

## Modelling Structures for Situated Discourse

**Nicholas Asher**

*CNRS, ANITI, IRIT Toulouse*

NICHOLAS.ASHER@IRIT.FR

**Julie Hunter**

*LINAGORA, Toulouse*

JHUNTER@LINAGORA.COM

**Kate Thompson**

*LINAGORA, Toulouse*

CTHOMPSON@LINAGORA.COM

**Editor:** Manfred Stede

Submitted 01/2019; Accepted 03/2020; Published online 03/2020

### Abstract

In this paper, we argue that modelling situated discourse requires not only allowing nonlinguistic events to enter into discourse relations with speech act contents, but also modelling semantic interactions between nonlinguistic events themselves. In an evolving nonlinguistic context, these interactions can give rise to a rich semantic, nonlinguistic structure that is relevant for the interpretation of conversational moves. Examining how these nonlinguistic structures interact with structures determined by dialogue moves reveals new types of discourse structure and a novel perspective on discourse threads and goals. We motivate our arguments with a study of a corpus of situated multiparty chats developed for the STAC project<sup>1</sup> and annotated for discourse structure in the style of *Segmented Discourse Representation Theory* (SDRT; Asher and Lascarides, 2003). The STAC corpus is not only a rich source of data on strategic conversation, but also the first corpus that we are aware of that provides discourse structures for multiparty dialogues situated within a virtual environment. The corpus was annotated in two stages: we initially annotated the chat moves only, but later decided to annotate interactions between the chat moves and non-linguistic events from the virtual environment. This two-step procedure allows us quantify various ways in which adding information from the nonlinguistic context affects dialogue structure.<sup>2</sup>

**Keywords:** discourse structure, situated discourse, nonlinguistic context, nonlinguistic events

### 1. Introduction

The study of discourse structure, in particular rhetorical structure, on texts is now a well entrenched cottage industry in computational linguistics. Several discourse annotated corpora exist including the Penn Discourse Treebank (PDTB; Prasad et al., 2008) and the RST Discourse Treebank (RST-DT; Carlson et al., 2002)—the only large corpus with full discourse structure for texts—as well as smaller annotated corpora such as DISCOR (Baldrige et al., 2007) or ANNODIS (Afantenos et al., 2012). The STAC project (Hunter et al., 2015b; Asher et al., 2016) extends this work to dialogue

1. *Strategic Conversation*, ERC grant n. 269427.

2. We gratefully acknowledge support from ERC grant 269427 (Strategic Conversation), BPI France grant P169201 (LinTO-Assistant vocal open-source respectueux des données personnelles pour l'entreprise), and 3IA grant ANITI ANR-19-PI3A-0004 (ANITI – Artificial and Natural Intelligence Toulouse Institute). We would also like to thank four anonymous reviewers and the editor for their extensive and very helpful comments.

annotation on a corpus of chats from an online version of the game *Settlers of Catan*. The STAC corpus is, as far as we know, the only corpus with full relational, discourse structure for dialogues, and the only corpus with a complete set of full discourse annotations for *situated* dialogues.

In this paper we characterize in a largely informal but precise way how the game events in our corpus interact with chat moves. To do so, we exploit two sets of annotations for our corpus, which we call the *chat-only annotations* and the *situated annotations*. The former contain discourse structures determined by only chat moves from the corpus, while the latter, which were begun only once the chat-only annotations were complete, contain full situated discourse structures that account for both game events and chat moves. In the situated annotations, both chat moves and game events contribute arguments to rhetorical relations, which allows us to account for the flexible dynamics of natural discourse situations. The two sets of annotations enable us to clarify and quantify the influence that a situated environment can have on a linguistic message and to make a detailed, empirical comparison of the structures in the chat-only and the situated annotations. As noted in Tenbrink et al. (2013), little information is available concerning the systematic annotation of situated dialogues, and there are few annotated corpora. This paper goes towards filling these gaps.

The analysis developed in this paper goes considerably beyond models of the nonlinguistic context in terms of deixis and reference, in which nonlinguistic entities are understood as crucial for the *interpretation* of discourse structures, but not for their *construction* (Kaplan, 1989; Rickheit and Wachsmuth, 2006; Kranstedt et al., 2004; Kruijff et al., 2010). It also extends Tenbrink et al. (2013)'s work on the NavSpace corpus, in which nonlinguistic actions are treated as contributing dialogue acts but the topic of their contributions to overall rhetorical structures is not broached. Our aim in this paper is to show that in modelling situated discourse, we need to do more than allow nonlinguistic events to contribute arguments to discourse relations. We need to model *structural relations between nonlinguistic events* as they dynamically develop, because these relations are often relevant for interpreting dialogue moves, and we need to model the *higher level structures* that these relations give rise to. Doing so, as we explain, uncovers new types of discourse structures and novel ways of thinking about discourse threads and discourse goals.

In positing that nonlinguistic events contribute to larger discourse structures, this work echoes claims in Lascarides and Stone (2009), in which coverbal gestures can have discourse functions that are parasitic on verbal dialogue acts. We consider a much wider range of discourse interactions, however, and consider the nature of situated discourse in more general terms. We also examine a very different kind of interaction in which nonlinguistic events drive discourse development.

We provide here strong empirical evidence for theoretical claims from Hunter et al. (2018), but we go further by examining in detail the influence of the surrounding environment on discourse structure, describing in particular its effects on global discourse structures rather than local discourse relations alone. Section 2 describes our corpus, the annotation process, and three points about multiparty discourse. Section 3 argues that the game events in our corpus determine a rich structure of their own and that bringing this structure together with discourse structures determined by chat moves gives rise to new types of discourse structures. Section 4 then quantifies the differences between the chat-only and situated annotations. This comparison gives a fuller picture of the role of nonlinguistic *structures* in the overall interpretation of our situated interactions.

## 2. The *Settlers* corpus

In this section, we give an overview of the *Settlers* corpus and certain features of the annotations. More quantitative details can be found in the appendix. Section 2.1 explains the difference between the two sets of annotations, while Section 2.2 lays out the general approach adopted to annotate the corpus. Section 2.3 discusses three complications for building discourse structures in the chat-only annotations and our approach to them. These problems result from the multiparty setup and are not normally accounted for by theories of rhetorical structure.

### 2.1 Building the corpus

The *Settlers* corpus consists of a series of chats taken from an online version of the game *The Settlers of Catan* that have been annotated for discourse structure in the style of *Segmented Discourse Representation Theory* or SDRT (Asher, 1993; Asher and Lascarides, 2003). *The Settlers of Catan* is a multiparty, win-lose game in which players use resources such as wood and sheep to build roads, settlements, and cities on a game board. Players acquire resources in various ways, including trading with other players and rolling the dice. As shown in Figure 1, the game board is divided into hexes, each associated with a certain type of resource and a number between 2-6 or 8-12. A dice roll of, say, a 4 and a 2 gives any player with a building on a hex marked “6” one or more resources associated with that hex. Rolling a 7 triggers a series of moves: the current player must move a game piece known as “the robber” to a hex of her choice and then steal a resource from a player with a building on that hex. The robber will stay on that hex until moved in another turn, and its presence will continue to impact the game by blocking resource distributions for the occupied hex.

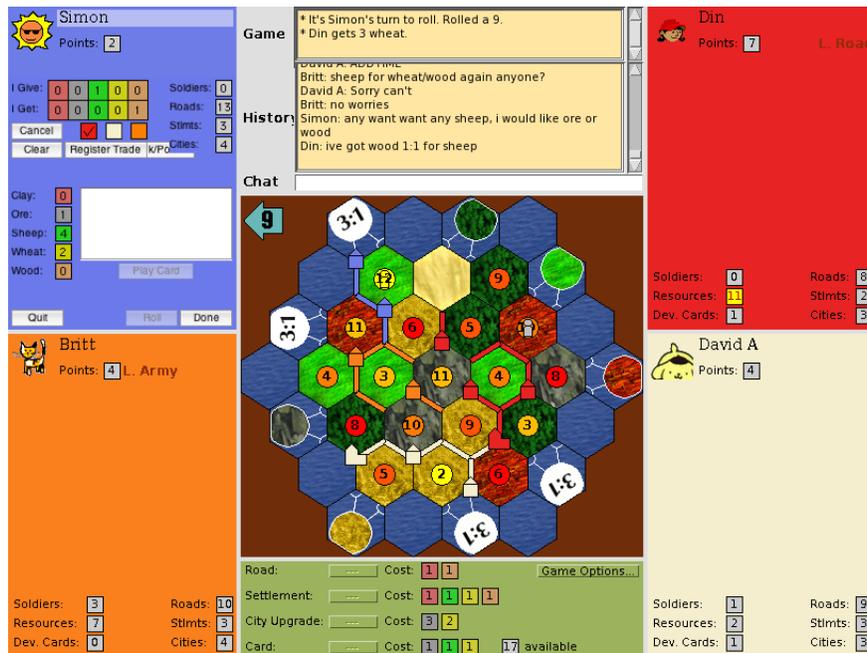


Figure 1: A snapshot of the game board for *The Settlers of Catan*

To construct the *Settlers* corpus, we modified an online, open-source version of *Catan* to include a chat window. Figure 1 illustrates the game interface and provides a snapshot of the way the game board looks to a particular player, Simon. Simon’s resources are shown in the upper left-hand rectangle, but as in the physical version of the game, Simon cannot see the resources of his opponents. Figure 1 shows that Simon is preparing to make a trade via a *Trade Panel*: he has prepared an offer to the (red) player Din, but has not yet clicked “Register Trade”. Once Simon registers his trade, Din’s response, whether he accepts or rejects the trade, will be described in the *Game* window, which records many game events that are public to all players. Finally, the *Chat* window allows players to chat, and prior chat is recorded in the *History* window. To encourage discussion, players were instructed to negotiate trades in the chat interface before executing an agreed trade through the Trade Panel.

**Chat-only annotations.** The *Settlers* corpus was developed as a part of a project whose original aim was to study the discourse structure of strategic dialogue, in which interlocutors can have divergent discourse goals and therefore fail to be entirely cooperative. As such, the annotations were originally limited to the chat moves of the corpus, as it is during trade negotiations that strategic reasoning was assumed to affect discourse structure. In total we annotated the complete chat transcripts of 46 games using the Glozz annotation tool (Widlöcher and Mathet, 2009) to build what we refer to as the *chat-only annotations*. To make the annotation process more manageable, we originally divided the chat history for each game into a number of (what we refer to as) *dialogues*, each of which involved one or more bargaining sessions. The criteria for dividing the dialogues were not precise, though there was the obvious goal of keeping each trading session intact. Typically dialogues contained just one negotiation session with one player leading the bargaining. However, occasionally annotators linked elements in one negotiation session with elements in another, and in that case, we considered the linked sessions as contributing to a single dialogue.

Players used the chat interface to discuss numerous aspects of the game state, and it ultimately became clear during the first annotation campaign that much of the chat conversation was related to the game board and game events in intricate and semantically significant ways that merited further study. This observation triggered a second round of annotations in which the chats were re-annotated in light of publicly observable game events, descriptions of which were extracted from the game logs. We refer to the outcome of this annotation campaign as the *situated annotations*.

**Situated annotations.** Figure 2 is an example from the situated annotations that illustrates how chat moves interact with game events ([CLICK FOR GRAPH](#)).<sup>3</sup> Every turn in our corpus, whether it is a chat move or a game event, is assigned a turn number, and all turns are automatically recorded and aligned in a game log for each game. Turn numbers are indicated in the left column of Figure 2. Game messages that were added in a later stage for the situated annotations were assigned sequence identifiers in order to preserve the numbering of the chat and game events that were present in the first stage. (Note that some game events were assigned numbers in the initial stage, although they were not shown to annotators.) Each turn is also identified with an agent, as shown in the second column of Figure 2. For chat moves, the agent is the player who typed the chat message (e.g., GWFS for turn 434);<sup>4</sup> visually observable game events and states are either described in Server messages,

3. In addition to the graphs provided in this paper, we add a link for each example that allows the viewer to see the example in a larger context and to compare the situated and chat-only graphs. Go to [https://www.irit.fr/STAC/stac\\_game\\_graphs/index.html](https://www.irit.fr/STAC/stac_game_graphs/index.html) for instructions on how to read the graphs (see the “Read Me” document) and to view the graphs for every game in the annotated corpus.

4. Small capitals indicate user names that have been abbreviated to save space or preserve anonymity.

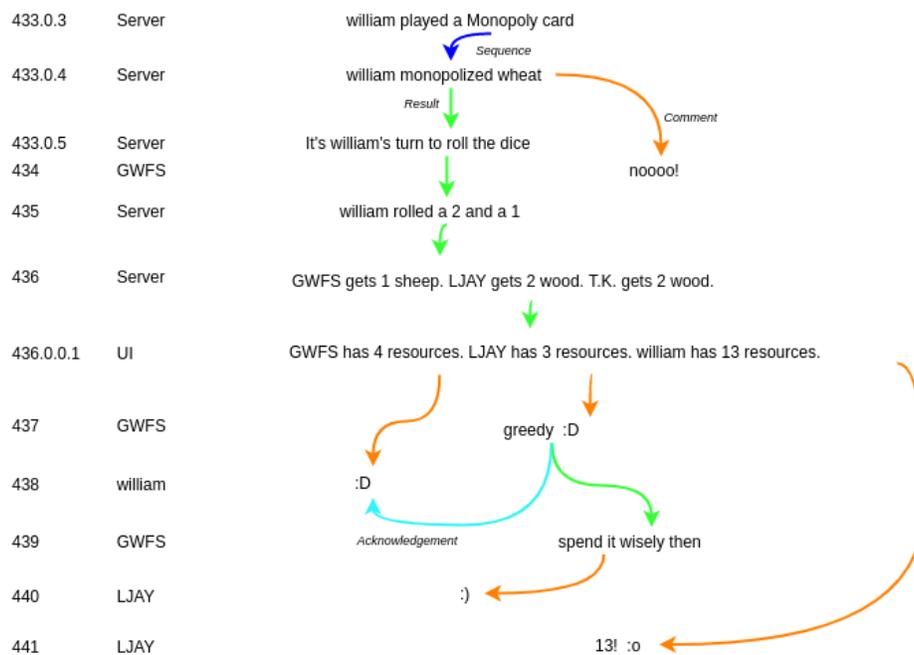


Figure 2: The edges in our graphs are color coded for relation labels; e.g., all of the green arrows above represent Result relations.

many of which were visible to all players via the Game window, or reconstructed (by our team) using information from the User Interface (UI). In Figure 2, William plays a Monopoly card, which allows him to steal all instances of a resource of his choice that are possessed by the other players. In turn 433.0.4, he steals all of the wheat, and GWFS comments on this move in turn 434. LJAY adds a smiley face in 440. There is some ambiguity as to whether the smile is a comment on the theft itself or on GWFS’ comment in 439; we opted for the latter interpretation, as the smile is added just after 439. In 441, LJAY comments on the result of the theft: that William has 13 resources.<sup>5</sup>

## 2.2 Annotating the corpus

The chats in our *Settlers* corpus were annotated for discourse structure in the style of SDRT, but there are numerous other theories of discourse structure for texts which might have been employed: *Rhetorical Structure Theory* (RST; Mann and Thompson, 1987), the *Linguistic Discourse Model* (LDM; Polanyi et al., 2004), the *Discourse Graphbank* model (Wolf and Gibson, 2005), *Discourse Lexicalized Tree Adjoining Grammar* (DLTAG; Forbes et al., 2003), and the *Penn Discourse Treebank* (PDTB; Prasad et al., 2008). Each one of these has, or at least can give rise to, an annotation model for discourse structure, and all of the theories agree on how the process of annotation for discourse structure should be approached.

5. For a more in-depth description of the corpus, see Asher et al. (2016) and Hunter et al. (2018), from which some of the foregoing description borrows. To see the webpage that was created to inform players about playing in our league, go to: <http://homepages.inf.ed.ac.uk/mguhe/socl/>. To view the annotations for the corpus, visit: <https://www.irit.fr/STAC/corpus.html>. The annotation manual for the chat-only annotations can be found here: <https://www.irit.fr/STAC/stac-annotation-manual.pdf>.

The annotation process begins by segmenting each text or dialogue to be annotated into a set of what we will call *elementary discourse units* or EDUs, which serve as the basic building blocks of discourse structures. EDUs are typically clauses but may also include material that is in the periphery of the main predication in a clause (cf. Afantenos et al., 2012). The next step in the process is to figure out how each EDU should be related to other discourse units in the discourse representation. This requires solving two interrelated problems: the *attachment problem* and the *labelling problem*. The attachment problem concerns *where* each EDU is attached in an incoming discourse structure. A typical case is that an EDU attaches to another discourse unit as the argument of a rhetorical relation, but it can also, at least in SDRT, first be added to a group of discourse units that *collectively* provide the argument to a discourse relation.<sup>6</sup> In this latter scenario, the discourse units that work together to provide a discourse argument form what is called a *complex discourse unit* or CDU in SDRT. The labelling problem then involves associating each discourse attachment with a label that denotes a discourse relation such as *Elaboration*, *Explanation*, *Narration*, *Question Answer Pair*, and so on.

In a coherent conversation, each EDU should serve a rhetorical function, which means that each EDU, apart from the first one, should be attached to the incoming discourse via a rhetorical relation; in other words, the discourse should have a *weakly connected* structure. SDRT thus represents the structure of a discourse as a weakly connected graph with *directed* edges—for a relation instance *Explanation*(EDU<sub>*n*</sub>,EDU<sub>*m*</sub>), for example, it is the content of EDU<sub>*m*</sub> that explains that of EDU<sub>*n*</sub>, not the other way around. SDRT further posits that graphs are *acyclic*, as every EDU is expected to advance the conversation, not make a loop in which it is ultimately related to itself. In contrast to many other theories, including DLTAG, LDM, and RST, however, SDRT does not require that its graphs have a tree structure.

To create our corpus, the chats were segmented automatically and manually corrected. Because of the nature of chat, this was a relatively simple task. The annotation of discourse structure, on the other hand, was a large effort. The first stage was carried out by four “naive” annotators who received training over 22 negotiation dialogues, which included 560 turns in all. Annotators were first asked to classify the segments as various kinds of dialogue acts. This involved labelling them according to surface form—QUESTION, COMMAND, or ASSERTION—and to categories more specific to the game—OFFER, COUNTEROFFER, ACCEPT (OFFER), REFUSAL (OFFER), and OTHER. These features proved to be important predictors for automatically learning discourse structure (Afantenos et al., 2015; Perret et al., 2016). Annotators were then asked to choose an attachment point for each chat move except the very first one—so long as they were able to find an intuitive attachment point—and to label the attachments with at least one out of 16 possible labels for rhetorical relation types. Further details on the annotation process can be found in the annotation manual (see footnote 5). After training, we evaluated our annotators on a small set of new dialogues in our pilot corpus. Using an exact match criterion of success, the inter annotator agreement score was a Kappa of 0.72 on attachment only on structures and 0.58 on attachment and labelling for doubly annotated dialogues—a respectable score given the complexity of handling both attachment and labelling together (Asher et al., 2016).

After these pilot annotations, four expert annotators continued the annotation process, and annotations of attachments and relations in the chat-only annotations went through a multi-step review process. The experts iterated passes over the annotations from the naive annotators as well as new

---

6. Not all theories of discourse structure explicitly countenance the construction of larger units; in RST (Mann and Thompson, 1987), the step of assigning nuclearity values to different nodes could probably be so extended, but the possibility is not explicitly discussed.



Figure 3: A truly non-treelike structure

annotations, improving the data and debugging it by checking for violations of constraints, such as acyclicity, in the annotated structures. We had at least five stages of revision. We used the Trello system (trello.com) to maintain our annotations, to keep track of the revisions, and to establish an agreement by consensus among four experts as to what should constitute our gold annotations.

### 2.3 Three points about multiparty dialogue

The multiparty nature of our corpora gave rise to various phenomena, some already noted in the literature on multiparty dialogue, that one finds rarely if at all in single authored text. We comment on three phenomena here: the presence of non-treelike discourse structures (cf. Afantenos et al., 2015; Perret et al., 2016; Asher et al., 2016), overlapping conversational threads (cf. Crystal, 2011; Bartlett, 2014; Afantenos et al., 2015; Asher et al., 2016), and subjective interpretations (cf. Lascarides and Asher, 2009; Ginzburg, 2012). The first two will be relevant to the characterization of our situated discourse structures in the sections that follow; the third we decided to ignore in our study for reasons that we explain below.

A non-treelike structure is one containing a discourse unit or DU with more than one incoming arrow, but within this category, we distinguish between *quasi-treelike* DUs and *truly non-treelike* DUs. A quasi-treelike DU has two incoming arrows (with two different labels) from the *same* source DU. For example, in uttering a sentence of the form “*p* but then *q*”, the speaker indicates that the content of *q* is linked to that of *p* by both a relation of contrast and a relation of sequence (Contrast and Narration in SDRT); the EDU determined by *q* is thus quasi-treelike. A truly non-treelike DU, on the other hand, has two or more incoming arrows from *different* source DUs. Turn 239 in Figure 3 is an example of a truly non-treelike DU ([CLICK FOR GRAPH](#)).

In turn 234, GWFS makes an offer, to which he receives three negative replies (235, 236, 238). He responds in 239 with the acknowledgement “kk” (=“okay cool”), which is intuitively aimed at all three negative responses. The label QAP is short for “Question-Answer Pair” and types the connections between 234 and each of the three replies. Truly non-treelike structures are fairly frequent in our corpora; out of a total of 12,588 DUs in the chat-only annotations, 928 or about 7% were truly non-treelike. Our annotation framework countenances both quasi and truly non-treelike structures, so this aspect of multiparty dialogue did not pose a problem for us, but we note that frameworks like that of RST (Mann and Thompson, 1987) and LDM (Polanyi et al., 2004) do not countenance such structures.

Dialogue threads are created when groups of two or more dialogue participants engage in two or more connected exchanges (connected via instances of discourse relations) that are developed

simultaneously but independently in the sense that they are not discourse connected. Figure 4 contains at least 3 threads, which we have represented with solid, dashed and dotted lines. Note that the lines in this graph do not reflect rhetorical relations but only thread membership. To see the full graph with relations, [CLICK HERE](#). As Figure 4 shows, the threads are developed simultane-

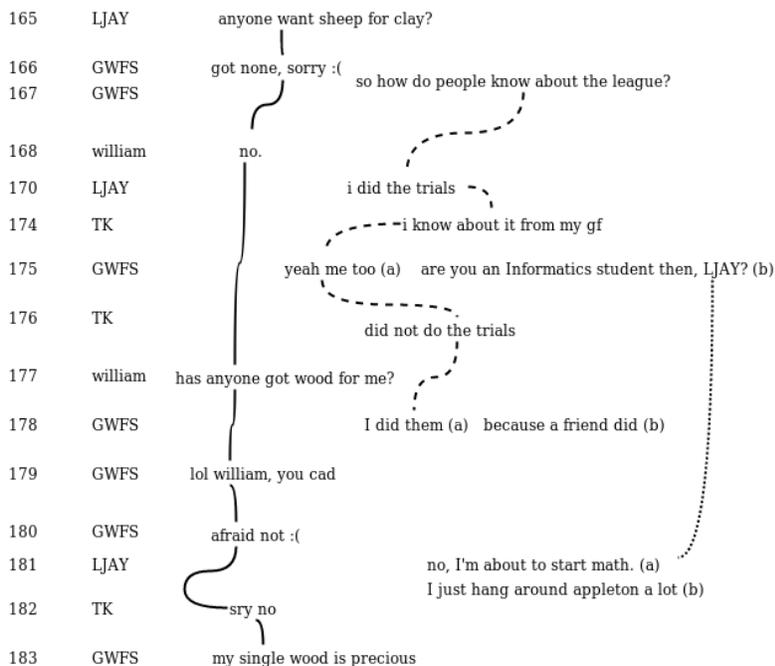


Figure 4: Three discourse threads

ously: for example, GWFS starts a new discussion in 167 that develops over seven segments, ending with 178b, and throughout this discussion, a sequence of trade negotiations initiated in 165—before GWFS started a new thread—develops and continues until well after 178b. The threads are independent in the sense that while two threads might share a common root (not pictured here), once they are created, there are no rhetorical relations that connect EDUs across threads. Threads call for a revision of assumptions about the projectivity of discourse structure (like those assumed in RST; Carlson and Marcu, 2001) and salience-driven constraints on discourse attachment as formulated for monologue (Polanyi, 1985; Asher, 1993), as developing a new thread does not block the accessibility of EDUs from a previously initiated thread.<sup>7</sup> Such constraints do not even hold for individual speakers: GWFS, for example, is free to engage in all three threads simultaneously. In Section 3 we will compare “standard” multi-party threads to similar structures generated in the situated annotations.

The third phenomenon that we address in this section, which we chose to ignore in our annotations, is the possibility that interlocutors might interpret a discourse differently, leading to differing, and even contradictory, representations of a conversation. Amplifying on Lascarides and Asher (2009); Venant et al. (2014); Venant and Asher (2015) or on the more fundamental theoretical analyses of Asher and Paul (2018), we should create an individual discourse graph for each participant

7. For more on how such salience constraints are affected in our multiparty chats, see Hunter et al. (2015a). See Hunter et al. (2018) for a discussion of how to extend such constraints to situated discourse.

(cf. Ginzburg, 2012). Doing this for our corpora, however, would have required building three or four discourse graphs for each game, which would have precluded a more extensive, global view of discourse structures across a significant corpus. We therefore decided to annotate the multiparty dialogue from the perspective of a third party observer who infers a structure based on players' public commitments to content derived from their contributions.

This choice had various consequences for our annotation decisions. It meant, for instance, that we ignored personal Server messages, including those displayed during trades and steals. When a player moves the robber and steals a resource from another player, for example, the History window will display a general message of the sort "william stole a resource from GWFS". At the same time, GWFS and william will each receive a personal and more detailed message, such as "william stole a wheat resource from you" and "You stole a wheat resource from GWFS," respectively. We disregarded all such personal messages.

Another effect of our decision to ignore subjectivity was that it left subjective contributions such as comments (linked by the Comment relation) incomplete and possibly ambiguous. Imagine that a player  $p$  utters a comment, say "thanks" or "sorry" or "ouch", with content  $c$ . Were we to have individual discourse graphs for each player,  $p$ 's graph would reflect commitment to  $c$ , while the other players' graphs would be updated to reflect (at most) their commitment to  $p$ 's commitment to  $c$ . As we opted for a single graph, we had to choose between having all players commit to  $c$ , or having all players—including  $p$ —commit to the content that  $p$  commits to  $c$ .<sup>8</sup> We chose the latter, weaker option, because it does not force us to ascribe attitudes to players without evidence that the players hold those attitudes. The downside, however, is that it yields an incomplete picture of players' commitments, as it prevents us from capturing how other players respond to  $p$ 's utterance of  $c$  and whether they accept  $c$ . This in turn yields a less satisfying analysis of acknowledgments of comments, in which the content of a comment is established as common ground.

Using only one graph had an even more significant impact on the treatment of Correction, and it led us to confine the use of Correction almost entirely to self-corrections or corrections by the Server of a player's attempted but forbidden action (as a player in this case is forced to correct her action if game play is to continue). Corrections between different players signal disagreements and thus leave room for differing points of view: a speaker who utters a correction with content  $q$  commits to the falsity of some content  $p$ , but the speaker of  $p$  might not be willing to accept the correction and hence might not commit to  $p$ 's being false. The closest that we could come to accurately representing such a disagreement between two different players without using individual graphs and commitment slates (Portner, 2004) was to use a Contrast relation that captures the fact that one player commits to  $p$  but another player commits to  $q$ . While this might be accurate, it, too, leads to incomplete representations: it misses the truth conditional effects of a correction on which the contents of the first term of the relation are taken, by at least one player, to be false.

Despite the incompleteness entailed by our choice to use a single discourse graph for each game in our corpus, it was the preferable option. In most cases, we simply lacked the information necessary to decide on each player's commitment to each chat move. In other words, not only would building individual graphs have greatly complicated the annotation process, but in many cases, it would have failed to solve the incompleteness problem. The result would not have justified the cost.

---

8. The semantics of most discourse relations, including Comment, is based on dynamic conjunction; this allows us to treat a comment as merely entailing common commitments to the content that the agent of the comment expresses a certain attitude.

### 3. Moving to situated dialogue

To make the situated annotations, we posited that game events can contribute what we call *elementary event units* or EEUs, whose contents serve as independent arguments to rhetorical relations just as the contents of EDUs or linguistic speech acts do (cf. also Tenbrink et al., 2013). Allowing EEUs to contribute arguments to discourse relations, however, opens up the possibility that two EEUs could be rhetorically related to each other. In fact, what we found in our corpus is that inferring such relations is often necessary for interpreting chat moves and the content of the overall game-chat exchanges. These relations between EEUs in turn determine higher-level structures of their own. Our central aim in this paper is to show the necessity of modelling these higher-level structures for (often nonlinguistic) events occurring in the larger situation and to explore how these structures interact with structures determined by conversational exchanges. This is what makes our task challenging: we are not simply trying to model the fact that *a single* nonlinguistic event can have an impact on discourse interpretation; we are trying to model interactions between two somewhat independent *structures*. Accordingly, this section addresses two questions: (i) What types of relations and structures do we see holding of EEUs in our corpus? (ii) What types of higher-level structures result from interactions between EEUs and EDUs in our corpus? Section 3.1 focuses on (i); Section 3.2, on (ii).

The setup of our corpus greatly facilitates study of these questions because a large portion of the game events are described in server messages that are temporally aligned with all of the chat moves. A major hurdle to studying the discursive contribution of nonlinguistic events is that the nonlinguistic context does not contain an analogue to a linguistic clause; nonlinguistic information is often presented in a steady stream, which means it is up to interpreters to determine how to individuate events. Moreover, given that the same event can generally be described in different ways depending on one's purposes, an interpreter must also determine what propositional content to associate with a nonlinguistic event. In other words, there is not only a more challenging segmentation problem for nonlinguistic information, but also a classification or *conceptualization* problem. Because the game logs for our corpus contain descriptions of many game events, however, our corpus allows us to largely bypass both the segmentation and conceptualization problems for the game events. This frees us up to focus on how such events influence the structure and interpretation of the larger chat exchanges.

The fact that the server provided descriptions of game events might suggest that EEUs are not so nonlinguistic after all. Hunter et al. (2018) address this concern in detail but we note here that some events were represented only visually to players. Information about these events—which included ending of turns and selection of an addressee for a trade among others—had to be extracted from the User Interface (UI) and assigned a content that could be subsequently annotated. A more general point, though, is that game events coming packaged with descriptions would be problematic if we were studying the conceptualization problem for nonlinguistic content. But in this empirical study we are not; we want to determine how the game events, *once they are segmented and conceptualized*, influence discourse construction and interpretation. For this it suffices that the interactions between discourse moves and game moves in our corpus mirror those that we could expect from people playing a nonvirtual version of *Settlers of Catan*.

### 3.1 The conceptualization and structure of game events

Nonlinguistic events, discounting gestural events produced as a part of a conversation, differ from speech acts in the sense that they must be associated with contents that are true in the external world; EEUs describe events that actually happened. As a result, the external world imposes a natural order over EEUs: if an EEU  $\epsilon_n$  appears before an EEU  $\epsilon_m$  in a linear ordering of EDUs and EEUs to be annotated for a situated discourse, then in general, we can infer that the event described in  $\epsilon_n$  happened before that described by  $\epsilon_m$ ; at least, we know that it didn't happen after it. In this, EEUs differ considerably from EDUs, whose linear presentation can depart from the order inferred over the events described in chat moves: in *Bill cried because Jane insulted him*, for example, the event that is described first is understood to have happened second.

The fact that the external world imposes its own temporal ordering over such nonlinguistic events might suggest that the structures that these events give rise to are more constrained than those for speech act events. This is to some extent true. Relations like Contrast, Conditional, Parallel, Clarification-Question, and Question-Elaboration, do not appear in our corpus between game events (though as indicated in Table 9 in the Appendix, some of them can relate EEUs to EDUs). This is to be expected: it is not clear how an event that is not a speech act or some other communicative act (such as a gestural event) could convey a conditional dependency, a contrast or parallel, or follow-up or clarification question. In a multimodal corpus containing communicative nonlinguistic events, the possibilities would be different: a puzzled facial expression can convey a request for clarification, and with conventional gestures, a wide range of relations may be available.

Nevertheless, the structures formed by game events in our corpus are significantly richer than a mere sequential ordering. Moreover, these structures can be influenced by chat moves—the chat moves providing “inverse information” in the sense of Katagiri et al. (2006) and Perry (1986) about the nonlinguistic context. Sorting out the structural possibilities was thus an imposing task, and all the more so given the sheer size of the situated annotations: our choices about which nonlinguistic event types to consider yielded 31,811 EEUs in the situated annotations, in contrast to 12,588 EDUs in the chat-only annotations. In what follows, we detail what we found, starting at the level of individual relations before moving on to larger structures.

#### 3.1.1 RELATIONS RELEVANT FOR THE INTERPRETATION OF THE SITUATED ANNOTATIONS

Many pairs of EEUs were in fact linked by the Sequence relation, where  $\text{Sequence}(\epsilon_n, \epsilon_m)$  holds just in case the event described by  $\epsilon_n$  took place before that described by  $\epsilon_m$ . But EEUs were also often linked by either Result, which indicates that the first EEU describes the cause of the second, or Continuation, which has the semantics of dynamic ‘&’ and simply requires that its arguments be true without commitment to temporal order. In Figure 5 ([CLICK FOR GRAPH](#)), J's comment in 207 is clearly a comment on 204, as it is about her dice roll, but it is also about 205 and, importantly, about the causal relationship between 204 and 205. Her dice roll sucked precisely because it *resulted* in a resource distribution for her opponents while yielding nothing for her, so a Result relation needs to be represented between 204, on the one hand, and the complex unit composed of the two segments for the two distribution events in 205, on the other. In addition, the two EEUs representing the distribution events in 205 need to be related via Continuation, as they occur simultaneously.

EEUs are also frequently arguments to Question-Answer Pair (QAP) and Elaboration relations. Figure 6 illustrates both cases ([CLICK FOR GRAPH](#)). Because the addressee of a trade offer (e.g., GWFS in 51.1 of Figure 6) was not identified in the Server messages (e.g., 51) for our game, we

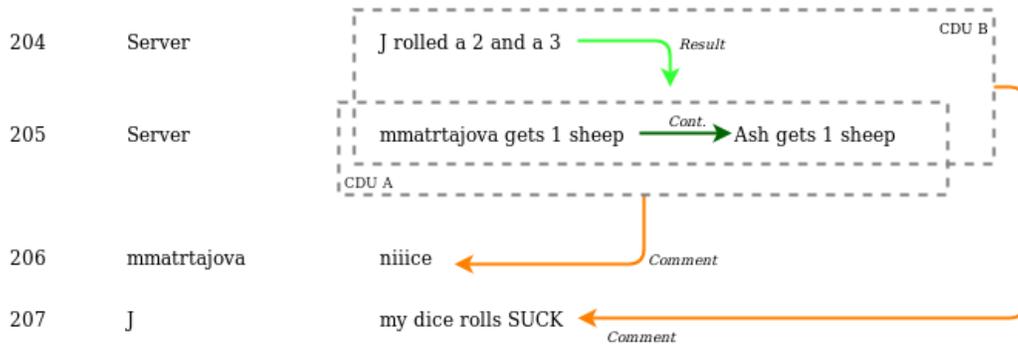


Figure 5: Result and Continuation relations between EEUs

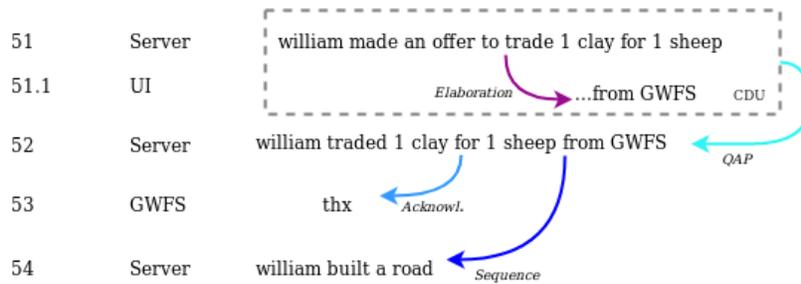


Figure 6: Elaboration and QAP relations between EEUs

had to extract this information separately, yielding a separate turn that more fully specifies an offer. We decided to link these pairs of EEUs via Elaboration, whose semantics in SDRT require that the second argument specify properties of the first. Thus in Figure 6, we get Elaboration(51, 51.1), and the fact that 51 and 51.1 work together to fully specify the offer is reflected by grouping them in a CDU. As Hunter et al. (2015a) argue, a natural way of understanding the relation between the CDU [51, 51.1] and the subsequent trade in 52 is as a QAP. The majority of linguistic offers in our corpus are in fact expressed as questions (see, for example, 165 in Figure 4), but even nonlinguistic offers in effect present a pair of alternatives: to trade or not to trade. A trade like that described in 52 then functions as a “yes”, and a refusal or rejection via the Trade Panel functions as a “no”.

Table 1 shows the distribution of relation labels for links between EEUs and/or CDUs containing only EEUs (i.e., with no EDUs at any level of constituency).

Type	Situated - EEUs only
Question_answer_pair	929
Continuation	8546
Elaboration	547
Result	12273
Correction	118
Background	45
Sequence	5695
Total	28153

Table 1: Relations between EEUs in the situated annotations

The few instances of Background in the corpus are very systematic: each time a player wins the game, the Server emits a message that reports the number of rounds that were played in the game and how long the game took. These messages were consistently attached with Background to the message announcing that a player had won the game (for games that were played through the end). Instances of Correction involve cases in which a player tries to make an illicit move and the system blocks the move, such as when a player tries to make a trade but lacks the necessary resources.

### 3.1.2 COMPLEX STRUCTURES RELEVANT FOR THE SITUATED ANNOTATIONS

A second way in which the structure over EEUs in our situated annotations departs from a mere sequential linear ordering, aside from involving relations other than Sequence, is that it was often natural or even necessary to group EEUs together as CDUs to capture the full content of a game. The exchanges in Figures 5 and 6, discussed above, provide two examples. Another case is when a player uses the trade interface to make successive offers to different players and then ends her turn after each of the offers results in a refusal. In such cases, annotators judged that the accumulation of refusals brought about the player’s decision to abandon her strategy and end her turn, and thus we grouped the series of failed offers in a CDU and related this CDU to the End Turn move via Result.

We also decided to use CDUs to group robber-related events, as players often commented on the complex events as a whole. As explained in Section 2.1, a roll of a 7 in *Settlers of Catan* brings out a game piece known as the “robber” and triggers a complex series of events that includes at least moving the robber and choosing a player to steal from, and depending on the configuration of the game board at that time, possibly other events as well. Figure 7 provides an example ([CLICK FOR GRAPH](#)). The decision to systematically represent robber events with CDUs, in which each sub-event causes the next sub-event, yielded the following relation instances for the robber events in Figure 7: Result(278,[278.1,278.2]) and Result(278.1,278.2). The situated annotations have a large number of CDUs, many of them quite large: 5777 EEU-only CDUs compared to 1450 EDU-only CDUs in the chat-only annotations. Table 8 in the Appendix provides more detail on the CDUs in both the chat-only and situated annotations.

While the larger discourse context can lead annotators to group EEUs into complex units, influence goes in the other direction as well: giving annotators access to the game events revealed that certain chat moves were working together in semantically significant ways that were not obvious in the absence of the nonlinguistic context. Turns 71-74 in Figure 8, for example, were not grouped

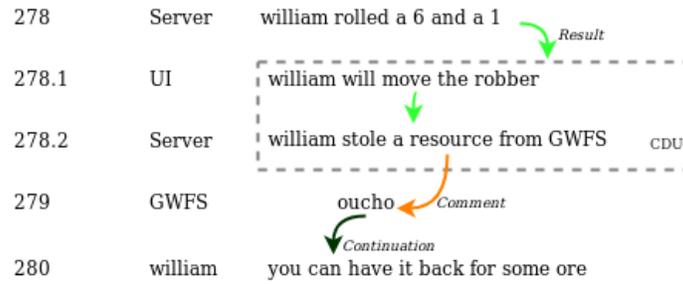


Figure 7: A typical robber sequence

in a CDU in the chat-only annotations, but had to be grouped together in the situated annotations ([CLICK FOR GRAPH](#)). In turn 71, T.K. expresses interest in trading with other players, but in turns

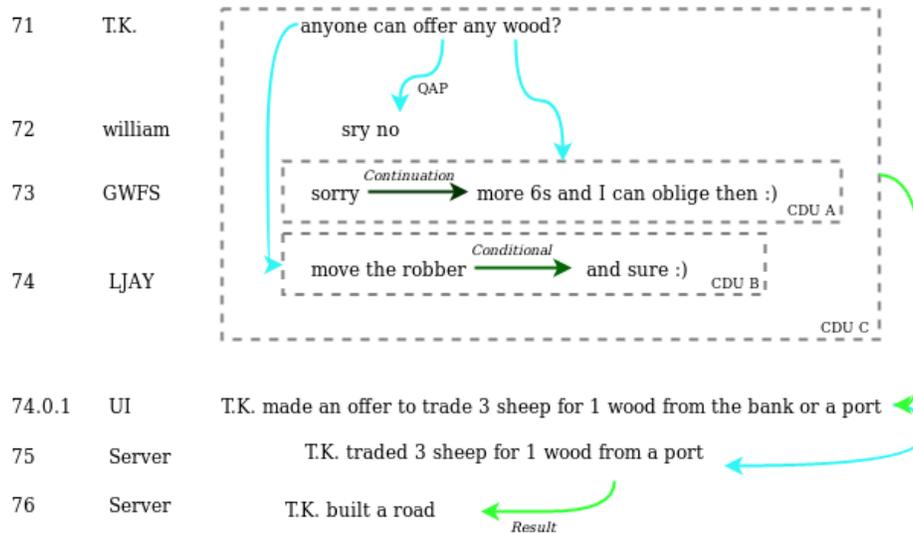


Figure 8: An EDU-only CDU introduced after consideration of game events

72-74, all of the other players reject his offer. In turn 75, T.K. goes to a port to get the wood he needs to build the road that he builds in 76. Because a trade from a port is far more “expensive” than a trade with another player—trades with other players are usually 1:1 or 2:1 at most, while a trade with a port is 3:1—annotators judged that T.K. only pursued this trade because his prior trading attempt was unsuccessful. That is, they judged that the trade was a *result* of the entire failed negotiation in 71-74.

### 3.2 Interactions between chat moves and game events

In the situated annotations, links between chat moves and game events run in both directions: the comment “oucho” in Figure 7 is a reaction to William’s stealing a resource in 280, but in Figure 8, T.K.’s trade with the port is a reaction to a failed negotiation exchange. In this subsection, we take a look at what kinds of structures arise from interactions between chat moves and game events and how they relate to discourse threads and discourse goals.

3.2.1 ASYMMETRIC AND INTERLEAVED STRUCTURES

Once we started to annotate the Server and UI messages, the conception of how dialogues should be individuated changed: the idea of using bargaining sessions as a rough guideline gave way to a turn-based criterion so that the situated exchanges in our corpus are broken down by turns that begin when a player gets the dice and end when she ends her turn and passes the dice to the next player. There are cases in which we find links across turn boundaries, as when a player comments on a move from the immediately preceding turn, but in general, breaking down the interactions in the games by turns worked out well. Figure 9, which represents the structure of an EEU only dialogue, gives an idea of what kind of structure a typical turn has (CLICK FOR GRAPH). Jon’s getting the dice in 238.0.2 results in his rolling a 6 and a 2 in 239, which in turn results in the set of resource distributions detailed in turn 240 (grouped into a CDU). The UI then updates information about the resources of all four players, and Jon ends his turn.

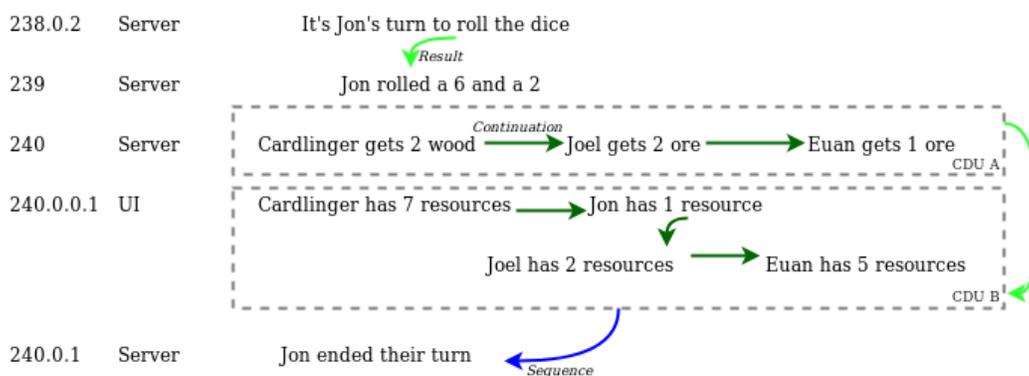


Figure 9: A standard EEU-only structure

The central role of game development often leads to an asymmetric semantic dependence of chat moves on the game moves in the situated annotations. This observation leads to the characterization of new types of discourse structure. Consider the example in Figure 10. The example begins with Dave getting the dice and ends with Dave finishing his turn. From the first to the last move, there is a continuous succession of game events: Dave rolls the dice, moves through a sequence of robber events, builds a road, tries to trade with the other players, buys a development card, and finally ends his turn. During all of this, he has a short exchange with Tomm about the resource that he stole (in 96, 98-100, and 102). This exchange with Tomm would not be interpretable without considering how the game is developing: we would be left to wonder what was unkind and what led to Dave’s getting one of Tomm’s resources. By contrast, were we to ignore this interaction, the development and interpretation of the game would remain completely intact and unchanged.

In Figure 10, we use a blue line to connect the moves central to game development and a magenta line to connect chat moves figuring in exchanges that asymmetrically depend on this development. The resulting subgraph indicated in blue is the backbone or what we will call the *core* of the structure in Figure 10. The set of outlying nodes connected in magenta determine two *peripheral structures*, the second of which is the exchange between Tomm and Dave described above. Clearly, ignoring the peripheral structures has no affect on the connectedness or the interpretation of the core, as the outlying nodes are all connected via outgoing edges, but ignoring the core leads to unconnectedness and uninterpretability for the peripheral structures.

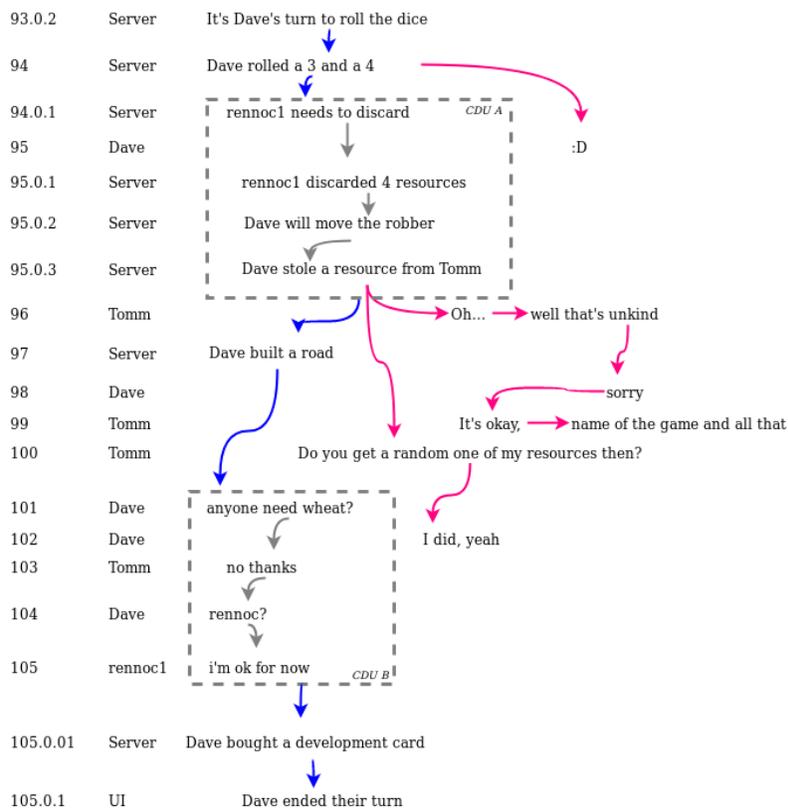


Figure 10: A dialogue and its asymmetric structure. The bold blue line traces the core, which runs through the EEUs contributed by Server and UI messages, including a CDU for the robber sequence in 94.0.1 and 95.0.1–95.0.3, and a CDU containing the chat moves 101 and 103–105 involved in a trade negotiation. The peripheral structures, indicated by magenta edges, consist of nodes which do not belong to the core and their incoming edges ([CLICK FOR GRAPH](#)).

Note that chat events that figure in trade negotiations are considered core moves in Figure 10. The core of this structure therefore forms what we call an *interleaved* structure. An interleaved structure is set off not by its structural properties but by the types of its nodes: it is simply a multimodal graph that contains nodes for both EDUs and EEUs. In this case, EEUs feed into EDUs, but unlike in the asymmetric structures in our corpus, EDUs also feed into EEUs.

In interleaved structures, chat moves break into the intuitive structure of game events, so the latter cannot be seen as forming an autonomous linear structure of their own. This exposes a further type of complication for the idea that game events can be represented with a temporal linear ordering to which the linguistic context can simply be appended. Moreover, interleaved structures can also figure in novel, non-treelike structures, as illustrated by Figure 11 ([CLICK FOR GRAPH](#)). In turn 256, Nelsen offers to trade wood for sheep or ore and both Kersti and Tyrant Lord respond with interest (turns 257/261 and 260, respectively). Their verbal acceptances result in two trade offers by Nelsen, represented with two CDUs, [262,262.0.1] and [264,264.0.1], whose members are related via Elaboration. The CDU [264,264.0.1] is related to Tyrant Lord’s acceptance in 263 via Sequence,

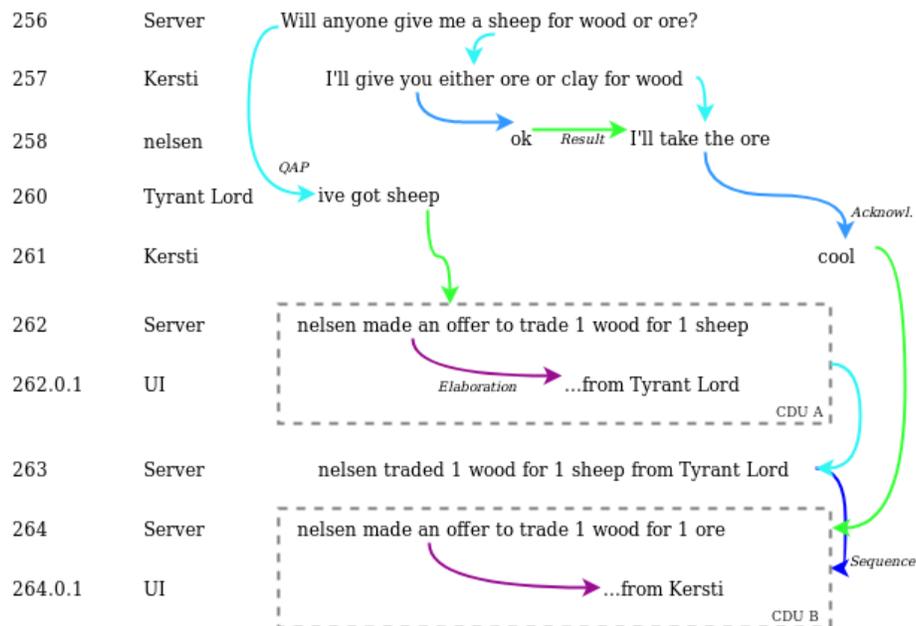


Figure 11: An example of a truly non-treelike CDU [264,264.0.1], with edges from 261 and 263

but it also has an incoming arrow from 261 labelled as a Result. The situated annotations add about 300 more non-treelike structures to the total for the chat-only annotations discussed in section 2.3.

Interactions between game events and chat moves also give rise to mixed CDUs, containing both EDUs and EEUs, that are interleaved into the larger game structure. One example is when a player reacts to a nonlinguistic trade offer with a chat move such as, “sorry, I don’t have wood” before rejecting the offer through the trade interface. Formal definitions for cores, peripheries, interleaved and asymmetric structures are given in the Appendix. In our two corpora, we have considered only maximal cores, also defined in the Appendix, that start with the initial DU of a dialogue and end with the last DU with respect to the textual, or in our case time stamp, ordering. Table 10 in the appendix gives statistical details concerning asymmetric and interleaved structures.

### 3.2.2 DISCOURSE THREADS AND GOALS

Asymmetric structures arise anytime a discourse splits into two threads that branch off of the same node.<sup>9</sup> This is a commonplace occurrence at dinner parties, for instance, when a group is chatting and then two people in the group decide to continue the conversation in slightly different directions, leading to a (perhaps temporary) split of the group into subgroups that each follow a different branch of the preceding conversation. In such a case, we could in principle choose either continuation of

9. For the purposes of this paper, we define asymmetric structures as involving two threads that never rejoin (which means that truly non-treelike structures do not count as asymmetric structures). This is a delicate point: in Hunter et al. (2018), we discuss an example that violates this assumption, and it might be that in face-to-face conversations speakers can use expressions such as “we were just discussing the same thing” to bring together two groups of conversationalists and two conversation threads. We suspect, however, that these violations are very limited and generally need to be marked explicitly, so we do not think that our simplified definition is problematic for our present purposes.

the conversation as defining the core of the discourse structure, yielding two possible asymmetric structures; the notion of a core is a thematic and functional one.

The graphs in our situated annotations reveal threads of a more constrained form. First, if we look at the graph for any one of our games, we find one thread that runs through the entire game—the thread that contains the game events and perhaps some interleaved chat moves. Second, the threads that make up the periphery contain typically 2.28 nodes on average, with a range of 50 nodes (see Table 10 for more details), whereas a split in a regular conversation can lead to an extended discussion.<sup>10</sup> Third, the structures in our peripheries contain no outgoing links into the core. This is a part of what it is to be a peripheral structure, but what is interesting about our corpus is that there are so many “bushes” of structure that make up the periphery for each game. Finally, while the interactions that make up the peripheries in our situated annotations are less extended than full conversations, they are longer than the peripheral structures that we find in single-authored text. Appositive relative clauses, for example, generally contribute peripheral structures of one to two discourse units that attach to the main discourse via a relation such as Background, but do not play a central role in the progression of a discourse (Venant et al., 2013).

These structural characteristics correlate with interesting features of the interactions in our corpus. The continuity of the series of game events as well as the brevity of the interactions in the periphery and the tendency of the players to return to focus on core events reflects the fact that playing and winning the game is clearly the leading goal and one that drives the interaction in the corpus; linguistic exchanges are secondary to this goal. Furthermore, if we look at the content of the chat exchanges and how they relate to the core, we find that they are largely reactive, being frequently attached to EEUUs via relations such as Comment (see Table 2 for details). These features correlate with the limited size of the peripheral structures and the fact that the structures do not feed back into the core—it is unsurprising that a commentary would fail to divert attention entirely away from the game and that it would be inert with regard to game development. Other types of threads might have a different discourse function that would be reflected through different structural features. A clarification request, for example, might temporarily take conversational participants away from a question that has been asked, but then feed back into it by having a direct impact on the nature of the answer that will ultimately be given to the main question (cf. Ginzburg, 2012).

While the goal of this paper is not to develop an account of exactly how the shape of a discourse structure or the distribution of various discourse relations relates to the nature of different discourse goals, we emphasize that there is an important connection between these topics that would be worthy of exploring in future work. To give another brief example, GWFS was the player who ultimately won the competition that we set up to build our corpus, and we suspect that part of his strategy for winning was to introduce conversational threads that were less closely related to the main events of the game than most of the other peripheral exchanges, especially those led by other players. It was almost as though he was trying to distract his competition. Table 3 shows that GWFS initiated more peripheral structures than other players. Figure 4, provides an illustration of GWFS’s behavior: LJAY

---

10. Conversational threads in face-to-face conversation often subdivide interlocutors into mutually exclusive groups. In our games, by contrast, two different threads can involve the same set of speakers. This is in part due to the task-based setup—even in face-to-face interactions, people can chat while also coordinating on a task that might require occasional discussion. It is also in part due to the chat environment—two people can easily carry on two conversational threads even in the absence of a task that they are jointly performing. It would be interesting in future work to see how different kinds of threads associate with different constraints on group membership, but we cannot go into that topic here.

relation_type	#	relation_type	#
Comment	1262	Continuation	40
Question_answer_pair	514	Result	38
Acknowledgement	355	Parallel	29
Elaboration	134	Background	25
Clarification_question	130	Sequence	15
Explanation	108	Correction	5
Contrast	58	Narration	3
Q_Elab	57	Alternation	1
		Conditional	1

Table 2: Breakdown of relations that connect asymmetric structures to a core in the situated annotations

initiates a trade negotiation, and GWFS replies to LJAY but then asks, “so how do people know about the league?,” setting off a discussion that is independent of the game at hand.

In the chat-only and situated annotations, we related GWFS’s question to his reply to LJAY via Background, but this is arguably unsatisfying. His question is not directly related to LJAY’s attempt to trade or even to the particular game that they are playing. Intuitively, he is “popping” up to a much larger, implicit topic that involves the information that they are playing this game and the other games as a part of a league (a pop that is signalled by his use of “so”). This is not an issue that we attempted to tackle when building our corpus, nor did it seem to us to be frequent enough to pose a significant problem for our annotation approach, but the hypothesis that GWFS’s strategy would be reflected by the way that his contributions influence the overall shape of the discourse would be an interesting topic for future investigation, and a potential point of contact with work on implicit topics and Questions Under Discussion (Ginzburg, 2012; Roberts, 2012).

emitter	#	emitter	#
GWFS	351	gramos	71
inca	164	raefbrisbin	67
LJAY	116	dmm	62
ztime	102	T.K.	61
skinnylinny	98	zorburt	57
sabercat	98	somdechn	51
william	81	Shawnus	51
Cardlinger	80	nelsen	51
		Tomm	51

Table 3: Number of peripheral structures started per speaker (cutoff at 50)

#### 4. Preservation of linguistic structure

We have emphasized that modelling the contribution of nonlinguistic events to a larger discursive interaction involves modelling the interaction between different *structures*. Not only does the non-linguistic context do more than add information via grounding or reference, it does more than merely contribute contents that function like dialogue acts. We must consider what kinds of structures non-linguistic events give rise to in their own right and how these structures interact with linguistic structures. Section 3.2.1 took on the first task and showed how the game events in our corpus form a rich structure of their own, while Section 3.2.2 tackled the second task and revealed new kinds of situated discourse structures that reveal information about how the different EDUS and EEUS involved contribute to discourse goals.

Section 4 aims to flesh out the claim about the importance of nonlinguistic structure by comparing the effects of adding game events in the situated annotations at the level not only of individual events and relation instances, but also at the structural level. While preservation of relation instances shows that our decisions about these instances were relatively robust even when limited to information present only in the linguistic turns, the figures for structural preservation show conclusively that we cannot take the situated annotations to be a conservative extension of the chat-only annotations. This is a familiar point when performing an error analysis on the output of a learning program for discourse structure: a respectable F1 score at the level of instances might not mean much in terms of structure preservation; the predicted structures might still be useless as discourse structures, as argued in Ferracane et al. (2019).

This comparison is made possible by the fact that the chat-only annotations in our corpus were completed in their entirety before we began the situated annotations and by the fact that the data set used for the situated annotations is a minimal extension of that used for the chat-only ones. That is, the two data sets differ only in that one includes game events and the other does not, enabling us to identify and circumscribe the semantic contribution of nonlinguistic events at the discourse level. The data sets thereby form something like a minimal pair, as illustrated by (1) and (2):

- (1) Louise read any book.
- (2) Louise read every book.

Minimal pairs are a useful tool in semantics for characterizing the semantic contribution of a certain type of expression. (1) and (2), for example, are useful for studying the behavior of polarity items, and their semantic value is only helped, not hindered, by the ungrammaticality of (1). While our pair of data sets differs importantly from minimal pairs used in formal semantics, as pointed out by one reviewer, a similar point can be made: rather than making it semantically irrelevant, the lack of game events in our chat-only data set renders it invaluable for our purposes.

All discourses take place in a larger context involving shared knowledge and presuppositions, if not a shared visual scene, and no corpus can account for all contextual information. Even our final corpus is impoverished in the sense that it does not include information on most game states because we could not find a way to extract information about them from the UI in a concise way that also would reflect their persistence rather than make them look like punctual events. What is important is that a corpus be adequate for the task for which it is employed. In our case, we do not need the most completely annotated corpus possible; we need one that characterizes methodically the influence of information about game events on the construction and interpretation of discourse structures for the chat moves. Our two data sets and their associated annotations are perfectly tailored for this task.

We begin in Section 4.1 with a simple comparison of the chat-only annotations and the situated annotations to show how much information was added in the extension and how many relations were preserved. Section 4.2 then looks at how structures were preserved—or not—as we moved from the chat-only annotations to the situated annotations.

#### 4.1 Preservation of relations

Both naive and expert annotators for the chat-only annotations attempted to relate each noninitial chat move to another part of the chat in order to capture the discursive function of the former. When they were unable to find a reasonable attachment point, the result was an “orphan” linguistic move; that is, a move with no incoming link. For example, GWFS’s comment “oucho” in Figure 7, repeated below as Figure 12, was left as an orphan in the chat-only annotations, but once the game events in Figure 12 were included it became clear that turn 279 was a reaction to and a commentary on William’s stealing a resource in 278.2 (CLICK FOR GRAPH).

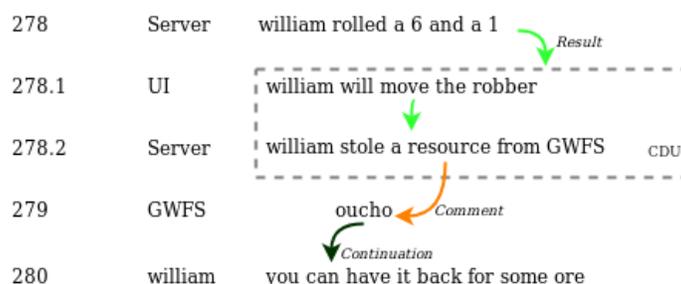


Figure 12: Turn 279 was an orphan in the chat-only annotations

In all, the chat-only annotations contained 1501 orphans, all of which were assigned incoming links after we introduced EEU’s into our annotations.<sup>11</sup> This led to a significant difference between the number of semantic links between the two sets of annotations: while the chat-only annotations contain 12,271 links, the situated annotations contain 3,591 links that relate an EDU (or a CDU containing at least one EDU) with an EEU (or a CDU containing at least one EEU).

In many cases, adding information about game events not only led to new relations, but it also led annotators to revise judgments about relations between chat moves. Table 4 shows how many relations from the chat-only annotations persisted when we moved to the situated annotations. As indicated in the table, around 20% of the relations in the chat-only annotations changed or disappeared in the situated annotations. In addition, relations were added between chat moves in the situated annotations, so that the persisting chat-only annotations made up 72% of the relations between chat moves in the situated annotations. The latter thus constitute a non-negligible revision of the chat-only annotations. Still, one might argue that linguistic structure was largely preserved; almost three quarters of the chat-only annotations were appropriate for the final corpus. Had we been

11. As explained in Section 2.1, the chats for the corpus were subdivided into dialogues (individuated roughly by player turns) in order to make annotation more tractable. While annotators were instructed to relate each non-initial dialogue chat move to another (dialogue internal or external) move where possible, they were also allowed to relate a dialogue initial chat move to another move if they felt it was the right thing to do. This instruction probably led to a bias towards finding antecedents for non-initial dialogue moves that was absent for initial moves. The number of orphans given above includes every dialogue-initial chat move of a non-chat-initial dialogue. If we consider only non-initial dialogue orphans, the total is 400.

Relations between chat moves (i.e., EDUs and CDUs composed only of EDUs)	#
Linguistic Games	12271
Situated Games	13605
Persisting	10206
Strictly Persisting (same relation type)	9829

Table 4: Persisting relations

gauging the success of a machine learning algorithm for learning discourse structure, this precision score would have counted as a success.

Figure 13 illustrates the difference between measuring preservation at the level of relation instances and the level of structures. The graph on the left is from the chat-only annotations and the graph on the right, from the situated annotations ([CLICK FOR FULL GRAPHS](#)).<sup>12</sup> Four out of five links from the left graph are preserved in the right graph; the only link that was removed in the shift to the situated annotations is a Continuation link between EDUs 425, the top EDU in the linguistic graph and EDU 427, represented in dark green. However, this link does not assign a real function to EDU 425, as it is naturally understood as a commentary on something, but not EDU 427. In the situated annotation, 425 comments on a CDU of game events or states (represented by the second red node) in the incoming context, and not 427. And this comment serves to explain why the author of 425 makes the offer he does in 427. This change breaks up a connected structure in the linguistic graph into two pieces in the situated annotation, and the two sub-structures figure in significantly different parts of the overall game structure. We also get information about the temporal relation between the top two nodes in the left graph and the bottom four: the blue arrow along the right side of the right graph represents Sequence, which imposes a temporal order on its arguments. This one small change in fact changes the structure *and* content of the conversation substantially.

## 4.2 Preservation of substructures

To measure structure preservation across annotations, we check for elementary embeddings of chat-only dialogue structures into the situated structures of the corresponding dialogues, where an elementary embedding preserves all relations, functions, and designated objects in  $\mathfrak{A}$  (Chang and Keisler, 1973). More precisely:

**Definition 1** *An elementary embedding is a one-to-one function  $f$  from the domain  $A$  of a structure  $\mathfrak{A}$  to the domain  $B$  of a structure  $\mathfrak{B}$  such that for any relation  $R$ , function  $g$  and designated object  $a$  in the signature of  $\mathfrak{A}$ ,  $\mathfrak{A} \models R(b, c)$  iff  $\mathfrak{B} \models R(f(b), f(c))$ ;  $\mathfrak{A} \models g(b) = c$  iff  $\mathfrak{B} \models g(f(b)) = f(c)$ , and  $f(a) = b$ , where  $b$  is the designated object of  $\mathfrak{B}$ .*

When such an embedding  $f$  exists,  $f(\mathfrak{A})$  is an *elementary substructure* of  $\mathfrak{B}$ . In fact, in our situation, there is a canonical embedding  $f$  that is the identity function on discourse units in the linguistic structure; so to say that  $f(\mathfrak{A})$  is an *elementary substructure* of  $\mathfrak{B}$  is just to say that  $\mathfrak{A}$  is an *elementary substructure* of  $\mathfrak{B}$ .

12. In Figure 13, we adopt the more minimalist format used for our clickable graphs to facilitate a visual comparison of the two structures under discussion. The more detailed graphs used throughout the rest of the paper would have been too large to place side by side.

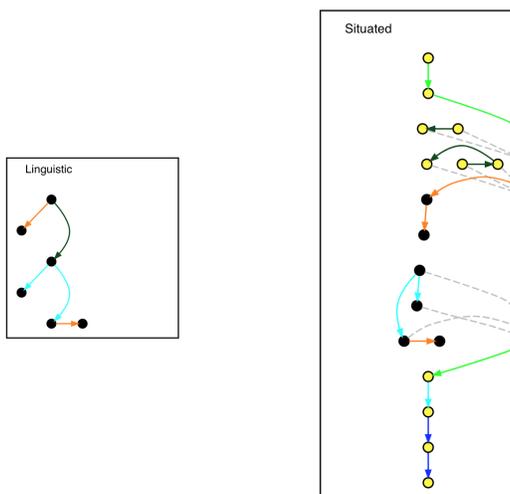


Figure 13: All but one of the relations in the chat-only structure (left) are preserved in the situated structure (right), but the former is broken into two pieces when we move to the situated annotations

Our discourse graphs are first-order structures with a designated object—in other words they are pointed models, as can be seen from Definition 2 for SDRT graphs.

**Definition 2** A discourse graph  $G$  is a tuple  $(V, E_1, E_2, \ell, Last)$ , where  $V$  is a set of EDUs and CDUs;  $E_1$ , a set of edges in  $V^2$  representing discourse attachments;  $E_2$ , a set of edges that relate each CDU to its members;  $\ell : E_1 \rightarrow \mathcal{P}(\mathcal{RL})$ , a function that labels the discourse attachments from  $E_1$  with sets of discourse relation labels taken from  $\mathcal{RL}$ ; and  $Last$ , a label for the last unit in  $V$  relative to textual order.

To check for the relevant elementary embeddings from our chat-only structures into situated structures, we ignore the element  $Last$  from Definition 2 so that we consider embeddings only on quadruples  $(V, E_1, E_2, \ell)$ .  $Last$  is used to define constraints on how a discourse can dynamically evolve, but in checking the preservation on a finished structure, we can (and must) ignore it.

Let  $\mathcal{L}$  be the weakly connected linguistic substructure for a dialogue  $d$  (henceforward our structures will pertain to whole dialogues), and let  $\mathcal{S}$  be the situated structure for  $d$  that includes all the nodes of  $\mathcal{L}$ . We say that  $\mathcal{L}$  is *elementarily preserved* in  $\mathcal{S}$  just in case  $\mathcal{L}$  is an elementary substructure of  $\mathcal{S}$ . Where major changes to the *chat-only annotations* occur in the situated annotation,  $\mathcal{L}$  will not survive as an elementary substructure of  $\mathcal{S}$ . Our results show that substructure preservation is a much stronger condition on persistence of information than the preservation of relation instances; we get 72% preservation of relation instances, but only 496 out of the 1137 dialogue structures in the chat-only annotations, or a little over 43%, are preserved as elementary substructures of the corresponding  $\mathcal{S}$  structures.

### 4.3 Preservation of subtypes

Building on discussions of discourse-central or *at-issue* content (Potts, 2005; Roberts, 2012; Simons et al., 2010), much current work in formal semantics and pragmatics asks how the use of a certain linguistic form or expression indicates the role of a particular discourse move in achieving

discourse goals. The results from our corpus study add a new dimension to this discussion because the information that is driving discourse development and the achievement of discourse goals in our corpus can often not be predicted at the level of individual linguistic moves. It is only by looking at the game-chat structures as a whole and the way that substructures like cores and peripheries function within these larger structures that we can see how individual moves inside of those structures address discourse goals. With this in mind, it is important to consider how *subtypes* of discourse structures—e.g., core and periphery types—are preserved between our two sets of annotations.

When a core structure is preserved by an  $\mathcal{S}$  embedding, this indicates that while the linguistic features of *individual* chat moves might not have been sufficient to reflect the role of these moves in achieving discourse goals, they were at least helpful in determining the role of these moves in linguistic substructures whose relevance to discourse development and goals was clear. Linguistic core structures that remain core structures in the situated annotations form an integral part of the game and address the main point of the overall discourse, which is to play and win the game. On the other hand, preservation of a peripheral structure indicates that linguistic information at the level of substructures was sufficient for determining the discursive functions of the individual moves in these substructures and that these moves were not integral to the main point.

More formally, we say that an  $\mathcal{L}$  substructure  $\mathfrak{A}$  of type  $\tau$  is  $\tau$  *preserved* under the canonical  $\mathcal{S}$  embedding  $f$  into an  $\mathcal{S}$  structure  $\mathfrak{B}$  just in case  $\mathfrak{A}$  is a substructure of a structure of type  $\tau$  in  $\mathfrak{B}$ . For example, if we consider the structure  $\mathfrak{A}$  for a core of a dialogue  $d$  in the chat-only annotations, then we say that  $\mathfrak{A}$  is core-preserved under an  $\mathcal{S}$  embedding  $f$  just in case  $\mathfrak{A}$  is a substructure of the core of  $\mathfrak{B}$ . When there is an  $\mathcal{S}$  embedding of an  $\mathcal{L}$  structure  $\mathfrak{A}$ , which may contain both core and periphery-type structures, such that all substructures of  $\mathfrak{A}$  maintain their type under the embedding, we call this *perfect type preservation*. Perfect type preservation strictly entails elementary preservation, core-type and periphery-type preservation, while core-type and periphery-type preservation together entail perfect type preservation.

In general, perfect type preservation is rare in our corpus. Tables 11 and 12 in the Appendix provide the details, but we summarize the main results here. Out of 296 core-only  $\mathcal{L}$  structures, which were in some sense the simplest case, only 23% had an  $\mathcal{S}$  embedding that was core-type preserving (and hence perfect-type preserving). Type preservation results were much worse for asymmetric structures; only 4% exhibited perfect type preservation; 7% of all  $\mathcal{L}$  structures exhibited periphery-type preservation and 11,5% exhibited core-type preservation embedding.

The general moral we draw from this study is that at least in our situated annotations, information relevant to the main goals of the conversational participants is not reliably signalled by linguistic means, either at the level of individual speech acts or at the level of relations. In fact our discussion in Section 3.2.2 suggests that linguistic chat in our corpus may reflect a secondary goal to distract other players from their main goals, something that the subtype preservation results support at least indirectly. In any case, elementary and subtype preservation will prove, we think, useful tools for analyzing the interplay between player goals, strategies and linguistic and nonlinguistic actions.

## 5. Conclusions

In this paper, we have surveyed and compared two sets of annotations of a corpus culled from an online version of the game *The Settlers of Catan*: the *chat-only annotations*, which take only chat moves into account, and the *situated annotations*, which also include descriptions of visually presented events from the games during which the players were chatting. Considering the nature

of the situated annotations and how they compare with the chat-only annotations allowed us to illuminate and measure a variety of ways in which information from the nonlinguistic context can influence the content and structure of a discourse.

Our study provides new data and statistics to substantiate claims from Hunter et al. (2018) that modelling discourse situated in a dynamically evolving nonlinguistic context requires attributing a rich structure to that context, and that this structure can interact with purely linguistic structures in new and interesting ways. The nature and extent of nonlinguistic influence evidenced in the comparison of our chat-only and situated annotations supports the claim that nonlinguistic information is relevant for far more than reference or domain restriction; nonlinguistic events can contribute entire propositions to discourse content without being picked out by a linguistic expression or deictic act. But we have also showed that we must go further than allowing nonlinguistic events to contribute arguments that link them rhetorically to the contents of speech acts: the structural and semantic relations that link nonlinguistic events to each other can also influence a linguistic message. What's more, as we showed in Section 3, looking at how the nonlinguistic and linguistic structures interact can give us a more complete understanding of what is driving discourse development and how speech acts and nonlinguistic events contribute to achieving overall discourse goals. Finally, Section 4 reinforced the importance of considering nonlinguistic structures by showing that when we consider structure, rather than merely relation instances, we get a different and more complete picture of how the game events in our corpus influenced the overall content of our annotations.

Future work will involve generalizing the results described in this paper to other types of corpora. The chats in our corpora are mostly directed towards the competitive task at hand (winning the game), the nonlinguistic events that show up in our discourse structures are highly standardized due to game rules, and the virtual environment leads to a relatively controlled and simplified nonlinguistic context. These features simplified the annotation of relations between the chats and the game state, which allowed us to carry out the comparisons described in this paper. Other situated conversations might take place in a much more complex and varied environment. Still, we believe that the basic points that we have used our corpora to make about situated discourse are largely general. In any case, at this point in empirical research on situated discourse, we suspect that a certain amount of standardization is a necessary feature of events that convey discourse functions beyond causal or sequential discourse relations, functions like answering or posing a question, for example. We would like to pursue this question in future research.

## References

- Stergos Afantenos, Nicholas Asher, Farah Benamara, Myriam Bras, Cécile Fabre, Lydia-Mai Ho-Dac, Anne Le Draoulec, Philippe Muller, Marie-Paule Péry-Woodley, Laurent Prévot, Josette Rebeyrolle, Ludovic Tanguy, Marianne Vergez-Couret, and Laure Vieu. An empirical resource for discovering cognitive principles of discourse organisation: the ANNODIS corpus. In *The Eight International Conference on Language Resources and Evaluation*, pages 2727–2734, Istanbul, Turkey, May 2012. European Language Resources Association (ELRA). URL <http://www.lrec-conf.org/proceedings/lrec2012/index.html>.
- Stergos Afantenos, Eric Kow, Nicholas Asher, and Jérémy Perret. Discourse parsing for multi-party chat dialogues. In *The 2015 Conference on Empirical Methods in Natural Language Processing*, pages 928–937, Lisbon, Portugal, September 2015. Association for Computational Linguistics. URL <http://aclweb.org/anthology/D15-1109>.

- Nicholas Asher. *Reference to Abstract Objects in Discourse*. Kluwer Academic Publishers, 1993.
- Nicholas Asher and Alex Lascarides. *Logics of Conversation*. Cambridge University Press, 2003.
- Nicholas Asher and Soumya Paul. Strategic conversation under imperfect information: epistemic Message Exchange games. *Logic, Language and Information*, 27(4):343–385, 2018.
- Nicholas Asher, Julie Hunter, Mathieu Morey, Farah Benamara, and Stergos Afantenos. Discourse structure and dialogue acts in multiparty dialogue: the STAC corpus. In *The Tenth International Conference on Language Resources and Evaluation*, pages 2721–2727, Portorož, Slovenia, May 2016. European Language Resources Association (ELRA). URL [http://www.juliejhunter.com/uploads/3/9/6/1/39617901/lrec2016\\_final.pdf](http://www.juliejhunter.com/uploads/3/9/6/1/39617901/lrec2016_final.pdf).
- Jason Baldridge, Nicholas Asher, and Julie Hunter. Annotation for and robust parsing of discourse structure on unrestricted texts. *Zeitschrift für Sprachwissenschaft*, 26:213–239, 2007.
- Tom Bartlett. *Analyzing Power in Language: A Practical Guide*. Routledge, London, 2014.
- Lynn Carlson and Daniel Marcu. Discourse tagging reference manual. *ISI Technical Report ISI-TR-545*, 54:56, 2001.
- Lynn Carlson, Mary Ellen Okurowski, and Daniel Marcu. *RST Discourse Treebank LDC2002T07*. Linguistic Data Consortium, University of Pennsylvania, 2002.
- C.C. Chang and H. Jerome Keisler. *Model Theory*. North Holland Publishing, 1973.
- David Crystal. *Internet Linguistics*. Routledge, London, 2011.
- Elisa Ferracane, Greg Durrett, Junyi Jessy Li, and Katrin Erk. Evaluating discourse in structured text representations. In *The 57th Annual Meeting of the Association for Computational Linguistics*, pages 646–653, Florence, Italy, July 2019. Association for Computational Linguistics. URL <https://www.aclweb.org/anthology/P19-1062>.
- Katherine Forbes, Eleni Miltsakaki, Rashmi Prasad, Anoop Sarkar, Aravind K. Joshi, and Bonnie L. Webber. D-LTAG system: Discourse parsing with a lexicalized tree-adjointing grammar. *Journal of Logic, Language and Information*, 12(3):261–279, 2003.
- Jonathan Ginzburg. *The Interactive Stance: Meaning for Conversation*. Oxford University Press, Oxford, 2012.
- Julie Hunter, Nicholas Asher, Eric Kow, Jérémy Perret, and Stergos Afantenos. Defining the right frontier in multi-party dialogue. In *The Semantics and Pragmatics of Dialogue (goDIAL)*, pages 95–103, Gothenburg, Sweden, August 2015a. URL [https://flov.gu.se/digitalAssets/1537/1537599\\_sem2015\\_godial\\_proceedings.pdf](https://flov.gu.se/digitalAssets/1537/1537599_sem2015_godial_proceedings.pdf).
- Julie Hunter, Nicholas Asher, and Alex Lascarides. Integrating non-linguistic events into discourse structure. In *The 11th International Conference on Computational Semantics (IWCS)*, pages 184–194, London, England, April 2015b. URL <http://www.aclweb.org/anthology/W15-0123>.
- Julie Hunter, Nicholas Asher, and Alex Lascarides. A formal semantics for situated conversation. *Semantics & Pragmatics*, 11(10), 2018. URL <https://doi.org/10.3765/sp.11.10>. early access.

- David Kaplan. Demonstratives. In J. Almog, J. Perry, and H. Wettstein, editors, *Themes from Kaplan*. Oxford, 1989.
- Yasuhiro Katagiri, Mayumi Bono, and Noriko Suzuki. Conversational inverse information for context-based retrieval of personal experiences. In *JSAI'05: Proceedings of the 2005 international conference on New Frontiers in Artificial Intelligence*, pages 365–376. JSAI, March 2006.
- Alfred Kranstedt, Peter Kühnlein, and Ipke Wachsmuth. Deixis in multimodal human computer interaction: An interdisciplinary approach. In A. Camurri and G. Volpe, editors, *Gesture-based Communication in Human-Computer Interaction*, pages 112–123. Springer, 2004.
- Geert-Jan M. Kruijff, Pierre Lison, Trevor Benjamin, Henrik Jacobsson, Hendrik Zender, Ivana Kruijff-Korbayová, and Nick Hawes. Situated dialogue processing for human-robot interaction. In H. I. Christensen, G. M. Kruijff, and J. L. Wyatt, editors, *Cognitive Systems*, pages 311–364. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. URL [http://dx.doi.org/10.1007/978-3-642-11694-0\\_8](http://dx.doi.org/10.1007/978-3-642-11694-0_8).
- Alex Lascarides and Nicholas Asher. Agreement, disputes and commitment in dialogue. *Journal of Semantics*, 26(2):109–158, 2009.
- Alex Lascarides and Matthew Stone. A formal semantic analysis of gesture. *Journal of Semantics*, 26(4):393–449, 2009.
- William C. Mann and Sandra A. Thompson. Rhetorical structure theory: A framework for the analysis of texts. *International Pragmatics Association Papers in Pragmatics*, 1:79–105, 1987.
- Jérémy Perret, Stergos Afantenos, Nicholas Asher, and Mathieu Morey. Integer linear programming for discourse parsing. In *the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 99–109, San Diego, California, June 2016. Association for Computational Linguistics. URL <http://www.aclweb.org/anthology/N16-1013>.
- John Perry. Perception, action, and the structure of believing. In R. Grandy and R. Warner, editors, *Philosophical Grounds of Rationality*, pages 333–361. Oxford, 1986.
- Livia Polanyi. A theory of discourse structure and discourse coherence. In W. H. Eilfort, P. D. Kroeber, and K. L. Peterson, editors, *Papers from the General Session at the 21st Regional Meeting of the Chicago Linguistics Society*. Chicago Linguistics Society, 1985.
- Livia Polanyi, Chris Culy, Martin van den Berg, Gian Lorenzo Thione, and David Ahn. A rule based approach to discourse parsing. In *the 5th SIGdial Workshop on Discourse and Dialogue*, pages 108–117, Cambridge, Massachusetts, USA, April 30 - May 1 2004. Association for Computational Linguistics. URL <http://www.aclweb.org/anthology/W04-2322>.
- Paul Portner. The semantics of imperatives within a theory of clause types. In *Semantics and Linguistic Theory 14*, pages 235–252, Evanston, Illinois, U.S.A., May 2004. URL <https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/2907>.
- Christopher Potts. *The Logic of Conventional Implicatures*. Number 7. Oxford University Press on Demand, 2005.

- Rashmi Prasad, Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Livio Robaldo, Aravind Joshi, and Bonnie Webber. The penn discourse treebank 2.0. In *The Sixth International Conference on Language Resources and Evaluation*, pages 2961 – 2968, Marrakech, Morocco, May 2008. ELRA. URL [http://www.lrec-conf.org/proceedings/lrec2008/pdf/754\\_paper.pdf](http://www.lrec-conf.org/proceedings/lrec2008/pdf/754_paper.pdf).
- Gert Rickheit and Ipke Wachsmuth. *Situated Communication*. Walter de Gruyter, 2006.
- Craige Roberts. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, 5(6):1–69, 2012.
- Mandy Simons, Judith Tonhauser, David Beaver, and Craige Roberts. What projects and why. In *Semantics and Linguistic Theory 20*, pages 309–327, Vancouver, British Columbia, Canada, April 29 - May 1 2010. URL <https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/2584>.
- Thora Tenbrink, Kathleen Eberhard, Hui Shi, Sandra Kuebler, and Matthias Scheutz. Annotation of negotiation processes in joint-action dialogues. *Dialogue and Discourse*, 4(2):185–214, 2013.
- Antoine Venant and Nicholas Asher. Dynamics of public commitments in dialogue. In *The 11th International Conference on Computational Semantics*, pages 272–282, London, England, 2015. Association for Computational Linguistics. URL <http://www.aclweb.org/anthology/W15/W15-0131>.
- Antoine Venant, Nicholas Asher, Philippe Muller, Pascal Denis, and Stergos D. Afantenos. Expressivity and comparison of models of discourse structure. In *Proceedings of Sigdial 2013*, pages 2–11, Metz, France, August 2013. URL <https://www.aclweb.org/anthology/W13-40.pdf>.
- Antoine Venant, Nicholas Asher, and Cedric Degremont. Credibility and its attacks. In *The 18th Workshop on the Semantics and Pragmatics of Dialogue*, pages 154–162, Edinburgh, Scotland, September 2014. URL <http://www.macs.hw.ac.uk/InteractionLab/Semdial/semddial14.pdf>.
- Antoine Widlöcher and Yann Mathet. La plate-forme GLOZZ : environnement d’annotation et d’exploration de corpus. In *la 16e Conférence Traitement Automatique des Langues Naturelles*, Senlis, France, June 2009. ATALA, LIPN. URL <http://talnarchives.atala.org/TALN/TALN-2009/taln-2009-court-023.pdf>.
- Florian Wolf and Edward Gibson. Representing discourse coherence: A corpus based study. *Computational Linguistics*, 31(2):249–287, 2005.

## 6. Appendix

### 6.1 Formal definitions for asymmetric structures (Hunter et al., 2018)

Let  $e(x, y)$  mean that the edge  $e$  connects its initial point  $x$  to its end point  $y$ .

**Definition 3** Let  $G = (V, E_1, E_2, \ell, Last)$  be a discourse graph and let  $i$  be the initial DU in  $V$  with respect to the textual order. A subgraph  $G' = (V', E'_1, E'_2, \ell^G \upharpoonright E'_1, Last)$  of  $G$  forms a **core**  $C$  just in case: (i)  $\{i, Last\} \subseteq V'$ ; (ii) the transitive closure of  $E'_1$  induces a transitive, asymmetric ordering  $R$  over  $V'$  in which for every element  $a$ , other than  $Last$  and  $i$ ,  $R(i, a)$  and  $R(a, Last)$ .

Note that any maximal chain over  $V$ , defined in the standard way, is a core, and any set of maximal chains over  $V$  forms a core as well. We call a core  $C$  of a graph  $G$  *maximal* just in case there is no substructure  $A$  of  $G$  such that  $P(A, C) \neq \emptyset$  and  $A$  is also a core of  $G$ .

**Definition 4** Let  $G$  be a discourse graph, and  $G'$  a subgraph of  $G$ . Let  $End(e)$  be the endpoints of an edge  $e$  and  $End(E) = \{x: \exists e \in E. x \in End(e)\}$ . Let  $\setminus$  stand for set-theoretic difference, and let  $End(E_1 \setminus E'_1) = V^P$  (for periphery). Then  $G - G' =_{defn} (V^P, E_1 \setminus E'_1, E_2 \upharpoonright V^P, \ell \upharpoonright (E_1 \setminus E'_1), x)$ , with  $x$  the last element in  $V \setminus V'$  ordered by a linear ordering  $\prec$  over  $V$ .

Note that  $G - G'$  may not be a discourse graph in our sense in that it is no longer weakly connected; nevertheless, it corresponds to a set of discourse graphs. Note also that for a given discourse graph  $G$  and substructure  $G'$ ,  $G'$  and  $G - G'$  may share nodes but form a partition over the set of relation instances or arcs in  $E_1$ .

**Definition 5**  $P(G, C)$ , the **periphery** of a structure  $G$  with respect to a core  $C$ , is such that  $P(S, C) = G - C$ .

**Definition 6** An **asymmetric structure**  $G$  is a graph with a core  $C$  such that  $C \neq G$ .

### 6.2 Additional Tables

Table 5: Table of truly non-treelike structures

	Linguistic	Situated
<b>Truly non-treelike structures</b>	928	1207
<b># truly non-treelike structures per dialogue</b>		
Mean	0.82	0.47
Max	13	14
Min	0	.tex 0

Table 6: Table of Basic Comparisons

	Linguistic	Situated
<b>overview</b>		
total Games	45	45
total Dialogues	1137	2593
total DUs (including CDUS)	14041	52050
<b># dialogues per game</b>		
max dialogues	45	120
min dialogues	2	13
mean dialogues	25.27	57.62
total Turns	10596	33940
<b># turns per dialogue</b>		
max turns	120	165
min turns	1	6
mean turns	9.88	14.08
<b># edus (or eeus) per turn</b>		
max per turn	4	4
min per turn	1	1
mean per turn	1.19	1.31
<b># edus (or eeus) per dialogue</b>		
max per dialogue	158	204
min per dialogue	1	5
mean per dialogue	11.07	17.12

Table 7: Table of argument types of relations

	Linguistic	Situated
edu - edu	9521	8741
edu - cdu	1454	2025
cdu - edu	995	1316
cdu - cdu	301	2776
edu - eeu	-	987
eeu - edu	-	462
eeu - eeu	-	19808
eeu - cdu	-	3324
cdu - eeu	-	3619

Table 8: Table of CDUs

	Linguistic	Situated
<b>total # cdus</b>	1450	7651
# depth-1	1394	7368
# depth-2	54	270
# depth-3	2	13
<b>cdu composition</b>		
edu only	1450	1825
eeu only	0	5777
both	0	49
<b># dus contained in cdus</b>		
mean	2.14	2.77
max	6	39
min	2	2

Table 9: The distribution of relation labels for links in the chat-only annotations and the situated annotations, with a breakdown for the situated annotations as follows: (i) relations between EEUs (and EEU-only CDUs), (ii) relations between EDUs (and EDU-only CDUs), and (iii) relations connecting chat and game moves.

Type	Linguistic	Situated	(i)	(ii)	(iii)
Question_answer_pair	2914	3865	929	2904	32
Comment	2037	2529	0	1644	881
Acknowledgement	1543	1771	0	1576	195
Continuation	1194	9391	8546	841	5
Elaboration	1044	1631	547	1054	30
Q_Elab	645	661	0	639	21
Contrast	537	539	0	532	7
Explanation	527	567	0	519	48
Clarification_question	456	529	0	404	125
Result	445	13741	12273	395	1074
Correction	232	293	118	169	7
Parallel	212	193	0	192	1
Conditional	155	154	0	153	0
Alternation	141	128	0	128	0
Narration	103	69	0	66	3
Background	86	134	45	86	3
Sequence	0	6863	5695	12	1159
Total	12271	43058	28153	11314	3591

As background to Table 10, around 38% of our 2595 situated structures are asymmetric; these structures average two peripheral structures attached to the core, with the maximum being 22 peripheral structures for one dialogue. The table makes clear the plethora of periphery type structures in the annotations. The remaining 1617 or 62% of our situated dialogues have what we call a *core-only* structure. Around 35% of the  $\mathcal{S}$  structures for dialogues in the situated annotations (902 out of 2595) are interleaved.

Table 10: Table of Asymmetric and Interleaved structures

	Linguistic	Situated
<b>Total periphery-type structures</b>	2129	2545
<b># periphery-type structures per dialogue</b>		
Mean	1.87	0.98
Max	18	22
Min	0	0
<b>Size of periphery-type structures</b>		
Max # nodes	39	51
Min # nodes	1	1
Mean # nodes	1.86	2.28
<b>Interleaved structures</b>		
Total # cores	1137	2593
Total # cores containing edus	1137	900
<b>% of core dus that are edus</b>		
Mean	1	0.29
Max	1	0.92
Min	1	0.02
<b>Percentage dialogue edus contained in core</b>		
Mean	0.68	0.21
Max	1	1
Min	0.02	0

Table 11: Table for Persisting core-only  $\mathcal{L}$  Structures

<b>Total persisting core-only <math>\mathcal{L}</math> structures</b>	161
<b>Perfect type preservation</b>	70
<b>Elementary preservation</b>	
core in $\mathcal{S}$ periphery	77
core in both $\mathcal{S}$ core and $\mathcal{S}$ periphery	14

Table 12: Table for Persisting Asymmetric  $\mathcal{L}$  Structures

<b>Total persisting asymmetric <math>\mathcal{L}</math> structures</b>	335
<b>Perfect type preservation</b>	95
<b>Core type preservation</b>	
$\mathcal{L}$ core in $\mathcal{S}$ core // $\mathcal{L}$ periph in $\mathcal{S}$ core & periph	33 or 28%
$\mathcal{L}$ core in $\mathcal{S}$ core // $\mathcal{L}$ periph in $\mathcal{S}$ -core	62
<b>Periphery type preservation</b>	
$\mathcal{L}$ core in $\mathcal{S}$ core & periph // $\mathcal{L}$ periph in $\mathcal{S}$ periph	30
$\mathcal{L}$ core in $\mathcal{S}$ periph and $\mathcal{L}$ periph in $\mathcal{S}$ periph	35
<b>Elementary preservation</b>	
$\mathcal{L}$ core in $\mathcal{S}$ core & periph // $\mathcal{L}$ periph in $\mathcal{S}$ core	28
$\mathcal{L}$ core in $\mathcal{S}$ periph // $\mathcal{L}$ periph in $\mathcal{S}$ core & periph	7
$\mathcal{L}$ core and $\mathcal{L}$ periph in $\mathcal{S}$ core and $\mathcal{S}$ periph	40
<b>Elementary Preservation with inverted types</b>	
$\mathcal{L}$ core in $\mathcal{S}$ periph // $\mathcal{L}$ periph in $\mathcal{S}$ core	5