Title
The Distribution of Repair in Dialogue

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 33(33)

ISSN
1069-7977

Authors
Colman, Marcus
Healey, Patrick

Publication Date
2011

Peer reviewed

The Distribution of Repair in Dialogue

Marcus Colman and Patrick G.T. Healey
{**marcus**,ph}@eecs.qmul.ac.uk
Queen Mary, University of London,
Interaction, Media and Communication Research Group,
School of Electronic Engineering and Computer Science,
London E1 4NS UK

Abstract

Repairs –the various ways in which people edit and reformulate conversational turns– are a characteristic feature of natural dialogue. However, relatively little is known about their overall frequency or distribution in conversation. We present a systematic, quantitative study of patterns of repair in two corpora: 'ordinary' dialogues from the British National Corpus (BNC) and task-oriented dialogues from the HCRC Map Task. We use this analysis to evaluate three hypotheses about patterns of repair 1) social 'preferences' 2) processing demands and 3) dialogue co-ordination. The results show that repair is more frequent in task-oriented dialogue, that use of repair is broadly unaffected by familiarity or mode of interaction but substantially affected by task roles. We argue that the complimentary patterns of repair used by conversational partners support the view of repairs as an integrated, cross-turn and cross-person, system for sustaining the mutual-intelligibility of dialogue.

Keywords: Communication; Dialogue; Repair; Conversation Analysis.

Introduction

Models of human communication in the cognitive sciences often idealise to a situation in which people are assumed to 'speak the same language' i.e. are assumed to be linguistically transparent to one another. However, in practice, "no two speakers of the same language ever speak exactly the same dialect of that language" (Fodor and Lepore, 1992, p.10). Dialogue always proceeds against a background of idiolect differences. Place of residence, cultural group, hobbies, occupation, education and age group all feed into differences in the way we use language (e.g., Clark; 1996). Detecting and dealing with differences in interpretation is thus a recurrent and routine problem in conversation.

Conversation analysts have described a structured set of repair mechanisms that people can use to signal problems, or possible problems, with interpretation (e.g. Schegloff, Jefferson and Sacks, 1977; Schegloff, 1987, 1992). The conversation analytic (CA) repair structures are built around three main points a) *initiation*: who signals a problem - whether it is the speaker of a problem turn ('self ') or a recipient of it ('other') b) *repair*: who actually makes a change to the problem turn (self or other) c) *position*: where in the conversational sequence these events occur; in the same turn as the problem, in the turn after the problem turn or in some subsequent turn.

The different possible trajectories from each potential 'problem' turn, through the initiation of a repair to its completion are part of an integrated repair 'space' that spans multiple turns and multiple participants (Schegloff 1992).

This paper addresses two basic questions about the structure of this repair space:

1. What is the typical distribution of repair in this space?
2. What factors explain this distribution?

The first question is about quantifying the distribution of different repair types. It has not, to our knowledge previously been addressed (although some studies have looked at the distribution of within-turn repairs e.g. Bortfeld, et al. 2001; Lickley and Bard, 1998). This gap is partly because CA focuses on the particular sequential consequences of repair in specific cases, not statistical distributions.[1]

The second question is about the kinds of explanation offered for the distribution of repairs. We identify three broad, non-exclusive, hypotheses.

Hypothesis 1. **Social Preference**: In CA, the structure of the repair space is primarily explained in terms of 'preference'; what a competent speaker might consider to be conversationally or socially appropriate. This leads to an ordering of different repair trajectories such that self-initiation is preferred over other-initiation and self-repair over other-repair (see examples below). This preference hierarchy operates independently of the type of conversational task or type of problem encountered (Schegloff, Jefferson and Sacks, 1977).

Hypothesis 2. **Processing Demands**: Repair phenomena may also be a side-effect of the difficulty of producing or comprehending language because of processing load. For example, Clark and Wasow (1998) show that self-repairs are more likely where people are producing more complex constituents. Most of the experimental evidence for this applies to first position self-repairs such as repeats and restarts (e.g. Branigan et al, 1999; Lickley and Bard, 1998) but could be generalized to at least some aspects of other forms. For example, the delay and markers that often precede a third position repair might simply reflect

[1] In fact, Schegloff (1993) specifically argues that CA phenomena are not suitable for quantitative analysis. For example, counting the frequency of occurrences of "what?", obscures differences in the complex variety of things that "what?" can be used to do in conversation (see Drew, 1997). This is true but it does not preclude the possibility that useful generalizations can be made by looking at the overall distribution of 'whats'. We also note that repair is unusual in that it can occur freely in any context, unlike e.g. laughter or sobbing (c.f. Schgeloff, 1993).

increased processing time. On this account repairs are essentially accidental 'disfluencies' or 'noise' although they may, nonetheless carry useful information for a hearer (c.f. Brennan and Schober, 2001).

Hypothesis 3. Dialogue Co-ordination: The distribution of repairs in dialogue might also be explained by the operation of structured co-ordination mechanisms. In this case, the structure of repair is not dependent on the social considerations or the ostensible problem per se, but on the joint activities people engage in to maintain the mutual-intelligibility of conversation in response to a problem. Bortfeld et al. (2001) present evidence for the role of position 1 self-repairs (repeats, restarts and fillers) in co-ordination. Here we consider the larger repair space. Healey (2008) has argued that the structure and operation of this space underpins all effective dialogue and directly constrains the form of shared symbol systems in general (see also Healey et al. 2007).

This paper evaluates these hypotheses through a quantitative, corpus analysis of the distribution of repair trajectories in two different dialogue corpora: the British National Corpus (BNC) and the HCRC Map Task Dialogues. These corpora make it possible to answer the following questions: How frequent is repair in ordinary conversation? What is the overall distribution of repair types? Do ordinary conversation and task-oriented dialogue differ in the frequency or distribution of repairs? Are patterns of repair affected by visual contact, by levels of familiarity or task roles?

Methods

To obtain reliable measures of the frequency of different repair types the repair protocol described in Healey et al. (2005) was used. This is designed to provide an index of the frequency of several basic CA categories of repair on the basis of surface characteristics alone. It consists of a binary branching decision tree of yes/no questions that are applied to each contribution to an interaction. These questions apply only to the surface form of the utterance. For example, the following questions are used to identify the position of a repair:

- Position 1: "Does the initiator edit, amend or reprise part of their contribution before another participant responds to it?"
- Position 2: "Is this contribution introduced to propose repetition or revision of another participant's contribution?"
- Position 3: "Is this contribution introduced to edit, amend or reprise a previous contribution by the initiator?"

We note that the repair protocol departs from the CA analysis in several respects. Categories such as embedded repair are not coded because they are not explicitly signaled as repairs and other categories such as position 4 repair are ignored because they are too rare to provide a useful comparative measure.

For ease of exposition and to facilitate comparison with previous work (e.g. Bortfeld et al. 2001) we provide glosses, detailed in

Table 1 for some of the CA terminology used in the protocol. These glosses are only intended as shorthand for the categories defined by the protocol and not intended as technical terms.

Table 1: Terminology

Gloss	**Repair Protocol Category**
Repeat	Position 1 Self-Initiated Self-repair 'Articulation'
Restart	Position 1 Self-Initiated Self-repair 'Formulation'
Transition	Position 1 Self-Initiated Self-Repair in Transition Space
Clarification Request (CR)	Position 2 Next Turn Repair Initiator (NTRI)
Correction	Position 2 Other-Initiated, Other-Repair
Follow-up	Position 3 Other-Initiated, Self-Repair
Reformulate	Position 3 Self-Initiated Self-Repair.

The following examples, taken from the HCRC map task dialogues (Anderson et al, 1991; see below for description) illustrate some of the phenomena of interest:

(1)Q3EC1. *Repeat* / P1 SI SR Articulation
Follower: which is due we-- … **due west?**

(2)Q1EC6. *Restart* / P1 SI SR Formulation
Giver: so you're underneath them … **between them**

(3)Q1EC4. *Transition* / P1 SI SR Transition Space
Giver: and, start going down Southeast … … … **you go past a pine forest on your right**

Position two corresponds to the first response, by another person, to a potentially problematic contribution (typically, but not always, in the next turn). The two most common types of event in these situations are Position 2 Next Turn Repair Initiators, glossed here as 'Clarification Requests' (CR), and Position 2 Other-Initiated, Other-Repairs which we gloss here as 'Corrections'.

(4)Q1EC6. *Clarification Request* / P2 NTRI
Giver: past a forge on your right?
Follower: **past a what?**

(5)Q1EC8. *Correction* / Position 2 OIOR.
Giver: right to the very end of … paper
Follower: **the very end of the map?**

(6)Q3EC1. *CR and Follow-up* / P2NTRI – P3OISR
Follower: so you want me to go … **east … then south?**
Giver: **no, south then east**, we may have a different map

Intuitively a Clarification Request (CR / P2NTRI) is likely to prompt a Follow-up (P3OISR)[2]. The final major repair type we consider is reformulation or position three self-initiated self-repair (P3SISR). This commonly occurs where someone recognizes something problematic about someone else's interpretation of one of their own preceding contributions. They then attempt to rephrase or repeat their original contribution to address this.

(7)Q1EC8. *Reformulate* / P3SISR

Giver: right … now, have you got the hot wells?
Follower: they're over a bit
Giver: **or hot springs?**

The British National Corpus.

To provide a baseline analysis of the relative frequency of different repair types in ordinary conversation a subset of dialogues from demographic portion of the British National Corpus (BNC) were used (Burnard, 2000). This consists of tape recordings of unscripted informal conversations in contexts ranging from business meetings to domestic interactions. The recordings are made by volunteers from a demographically balanced sample of ages, regions and social classes.

Thirty dialogues were randomly selected for analysis from the BNC. These consisted of 1934 utterances, 14,034 words produced by 41 people.

The HCRC Map Task Corpus

The HCRC map task corpus (Anderson et al, 1991) consists of dialogues between pairs, seated on opposite sides of a table who each have a copy of a schematic map (their view of each other's maps is obscured). On one person's map, the 'Giver's, there is a route marked which they attempt to describe to their partner the 'Follower' in enough detail to allow them to draw it onto their own map. Some landmarks on the maps are shared and some are not.

In half of these dialogues eye contact is possible between the participants; in the other half a screen is placed between them. In addition familiarity is manipulated so that half the pairs are familiar and half unfamiliar. An advantage of this task is that it provides a sensitive measure of task outcome – the difference or deviation score between the original given route and the one drawn by the Follower. All 128 dialogue transcripts from the published HCRC map task corpus were coded for repair. This corpus consists of 19,133 turns and 156,315 words.

Results

The results are reported in four stages; first the generic comparison of patterns of repair in ordinary and task-oriented dialogues. Secondly, comparison of the effects of modality (+/- Eye Contact) and familiarity on repair in the task-oriented dialogues. Thirdly a comparison of the effects of task-role. Lastly, an analysis of the effects of repair on outcome.

Repair in Ordinary and Task-Oriented Dialogue:

Figure 1 illustrates the baseline distribution of repair types found in ordinary dialogue (BNC). In the 1934 turns (14,034 words) in this sample 390 repairs were coded. This gives an approximate frequency of repair in ordinary dialogue of 1 event every 36 words or every 5 turns.

The pattern in Figure 1 is consistent with the general CA claim (Hypothesis 1) that earlier repair initiation and repair are preferred, that self-initiation is preferred over other-initiation and that self-repair is preferred over other-repair.

Although conversation lengths vary more in the BNC than the HCRC Map Task corpus average turn length is broadly comparable (BNC: 7.2 words, Map Task, 8.2 words).

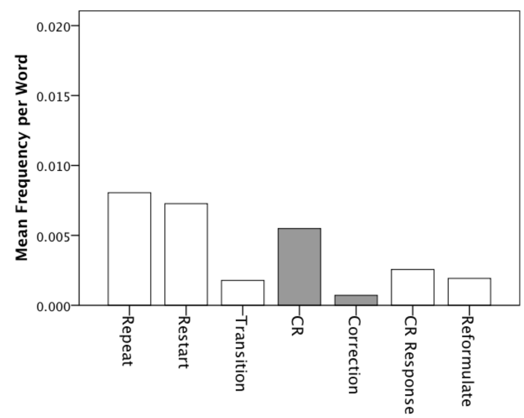

Figure 1: Repair Distribution in Ordinary Conversations (BNC). Light Bars = 'Self', Grey Bars = 'Other'

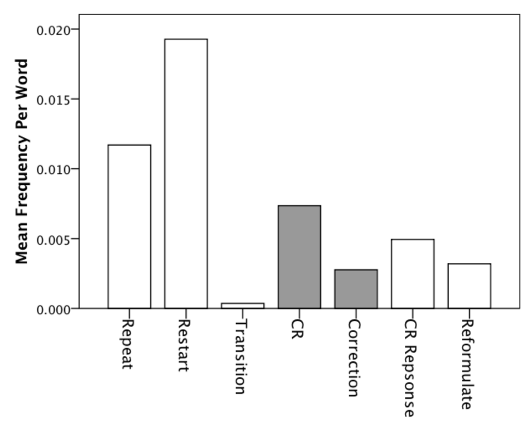

Figure 2: Repair Distribution in Task-Oriented Dialogues (Map Task). Light Bars = 'Self', Grey Bars = 'Other'

Figure 2 illustrates the distribution of repair types found in the Map Task. Comparison with Figure 1 suggests several contrasts. First, repairs in the Map task are approximately twice as frequent as in the BNC. A total of 7,977 repair events were coded giving an average frequency of more than 1 event every 2.5 turns or one every 20 words. This suggests that the arguably greater processing demands of the

[2] Note that this diverges from Schegloff's terminology who does not count the response to NTRI as a position 3 repair.

task-oriented dialogues are reflected in more repair (Hypothesis 2). Although there is a higher overall level of repair, relative use of different positions in the 'repair space' (positions 1, 2 and 3) is not reliably different in the ordinary conversations and the task-oriented dialogues $Chi^2_{(2)}= 1.01$, p = 0.32[3].

Three focused Chi^2 analyses tested for differences in the use of specific repair types. The relative balance of repeats and restarts (P1SISR) is reliably different: $Chi^2_{(1)}= 14.2$, p<0.01. Restarts are used more often in the task-oriented dialogues. People also make proportionally greater use of direct correction (i.e. P2OIOR) than requests for clarification (P2NTRI) in the task oriented dialogues: $Chi^2_{(1)}= 11.5$, p<0.01. However, the relative proportion of CR's to Follow-ups is not reliably different $Chi^2_{(1)}= 3.14$, p=0.07.

To summarise, people do substantially more repair in the task-oriented dialogues and make more use of restarts and direct corrections. However, they are equally likely to respond to a clarification request in either case.

Effects of Modality and Familiarity:

The distribution of repairs in the familiar and unfamiliar pairs and in the Face-to-Face vs. No Eye Contact dialogues is illustrated in Figure 3.

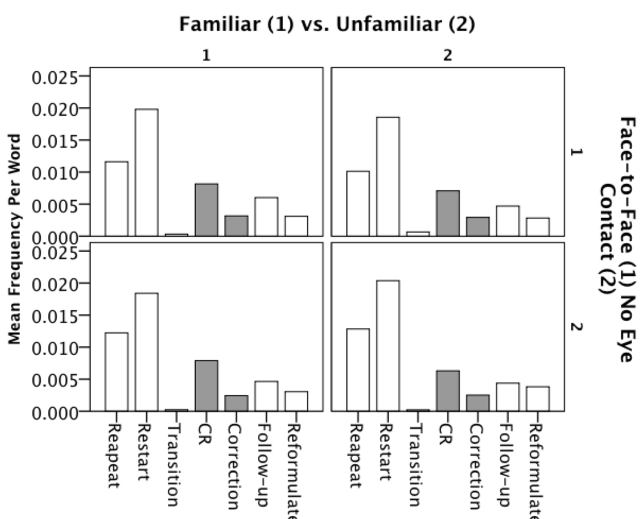

Figure 3: Repair Distribution by Modality and Familiarity.

As Figure 3 suggests, the level and distribution of repair types is relatively unaffected by familiarity and by mode of interaction. Mann-Whitney U Tests on each of the repair categories in Figure 3 comparing Familiar vs. Unfamiliar pairs shows only one reliable difference (see Table 2). Familiar pairs are more likely to use clarification requests (NTRI's) than unfamiliar pairs, consistent with Hypothesis 1.

[3] A criterion level of p < 0.05 is adopted for all statistical tests. Exact probabilities are reported for completeness.

The parallel analysis of Face-to-Face vs. No Eye Contact shows only one reliable difference, in this case in the use of transition space repairs. Participants in the Face-to-Face condition make more use of transition space repairs than those in the no eye-contact condition. Intuitively this makes sense since when people can see each other they may be able to detect some problems with interpretation by looking at their partners when they have finished their turn. However, the absolute frequencies are small (face-to-face n=31, no eye contact n=12) and therefore this result must be treated with caution.

Table 2: Repairs by Familiar and Unfamiliar Pairs (Mann Whitney U)

Repair Type	P
Repeat (P1SISR: Articulation)	0.50
Restart (P1SISR: Formulation)	0.82
Transition (SISR: Transition Space)	0.94
CR (P2 NTRI)	0.04 *
Correction (P2 OIOR)	0.82
Follow-up (P3 OISR)	0.19
Reformulate (P3 SISR)	0.68

Table 3: Repairs by Face-to-Face and No Eye Contact Pairs (Mann Whitney U)

Repair Type	P
Repeat (P1SISR: Articulation)	0.50
Restart (P1SISR: Formulation)	0.76
Transition (SISR: Transition Space)	0.00 *
CR (P2 NTRI)	0.54
Correction (P2 OIOR)	0.19
Follow-up (P3 OISR)	0.15
Reformulate (P3 SISR)	0.21

In summary neither familiarity nor modality have strong overall effects on the use of repair by participants. Familiar pairs show a tendency to use more clarification requests, possibly indicating a greater willingness to signal problems i.e. initiate repairs. Face-to-face pairs make more use of transition space repairs suggesting some advantage to being able to visually monitor how a contribution is received.

Effects of Task Role:

As noted, the Map Task has two task roles. The Giver has a route that needs to be described so the Follower can draw it. Roles thus index a basic informational asymmetry that is useful for testing Hypothesis 2. As Figure 4 and Figure 5 show, this difference in role is reflected in different distributions of repair types used.

Wilcoxon tests show that only frequency of repeats is not reliably different across roles (p=0.13). For Clarification Requests (p<0.01) and Corrections (p<0.01) the Follower's mean frequency is significantly higher. For all the other categories: Restarts (p<0.01), Transition (p=0.02), Follow-ups (p<0.01) and Reformulations (p<0.01) the Giver's mean frequency is higher.

As Figure 4 and Figure 5 indicate there is a significant complementarity in the use of repairs in the two roles. This is confirmed by the inter-correlations between Giver and Follower repair types for word count ratios given in Table 4 (only significant associations are reported).

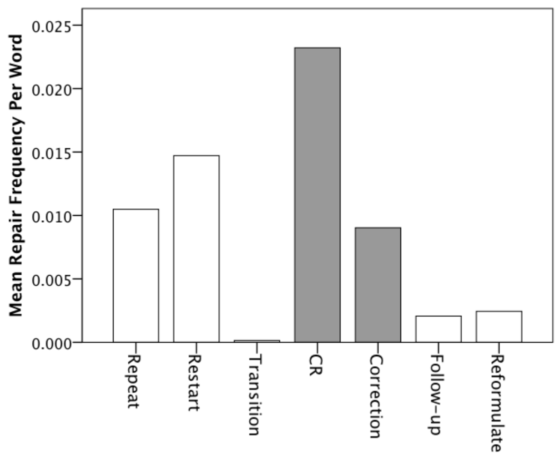

Figure 4: Distribution of Repairs Used by Route Followers

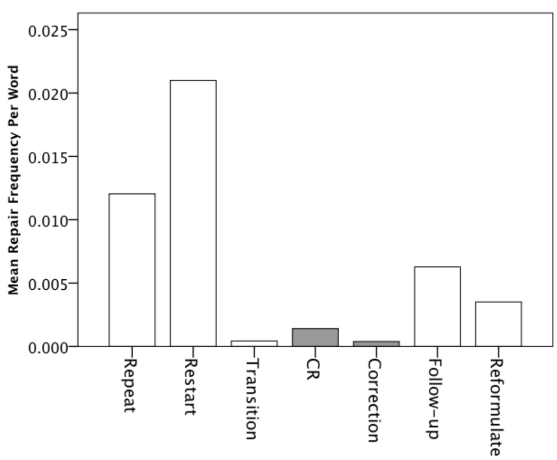

Figure 5: Distribution of Repairs Used by Route Givers

Table 4: Inter-correlations in Use of Repair Types.

Repair types	rho	p
G Restart – F Restart	.293	.001
G Transition – F Transition	.310	.000
G Correction – F Repeat	.189	.033
G Correction – F CR	-.261	.003
G Correction – F Follow-up	.308	.000
G CR – F Repeat	.174	.049
G CR – F CR	-.266	.002
G CR – F Follow-up	.790	.000
G Follow-up – F CR	.655	.000
G Follow-up – F Correction	.214	.015

Task Performance:

Step-wise linear regression provides an assessment of the extent to which the different repair types independently impact on task performance. Route deviation score is the dependent variable and familiarity (Y/N), modality (+/- Eye Contact), and frequency of Giver and Follower repair types results in a model with three predictors: 1. Familiarity (Standardised Beta=.688, t=6.95, p<0.01) with more familiar pairs tending to produce more accurate routes (64cm^2 deviation vs. 79cm^2). 2. Medium (Standardised Beta= -.464, t=-5.03, p<0.01) with audio only pairs tending to produce more accurate routes than face-to-face pairs (69 cm^2 vs. 74cm^2) and 3. Giver's restart repairs (*Formulation* / P1SISR: Standardised Beta=-.344, t=3.26, p=0.01) with more restarts producing more accurate routes.

Discussion

The results provide some support for the Social Preference explanation of repair distribution (Hypothesis 1). The most frequent repair types in both corpora are position 1 self-initiated repairs (repeats and restarts). There is also a clear statistical preference for other-initiation (CR /NTRI) over other-repair (Correction/OIOR) in position two. The quantitative analysis also reveals that this difference is moderated by familiarity. Where the Map Task participants know each other they use reliably more explicit prompting and clarification (NTRI).

One clear source of support for the processing demands explanation (Hypothesis 2) is the finding that the frequency of repairs in the task-oriented dialogues is approximately double that in the ordinary conversations in the BNC. It seems likely that the need to communicate relatively fine-grained information about routes and landmarks places significant demands on participants. This produces the predicted increase in number of repairs. This finding is difficult to explain in terms of social preferences although it might be argued that the experimental situation of the map task itself changes what people feel is or is not appropriate.

The more specific finding that transition space repairs are used more in the eye contact condition suggests that people take advantage of the available non-verbal signals to judge whether they are making themselves clear or not, and adjust their contribution accordingly. Again, this seems more consistent with the management of processing demands (Hypothesis 2) than the operation of social preferences.

The main support for the explanation of repair structure in terms of dialogue coordination mechanisms (Hypothesis 3) comes from the more detailed evidence of inter-dependencies in choice of repair type across the two different roles. It is worth noting that all participants acted both as Givers and Follower in the HCRC Map task so individual differences are directly controlled for.

It is clear that these two different task roles are associated with divergent patterns of repair. The Giver primarily works to produce intelligible descriptions (Position 1 repairs) and the Follower focuses on clarifying them (Position 2 repairs). Social preferences ought to be the same for both roles. Instead the results suggest that the informational asymmetry is directly reflected in the patterns of repair. This supports Hypothesis 2 (processing demands)

However, the complimentary nature of repair patterns in Figure 4 and Figure 5 provides strongest support for Hypothesis 3; dialogue co-ordination. There is a close inter-dependency between the different trajectories in the repair space that conversational partners use as shown in the correlations in Table 4. For example, the more one participant requests clarification (NTRI) the less the other does. Also, the more the Giver offers possible corrections (P20I0R) the less often the Follower needs to clarify. This suggests that participants are making systematic and coordinated trade-offs in their use repair, thereby reducing the joint effort needed to communicate effectively (Clark, 1996).

Conclusion

The results of this study show that repairs are a pervasive feature of dialogue. They also provide useful quantitative evidence about the basic structure and function of the repair space. Repairs are not just noise in the signal or 'disfluencies'. They pattern in systematic ways and these patterns are not easily explained by appeal to social preferences or processing demands alone. As previous work has shown, the surface form of self-repairs can provide useful information for co-ordination in dialogue. The results reported here show the larger scale cross-person and cross-turn patterns that structure the 'repair-space'. This space, we have argued, is used as a coordinated, integrated system by all the parties to an interaction and, we claim, is critical to maintaining the mutual-intelligibility of dialogue.

Acknowledgments

This work was supported by the EPSRC DiET Project (EP/D057426/1). Thanks especially go to Amy Isard, Jean Carletta, Arash Eshghi and Greg Mills for their correspondence and help.

References

Anderson, A.; Bader, M.; Bard, E.; Boyle, E.; Doherty, G. M.; Garrod, S.; Isard, S.; Kowtko, J.; McAllister, J.; Miller, J.; Sotillo, C.; Thompson, H. S. & Weinert, R. (1991) The HCRC Map Task Corpus. In: Language and Speech 34351-366.

Bortfeld, H. Leon, S.D., Bloom, E., Schober, M.F., Brennan, S. (2001) Disfluency Rates in Conversation: Effects of Age, Relationship, Topic, Role, and Gender. *Language and Speech,* 434 (23), 12293–124579.

Branigan, H., Lickley, R. and McKelvie, D. (1999). Non-linguistic influences on rates of disfluency in spontaneous speech. In *Proceedings of the 14th International Conference of Phonetic Sciences*, 1999.

Brennan, S and Schober M. F. (2001) How Listeners Compensate for Disfluencies. Journal of Memory and Language, 44, pp. 274-296.

Burnard (2000). Reference Guide for the British National Corpus (World Edition). Oxford University Computing Services.

Carletta, J. C., Isard, A., Isard, S., Kowtko, J., Doherty-Sneddon, G., & Anderson, A. (1996). *HCRC Dialogue Structure Coding Manual (HCRC/TR-82),* Human Communication Research Centre, University of Edinburgh.

Clark, H.H. and Wasow, T. (1998) Repeating Words in Spontaneous Speech. *Cognitive Psychology*, 37, 201–242.

Clark, H.H. (1996) *Using Language*. Cambridge University Press.

Drew, P. (1997). "'Open' class repair initiators in response to sequential sources of trouble in conversation." *Journal of Pragmatics* 28: 69-101.

Fodor, J. and Lepore, E. (1992) "E. Lepore, *Holism: A Shopper's Guide*. Blackwell.

Healey P.G.T. (2008) "Interactive Misalignment: The Role of Repair in the Development of Group Sub-languages" in Cooper R. and and Kempson R. (eds) *Language in Flux*. Palgrave-McMillan. p 13-39

Healey, P.G.T., Swoboda, N., Umata I. and King, J. (2007) "Graphical Language Games: Interactional constraints on representational form" Cognitive Science, 31, 285-309.

Healey, P. G. T., Colman, M. and Thirlwell, M. (2005). Analysing Multi-Modal Communication: Repair-Based Measures of Human Communicative Co-ordination. in *Natural, Intelligent and Effective Interaction in Multimodal Dialogue Systems*, Dordrecht: Kluwer Academic, 2005.

Kowtko, J. C., Isard, S. D. and Doherty, G. M. (1991). Conversational games within dialogue. In *Proceedings of the espirit workshop on discourse coherence.*

Lickley, R and Bard, E. (1998). When can listeners detect disfluency in spontaneous speech? *Language and Speech* 41.

Schegloff, E. A. (1987). Some Sources of Misunderstanding in Talk-in-Interaction, *Linguistics*, 25 (1987), 201-218.

Schegloff, E. (1992). Repair after the next turn: the last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology*, 97(5), 1295-1345.

Schegloff, E. A. (1993). Reflections on Quantification in the Study of Conversation. *Research on Language and Social Interaction,* 26:1, 1993, 99-128.

Schegloff, E. A. (1997). "Third Turn Repair," in G. R. Guy, C. Feagin, D. Schiffrin and J. Baugh (eds.), *Towards a Social Science of Language: Papers in honor of William Labov. Volume 2: Social Interaction and Discourse Structures*. Amsterdam: John Benjamins.

Schegloff, E., Jefferson, G. and Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, 53, 361-382.