997.
- Ginzburg, J. *The Interactive Stance: Meaning for Conversation*. Oxford University Press, in press.
- Ginzburg, J., Fernández, R., and Schlangen, D. Unifying self- and other-repair. In *Proceedings of the 11th Workshop on the Semantics and Pragmatics of Dialogue (DECA-LOG)*. 2007.
- Goffman, E. *Forms of Talk*. University of Pennsylvania Press, 1981.
- Goodwin, C. The interactive construction of a sentence in natural conversation. In Psathas, G., editor, *Everyday Language: Studies in Ethnomethodology*, pages 97–121. Irvington Publishers, New York, 1979.
- Gregoromichelaki, E. *Conditionals: A Dynamic Syntax Account*. Ph.D. thesis, King's College London, 2006.
- Gregoromichelaki, E., Sato, Y., Kempson, R., Gargett, A., and Howes, C. Dialogue modelling and the remit of core grammar. In *Proceedings of IWCS*. 2009.
- Gregoromichelaki, E., Kempson, R., Purver, M., Mills, G. J., Cann, R., Meyer-Viol, W., and Healey, P. G. T. Incrementality and intention-recognition in utterance processing. *Dialogue and Discourse*, 2(1):199–233, 2011.
- Grenoble, L. A. Conversational structure, syntax and the clause in Russian. In *The First Conference of The Slavic Linguistics Society*. 2006.
- Grice, H. Logic and Conversation. *Syntax and Semantics*, 3(S 41):58, 1975.



- Grice, H. P. Utterer's meaning and intentions. *The Philosophical Review*, 68:147–177, 1969.
- Guhe, M. *Incremental Conceptualization for Language Production*. NJ: Lawrence Erlbaum Associates, 2007.
- Hadelich, K., Branigan, H., Pickering, M., and Crocker, M. Alignment in dialogue: Effects of visual versus verbal-feedback. In *Proceedings of the 8th Workshop on the Semantics and Pragmatics of Dialogue*, pages 35–40. 2004.
- Hartsuiker, R., Pickering, M., and Veltkamp, E. Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science*, 15:409–414, 2004.
- Hayashi, M. Where grammar and interaction meet: A study of co-participant completion in Japanese conversation. *Human Studies*, 22(2):475–499, 1999.
- Healey, P. G. T. Interactive misalignment: The role of repair in the development of group sub-languages. In Cooper, R. and Kempson, R., editors, *Language in Flux*. College Publications, 2008.
- Healey, P. G. T., Purver, M., King, J., Ginzburg, J., and Mills, G. Experimenting with clarification in dialogue. In *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*. Boston, Massachusetts, 2003.
- Healey, P. G. T., Howes, C., and Purver, M. Does structural priming occur in ordinary conversation? In *Proceedings of Linguistic Evidence 2010*. Tübingen, 2010.
- Helasvuo, M.-L. Shared syntax: The grammar of co-constructions. *Journal of Pragmatics*, 36(8):1315–1336, 2004.
- Hjalmarsson, A. The additive effect of turn-taking cues in human and synthetic voice. *Speech Communication*, 53(1):23 – 35, 2011.
- Horton, W. and Gerrig, R. Conversational common ground and memory processes in language production. *Discourse Processes*, 40(1):1–35, 2005.

- Hough, J. Incremental semantics driven natural language generation with self-repairing capability. In *Recent Advances in Natural Language Processing (RANLP)*, pages 79–84. Hissar, Bulgaria, 2011.
- Howes, C., Healey, P. G. T., and Mills, G. A: An experimental investigation into... B:... Split utterances. In *Proceedings of the 10th Annual SIGDIAL Meeting on Discourse and Dialogue (SIGDIAL 2009 Conference)*, pages 79–86. Association for Computational Linguistics, 2009.
- Howes, C., Healey, P. G. T., and Purver, M. Tracking lexical and syntactic alignment in conversation. In *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*. Portland, OR, 2010.
- Howes, C., Purver, M., Healey, P. G. T., Mills, G. J., and Gregoromichelaki, E. On incrementality in dialogue: Evidence from compound contributions. *Dialogue and Discourse*, 2(1):279–311, 2011.
- Kamide, Y. and Mitchell, D. Incremental pre-head attachment in Japanese parsing. *Language and Cognitive Processes*, 14:631–662, 1999.
- Kamide, Y., Altmann, G., and Haywood, S. The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49(1):133–156, 2003.
- Kangasharju, H. Aligning as a team in multiparty conversation. *Journal of Pragmatics*, 26(3):291–319, 1996.
- Kelleher, J. and Kruijff, G. Incremental generation of spatial referring expressions in situated dialog. In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics*, pages 1041–1048. Association for Computational Linguistics, 2006.
- Kempson, R., Meyer-Viol, W., and Gabbay, D. *Dynamic Syntax: The Flow of Language Understanding*. Blackwell, 2001.
- Keysar, B. Communication and miscommunication: The role of egocentric processes. *Intercultural Pragmatics*, 4(1):71–84, 2007.

- Kleist, H. and Hamburger, M. On the gradual construction of thoughts during speech. *German Life and Letters*, 5(1):42–46, 1951.
- Kruger, J., Epley, N., Parker, J., and Ng, Z. Egocentrism over e-mail: Can we communicate as well as we think? *Journal of Personality and Social Psychology*, 89(6):925–2005.
- Lerner, G. Collectivities in action: Establishing the relevance of conjoined participation in conversation. *Text-Interdisciplinary Journal for the Study of Discourse*, 13(2):213–246, 1993.
- Lerner, G. H. On the syntax of sentences-in-progress. *Language in Society*, pages 441–458, 1991.
- Lerner, G. H. On the “semi-permeable” character of grammatical units in conversation: Conditional entry into the turn space of another speaker. In Ochs, E., Schegloff, E. A., and Thompson, S. A., editors, *Interaction and Grammar*, pages 238–276. Cambridge University Press, 1996.
- Lerner, G. H. Collaborative turn sequences. In *Conversation Analysis: Studies from the First Generation*, pages 225–256. John Benjamins, 2004.
- Lerner, G. H. and Takagi, T. 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.
- Levelt, W. *Speaking: From Intention to Articulation*. MIT Press, 1989.
- Levinson, S. Putting linguistics on a proper footing: Explorations in Goffman’s concepts of participation. In Drew, P. and Wootton, A., editors, *Erving Goffman: Exploring the Interaction Order*, pages 161–227. Oxford, England: Polity Press, 1988.
- Liang, K.-Y. and Zeger, S. L. Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1):13, 1986.
- Loebell, H. and Bock, K. Structural priming across languages. *Linguistics*, 41(5):791–824, 2003.

- Metzing, C. and Brennan, S. When conceptual pacts are broken: Partner-specific effects on the comprehension of referring expressions. *Journal of Memory and Language*, 49(2):201–213, 2003.
- Newlands, A., Anderson, A., and Mullin, J. Adapting communicative strategies to computer-mediated communication: An analysis of task performance and dialogue structure. *Applied Cognitive Psychology*, 17(3):325–348, 2003.
- Niv, M. A psycholinguistically motivated parser for CCG. In *Proceedings of the 32nd Annual Meeting on Association for Computational Linguistics*, pages 125–132. Association for Computational Linguistics, 1994.
- Ono, T. and Thompson, S. A. What can conversation tell us about syntax. In Davis, P., editor, *Alternative Linguistics: Descriptive and Theoretical Modes*. Benjamin, 1993.
- Pickering, M. and Garrod, S. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27:169–226, 2004.
- Pickering, M. and Garrod, S. Alignment as the basis for successful communication. *Research on Language and Computation*, 4:203–228, 2006.
- Pickering, M. and Garrod, S. Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, 11(3):105–110, 2007.
- Pickering, M. and Garrod, S. Prediction and embodiment in dialogue. *European Journal of Social Psychology*, 39:1162–1168, 2009.
- Poesio, M. and Rieser, H. Completions, coordination, and alignment in dialogue. *Dialogue and Discourse*, 1:1–89, 2010.
- Purver, M. *The Theory and Use of Clarification Requests in Dialogue*. Ph.D. thesis, University of London, 2004.
- Purver, M., Ginzburg, J., and Healey, P. G. T. On the means for clarification in dialogue. In Smith, R. and van Kuppevelt, J., editors, *Current and New Directions in Discourse & Dialogue*, pages 235–255. Kluwer Academic Publishers, 2003.
- Purver, M., Cann, R., and Kempson, R. Grammars as parsers: Meeting the dialogue challenge. *Research on Language and Computation*, 4(2-3):289–326, 2006.

- Purver, M., Howes, C., Gregoromichelaki, E., and Healey, P. G. T. Split utterances in dialogue: A corpus study. In *Proceedings of the 10th Annual SIGDIAL Meeting on Discourse and Dialogue (SIGDIAL 2009 Conference)*, pages 262–271. Association for Computational Linguistics, London, UK, 2009.
- Purver, M., Gregoromichelaki, E., Meyer-Viol, W., and Cann, R. Splitting the ‘I’s and crossing the ‘You’s: Context, speech acts and grammar. In Łupkowski, P. and Purver, M., editors, *Aspects of Semantics and Pragmatics of Dialogue. SemDial 2010, 14th Workshop on the Semantics and Pragmatics of Dialogue*, pages 43–50. Polish Society for Cognitive Science, Poznań, 2010.
- Reddy, M. The conduit metaphor: A case of frame conflict in our language about language. *Metaphor and thought*, pages 284–324, 1979.
- Reitter, D., Moore, J., and Keller, F. Priming of syntactic rules in task-oriented dialogue and spontaneous conversation. In Sun, R., editor, *Proceedings of the 28th Conference of the Cognitive Science Society*, pages 685–690. 2006.
- Rosé, C. P., Litman, D., Bhembé, D., Forbes, K., Silliman, S., Srivastava, R., and VanLehn, K. A comparison of tutor and student behavior in speech versus text based tutoring. In *Proceedings of the HLT-NAACL 2003 Workshop on Building Educational Applications Using Natural Language Processing*, page 37. Association for Computational Linguistics, 2003.
- Rühlemann, C. *Conversation in Context: A Corpus-Driven Approach*. Continuum, 2007.
- de Ruiter, J., Mitterer, H., and Enfield, N. Projecting the end of a speaker’s turn: A cognitive cornerstone of conversation. *Language*, 82(3):515–535, 2006.
- Sacks, H. *Lectures on Conversation*. Blackwell, 1992.
- Sacks, H., Schegloff, E. A., and Jefferson, G. A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4):696–735, 1974.
- Sag, I. A., Wasow, T., and Bender, E. M. *Syntactic Theory: A Formal Introduction*. CSLI Publications, Stanford, 2003.

- Sato, Y. Local ambiguity, search strategies and parsing in Dynamic Syntax. In Gregoromichelaki, E., Kempson, R., and Howes, C., editors, *The Dynamics of Lexical Interfaces*. CSLI Publications, 2011.
- Schachter, S., Christenfeld, N., Ravina, B., and Bilous, F. Speech disfluency and the structure of knowledge. *Journal of Personality and Social Psychology*, 60(3):362–367, 1991.
- Schegloff, E. A. 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.
- Schegloff, E. A. Turn organization: One intersection of grammar and interaction. In Ochs, E., Schegloff, E. A., and Thompson, S. A., editors, *Interaction and Grammar*, pages 52–133. Cambridge University Press, 1996.
- Schegloff, E. A., Jefferson, G., and Sacks, H. The preference for self-correction in the organization of repair in conversation. *Language*, 53(2):361–382, 1977.
- Schlangen, D. From reaction to prediction: Experiments with computational models of turn-taking. In *Proceedings of the 9th International Conference on Spoken Language Processing (INTERSPEECH)*. Pittsburgh, PA, 2006.
- Schwanenflugel, P. and Shoben, E. The influence of sentence constraint on the scope of facilitation for upcoming words. *Journal of Memory and Language*, 24(2):232–252, 1985.
- Shannon, C. and Weaver, W. *A Mathematical Model of Communication*. University of Illinois Press, 1949.
- Shriberg, E., Stolcke, A., and Baron, D. Observations on overlap: Findings and implications for automatic processing of multi-party conversation. In *Seventh European Conference on Speech Communication and Technology*. 2001.
- Skantze, G. and Schlangen, D. 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. Association for Computational Linguistics, Athens, Greece, 2009.

- Skuplik, K. Satzkooperationen. definition und empirische untersuchung. SFB 360 1999/03, Bielefeld University, 1999.
- Smith, N. *Chomsky: Ideas and Ideals*. Cambridge University Press, 1999.
- Steedman, M. *The Syntactic Process*. MIT Press, Cambridge, MA, 2000.
- Steedman, M. and Baldridge, J. Combinatory categorial grammar, 2003. Unpublished tutorial (Chapter 5).
- Sun, H. Pauses and co-construction in Chinese peer review discussions. *Pragmatics and language learning*, pages 121–141, 1995.
- Szczepek, B. Formal aspects of collaborative productions in English conversation. *Interaction and Linguistic Structures (InLiSt)*, <http://www.uni-potsdam.de/u/inlist/issues/17/>, 2000a.
- Szczepek, B. Functional aspects of collaborative productions in English conversation. *Interaction and Linguistic Structures (InLiSt)*, <http://www.uni-potsdam.de/u/inlist/issues/21/>, 2000b.
- Tannen, D. *Talking Voices: Repetition, Dialogue and Imagery in Conversational Discourse*. Cambridge: Cambridge University Press, 2007. Second Edition.
- Tao, H. and McCarthy, M. J. Understanding non-restrictive which-clauses in spoken English, which is not an easy thing. *Language Sciences*, 23(6):651–677, 2001.
- Taylor, T. J. *Mutual Misunderstanding: Scepticism and the Theorizing of Language and Interpretation*. Routledge, London, 1992.
- Thothathiri, M. and Snedeker, J. Give and take: Syntactic priming during spoken language comprehension. *Cognition*, 108(1):51–68, 2008.
- Traum, D. *A Computational Theory of Grounding in Natural Language Conversation*. Ph.D. thesis, University of Rochester, 1994.
- Walker, G. On some interactional and phonetic properties of increments to turns in talk-in-interaction. In Couper-Kuhlen, E. and Ford, C. E., editors, *Sound Patterns in*

*Interaction: Cross-linguistic Studies from Conversation*. John Benjamins Publishing Company, 2004.

Wilson, D. and Sperber, D. Truthfulness and relevance. *Mind*, 111(443):583, 2002.

Wright, B. and Garrett, M. Lexical decision in sentences: Effects of syntactic structure. *Memory & Cognition*, 12(1):31–45, 1984.



## Appendix A

### Balloon Task

---

Please choose a nickname and then read ALL of this text before starting.

The task is to collaborate with your partner(s) to resolve a dilemma. To do this you will be using an online chat tool (the experimenter will explain how this works) which you use to communicate.

#### **The situation**

Three people are in a hot air balloon. The balloon is losing height and about to crash into the mountains. Having thrown everything imaginable out of the balloon, including food, sandbags and parachutes, their only hope is for one of them to jump to their certain death to give the balloon the extra height to clear the mountains and save the other two. But who is it to be?

#### **The three people are:**

*Dr. Nick Rivers* – a cancer research scientist who believes he is on the brink of discovering a cure for most common types of cancer. He is a good friend of Tom and Susie Derkins.

*Mrs. Susie Derkins* – a primary school teacher. She is over the moon because she is 7 months pregnant with her second child.

*Mr. Tom Derkins* – the balloon pilot. He is the husband of Susie, who he loves very much. He is also the only one with any balloon flying experience.

#### **Your task**

You must discuss the pros and cons of keeping each of the three people in the balloon with your partner, and come to an agreement about who should be thrown out...

## Appendix B

### Subarctic Survival Task

---

You and your companions have just survived the crash of a small plane. Both the pilot and co-pilot were killed in the crash. It is mid-January, and you are in Northern Canada. The daily temperature is  $-25^{\circ}\text{C}$  and the night time temperature is  $-40^{\circ}\text{C}$ .

There is snow on the ground, and the countryside is wooded with several creeks criss-crossing the area. The nearest town is 20 miles away. You are all dressed in city clothes appropriate for a lecture.

Your group of survivors managed to salvage 12 items (each of you can see descriptions of three of these):

- A ball of steel wool
- A small axe
- A loaded .45-calibre pistol

Your task as a group is to introduce your own items and discuss ALL 12 to decide the order of importance for your survival, and why. You **MUST** come to agreement as a group.

#### **Other participants items:**

- A sectional air map made of plastic
- Can of Crisco shortening (semi-solid cooking fat in a metal tin)
- Newspapers (one per person)
  
- A compass
- Cigarette lighter (without fluid)
- One litre of whisky
  
- Family-size chocolate bars (one per person)
- Extra shirt and trousers for each survivor
- 20 x 20 ft. piece of heavy-duty canvas

#### **Subarctic survival... What you should have chosen...**

Mid-January is the coldest time of year in Northern Canada. The first problem the survivors face is the preservation of body heat and the protection against its loss. This problem can be solved by building a fire, minimizing movement and exertion, using as much insulation as possible, and constructing a shelter. The participants have just crash-landed. Many individuals tend to overlook the enormous shock reaction this has on the human body and the deaths of the pilot and co-pilot increases the shock. Decision-making under such circumstances is extremely difficult. Such a situation requires a strong emphasis on the use of reasoning for making decisions and for reducing fear and panic. Shock would be shown in the survivors by feelings of helplessness, loneliness, hopelessness, and fear. These feelings have brought about more fatalities than perhaps any other cause in survival situations. Certainly the state of shock means the movement of the

survivors should be at a minimum, and that an attempt to calm them should be made. Before taking off, a pilot has to file a flight plan which contains vital information such as the course, speed, estimated time of arrival, type of aircraft, and number of passengers. Search-and-rescue operations begin shortly after the failure of a plane to appear at its destination at the estimated time of arrival.

The 20 miles to the nearest town is a long walk under even ideal conditions, particularly if one is not used to walking such distances. In this situation, the walk is even more difficult due to shock, snow, dress, and water barriers. It would mean almost certain death from freezing and exhaustion. At temperatures of minus 25 to minus 40, the loss of body heat through exertion is a very serious matter. Once the survivors have found ways to keep warm, their next task is to attract the attention of search planes. Thus, all the items the group has salvaged must be assessed for their value in signalling the group's whereabouts.

The ranking of the survivors items was made by Mark Wanvig, a former instructor in survival training for the Reconnaissance School of the 101st Division of the U.S. Army. Mr. Wanvig currently conducts wilderness survival training programs in the Minneapolis, Minnesota area. This survival simulation game is used in military training classrooms.

### Rankings

1. *Cigarette lighter (without fluid)* The gravest danger facing the group is exposure to cold. The greatest need is for a source of warmth and the second greatest need is for signalling devices. This makes building a fire the first order of business. Without matches, something is needed to produce sparks, and even without fluid, a cigarette lighter can do that.

2. *Ball of steel wool* To make a fire, the survivors need a means of catching the sparks made by the cigarette lighter. This is the best substance for catching a spark and supporting a flame, even if the steel wool is a little wet.

3. *Extra shirt and trousers for each survivor* Besides adding warmth to the body, clothes can also be used for shelter, signalling, bedding, bandages, string (when unravelled), and fuel for the fire.

4. *Can of Crisco shortening* This has many uses. A mirror-like signalling device can be made from the lid. After shining the lid with steel wool, it will reflect sunlight and generate 5 to 7 million candlepower. This is bright enough to be seen beyond the horizon. While this could be limited somewhat by the trees, a member of the group could climb a tree and use the mirrored lid to signal search planes. If they had no other means of signalling than this, they would have a better than 80% chance of being rescued within the first day. There are other uses for this item. It can be rubbed on exposed skin for protection against the cold. When melted into oil, the shortening is helpful as fuel. When soaked into a piece of cloth, melted shortening will act like a candle. The empty can is useful in melting snow for drinking water. It is much safer to drink warmed water than to eat snow, since warm water will help retain body heat. Water is important because dehydration will affect decision-making. The can is also useful as a cup.

5. *20 x 20 foot piece of canvas* The cold makes shelter necessary, and canvas would protect against wind and snow (canvas is used in making tents). Spread on a frame made of trees, it could be used as a tent or a wind screen. It might also be used as a ground cover to keep the survivors dry. Its shape, when contrasted with the surrounding terrain, makes it a signalling device.

6. *Small axe* Survivors need a constant supply of wood in order to maintain the fire. The axe could be used for this as well as for clearing a sheltered campsite, cutting tree branches for ground insulation, and constructing a frame for the canvas tent.

7. *Family size chocolate bars (one per person)* Chocolate will provide some food energy. Since it contains mostly carbohydrates, it supplies the energy without making digestive demands on the body.

8. *Newspapers (one per person)* These are useful in starting a fire. They can also be used as insulation under clothing when rolled up and placed around a person's arms and legs. A newspaper can also be used as a verbal signalling device when rolled up in a megaphone-shape. It could also provide reading material for recreation.

9. *Loaded .45-calibre pistol* The pistol provides a sound-signalling device. (The international distress signal is 3 shots fired in rapid succession). There have been numerous cases of survivors going undetected because they were too weak to make a loud enough noise to attract attention. The butt of the pistol could be used as a hammer, and the powder from the shells will assist in fire building. By placing a small bit of cloth in a cartridge emptied of its bullet, one can start a fire by firing the gun at dry wood on the ground. The pistol also has some serious disadvantages. Anger, frustration, impatience, irritability, and lapses of rationality may increase as the group awaits rescue. The availability of a lethal weapon is a danger to the group under these conditions. Although a pistol could be used in hunting, it would take an expert marksman to kill an animal with it. Then the animal would have to be transported to the crash site, which could prove difficult to impossible depending on its size.

10. *Litre of whisky* The only uses of whiskey are as an aid in fire building and as a fuel for a torch (made by soaking a piece of clothing in the whiskey and attaching it to a tree branch). The empty bottle could be used for storing water. The danger of whiskey is that someone might drink it, thinking it would bring warmth. Alcohol takes on the temperature it is exposed to, and a drink of minus 30 degree whiskey would freeze a person's oesophagus and stomach. Alcohol also dilates the blood vessels in the skin, resulting in chilled blood being carried back to the heart, resulting in a rapid loss of body heat. Thus, a drunken person is more likely to get hypothermia than a sober person is.

11. *Compass* Because a compass might encourage someone to try to walk to the nearest town, it is a dangerous item. Its only redeeming feature is that it could be used as a reflector of sunlight (due to its glass top).

12. *Sectional air map made of plastic* This is also among the least desirable of the items because it will encourage individuals to try to walk to the nearest town. Its only useful feature is as a ground cover to keep someone dry.

## Appendix C

### Balloon Task Questionnaire

---

**Please complete this form after you have finished the Balloon Task**

Name used in chat tool \_\_\_\_\_

Who did you agree should be thrown off the balloon? \_\_\_\_\_

Do you think this was the correct decision? yes/no

If no, who did you think it should be? \_\_\_\_\_

How easy did you feel it was to come to an agreement?

Very easy			Neither easy nor difficult			Very difficult
1	2	3	4	5	6	7

How easy did you feel the conversation was to understand?

Very easy			Neither easy nor difficult			Very difficult
1	2	3	4	5	6	7

Compared to a face-to-face conversation, how smoothly did you feel the conversation went?

Much more smoothly			The same			Much less smoothly
1	2	3	4	5	6	7

Please enter any other comments you may have about the study below

---



---



---

**Thank you for taking part in this study!**

## Appendix D

### Arctic Survival Task Questionnaire

---

Name used in chat tool \_\_\_\_\_

Please put the objects in order of importance (1=most important)

1. \_\_\_\_\_ 2. \_\_\_\_\_ 3. \_\_\_\_\_  
 4. \_\_\_\_\_ 5. \_\_\_\_\_ 6. \_\_\_\_\_  
 7. \_\_\_\_\_ 8. \_\_\_\_\_ 9. \_\_\_\_\_  
 10. \_\_\_\_\_ 11. \_\_\_\_\_ 12. \_\_\_\_\_

How easy did you feel it was to come to an agreement?

Very easy		Neither easy nor difficult		Very difficult		
1	2	3	4	5	6	7

How easy did you feel the conversation was to understand?

Very easy		Neither easy nor difficult		Very difficult		
1	2	3	4	5	6	7

Compared to a face-to-face conversation, how smoothly did you feel the conversation went?

Much more smoothly		The same		Much less smoothly		
1	2	3	4	5	6	7

Please enter any other comments you may have about the study below

---



---



---



---

## Appendix E

### Questionnaire Results

---

Task	Interface	Agreement		Understanding		Smoothness		N
		Mean	(s.d.)	Mean	(s.d.)	Mean	(s.d.)	
Balloon Exp 1	LbyL	3.617	(1.823)	3.183	(1.610)	4.008	(1.530)	60
Balloon Control	CbyC	3.583	(1.564)	2.500	(1.314)	3.750	(1.712)	12
Balloon Exp 3	CbyC	4.139	(1.624)	2.639	(1.496)	4.639	(1.246)	36
Balloon Task	Total	3.787	(1.735)	2.926	(1.557)	4.190	(1.486)	108
Arctic Exp 2	CbyC	4.211	(1.543)	2.987	(1.194)	5.369	(1.153)	76
Arctic Control	CbyC	4.000	(1.344)	3.094	(1.510)	5.656	(1.096)	32
Arctic Task	Total	4.148	(1.484)	3.019	(1.290)	5.454	(1.139)	108

## Appendix F

### Experiment 1: Additional Statistical Analyses

---

IV	F	p
Same/Other	0.045	0.831
Floor Change	3.637	0.057
Within/Between Constituent	2.970	0.086
Participant	4.826	<0.001**
Same/Other $\times$ Floor Change	0.847	0.358
Floor Change $\times$ Within/Between	0.026	0.872
Same/Other $\times$ Within/Between	0.074	0.785
Same/Other $\times$ Floor Change $\times$ Within/Between	0.144	0.705

Table F.1: ANOVA on log transformed typing time of turn including within/between constituent

Note that antecedent and continuation start completeness and constituency of split point were analysed in separate models due to the fact that the measures are not independent.

IV	F	p
Same/Other	1.581	0.209
Floor Change	0.481	0.488
Within/Between Constituent	3.069	0.081
Participant	4.269	<0.001**
Same/Other $\times$ Floor Change	0.003	0.957
Floor Change $\times$ Within/Between	0.405	0.525
Same/Other $\times$ Within/Between	0.314	0.576
Same/Other $\times$ Floor Change $\times$ Within/Between	0.134	0.715

Table F.2: ANOVA on log transformed number of characters including within/between constituent



IV	Model effects		Parameter Estimates		
	Wald $\chi^2$	p	B	Wald $\chi^2$	p
Same/Other (SO)	8.355	0.004**	-0.450	5.200	0.023*
Floor Change (FC)	3.498	0.061	0.074	0.071	0.790
Within/Between (WB)	0.101	0.750	-0.128	0.486	0.485
SO $\times$ FC	0.029	0.864	0.269	0.409	0.523
FC $\times$ WB	0.136	0.712	0.064	0.036	0.849
SO $\times$ WB	0.273	0.601	0.313	0.898	0.343
SO $\times$ FC $\times$ WB	0.681	0.409	-0.441	0.681	0.409

Table F.3: GEE on deletes per characters including within/between constituent

The model in table F.3 (QICC = 92.780) was not as good a fit to the data as the simple model including only same/other, floor change and their interaction (QICC = 85.044).

## Appendix G

### Experiment 2: Additional Statistical Analyses

---

	AA		BA		AB		BB		fake sender	
	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>
AA			-0.627	0.265	0.198	0.422	0.322	0.373	-0.433	0.332
BA	-0.627	0.265			0.793	0.214	0.898	0.184	0.303	0.381
AB	0.198	0.422	0.793	0.214			0.125	0.450	-0.634	0.263
BB	0.322	0.373	0.898	0.184	0.125	0.450			-0.759	0.224
fake	-0.433	0.332	0.303	0.381	-0.634	0.263	-0.759	0.224		

Table G.1: Z-tests of proportion of responses by intervention type

IV	Model effects		Parameter Estimates		
	Wald $\chi^2$	<i>p</i>	B	Wald $\chi^2$	<i>p</i>
Same/Other (SO)	0.065	0.798	-0.121	0.503	0.478
Floor Change (FC)	1.271	0.260	-0.342	1.879	0.170
Antecedent (Ant)	0.078	0.779	-0.401	3.247	0.072
Continuation (Cont)	0.047	0.828	0.378	0.675	0.411
SO $\times$ FC	0.129	0.720	0.931	2.987	0.084
SO $\times$ Ant	0.960	0.327	0.125	0.149	0.699
SO $\times$ Cont	<0.001	0.993	-0.750	1.721	0.190
FC $\times$ Ant	1.098	0.295	1.338	4.193	0.041
FC $\times$ Cont	1.768	0.184	0.033	0.002	0.961
Ant $\times$ Cont	0.226	0.635	-0.366	0.440	0.507
SO $\times$ FC $\times$ Ant	4.385	0.036*	-1.439	2.163	0.141
SO $\times$ FC $\times$ Cont	0.119	0.730	-0.303	0.111	0.739
SO $\times$ Ant $\times$ Cont	7.649	0.006**	1.739	6.393	0.011*
FC $\times$ Ant $\times$ Cont	1.018	0.313	-0.720	0.526	0.469
SO $\times$ FC $\times$ Ant $\times$ Cont	0.008	0.927	0.144	0.008	0.927

Table G.2: GEE onset delay - complete model

Pairwise comparisons of the significant 3-way interactions also did not result in any significant effects.

IV	Model effects		Parameter Estimates		
	Wald $\chi^2$	p	B	Wald $\chi^2$	p
Same/Other (SO)	0.102	0.750	-0.687	2.729	0.099
Floor Change (FC)	0.083	0.773	-0.451	1.831	0.176
Other Same/Other (OthSO)	12.856	<0.001**	-0.741	3.767	0.052
Other Floor Change (OthFC)	13.567	<0.001**	0.669	2.199	0.138
SO $\times$ FC	0.023	0.880	0.680	2.265	0.132
SO $\times$ OthSO	2.740	0.098	0.853	2.717	0.099
SO $\times$ OthFC	0.759	0.384	0.481	0.478	0.489
FC $\times$ OthSO	2.135	0.144	0.835	2.612	0.106
FC $\times$ OthFC	0.849	0.357	0.104	0.026	0.873
OthSO $\times$ OthFC	5.933	0.015*	-0.524	1.040	0.308
SO $\times$ FC $\times$ OthSO	1.398	0.237	-0.936	1.792	0.181
SO $\times$ FC $\times$ OthFC	0.384	0.536	-0.560	0.400	0.527
SO $\times$ OthSO $\times$ OthFC	0.008	0.929	-0.158	0.046	0.830
FC $\times$ OthSO $\times$ OthFC	0.032	0.859	-0.315	0.149	0.699
SO $\times$ FC $\times$ OthSO $\times$ OthFC	0.157	0.692	0.411	0.157	0.692

Table G.3: GEE onset delay by non-responding recipient - complete model

IV	Model effects		Parameter Estimates		
	Wald $\chi^2$	p	B	Wald $\chi^2$	p
Same/Other (SO)	7.333	0.007**	-0.586	6.386	0.012*
Floor Change (FC)	2.356	0.125	-0.242	0.728	0.394
Within/Between (WB)	2.108	0.147	-0.212	0.145	0.703
SO $\times$ FC	0.001	0.976	0.014	0.001	0.971
SO $\times$ WB	0.001	0.978	0.016	0.001	0.980
FC $\times$ WB	0.408	0.523	-0.244	0.155	0.694
SO $\times$ FC $\times$ WB	0.005	0.945	-0.057	0.005	0.945

QIC = 132.301; QICC = 132.910, unstructured correlation matrix

Table G.4: GEE on typing time of response with constituency

IV	Model effects		Parameter Estimates		
	Wald $\chi^2$	p	B	Wald $\chi^2$	p
Same/Other (SO)	7.970	0.005**	-0.568	5.590	0.012
Floor Change (FC)	0.464	0.496	0.006	0.001	0.982
Within/Between (WB)	0.532	0.466	0.079	0.020	0.888
SO $\times$ FC	0.003	0.960	-0.102	0.067	0.796
SO $\times$ WB	0.101	0.750	-0.246	0.147	0.702
FC $\times$ WB	0.367	0.545	-0.374	0.316	0.574
SO $\times$ FC $\times$ WB	0.024	0.878	0.147	0.024	0.878

QIC = 138.200; QICC = 135.581, unstructured correlation matrix

Table G.5: GEE on number of characters with constituency

IV	Model effects		Parameter Estimates		
	Wald $\chi^2$	p	B	Wald $\chi^2$	p
Same/Other (SO)	0.017	0.897	0.053	0.045	0.832
Floor Change (FC)	2.471	0.116	0.218	0.450	0.502
Within/Between (WB)	0.723	0.395	0.047	0.028	0.867
SO $\times$ FC	0.412	0.521	-0.305	0.493	0.482
SO $\times$ WB	0.462	0.497	-0.252	0.397	0.529
FC $\times$ WB	1.476	0.224	-0.056	0.013	0.909
SO $\times$ FC $\times$ WB	2.053	0.152	0.997	2.053	0.152

QIC = 70.485; QICC = 75.241, exchangeable correlation matrix

Table G.6: GEE on onset delay with constituency

IV	Model effects		Parameter Estimates		
	Wald $\chi^2$	p	B	Wald $\chi^2$	p
Onset delay; QIC = 55.998; QICC = 69.142					
Oth Same/Other (SO)	13.130	<0.001**	0.953	10.139	0.001**
Oth Floor Change (FC)	16.720	<0.001**	0.995	13.490	<0.001**
Antecedent	0.508	0.476	-0.358	2.118	0.146
Continuation	1.041	0.308	0.372	1.956	0.162
Oth SO $\times$ Oth FC	7.016	0.008**	-0.793	7.016	0.008**
Oth SO $\times$ Ant	0.659	0.417	0.231	0.659	0.417
Oth SO $\times$ Cont	1.427	0.232	-0.410	1.427	0.232
Oth FC $\times$ Ant	0.183	0.669	0.106	0.183	0.669
Oth FC $\times$ Cont	0.540	0.462	-0.210	0.540	0.462
Ant $\times$ Cont	0.392	0.531	0.155	0.392	0.531

Table G.7: GEEs of onset delay by other same/other, other floor change, antecedent end-completeness and continuation start-completeness

## Appendix H

### Entropy values

---

POS Tag	Description	POS Entropy	Lexical Entropy	Number of tokens
CC	Coordinating conjunction	3.83	7.07	1624
DT	Determiner	2.40	8.14	4658
IN	Preposition or subordinating conjunction	3.45	6.65	4435
JJ	Adjective	3.29	8.06	2461
MD	Modal	1.44	6.28	1677
NN	Noun, singular or mass	3.72	7.45	8515
NNP	Proper noun, singular	3.97	6.81	1479
NNS	Noun, plural	3.86	6.85	1429
PRP	Personal pronoun	3.41	6.88	5019
RB	Adverb	4.16	7.87	4137
RP	Particle	3.40	5.59	392
TO	to	1.59	6.81	1203
VB	Verb, base form	3.99	7.10	3890
VBD	Verb, past tense	4.02	6.85	748
VBG	Verb, gerund or present participle	3.82	6.34	1081
VBN	Verb, past participle	3.82	6.27	660
VBP	Verb, non-3rd person singular present	4.03	7.27	2445
VBZ	Verb, 3rd person singular present	3.99	7.32	2327

Table H.1: POS and lexical entropy values used in experiment 3

## Appendix I

### Experiment 3: Additional Statistical Analyses

In simple slopes analyses, the model is rerun with covariates held at different values to see if the interaction is significant for another main effect at these different levels. For these purposes, low entropy is 1 s.d. below the mean and high is 1 s.d. above. As can be seen in table I.1, higher order interaction effects and effects of the raised or lowered variable do not change. These will be omitted from subsequent tables.

Condition	IV	Model effects	
		Wald $\chi^2$	p
High lex	Lexical entropy	2.721	0.099
	POS entropy	7.446	0.006**
	Lex $\times$ POS entropy	5.893	0.015*
Low lex	Lexical entropy	2.721	0.099
	POS entropy	0.040	0.842
	Lex $\times$ POS entropy	5.893	0.015*
High POS	Lexical entropy	5.840	0.016*
	POS entropy	2.778	0.096
	Lex $\times$ POS entropy	5.893	0.015*
Low POS	Lexical entropy	0.692	0.406
	POS entropy	2.778	0.096
	Lex $\times$ POS entropy	5.893	0.015*

Table I.1: Simple slopes analysis of response or not by lexical entropy, POS entropy

Condition	IV	B	Model effects	
			Wald $\chi^2$	p
High POS	Lex entropy	-0.012	2.136	0.144
Low POS	Lex entropy	0.163	4.720	0.030*
High lex	Antecedent Completeness	0.010	0.022	0.882
	POS entropy	-0.032	0.127	0.722
Low lex	Antecedent Completeness	0.480	9.189	0.002*
	POS entropy	-0.207	3.626	0.057
Ant Comp = y	Lex entropy	0.075	6.445	0.011*
Ant Comp = n	Lex entropy	-0.046	1.873	0.171

Table I.2: Simple slopes analysis of onset delay by lexical entropy, POS entropy

Condition	IV	Parameter Estimates		
		B	Wald $\chi^2$	p
High lex	Antecedent Completeness	-2.015	8.728	0.003**
Low lex	Antecedent Completeness	0.480	0.689	0.407
Ant Comp = y	Lex entropy	0.955	11.073	0.001**
Ant Comp = n	Lex entropy	-0.293	3.075	0.080

Table I.3: Simple slopes analysis of CC responses by lexical entropy, POS entropy

Condition	IV	Model effects		
		B	Wald $\chi^2$	p
Basic Model	Line Number		3.154	0.076
	Topic		8.784	0.003*
	Lex entropy		1.460	0.227
	POS entropy		0.540	0.463
	Topic $\times$ Lex entropy		4.061	0.044*
	Topic $\times$ POS entropy		3.044	0.081
	Lex entropy $\times$ POS entropy		1.413	0.235
	Topic $\times$ Lex $\times$ POS		5.634	0.018*
High POS	Topic $\times$ lex entropy	-0.061	0.014	0.905
	Lex entropy	0.028	<0.001	0.992
Low POS	Topic $\times$ lex entropy	-1.470	10.794	0.001**
	Lex entropy	0.372	4.054	0.044*
High lex	Topic $\times$ POS entropy	1.540	6.462	0.011*
	POS entropy	-0.403	1.191	0.275
Low lex	Topic $\times$ POS entropy	0.131	0.064	0.800
	POS entropy	-0.059	0.001	0.979
High POS, high lex	Topic	-0.376	3.154	0.076
High POS, low lex	Topic	-0.253	0.211	0.646
Low POS, high lex	Topic	-3.455	13.634	<0.001**
Low POS, low lex	Topic	-0.514	0.455	0.500

Table I.4: Simple slopes analysis of onset delay by lexical entropy, POS entropy and topic

## Appendix J

### Dynamic Syntax Rules

(J.1) AXIOM

$$\{\dots?Ty(t)\dots\diamond\}$$

(J.2) INTRODUCTION

$$\frac{\{\dots\{\dots?Ty(Y)\dots\diamond\}\dots\}}{\{\dots\{\dots?Ty(Y),?\langle\downarrow_0\rangle Ty(X),?\langle\downarrow_1\rangle Ty(X \rightarrow Y),\dots\diamond\}\dots\}}$$

(J.3) INTRODUCTION - SUBJECT AND PREDICATE

$$\frac{\{\dots\{Tn(n),?Ty(t),\diamond\}\}}{\{\dots\{Tn(n),?Ty(t),?\langle\downarrow_0\rangle Ty(e),?\langle\downarrow_1\rangle Ty(e \rightarrow t),\diamond\}\}}$$

(J.4) PREDICTION

$$\frac{\{\dots\{Tn(n),\dots,?\langle\downarrow_0\rangle\phi,?\langle\downarrow_1\rangle\psi,\diamond\}\dots\}}{\{\dots\{Tn(n),\dots,?\langle\downarrow_0\rangle\phi,?\langle\downarrow_1\rangle\psi\},\{\langle\uparrow_0\rangle Tn(n),?\phi,\diamond\},\{\langle\uparrow_1\rangle Tn(n),?\psi\}\dots\}}$$

(J.5) PREDICTION - SUBJECT AND PREDICATE

$$\frac{\{\dots\{Tn(0),?\langle\downarrow_0\rangle Ty(e),?\langle\downarrow_1\rangle Ty(e \rightarrow t),\diamond\}\}}{\{\dots\{Tn(0),?\langle\downarrow_0\rangle Ty(e),?\langle\downarrow_1\rangle Ty(e \rightarrow t)\},\{\langle\uparrow_0\rangle Tn(0),?Ty(e),\diamond\},\{\langle\uparrow_1\rangle Tn(0),?Ty(e \rightarrow t)\}\}}$$

(J.6) THINNING

$$\frac{\{\dots\{\dots,\phi,\dots,?\phi,\dots,\diamond\}\dots\}}{\{\dots\{\dots,\phi,\dots,\diamond\}\dots\}}$$

(J.7) COMPLETION

$$\frac{\{\dots\{Tn(n),\dots\},\{\langle\mu^{-1}\rangle Tn(n),\dots,Ty(X),\dots\diamond\}\dots\}}{\{\dots\{Tn(n),\dots,\langle\mu\rangle Ty(X),\dots,\diamond\},\{\langle\mu^{-1}\rangle Tn(n),\dots,Ty(X),\dots\}\dots\}} \\ \mu^{-1} \in \{\uparrow_0, \uparrow_1, \uparrow_*, L^{-1}\}, \mu \in \{\downarrow_0, \downarrow_1, \downarrow_*, L\}.$$



(J.8) ANTICIPATION

$$\frac{\{\dots\{Tn(n), \dots, \diamond\}, \{\langle \uparrow \rangle Tn(n), \dots, ?\phi, \dots\} \dots\}}{\{\dots\{Tn(n), \dots\}, \{\langle \uparrow \rangle Tn(n), \dots, ?\phi, \dots, \diamond\} \dots\}}$$

(J.9) ELIMINATION

$$\frac{\{\dots\{Tn(n), \dots, ?Ty(X), \langle \downarrow_0 \rangle (Fo(\alpha), Ty(Y)), \langle \downarrow_1 \rangle (Fo(\beta), Ty(Y \rightarrow X)), \dots, \diamond\}, \dots\}}{\frac{\{\dots\{Tn(n), \dots, ?Ty(X), Fo(\beta(\alpha)), Ty(X), \langle \downarrow_0 \rangle (Fo(\alpha), Ty(Y)), \langle \downarrow_1 \rangle (Fo(\beta), Ty(Y \rightarrow X)), \dots, \diamond\}, \dots\}}{\text{Condition: } \langle \downarrow_i \rangle ?\phi, i \in \{1, 0\}, \text{ does not hold.}}}$$

(J.10) \*ADJUNCTION:

$$\frac{\{\dots\{\{Tn(a), \dots, ?Ty(t), \diamond\} \dots\}}{\{\dots\{\{Tn(a), \dots, ?Ty(t), \}, \{\langle \uparrow_* \rangle Tn(a), ?\exists \mathbf{x}. Tn(\mathbf{x}), \dots, ?Ty(e), \diamond\} \dots\}}}$$

(J.11) MERGE:

$$\frac{\{\dots\{\dots \mathbf{DU}, \mathbf{DU}' \dots\} \dots\}}{\{\dots\{\dots \mathbf{DU} \sqcup \mathbf{DU}' \dots\} \dots\}} \\ \diamond \in \mathbf{DU}'$$

(J.12) LATE \*ADJUNCTION:

$$\frac{\{Tn(n), \dots, \{\uparrow_* Tn(n), Tn(a), \dots, Ty(X), \diamond\}, \dots\}}{\{Tn(n), \dots, \{\uparrow_* Tn(n), Tn(a), \dots, Ty(X)\}, \{\langle \uparrow_* \rangle Tn(a), ?Ty(X), \exists \mathbf{x}. Tn(\mathbf{x}), \diamond\}, \dots\}}$$

(J.13) LINK ADJUNCTION:

$$\frac{\{\dots\{\{Tn(a), Fo(\alpha), Ty(e), \diamond\} \dots\}}{\{\dots\{\{Tn(a), Fo(\alpha), Ty(e)\}, \{\langle L^{-1} \rangle Tn(a), ?Ty(t), \langle \downarrow_* \rangle Fo(\alpha), \diamond\} \dots\}}}$$

(J.14) LINK EVALUATION (Non-restrictive construal<sup>1</sup>):

$$\frac{\{\dots\{Tn(a), \dots, Fo(\phi), Ty(t), \diamond\}, \{\langle L^{-1} \rangle MOD(Tn(a)), \dots, Fo(\psi), Ty(t)\}}{\{\dots\{Tn(a), \dots, Fo(\phi \wedge \psi), Ty(t), \diamond\}, \{\langle L^{-1} \rangle MOD(Tn(a)), \dots, Fo(\psi), Ty(t)\}} \\ MOD \in \{\langle \uparrow_0 \rangle, \langle \uparrow_1 \rangle\}^*$$

---

<sup>1</sup>See Cann et al. (2005), Chapter 3, for explanation and details of restrictive relative clauses.