The Nature of Common Ground Units: an Empirical Analysis Using Map Task Dialogues

Lesley Stirling*, Ilana Mushin*†, Janet Fletcher*, Roger Wales†

*University of Melbourne, Melbourne 3010, Australia
{l.stirling, i.mushin, j.fletcher}@linguistics.unimelb.edu.au
†La Trobe University, Melbourne 3083, Australia
r.wales@latrobe.edu.au

Abstract

This paper evaluates the meso-level 'Common Ground Units' (CGUs) proposed by Nakatani & Traum (1999) by applying this category to four dialogues from the Australian map task corpus. We first consider two formal characteristics of CGU boundaries: associations with following turn boundaries and associations with Intermediate and Intonational Phrases. We then give a profile of CGU initiating, grounding and final elements in terms of DAMSL dialogue act codes. We discuss some of the problems which arise in applying the category of CGU and conclude by proposing some parameters for consideration in a typology of CGUs.

1. Introduction

In this paper we consider the nature of 'Common Ground Units' (CGUs) as part of higher-level dialogue structure, by examining CGUs in four dialogues from the MAP TASK section of the Australian National Database of Spoken Language (ANDOSL). The work is part of a larger on-going study of the relationship between dialogue structure and prosodic structure, where we consider dialogue structure at different levels (micro-, meso- and macro-) and investigate whether and how dialogue segmentation is associated with various correlates of prosodic structure.

The ANDOSL Map Task (Millar et al., 1994) is closely modelled on the HCRC Map Task (Anderson et al., 1991). Participants worked in pairs, each with a map in front of them that the other could not see. One participant (the 'instruction-giver' IG) had a route marked on their map and was required by the task to instruct the other (the 'instruction-follower' IF) in drawing the correct route onto their own map. The maps differed to some degree in presence, position and names of landmarks. The four dialogues considered for this study were from two mixed-gender pairs, one pair who previously knew each other well and one pair who had never met before. Each pair produced one dialogue with the female as IG and one with the male.

'Grounding' is the process whereby information contributed by participants in a communicative interaction is mutually acknowledged as having entered the 'common ground', or shared knowledge of the participants (Clark & Marshall, 1981; Clark & Schaefer, 1989; Traum, 1994). This process takes place by virtue of a contribution being proposed by one participant and then evidence being given by the other that they have perceived and understood it: the evidence may be as minimal as 'proceeding as usual', may consist in a head-nod or other non-verbal cue, or may be an overt verbal acknowledgement or response. It is a dynamic process which can be modelled as a series of collaborative negotiations interspersed with 'moments' of grounding leading to change in the pragmatic and semantic status of the information considered to that point, as 'grounded' pragmatic and semantic information is added to the participants' assumed common ground (cf. Poesio & Traum, 1997).

Recently attention has been focused on the way in which the grounding process may interact with higher-level dialogue structure. Nakatani & Traum (1999) proposed a new coding scheme that applies simplified principles of grounding theory at the 'meso-' level of dialogue structure. 'Common Ground Units' (CGU's) are defined which they hypothesise might function as the minimal units for even higher-level dialogue units based on intentional/informational structure.

The status of CGUs is still under discussion (e.g. Core et al., 1999). In this paper we first give a formal profile of CGUs by examining how CGU beginning and end points correspond with measures of discontinuity in the speech signal such as turn boundaries and intonational phrase boundaries. We then give a functional profile of CGUs by examining the mapping between CGUs and dialogue acts labelled according to a version of DRI/DAMSL. These issues should contribute to our understanding the nature of CGUs, their status as dialogue units and their relationship to other kinds of dialogue structure.

2. Method

The dialogues were digitised for analysis at 22 kHz using Entropic's ESPS / Waves + speech analysis software running on a Sun Workstation in the Phonetics Laboratory of the University of Melbourne. A complete orthographic transcription of the dialogues was carried out.

As part of the larger study described above, a range of prosodic characteristics of the dialogues had been independently coded. Turn start and endpoints were labelled; in cases of speaker overlap, it was still possible to clearly mark a turn beginning and a turn end because the original ANDOSL recordings were dual channel files. The speech data were also annotated for Break Indices 3 and 4 (corresponding to intermediate and intonational phrase boundaries, respectively) according to the ToBI (Tones and Break Indices) prosodic transcription conventions for Australian English detailed in Fletcher and Harrington (1996).

The dialogues had also been independently coded for dialogue act using the 'Switchboard' version of DAMSL (SWBD-DAMSL) described in Jurafsky et al. (1997). Stirling et al. (Forthcoming) give more details and discussion of this coding. Since a major goal of the project was precisely to investigate associations between prosodic characteristics and functional discourse categories, we

coded for discourse categories independently from dividing the speech signal into prosodic units: for example, dialogue act coding did not presume prior division of the speech into intonational units (cf. the discussion in Zollo & Core (1999)).

2.1. CGU coding

We followed Nakatani & Traum (1999) in our identification of CGUs in the four dialogues. Their CGUs are similar, but not identical, to Clark & Schaeffer's (1989) 'Contributions' and to the 'Discourse units' defined in Traum & Hinkelman (1992), Traum (1994). Like these other units of grounding, a CGU is considered to consist of all the linguistic material involved in achieving grounding of an initial contribution of information by one participant. Thus the prototypical CGU in our dialogues was an initial contribution by either leader or follower in the map task, and a response to this by the other participant which indicated that they had heard and understood this contribution (though not necessarily agreed with it). An example is given in (1).

(1) IG: just cross that river
IF: yes

Note that there is no requirement on CGUs to be informationally or intentionally coherent: providing that the information is grounded (or groundable) in the same way, it is considered to belong to the same CGU. Thus the initial contribution to the CGU may be complex in consisting of several dialogue acts of the same or different kinds, as long as they can all be grounded by the same response.

CGUs may also be complex in containing repair or clarification sequences which need to be negotiated prior to grounding being possible. While Clark & Schaeffer treated such sequences as embedded contributions, and therefore modelled grounding units as highly recursive tree structures, Nakatani & Traum avoid coding such complexity in terms of embedded CGUs, and we have followed this practice. An example of a CGU containing a repair sequence is given in (2).[1]

(2) IG: now come UNDER the Brownwood S~~
IF: [Branded Steers yep]
IG: [Steers just] about catching his tail
IF: right

Like earlier writers, however, Nakatani & Traum do allow for overlapping CGUs. This is a function of the fact that an utterance may simultaneously be used to ground (and conclude) an already open CGU, and to make a contribution which initiates a new CGU and itself requires grounding. An example is given in (3).

(3) IG: now from that cross can you see the Galah Open-cut Mine?
IF: I've got a Dingo Open-cut Mine *[grounds first CGU / initiates second* CGU*]*
IG: right

CGUs as defined by Nakatani & Traum have several additional interesting characteristics.

First, CGUs can be discontinuous if the same information treated in one grounding negotiation is revisited later (for further confirmation or repair, for example). We adopted Nakatani & Traum's heuristic of allowing later mentions to be included in a previous CGU only if no more than three CGUs intervened. Both in deciding whether to treat an interaction as a repair sequence within a CGU and in identifying discontinuous CGUs we found that it was a nontrivial matter in the map task dialogues to distinguish between negotiation over grounding of a contribution and negotiation at a deeper level of agreement or understanding. Some of the dialogues we examined did contain rather complex discussions of information concerning the position of landmarks and the direction of route segments being described by the leader; in some cases it was difficult to decide whether to handle these discussions as part of a discontinuous grounding negotiation or at a larger level of structure (such as Nakatani & Traum's IUs). We took a relatively conservative approach in these decisions and ended up with few instances of discontinuous CGUs.

Second, CGUs need not segment the speech signal exhaustively, since speech material which does not contribute to the grounding of information may be omitted from the coding of CGUs. This includes such material as clear false starts and other disfluencies, 'self-talk', and 'phatic' communion such as the establishment of the channel at the beginning of an interaction.

Material which is excluded from CGUs in this way can be distinguished from instances where one of the participants attempts to begin a grounding negotiation which is subsequently cancelled or dropped before the information is grounded: following Nakatani & Traum, abandonned CGUs of this kind were coded but starred, and in fact are excluded from the quantitative results presented below. Abandonned CGUs included cases where one participant began a contribution, which was then interrupted either by the other participant or by the speaker herself, and which were not resumed at least within the three CGU limit mentioned above.

One further issue requires elaboration. This concerns the criteria used for determining the end point for an open CGU and for recognising the beginning of a new CGU.

First note that grounding may consist in nonverbal signals such as head nods, smiles or laughter, or in the most minimal case simply continued attention. It is usually not possible to identify such signals in audio data, although we did in a very few cases count audible laughter as grounding a CGU, and in one case the context made it clear that a non-verbal signal must have been present to allow the dialogue to proceed. In all other cases, we expected the grounding element to be an utterance of some kind.

Second, Clark & Schaeffer argued that every utterance (indeed, every signal – p. 266) is a contribution which requires grounding, including minimal acknowledgements of other contributions. It was assumed that the kind of grounding evidence required in such cases would be the most minimal possible: continued attention and/or moving on to the next relevant contribution. Nakatani & Traum, following others such as Traum & Allen (1992), adopted an approach where minimal acknowledgements made without any new information being introduced were simply included in the CGU they grounded without themselves setting up a new CGU: this makes for considerably less complexity in the coding. We followed this approach.

[1] Square bracketting indicates overlap between utterances.

We also adopted the perhaps more controversial approach of including minimal answers to questions, especially positive answers to check questions, just in the CGU initiated by the question. We will return to this point later.

Third, we noted above that in some cases CGUs overlap, with the same utterance acting as a response and grounding element for one CGU and as the initiating element of the next. In some cases it is difficult to decide on what should be included in which CGU in such cases. We followed the principle that if we could distinguish a grounding element from a continuation which provided new information, we did so, and included just the grounding element in the first CGU. If this was not possible, we included just the first utterance of the new information chunk in both CGUs, on the principle that after the first utterance the former CGU was clearly grounded.

As indicated above, Nakatani & Traum advocate including in a CGU any material produced in the negotiation over grounding including questions designed to achieve clarity or understanding by the ultimate grounder. We followed this practice by making a distinction between clarifying questions which did not ask for extra information and those which did; the former were coded as part of the previous CGU and the latter were treated as both grounding the previous CGU and initiating a new CGU.

Finally, note that, as a number of other authors such as Clark & Schaffer (1989), Nakatani & Traum (1999) have pointed out, initiating contributions for CGUs need not be complete illocutionary acts. It was quite common in our dialogues for participants to cooperatively divide up complex material into small chunks, for example negotiating the grounding of a reference to a landmark or to the source for a route before going on to deal with more complex actions concerning it. An example is given in (4). We treated these units as separate CGUs, thus effectively treating virtually all 'backchannels' as grounding elements marking a CGU boundary, as long as the expression being grounded was in some way informationally complete. However note that this decision fails to address the issue of grounding by non-verbal signals such as head-nods, which may also occur throughout an utterance. (For a discussion of this issue see Core et al. (1999).)

(4) CGU1 IG: now from there
CGU1 IF: mhm
CGU2 IG: underneath that you should have
Consumer Trader Fair
is it there?
IF: yes I've got that
IG: oh right

Two coders independently labelled the dialogues for CGUs then resolved their differences to produce a consensual version.[2] CGUs were numbered and their beginning and end points were entered into a separate xwaves label tier for each participant in the dialogue: coding CGUs in xwaves label tiers in this way fails to adequately reflect the potential and actual complexity of some units, since although overlaps can be coded (by including material within the boundaries of two CGUs), discontinuities can be marked (by numbering the discontinuous chunks the same), and excluded material can in principle be indicated, the result is a complex set of codes which is difficult to compile for analysis. Nevertheless, we found it to be a useful way to associate the boundaries of CGUs with other discourse and prosodic phenomena.

While various divisions have been made of the internal structure of grounding units and the subfunctions of the elements within the unit (for example see Clark & Schaeffer (1989), Traum (1994)), we will be most concerned here with characteristics of the initiating element, the grounding element, and the final element.

3. Results

The total number of grounded CGUs in the corpus was 412. In what follows we give a profile of these grounded CGUs from a formal perspective, in considering turn taking and prosodic features of their boundaries, and from a functional perspective, in considering their dialogue act profile. In this way we can contribute to an evaluation of CGUs as defined above in terms of their validity as discourse units and the way in which they interact with other discourse phenomena.

In the map task domain the Instruction-Giver has a privileged position in terms of the information available to him/her. This is reflected in the fact that while leader and follower produce roughly the same number of turns – in fact, leaders produced 308 turns overall while followers produced 335 turns – the average turn length was shorter for the follower in each dialogue (averages overall were 6 seconds for leader turns (range 4 – 11.85 seconds) and 1.8 seconds for follower turns (range 1.38 – 2.67 seconds)). While it is virtually always the case that the grounding element in a CGU is produced by the participant who did not produce the initiating unit, deriving initiative data about grounding is complicated by the fact that in 13% of cases the final unit in a CGU was not the grounding element. Taking this into account, we did in fact observe a pattern in which leaders produced relatively more contributions to be grounded and followers produced relatively more grounding units: overall, 64% of the CGUs were grounded by the follower, and 36% by the leader, with little variation across dialogues.

3.1. Formal profile of CGUs: turn taking and prosodic characteristics

It is well-accepted that aspects of prosodic structure are often associated with discourse segment boundaries (for example see Grosz & Hirschberg, 1992; Nakatani et al., 1995; Hirschberg & Nakatani, 1996; Swerts, 1997). However most studies have focused on the interaction between various correlates of prosodic structure and discourse segmentation at the micro- (e.g. dialogue act) level. The question of how to calculate prosodic associations with larger level discourse units, especially when they may not be temporally sequential, is nontrivial and is an ongoing concern of our project. In this study we make a start by measuring the association between the end

[2] Intercoder reliability was not a focus of this study, however it is worth noting both that it is difficult to devise an appropriate measure of this for CGU coding, and that other work has suggested that CGU coding in general does not exhibit high degrees of intercoder agreement (cf. Core et al., 1999).

of the final utterance in the CGU and turn boundaries and Break Index.[3]

42% of grounded CGUs overall were followed by a turn boundary, that is, a change of speaker (range 26%-50% across the four dialogues). More interestingly, CGUs followed by turn boundaries tended to be grounded by IFs (76%, as compared with 64% of CGUs grounded by IFs overall, as mentioned above). This is because turn-bounded CGUs tended to be grounded by simple acknowledgements. These results are summarised in Table 1.[4]

CGUs followed by turn boundaries	Total for 4 dialogues
Grounded by IF, no turn boundary	131 (54.8%)
Grounded by IG, no turn boundary	108 (45.2%)
Total no turn boundary	239 (58%)
Grounded by IF, with turn boundary	132 (76.3%)
Grounded by IG, with turn boundary	41 (23.7%)
Total with turn boundary	173 (42%)
Total CGUs	412 (100%)

Table 1: Associations between endpoints of grounded CGUs, turn boundaries, and initiation of grounding

There was a much stronger association between CGU endpoints and Break Indices. 79.6% of CGU final boundaries coincided with a BI 4, or Intonational Phrase.[5] This is consistent with other unpublished work which has shown that association with BI 3 or 4 is a good indicator not just of dialogue act boundaries but also of the boundaries of larger-level units. These results are summarised in Table 2.[6]

BI value for final unit of CGU	Total for 4 dialogues
BI 4	328 (79.6%)
BI 3	62 (15%)
Neither BI 3 nor 4	22 (5.4%)
Total grounded CGUs	412 (100%)

Table 2: Associations between endpoints of grounded CGUs and Break Indices

The 22 cases where the CGU boundary did not coincide with a major prosodic unit were those where coders felt it was possible to make a functional division between an element which grounded the previous CGU (usually a yes/no or discourse particle such as *Okay*, *Allright*), but where speakers did not make an intonational break between this element and the following material.

We also found that a third of the time (in 33% of cases) the grounding units in CGUs overlapped with the contribution being grounded – in other words, the new speaker didn't wait until the previous speaker had finished their contribution before grounding it.

3.2. Functional profile of CGUs: dialogue act types

Recently, considerable interest has been expressed about the relationship between micro-level dialogue act coding such as DRI/DAMSL and grounding phenomena. For example Core et al. (1999) noted as questions for further consideration the relationship between backward-looking functions and grounding and the question of whether grounding units such as CGUs should be composed of dialogue acts as their minimal units; Nakatani & Traum (1999) on the other hand assume that both are separately coded from minimal units such as intonational phrases. Poesio & Traum (1998) define different kinds of conversational acts including both core speech acts and grounding acts, and divide the backward-looking functions of DRI/DAMSL between the two, with Understanding BFs as grounding acts. Zollo & Core (1999) describe a method of automatically extracting grounding features and identifying grounding units once a dialogue has been coded with the DRI backward- and forward-looking tags.

While it would be informative to consider the pattern of dialogue act coding across CGUs in some detail, space considerations preclude this here, so we will summarise our findings only.[7] Tables 3 and 4 show the distribution of CGU initial elements and CGU grounding elements between DAMSL forward-looking functions and backward-looking functions.[8]

[3] We initially also considered associations between CGU beginnings and ends and silent pause location, where a silent pause was defined as a break in the acoustic waveform of more than 150ms that was not part of a stop closure phase. However CGU boundaries were only randomly associated with silent pause location, which is consistent with findings reported elsewhere that silent pause location and duration are not reliable indicators of dialogue segment boundaries (cf. Stirling et al.,. Forthcoming) although they have been found to be in monologue (e.g. Swerts, 1997; Nakatani et al., 1995).

[4] In deciding whether the final element of a CGU was followed by a turn boundary we considered only material which had been coded as belonging to some CGU, whether grounded or abandonned.

[5] Endpoints were measured at the righmost boundary of the final word in the CGU, including the ends of utterances performing 'double' function as ending one CGU and beginning the next.

[6] A chi square analysis was done to check on the possible relation between the four dialogues and the sub-categories of BI analysis: the results of this analysis were chi square =4.7; df, 6; significance, 0.5. Thus there is no statistical relation between the BI categories and the different dialogues. As a consequence it is legitimate to pool the data across the four dialogues, which is the form of the data that is reported below.

[7] The full set of SWBD-DAMSL labels can be found in Jurafsky et al. (1997) or Stirling et al (forthcoming); they are very similar to the DRI/DAMSL labels of Allen & Core (1997). Here we mostly refer to the major top level categories of forward-looking and backward-looking function types: 'Statement', 'Information-request' and 'Action-directive' (actually subcategories of 'Influencing-addressee-future-action', 'Agreement', 'Understanding', and 'Answer'.

[8] As noted above, CGUs may be initiated by a turn consisting of more than one dialogue act: here we give figures for the first act in the CGU only. While as we shall see not all grounding elements were final elements in the CGU, at this level of generality

Dialogue act type	Total for 4 dialogues
FF	348 (84.5%)
BF	63 (15.3%)
Other	1 (0.2%)
Total	412 (100%)

Table 3: DAMSL Dialogue act type of initial elements of CGUs

Dialogue act type	Total for 4 dialogues
FF	59 (14.3%)
BF	347 (84.2%)
Other	6 (1.5%)
Total	412 (100%)

Table 4: DAMSL Dialog act type of grounding elements of CGUs

Unsurprisingly, CGUs were most likely to be initiated by FFs – 'Statements', 'Information-requests' or 'Action-directives' - and most likely to be grounded by BFs. 32% of grounding elements (n=132) were simple 'Acknowledgements'. Most of the BFs which initiated CGUs were those with a 'double' function (52% – chiefly extended answers to initiating questions) or were continuations of extended answers. Similarly, most cases of FFs grounding or concluding CGUs were also 'doubles' (71% – statements or questions made in response to an initiator's original contribution). There were 75 cases of 'double' function elements (thus, 150 CGUs or 36% began or ended with a shared element). Table 5 shows the breakdown of these units by dialogue act type.

Dialogue act type	Total for 4 dialogues
Statement	21
Information-request	20
Action-directive	3
Other FF	1
Agreement	6
Understanding	7
Answer	20
Total FF	42
Total BF	33
Total	75

Table 5: Dialogue act type of 'double' function elements

Zollo & Core (1999) identified as problematic for their automatic extraction method the class of cases where the grounding element also initiated a new contribution and therefore would not be coded with a BF: such cases would be missed as grounding elements. From our data this is a problem which would affect a sizeable 18% of CGUs. Had we followed Clark & Schaefer's practice of identifying a new grounding unit for each acknowledgement, grounded for the most part by the occurrence of a next relevant utterance, the proportion of grounding units which were forward-looking functions would have been much higher.

In 54 CGUs (13% of the total) the grounding element was not the final element in the CGU. Upon further analysis these cases were of three main kinds:

(i) The first case, which accounted for 26 of the 54, was an artefact of our decision to include simple responses, especially positive responses to check questions, just within the CGU initiated by the question: in 32% of cases these responses were followed by an acknowledgement by the CGU initiator.

(ii) The second type were cases where the grounding element was an acknowledgement in the form of a repetition. 10 out of 22 (45%) repetition-acknowledgements in the data were themselves followed by an acknowledgement by the CGU initiator.

(iii) Thirdly were cases where a simple 'Acknowledgement' or 'Accept' functioning as grounding element was itself followed by a further 'Acknowledgement' or 'Accept' from the CGU initiator.

Finally, CGUs did not necessarily preserve DAMSL dialogue act boundaries. This occurred in two cases. First, as mentioned in section 2, some dialogue acts were grounded in installments. Second, again because of the overlapping CGUs already discussed, what had been coded as a single dialogue act was in some cases divisible into a grounding unit followed by the initiating unit of a new CGU.

4. Discussion

A number of authors have proposed taxonomical distinctions according to which grounding units can be characterised. Clark & Schaefer (1989) listed different types of presentations in terms of how they corresponded to complete utterances or turns and different types of acceptances in terms of the kind of evidence they provide for grounding. Traum (1994) distinguished different types of 'grounding acts' according to the subfunctions of utterances within a grounding negotiation. Core et al. (1999) include a proposal for distinguishing CGUs in terms of three types of acknowledgement. The results we have presented are less concerned with the internal structure of CGUs but suggest two additional parameters along which CGUs can be distinguished:

1. Turn properties of final unit in CGU

- The final unit may be a complete turn by one speaker (in these cases the CGU normally ends with a BI 4)
- The final unit may be part of a turn, with a later part of the turn functioning to initiate the next CGU (in some cases the CGU then ends in a BI 3 or even less; there is a problem in deciding when you divide a turn into a CGU completion part and a CGU initiation part – sometimes it seems the turn can be divided functionally, but there is no division in intonation unit)
- The final unit may be part of a turn, where the segment cannot be divided into CGU completion and

the figures are essentially the same whether we consider grounding elements or final elements.

CGU initiation elements (these are the 'double' function cases; here the CGU may well end in a BI 4, but by virtue of the fact that it is a whole unit which is doing double duty)

2. Grounding properties of final unit in CGU

- The grounding element is also the final element (the simplest case)
- The grounding element is followed by further material from the initiating speaker, acknowledging / grounding the grounding element

One of the characteristics of our data was the occurrence of complexity in what Clark & Schaefer described as the 'acceptance' phase of a grounding unit: the part where grounding is negotiated, after the initial contribution of information under consideration. In addition to expected complexity where repair or clarification sequences occur, we found that complexes of grounding units or of multiple acknowledgements made by both participants without any new information being introduced occurred quite commonly. Clark & Schaefer (1989), Nakatani & Traum (1999) and Core et al. (1999) all note that it sometimes takes several acknowledgements back and forth by the two speakers to establish common ground sufficient for current purposes: where the requirement on mutual understanding is high, as in task-based interactions of the kind we consider, this may be particularly the case. The following are representative examples.

(5) is a case where there were simply a number of acknowledgements provided by the grounding participant.

(5) IG: now we're going to start you off on tour one
IF: thank you darling
allright we're starting on tour one

(6) illustrates the case where an acknowledgement was repeated after overlap during an initial grounding unit; this occurred very commonly in our data.

(6) IG: uh about an inch below the bottom branch of the weste[rn tr]ee
IF: [yeh]
yeh

(7) is the case where a simple response by one speaker is followed by a simple acknowledgement by the other, which has a grounding function but perhaps also other interactive functions such as priming the listener for new information to come.

(7) IG: no you're above the Statistics Centre
IF: right
IG: allright

(8) illustrates the case mentioned in section 3.2 where a grounding element in the form of a repetition elicits further acknowledgement.

(8) IG: you head south
IF: um right head south
IG: yeh

And finally (9) is a case where there were multiple acknowledgements by both parties, perhaps reflecting a particularly difficult sequence of negotiations.

(9) IG:not that far
you're only supposed to go about a centimetre
IF: oh
IG:[yeh]
IF: [OK]
right
O[K]
IG: [O]K
IF: allright [so~~]
IG: [um]
IF: only a centimetre OK

5. Conclusion

The notion of grounding is well motivated but how well does it translate into formal definitions of units in dialogue which can be measured for associations with other dialogue phenomena? For instance, examples such as those in the previous section raise questions about the validity of attempting to identify a single 'grounding element' at which point the information under consideration is taken to enter the Common Ground. More generally, they raise the question for us of what is being measured by CGUs as defined here and in Nakatani & Traum (1999). The decision not to open new CGUs for 'simple' responses of various kinds in effect means that we are not looking at grounding per se here, but at something like the subset of 'more contentful' grounding units: this makes the meso-level units defined much easier to work with especially with large amounts of data, and more akin to other meso-level units such as adjacency pairs or games, but arguably misrepresents the grounding profile of the dialogue and its participants.

Nevertheless, we found that coding the map task dialogues for CGUs has enriched our understanding of what is going on in the dialogues and raised questions which further work will pursue. These will include the relation of grounding to higher-level discourse structure, and the development of a metric of complexity of CGUs to correlate with stylistic characteristics of particular dialogues, including degree of perceived dysfluency (evident in particular in dialogues between pairs who were previously unknown to one another).

6. References

Allen, J. & Core, M. (1997). Draft of DAMSL: Dialog Act Markup in Several Layers. Draft contribution for the Discourse Resource Initiative. Also available at: http://www.cs.rochester.edu/trains/research/annotation.

Anderson, A., Bader, M., Bard, E., Boyle, E., Doherty, G., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., Thompson, H. & Weinert, R. (1991). The HCRC map task corpus. *Language and Speech*, 34, 351--366.

Clark, H. & Marshall, C. (1981). Definite reference and mutual knowledge. In A. Joshi, B. Webber, & I. Sag (Eds.), *Elements of Discourse understanding* (pp. 10--63). Cambridge: Cambridge University Press.

Clark, H. & Schaefer, E. (1989). Contributing to discourse. *Cognitive Science*,13, 259-294.

Core, M., Ishizaki, M., Moore, J., Nakatani, C., Reithinger, N., Traum, D. & Tutiya, S. (1999). The report of the Third Workshop of the Discourse Resource Initiative, Chiba University and Kazusa Academia Hall.

Fletcher, J. & Harrington, J. (1996). Timing of intonational events in Australian English. In P. McCormack & A. Russell (Eds.), *Proceedings of the Sixth Australian International Conference on Speech Science and Technology* (pp. 611—615).

Grosz, B. & Hirschberg, J. (1992). Some intonational characteristics of discourse structure. In *Proceedings of the International Conference on Spoken Language Processing*, Banff (pp. 429--432).

Hirschberg, J. & Nakatani, C. (1996). A prosodic analysis of discourse segments in direction-giving monologues. In *Proceedings of the 34th Annual Meeting of the ACL, Santa Cruz* (pp. 286—293).

Jurafsky, D., Schriberg, L. & Biasca, D. (1997). Switchboard SWBD-DAMSL Shallow-Discourse-Function-Annotation Coder's Manual, Draft 13. Technical Report TR 97-02, Institute for Cognitive Science, University of Colorado at Boulder. Also available at:
http://stripe.colorado.edu/~jurafsky/manual.august1.html.

Millar, J., Vonwiller, J., Harrington, J. & Dermody, P. (1994). The Australian National Database of Spoken Language. In *Proc. ICASSP-94* (pp. 197—100).

Nakatani, C., Grosz, B., & Hirschberg, J. (1995). Discourse structure in spoken language: studies on speech corpora. In *Proceedings of the AAAI-95 Spring Symposium on Empirical Methods in Discourse Interpretation and Generation*.

Nakatani, C. & Traum, D. (1999). Coding discourse structure in dialogue (Version 1.0). University of Maryland Institute for Advanced Computer Studies Technical Report UMIACS-TR-99-03.

Poesio, M. & Traum, D. (1997). Conversational acts and discourse situations. *Computational Intelligence*, 13(3), 309--347.

Poesio, M. & Traum, D. (1998). Towards an axiomatizatoin of dialogue acts. In *Proceedings of the Twente Workshop on the Formal Semantics and Pragmatics of Dialogues* (pp. 207—222).

Stirling, L., Fletcher, J., Mushin, I. & Wales, R. (Forthcoming.) Representational issues in annotation: using the Australian map task corpus to relate prosody and discourse structure. *Speech Communication*.

Swerts, M. (1997). Prosodic features at discourse boundaries of different strength. *Journal of the Acoustical Society of America* 101(1), 514--521.

Traum, D. (1994). *A computational theory of grounding in natural language conversation*. PhD thesis, Department of Computer Science, University of Rochester. Also available as TR 545, Department of Computer Science, University of Rochester.

Traum, D. & Allen, J. (1992). A speech acts approach to grounding in conversation. In *Proceedings of the 2nd International Conference on Spoken Language Processing* (ICSLP-92) (pp. 137—140).

Traum, D. & Hinkelman, E. (1992). Conversation acts in task-oriented spoken dilaogue. *Computational Intelligence*, 8(3). Special issue on Non-literal Language.

Zollo, T. & Core. M. (1999). Automatically extracting grounding tags from BF tags. In *Proceedings of the 37th Meeting of the ACL*, Workshop on Standards and Tools for Discourse Tagging, College Park.