Comparing discourse structures between purely linguistic and situated messages in an annotated corpus

Nicholas Asher NICHOLAS.ASHER@IRIT.FR
IRIT-CNRS, Toulouse
Affiliation1 second line

Julie Hunter JULIEHUNTER@GMAIL.COM
Universitat Pompeu Fabra, Barcelona
IRIT, Université Paul Sabatier, Toulouse

Kate Thompson ULLALUME@GMAIL.COM
IRIT, Université Paul Sabatier, Toulouse
Affiliation3 second line

Editor: Name Surname

Submitted MM/YYYY; Accepted MM/YYYY; Published online MM/YYYY

Abstract

This paper describes a corpus of situated multiparty chats developed for the STAC project[1] 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 has allowed us quantify various ways in which adding information from the nonlinguistic context affects dialogue structure. In this paper, we look at how annotations based only on linguistic information were preserved once the nonlinguistic context was factored in. We explain that while the preservation of relation instances is relatively high when we move from one corpus to the other, there is little preservation of higher order structures that capture "the main point" of a dialogue and distinguish it from peripheral information.

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

1. Introduction

The study of discourse structure on texts is now a well entrenched cottage industry in computational linguistics. Several 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 several smaller annotated corpora such as DISCOR (Baldridge et al., 2007) or ANNODIS (Afantenos et al., 2012a). The STAC project (Hunter et al., 2015b; Asher et al., 2016) extends this work to dialogue annotation, and the STAC corpus is, as far as we know, the only corpus with full discourse structure for dialogues. What's more, it is the only corpus with a complete set of full discourse annotations for *situated* dialogues. The corpus,

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

which consists of multiparty chats extracted from on-line game sessions of *Settlers of Catan*, was annotated in two stages: we first considered all of the chat moves in the corpus, yielding what we call the *linguistic corpus*, and we later extended and revised the linguistic corpus by taking into account game states and events that were observable through the game interface, yielding what we call the *situated corpus*.

The aim of this paper is to provide a detailed, empirical comparison of the linguistic and situated STAC corpora in order to bring out significant differences between situated and non-situated discourse and to clarify and quantify the influence that a situated environment can have on a linguistic message. Hunter et al. (2018) argue that modelling situated dialogue requires substantial revisions to models of discourse structure and semantics, because a situated environment allows interlocutors to exploit events in the surrounding situation that are independent of the discourse—in the sense that they were not produced with an intent to build on the discourse at hand—in the same way that they use the contents of speech acts to develop a discourse. In other words, the contents of discourse independent events can serve as arguments to discourse or rhetorical relations, and as argued in Hunter et al. (2018), they can do so even when the surrounding discourse context contains no referential or deictic expressions, meaning that the use of discourse independent events goes beyond the various forms of deixis and reference linguists and computer scientists typically study (Kaplan, 1989; Rickheit and Wachsmuth, 2006; Kranstedt et al., 2004; Kruijff et al., 2010). Such events do more than influence discourse interpretation; they play a constitutive role in its construction and development.

The study described in this paper provides strong empirical evidence for many of the theoretical claims of Hunter et al. (2018). It also goes beyond Hunter et al. (2018) by looking in more detail at the influence of the surrounding environment on discourse structure and by arguing for the importance of considering effects on global discourse structure rather than local discourse relations alone. Section 2 introduces the linguistic and situated corpora. Section 3 details the motivation for annotating game events, the complications that arose in annotating them, and the influence that adding them had on the annotations from the linguistic corpus. Section 4 then argues that to fully appreciate the impact that discourse independent events can have on a linguistic message, we need to consider the higher-order shape of a discourse structure, rather than individual relation instances, and examine how the overall shape of a structure built based on discursive cues alone can be affected once we take discourse independent events into account.

2. The *Settlers* corpora

In this section, we give an overview of the *Settlers* corpora and certain features of their annotations. More quantitative details can be found in the appendix. Section 2.1 explains the difference between the two corpora and provides an example from the situated corpus. Section 2.2 lays out the general approach adopted to annotate the corpora. Section 2.3 discusses three complications for building discourse structures in our corpora 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 linguistic and situated corpora

The *Settlers* corpora consist 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*

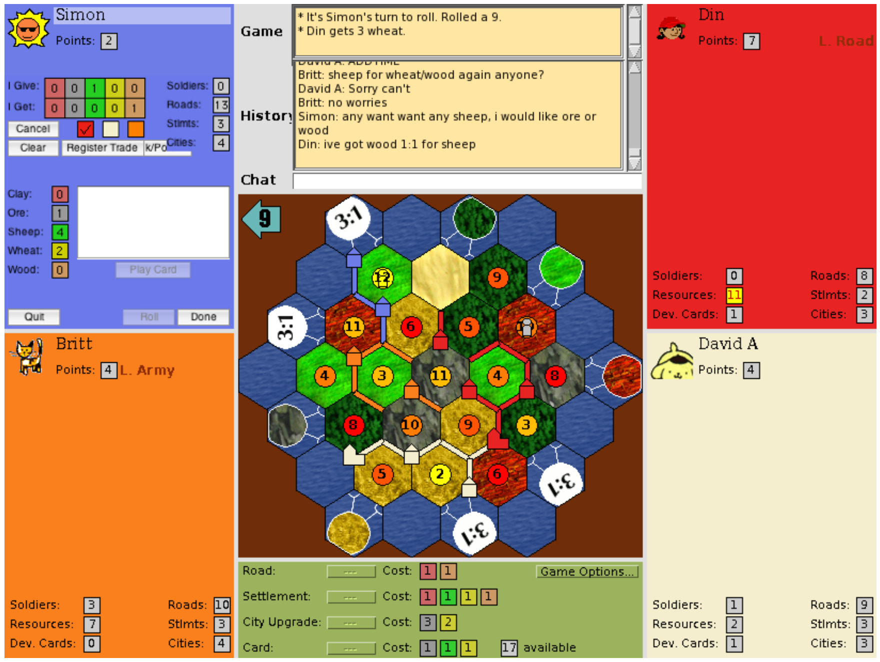

Figure 1: A snapshot of the game board for *The Settlers of Catan*

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, for example, 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.

To construct the *Settlers* corpora, we modified an online, open-source version of *Catan* to include a chat window. Figure 1 illustrates the game interface (Afantenos et al., 2012b) 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 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.

Linguistic Corpus. The *Settlers* corpora were 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 (Mathet and Widlöcher, 2009) to build what we refer to as the *linguistic corpus*. To make the annotation process more manageable, we originally divided the chat history for each game in the linguistic corpus 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 in intricate and semantically significant ways to the virtual gameboard and the game events. In some cases, for example, the absence of game-related information left annotators unable to connect (non-discourse-initial) anaphoric or fragmentary chat moves to any preceding discourse moves, meaning that some dialogues did not have a weakly connected structure, which should not happen in a representation of a coherent conversation. The incompleteness of the linguistic corpus 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 corpus*.

Situated corpus. (1) is an example from the situated corpus that illustrates how chat moves can interact with game events. (CLICK FOR GRAPH)[2]

(1)

433.0.3	Server	william played a Monopoly card.
433.0.4	Server	william monopolized wheat.
433.0.5	Server	It's william's turn to roll the dice.
434	GWFS	noooo!
435	Server	william rolled a 2 and a 1.
436	Server	GWFS gets 1 sheep. LJAY gets 2 wood. T.K. gets 2 wood.
436.0.0.1	UI	GWFS has 4 resources. LJAY has 3 resources. william has 13 resources.
437	GWFS	greedy :D
438	william	:D
439	GWFS	spend it wisely then
440	LJAY	:)
441	LJAY	13! :o

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. The turn numbers for (1) are indicated in the left column. Game messages that were added in a later stage to build the situated corpus were assigned sequence identifiers in order to preserve the original 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

2. Go to 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.

also identified with an agent, as shown in the middle column of (1). For chat moves, the agent is the player who typed the chat message (e.g., GWFS for turn 434);[3] visually observable game events and states are either described in Server messages, many of which were visible to all players in the Game window, or reconstructed (by our team) using information from the User Interface (UI). In (1), William plays a Monopoly card, which allows him to steal all instances of a particular resource of his choice that are possessed by the other players. In turn 433.0.4, he steals all of the wheat. Both GWFS and LJAY comment on William's move. There is some ambiguity as to whether LJAY in 440 comments on the theft itself or on GWFS' comments in 437. Immediately after, in 441, LJAY comments on the result of the theft: that William has 13 resources.[4]

2.2 Annotating the corpora

The chats in our *Settlers* corpora 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.

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.[5] 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.[6] 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 SDRT, the annotation process should, when a discourse is coherent, yield a weakly-connected, directed acyclic graph of the following sort.[7]

3. Small capitals indicate user names that have been abbreviated to save space or preserve anonymity.
4. 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 view 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.
5. See the ANNODIS manual for details on the latter (Muller et al., 2012; Afantenos et al., 2012a).
6. We note that 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 extended to allow for this, but the possibility is not explicitly discussed.
7. The formulation of Definition 1 differs slightly from that in Hunter et al. (2018) in order to allow multiple arrows between the same pair of discourse units, which is a possibility that arises in our corpora.

Definition 1 *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.*

Some theories, including DLTAG, LDM, and RST, restrict discourse representations to trees,[8] while the PDTB annotation approach does not aim to build full discourse structures for texts and thus remains agnostic as to what form such representations could take.[9]

2.3 Three points about multiparty dialogue

The multiparty nature of our corpora forced us to confront various issues that either do not arise or arise far less frequently in single-authored text. In this subsection, we discuss what we judge to be the three most significant issues: that the multiparty setup required us to countenance non-treelike discourse structures (cf. Afantenos et al., 2015; Perret et al., 2016; Asher et al., 2016), that it led to numerous overlapping conversational threads, and that it introduced the possibility of subjective discourse interpretations. The need to represent the first two phenomena in our annotations was undeniable; here we take a look at these structures in the linguistic corpus, and in Section 3, we extend the discussion to the situated corpus. As for the third point, we opted not to build subjective discourse structures. We explain this decision 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 *truly non-treelike* DUs and *quasi-treelike* DUs. A truly non-treelike DU has two or more incoming arrows from different source DUs, while a quasi-treelike DU has two incoming arrows (with two different labels) from the *same* source DU. Turn 239 in (2) is an example of a truly non-treelike DU, as illustrated in Figure 2. (CLICK FOR GRAPH)

(2)	234	GWFS	anyone got wheat for a sheep?
	235	inca	sorry, not me
	236	CCG	[nope.]$_a$ [you seem to have lots of sheep!]$_b$
	237	GWFS	yup baaa
	238	dmm	i think i'd rather hang on to my wheat i'm afraid
	239	GWFS	kk I'll take my chances then...

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 replies. This exchange yields a diamond-shaped graph, shown in Figure 2 (ignoring 236b and 237). The label QAP is short for "Question-Answer Pair" and types the connections between 234 and each of the three replies; Ack is short for "Acknowledgement" and types the connection between each reply and the acknowledgement in 239.

8. SDRT posits several constraints on discourse graphs, including acyclicity and weak connectedness (Perret et al., 2016), but they need not be trees.
9. In addition to annotation for discourse structure, the chats in our *Settlers* corpora were also annotated for dialogue acts, which proved important predictors for automatically learning discourse structure (Afantenos et al., 2015; Perret et al., 2016).

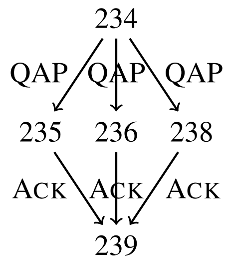

Figure 2: Discourse structure for Example (2). This structure is truly non-treelike, as node 239 has three incoming edges, each from a different source.

Truly non-treelike structures are fairly frequent in our corpora; out of a total of 12,588 DUs in the linguistic corpus, 928 or about 7% were truly non-treelike.

Another result of multiparty dialogue exhibited by our corpus is the presence of overlapping threads, also noted by Afantenos et al. (2015) and Asher et al. (2016). Groups of two or more dialogue participants can create subdialogues and threads of conversation that divide, merge, or get dropped as the dialogue proceeds. Example (3) contains at least 3 threads, which we have represented with different fonts. (CLICK FOR GRAPH)

(3)

165	LJAY	anyone want sheep for clay?
166	GWFS	got none, sorry :(
167	GWFS	**so how do people know about the league?**
168	william	no
170	LJAY	**i did the trials**
174	TK	**i know about it from my gf**
175	GWFS	[**yeah me too,**]$_a$
		[*are you an Informatics student then,* LJAY*?*]$_b$
176	TK	**did not do the trials**
177	william	has anyone got wood for me?
178	GWFS	[**I did them**]$_a$ [**because a friend did**]$_b$
179	GWFS	lol william, you cad
180	GWFS	afraid not :(
181	LJAY	[*no, I'm about to start math.*]$_a$
		[*I just hang around appleton a lot*]$_b$
182	TK	sry no
183	GWFS	my single wood is precious
184	william	what's a cad?

Note that turns can be complex, involving several discourse units (usually clauses) that serve distinct discourse functions, and threads can cross over one another. These crossing patterns yield discourse connections that challenge assumptions about the projectivity of discourse structure (like those assumed in RST; Carlson and Marcu, 2001) and violate salience-driven constraints on discourse attachment as formulated for monologue (Polanyi, 1985; Asher, 1993).[10] To some extent, the presence of overlapping threads is an artefact of the design of our game interface. Players can

10. 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.

type messages simultaneously without realizing what messages other players are composing, and sometimes, a player will enter a message so quickly after another appears that the author of the second message could not have had time to read the content of the immediately preceding message. This gives rise to sequences of chat moves that might be unnatural in face-to-face conversation. Moreover, because the History window records the most recent set of chat moves, the form and content of messages remains accessible for longer than it does in spoken conversation, perhaps making it easier for players to comment on slightly older messages. Still, threads are possible in face-to-face conversation, especially when conversational participants are jointly working on a task while conversing. We will see how structures similar to threads are generated in the situated corpus and study them in Section 4.

The third feature of our multiparty corpus that we had to confront was the possibility of subjective discourse interpretation. When annotating a single-authored text for discourse structure, it makes sense to construct a single discourse graph. Moving to dialogue complicates this picture because interlocutors can interpret the discourse differently, leading to differing, and even contradictory, representations of the conversation (Asher and Fernando, 1997). Amplifying on Lascarides and Asher (2009); Venant et al. (2014); Venant and Asher (2015), ideally one would create an individual discourse graph for each participant (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. It also would have made a process with undeniably subjective elements even more subjective. 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 has immediate consequences for the annotation of certain relations, most notably the relation of Correction. Typically, Corrections signal disagreements and thus convey different points of view; the speaker of β commits to α being false, but the speaker of α might not be willing to accept the correction and commit to α's being false. With only one annotated structure for the dialogue agents, we cannot represent this difference of opinion easily and have therefore opted to confine Correction almost entirely to self-corrections or corrections by the Server of a player's attempted but forbidden action, which the player perforce has to accept, since the Server will not allow him to perform the desired action.[11]

3. Moving to situated dialogue

To build the situated corpus, we posited that game events and states contribute units analogous to EDUs, which can be incorporated into the representations of discourse structures for our chats. Because these events and states are generally represented nonlinguistically, and are in any case discourse independent is the sense described in the introduction, we refer to the units that they contribute as *elementary event units* or EEUs.[12] In this section, we delve into the motivations for,

11. Our decision to adopt the perspective of a third party observer made us unable to show that players had different information about game events. When a player steals a resource from another player, for example, the player who has been robbed knows what type of resource she lost as does the one who gets the resource, but the other players are only told by the Server that *some* resource was taken.

12. See Hunter et al. (2018) for a discussion of the nonlinguistic status of EEUs in our situated corpus and of the difference between discourse independent and nonlinguistic events. Briefly, the class of nonlinguistic events and discourse independent events overlap, but neither is properly included in the other. Co-verbal gestures, for instance, arguably

and complications associated with, this assumption. We begin in Section 3.1 by arguing for the need to introduce EEUs and to allow them to contribute to discourse relations. Section 3.2 then argues that this need goes beyond allowing EEUs to interact with chat moves—our data motivates a rich conceptualization of the game state itself, including semantic relations between game events and non-trivial groupings of game events. Finally, Section 3.3 shows how adding EEUs to a discourse representation can influence judgments about even linguistic-only moves; that is, in many cases, moving to the situated corpus led annotators to revise judgments about a relation that had been inferred between two chat moves in the linguistic corpus.

To our knowledge, allowing discourse independent events to contribute content to a discourse in the way that we propose is novel and treated theoretically only in Hunter et al. (2018). In models of deixis and reference, nonlinguistic entities are understood as crucial for the *interpretation* of discourse, but not for the *construction* of it. This section goes beyond Hunter et al. (2018) by examining in more detail how adding EEUs influenced the construction of discourse representations in our situated corpus and by providing empirical support for claims left at a theoretical level in Hunter et al. (2018).

3.1 The incompleteness of the linguistic corpus

Annotators for the linguistic corpus were assigned the task of relating each noninitial chat move to another part of the chat in order to capture the discursive function of the former. In many cases, however, annotators were unable to determine the attachment point for a chat move, and so its role in the coherence of the overall chat was left unaccounted for. Such moves were represented as "orphans" in the discourse graphs for our chats, meaning that they had no incoming links. This was the case for "oucho" in (4), which has no incoming arrow in the linguistic corpus. (CLICK FOR GRAPH)

	278	Server	william rolled a 6 and a 1.
	278.1	Server	william will move the robber.
(4)	278.2	Server	william stole a resource from GWFS
	279	GWFS	oucho
	280	william	you can have **it** back for some ore

The finalized linguistic corpus contains 1277 orphans.[13] In addition to having no incoming links, 60 of these have no outgoing links, which means that the graphs for the dialogues in which they figure are not even weakly connected.

count as nonlinguistic but discourse *dependent* to the extent that they are produced as a part of a larger communicative act meant to contribute to the discourse at hand. On the other hand, speech acts from one conversation that are exploited by eavesdroppers engaged in a separate conversation arguably count as linguistic but discourse independent. In this paper, we often treat EEUs as the contribution of discourse independent events in order to avoid the discussion of the nonlinguistic status of EEUs; however, for reasons explained in Hunter et al. (2018), we think that the majority of game events do deserve to be counted as nonlinguistic, and we describe them as such when doing so seems innocuous and more intuitive.

13. 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.

In many cases, orphaned chat moves take the form of elliptical or otherwise fragmentary utterances. Such anaphoric chat moves require that we incorporate EEUs into our discourse representations not only to capture the discursive contribution of these chat moves, but also to determine their content. In turn 280 from Example (4), William's use of the third-person pronoun *it* refers back to the resource that he stole from GWFS in turn 278.2. And in turn 57 of (5), GWFS's utterance of *fast mover!* is clearly a response to William's building success. Thus *fast mover!* conveys the full content, *you're a fast mover!*, with *you* referring to William. (CLICK FOR GRAPH)

(5)

51	Server	william made an offer to trade 1 clay for 1 sheep.
52	Server	william traded 1 clay for 1 sheep from GWFS.
53	GWFS	thx
54	Server	william built a road.
55	Server	william built a settlement.
57	GWFS	fast mover!

Semantic dependencies between chat moves and game moves go in the opposite direction as well: sometimes a chat move serves to bring about a game event, so that some structures for our situated corpus contain edges that extend from a chat move to a game move. One common scenario is for a verbal trade negotiation to result in a trade offer made via the game interface. In (6), the offer in 250 arguably results from Ash's acceptance in 249 of the verbal offer negotiated in the preceding moves. (CLICK FOR GRAPH)

(6)

244	mmatrtajova	anyone will trad wheat or sheep?
245	Ash	yes for wood
246	J	nopes
247	mmatrtajova	okay wood for wheat?
248	mmatrtajova	and sheep for ore?
249	Ash	ok
250	Server	mmatrtajova made an offer to trade 1 ore, 1 wood for 1 sheep, 1 wheat.

Allowing EEUs to contribute to representations of discourse structure greatly increases the number of semantic links to and from chat moves (EDUs); the linguistic corpus contains a total of 12,271 links, while the situated corpus contains 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). This shows that the annotations from the linguistic corpus alone fell well short of capturing the semantic content of the discourses in our games. Table 1 shows the number of links from the situated corpus that hold between an EDU and a CDU containing at least one EEU (labelled cdu_e), and vice versa (i/ii); the number of links that hold between a CDU and another CDU, where the union of the two CDUs contains at least one EDU and one EEU (iii); the number of links from an EDU to an EEU, and vice versa (iv/v); and the number of links from an EEU to a CDU containing at least one EDU (labelled cdu_d), and vice versa (vi/vii).

Table 1: Table of argument types for links relating chat moves (an EDU or CDU containing an EDU) with game moves (an EEU or CDU containing at least one EEU)

	Situated	
(i)	edu - cdu$_e$	381
(ii)	cdu$_e$ - edu	676
(iii)	cdu - cdu	528
(iv)	edu - eeu	462
(v)	eeu - edu	987
(vi)	eeu - cdu$_d$	178
(vii)	cdu$_d$ - eeu	379

In the end, our choice to treat game events as contributing EEUs that can be incorporated into discourse graphs was driven by the incompleteness of the linguistic corpus as well as the fact that game events can support subsentential elliptical or anaphoric constructions, provide antecedents for anaphoric pronouns, and intuitively serve as terms of discourse relations. And indeed, once we introduced EEUs into our annotations, all 1501 orphans from the linguistic corpus were assigned incoming links, and all of our graphs became weakly connected. The fact that an EEU can enter into a discourse relation or provide an antecedent for an anaphor, however, entails that an EEU, like an EDU, has something like a propositional type and comes with a descriptive or conceptual content that characterizes it. But what determines this conceptual content when an EEU is contributed by a nonlinguistic event?[14] And if EEUs contribute proposition-like contents, can they enter into semantic relations with each other? If so, what is the nature of these interactions? In the next subsection, we take a look at some of the difficulties that we encountered in trying to model the content and structure of the nonlinguistic context in our situated corpus.

3.2 The conceptualization and structure of game events

Both theoretical work on discourse structure and machine learning methods designed to advance automatic parsing of discourse structure have made strides in solving the segmentation, attachment, and labelling problems for text and even dialogue. The solutions adopted for linguistic discourse, however, do not translate immediately to the nonlinguistic context. A linguistic clause provides a rough guideline for the segmentation of text or dialogue, and more or less determines the content of a discourse segment. There is no natural analogue to a clause in the nonlinguistic context, and nonlinguistic information is often presented in a steady stream, which means it is up to interpreters to determine how to individuate events and what propositional contents to associate with them. In other words, there is not only a more challenging segmentation problem for nonlinguistic information, but also a *conceptualization* problem for discourse segments. Moreover, while in a coherent discourse, we can generally assume that every clause uttered has a semantically important role to play, the nonlinguistic context may contain many perceptible events—perhaps even the majority of them—that are semantically irrelevant. Interpreters therefore have the extra work of deciding which part of the nonlinguistic context to even consider for discourse purposes. Finally, features such as syntactic

14. EEUs will generally be contributed by nonlinguistic events, though see footnote 12.

structure, the category and form of a predicate, or the type of discourse marker chosen can help to guide attachment and labelling of discourse segments. These features are not to be found in the nonlinguistic context. The world might determine a natural temporal ordering of events, but it is unclear what features guide inferences to any richer structure over the nonlinguistic context.

One of the great advantages of our situated corpus is that it allows us to largely bypass the question of how to segment the flow of nonlinguistic information as well as the question of what propositional contents to assign to event units. As previously explained, the game logs for our corpus contain descriptions of all game events, including dice rolls, card plays, resource distributions, and building events. Furthermore, these events are temporally ordered and aligned with all of the chat moves. This facilitates the study of how the addition of EEUs affects current models of discourse structure.

That said, we note that the setup of our corpus does not entirely solve the segmentation and conceptualization problems for EEUs. Information such as who is sitting where, when a turn is ended, and to whom a Trade Panel offer is made, is given visually to players, and therefore encoded in some form in the game log, but it is not described in Server messages of the sort seen by players. This information had to be extracted from the User Interface (UI) and assigned a content that could be subsequently annotated. Sometimes, UI information about game states was too difficult to extract in an annotatable form; we opted to ignore these states and to relate any chat utterances relevant to such a state to the most recent game event that brought it about. Another choice that we had to make when determining EEU content was whether to consider types of events whose tokens were described in the game log but to which references were either nonexistent in our corpus, or nearly so. One example is when players would change their avatars, which they sometimes did numerous times in a row. Because our decision to include UI or Server messages was based on event-types—either all events of a certain type were included or none were—including such events ran the risk of significantly increasing the number of EEUs in the corpus without shedding light on semantically significant interactions. We opted to ignore these events and focus on the types of events that players were likely to incorporate in their conversations.

Despite the complications described above, segmentation and conceptualization were fairly straightforward for the game events. The attachement and labelling problems, however, were a different matter. As the events are temporally ordered, a natural hypothesis would be that the structure of game events forms a strict linear order in which each pair of successive EEUs, ϵ_n and ϵ_{n+1}, figures in an instance of the discourse relation Sequence, such that Sequence($\epsilon_n, \epsilon_{n+1}$).[15] The idea would then be that we could simply connect chat moves to EEUs in the game structure without influencing this ordering. What we found, however, was that the game events have a rich structure of their own and also that this structure could be influenced by chat moves. Sorting out this structure was an imposing task given the sheer size of the situated corpus: our choices about which nonlinguistic event types to consider yielded 31,811 EEUs in the situated corpus, in contrast to 12,588 EDUs in the linguistic corpus.

15. Sequence is semantically equivalent to the relation called 'Narration' in the linguistic corpus (Asher and Lascarides, 2003). We opted to use the name 'Sequence' in the situated corpus because we found it less confusing, given that EEUs do not form a narrative in the traditional sense.

3.2.1 BEYOND SEQUENCE

Causal relations between EEUs were often semantically relevant, meaning that Sequence alone did not adequately capture the links between all EEUs.

(7)

204	Server	J rolled a 2 and a 3.
205	Server	mmatrtajova gets 1 sheep. Ash gets 1 sheep.
206	mmatrtajova	nicee
207	J	my dice rolls SUCK

In (7), (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; 207 is intuitively related to a complex unit ϵ, whose content is Result(204,205).

Of course, Result(204,205) entails Sequence(204,205): a cause must precede its effect, and when both cause and effect are singular events, this enforces a sequential order between them (Asher and Lascarides, 2003). Ultimately, then, all pairs of events related by Result are also related sequentially. Other events described in our corpus, however, cannot be even sequentially ordered. As indicated in Table 2, EEUs in our corpus are frequently related by Continuation, which has the semantics of Boolean '&' and is used to relate events or states without a commitment to temporal order. This relation was used when, for example, a dice roll led to multiple resource distributions. Turn 205 in (7) contains two segments—*mmatrtajova gets 1 sheep* and *Ash gets 1 sheep*—that are related via Continuation.

Table 2: 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

As Table 2 shows, Sequence, Result, and Continuation make up the majority of the relations that structure the game state. The sum for the discourse relation Sequence does not include edges labelled with Result; when two arguments are related by Result, the sequential relation between them is not explicitly marked, but left as an entailment.

Table 2 indicates that EEUs are also frequently arguments to Question-Answer Pair (QAP) and Elaboration relations. (8) illustrates both cases. (CLICK FOR GRAPH)

(8)	51	Server	william made an offer to trade 1 clay for 1 sheep.
	51.1	UI	...from GWFS
	52		william traded 1 clay for 1 sheep from GWFS.
	53	GWFS	thx
	54	Server	william built a road.

Because the addressee of a trade offer (e.g., GWFS in 51.1) was not identified in the Server messages (e.g., 51) for our game, we 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 (8), 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 the two units 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, 244 in (6)), 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".

The 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. (Note that in a few games, players ran out of time before anyone won, so no such messages were emitted.) 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. Other relations, such as Contrast, Conditional, Parallel, Clarification-Question, and Question-Elaboration, do not appear in our corpus between pairs of EEUs (or CDUs made from only EEUs), though as indicated in Table 10 in the Appendix, some of them can relate EEUs to EDUs. Given that the EEUs in our situated corpus come from game events—rather than communicative nonlinguistic events such as gestures—this is to be expected; it is not clear how a real-world event that is not a speech or communicative act could convey a conditional dependency, a contrast or parallel, or a dependent question. In a multimodal corpus containing communicative nonlinguistic events, the possibilities would be different: a puzzled facial expression can convey a clarification question at least indirectly, and with conventional gestures, a wide range of relations may be available.

3.2.2 GROUPING GAME EVENTS

The discussions of (7) and (8) highlight a second way in which the structure over EEUs in our situated corpus departs from a sequential linear ordering, aside from involving relations other than Sequence. In particular, it was often natural or even necessary to group EEUs together as CDUs to capture the full content of a game. One common case is when a single dice roll leads to resource distributions for multiple players. In these examples, we segmented each resource distribution (as explained above for turn 205 in (7)), related the segments with instances of Continuation, grouped the segments together in a CDU, and then treated the CDU as a collective result of the dice roll. Thus for turns 204 and 205 in (7), given below as (9), we get Result(204,[205a, 205b]), where [205a, 205b] is a CDU composed of the two segments in 205. (CLICK FOR GRAPH)

(9)	204	Server	J rolled a 2 and a 3.
	205	Server	mmatrtajova gets 1 sheep. Ash gets 1 sheep.

Another situation that calls for CDUs of game events 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 ends 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 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, but, depending on the configuration of the game board at that time, possibly other events as well. (4), repeated below, provides an example. (CLICK FOR GRAPH)

	278	Server	william rolled a 6 and a 1.
	278.1	Server	william will move the robber.
(4)	278.2	Server	william stole a resource from GWFS
	279	GWFS	oucho
	280	william	you can have **it** back for some ore

Because we viewed the sequences of events triggered by the roll of a 7 as sub-events of larger "robber events", we decided to systematically represent them as CDUs, with each sub-event causing the next sub-event. This decision yields the following relation instances for the robber events in (4): Result(278,[278.1,278.2]) and Result(278.1,278.2).

The situated corpus has a large number of CDUs: 5777 EEU-only CDUs compared to 1450 EDU-only CDUs in the linguistic corpus. Many of the EEU-only CDUs contain numerous units (maximum 39, mean 2.77), especially in comparison to the linguistic corpus, in which CDUs are small and relatively less frequent (max 6, mean 2.14).[16] Table 9 in the Appendix provides more detail on the CDUs in both the linguistic and situated corpora.

We also note in passing that in addition to the 5777 EEU-only CDUs, the situated corpus contains CDUs composed of at least some EDUs, which are formed when a group of chat moves together gives rise to a game event like a nonlinguistic trade (see Table 9). Because the chat moves together result in a nonlinguistic event, we group them together in a CDU to serve as an argument of a Result relation. While this subsection is dedicated to the structural relations between EEUs (and CDUs containing them), we mention these EDU-containing CDUs because they expose 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. In such cases, 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.

3.2.3 NOVEL NON-TREELIKE STRUCTURES

The situated corpus also contains a new kind of non-treelike structure that again prevents a linear representation of events. (10) provides an illustration. (CLICK FOR GRAPH)

16. It would be interesting in future work to see how the size and number of CDUs varies between monomodal and multimodal discourse and across genres. Another point of comparison would be between single authored text and multilogue. Afantenos et al. (2012a) provides figures for a corpus of single authored texts annotated with SDRT's discourse structures, in which CDUs make up over 30% of the discourse units and are relatively large with richly recursive structure. Table 9 shows that in the linguistic STAC corpus, by contrast, CDUs are only around 11% of the discourse units, and they contain on average a bit more than 2 EDUs. The corpora, however, concern different genres, so these differences require deeper investigation.

(10)

256	nelsen	Will anyone give me a sheep for wood or ore?
257	Kersti	I'll give you either ore or clay for wood
258	nelsen	I'll take the ore
260	Tyrant Lord	ive got sheeop
261	Kersti	cool
262	Server	nelsen made an offer to trade 1 wood for 1 sheep.
262.0.1	UI	... from Tyrant Lord
263	Server	nelsen traded 1 wood for 1 sheep from Tyrant Lord
264	Server	nelsen made an offer to trade 1 wood for 1 ore.
264.0.1	UI	... from Kersti

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, but it also has an incoming arrow from 261 labelled as a Result. These relations yield the structure in Figure 3.[17]

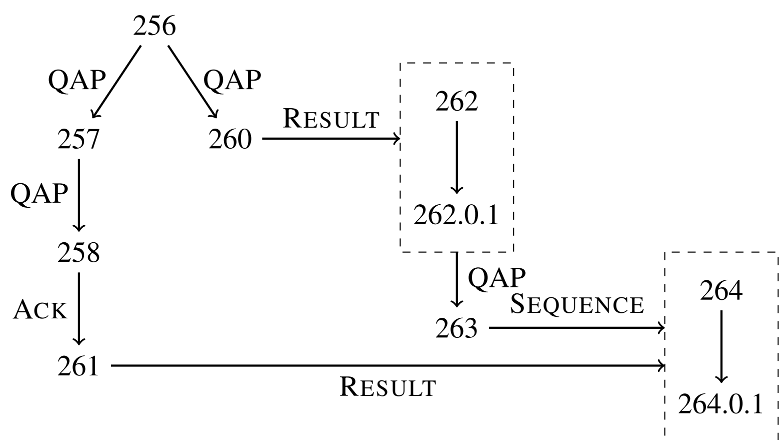

Figure 3: Discourse structure for (10) with a truly non-treelike CDU [264,264.0.1]

The situated corpus adds about 300 more non-treelike structures to the total for the linguistic corpus discussed in section 2.3.

3.2.4 INTERRUPTIONS

The last type of complication that we will discuss for representing the structure of game events arises from events that "interrupt" the expected structure of a game. There are many events that occur in the game environment that are intuitively not a part of the main structure of the game. Every once in a while, for example, someone who is not a player in the current game will accidentally log in and his or her presence will be signalled in the Game window; eventually the person in question

17. In Figure 3, we follow the convention from SDRT of representing subordinating relations with (roughly) vertical arrows and coordinating relations with (roughly) horizontal arrows to improve readability. In terms of content, this graph is the same as the graph linked in the text above. (Result relations are represented in green, Elaboration in purple, QAP and Acknowledgments in turquoise, and Sequence in dark blue.)

will realize the mistake and exit the game, and the exit will also be recorded in the Game window. Sometimes other players comment on such interruptions, and sometimes they don't. These events recall real-life situations in which someone walks into the room where a game is being played and then walks out, or even a situation in which a player coughs while a conversation is going on. The other conversationalists might take notice of such an event and comment on it or not. If they ignore it, then it is natural to think that we should also ignore it in a representation of the nonlinguistic context for the discourse, as it is not relevant to discourse interpretation. For our situated corpus, we decided to systematically include all game-joining and game-exiting moves, having judged the vast majority of them to be discourse-relevant. For the few that were entirely irrelevant, we treated them as separate threads that branch off of the nonlinguistic context and do not feed back into it; that is, there are no outgoing edges from these interruptions back to the sequence of "main" events. For our corpus, this was not a costly decision, as such events were rare, but in a real-world nonlinguistic context they would be everywhere present, and choices would have to be made about whether and how to represent them.

Some interruptions, such as that in (11) (CLICK FOR GRAPH), were more complicated to handle:

(11)

15	Server	Ash rolled a 2 and a 5.
15.0.1	Server	Ash will move the robber.
15.0.2	UI	Ash ended their turn.
15.0.3	Server	You can't end your turn yet.
15.0.4	UI	Ash ended their turn.
15.0.5	Server	You can't end your turn yet.
18	Ash	nothing
19	mmatrtajova	well then choose a number that actually is on the map!
20	mmatrtajova	:)
20.0.1	Server	Ash moved the robber, must choose a victim.
20.0.2	Server	Ash stole a resource from J

As explained above, robber events have a regular structure, which the game interface forces players to respect. In 15.0.1, the Server demands that Ash move the robber, but rather than moving the robber, Ash attempts to end his turn in 15.0.2. The Server rejects this move in 15.0.3, and then Ash tries again to end his turn, only to be corrected again by the Server. Ash is forced to proceed through the requisite sequence of subevents that are triggered by his roll of a 7. Unlike the interruption of someone accidentally logging in to an ongoing game and then exiting without comment, such interruptions of robber event sequences are discourse relevant. In turn 18, Ash comments on his lack of success and Mmatrtrajova (incorrectly) diagnoses the problem and offers a suggestion. In other words, there are explicit linguistic turns that attach to these interruptions, so they need to be represented in our discourse graphs.

Annotating these interruptions poses a problem for our annotation framework, however. Venant et al. (2013) postulated that CDUs have an integrity constraint that does not allow an arrow to extend from an element outside of a CDU to an element inside of the CDU; an arrow may not "puncture" a CDU (although arrows that point in the other direction—those that extend from an element inside of a CDU to an element outside of it—are allowed). This constraint reflects the intuition that discourse structures should have a clean and relatively simple recursive structure. Given that discourse units should have a natural expression in terms of logical formulas, punctures make these formulas highly complex; they require quantifying over DUs and into propositional structures or unnatural reduplications of DUs (Venant et al., 2013).

The interruptions in (11) threaten to violate the no-puncture constraint. We assume that something about the sequence of events in 15.0.2-20 prompts Ash to finally continue with the robber sequence by moving the robber in turn 20.0.1, so there should be an arrow from some element of 15.0.2-20 (or perhaps the whole series of events) to 20.0.1. But 20.0.1 is a part of a CDU representing the larger robber event; if we do not include 15.0.2-20 in this CDU, the arrow from this exchange to 20.0.1 will puncture the robber CDU. To preserve the integrity constraint, we decided to include discourse-relevant interruptions like those in (11) in the CDUs for robber events; while the *event type* for robber events does not contain interruptions, a particular realization of a robber event type arguably can, and these interruptions should be noted as such before the robber event can be completed. Including such interruptions in the representation of a particular robber event is akin to including interruptions in one's sleep in a representation of a bad night's sleep.

3.3 Consequences for the interpretation of linguistic moves

In Section 3.1, we argued that situated conversation, like that in the chats for our *Settlers* corpus, sometimes demands that we treat nonlinguistic events as contributing propositional contents that in effect function like speech act contents. In Section 3.2, we showed that discourse interpretation often requires attributing a rich structure to the nonlinguistic or discourse independent context; modelling situated conversation involves modelling the interaction between two independently evolving *structures*. In this section, we turn back to the interpretation of linguistically-expressed contents and argue that the structure and content of the nonlinguistic (or discourse independent) context can influence the way that we conceptualize even the relations between *linguistic* (or discourse internal) moves. We motivate this claim by showing how the addition of game events in the creation of our situated corpus led annotators to revise certain decisions that had been made for the linguistic corpus concerning how chat moves should be attached to one another, how these attachments should be labelled, and how chat moves should be grouped together into CDUs.

One immediate way in which adding EEUs shifted the conception of the linguistic structure is by more clearly delineating dialogues. Once we started to annotate the Server and UI messages, the entire 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 when we talk of a dialogue in the situated corpus, we are really talking about a sequence that begins when a player gets the dice and ends when she ends her turn and passes the dice to the next player. Bargaining sessions are then subordinate to these turn-based structures, as it is the person who holds the dice that is in charge of trades for that turn.

In our situated corpus, it is the game structure that drives discourse development. It can happen that chat moves link to chat or game moves that cross turn boundaries, as when a player comments on a move from the immediately preceding turn, but in general, the structure of the game events in the situated corpus provides the backbone for the chat-game interactions. This means that this backbone has a rather systematic structure. Figure 4, which represents the structure of an EEU only dialogue in the situated corpus, gives an idea of what kind of structure a typical turn has. Jon's getting the dice in 238.0.2 (first yellow node) results in his rolling an 8 in 239 (second yellow node), which in turn results in the set of resource distributions detailed in turn 240 (grouped into a CDU represented by the first red node). The UI then updates information about the resources of all four players (second CDU/red node). Jon then ends his turn in 240.0.1 (bottom yellow node).

238.0.2	Server	It's Jon's turn to roll the dice.
239	Server	Jon rolled a 6 and a 2.
240	Server	Cardlinger gets 2 wood. Joel get 2 ore. Euan gets 1 ore.
240.0.0.1	UI	Cardlinger has 7 resources. Jon has 1 resource. Joel has 2 resources. Euan has 5 resources.
240.0.1	UI	Jon ended their turn.

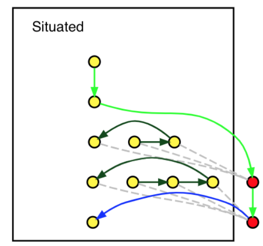

Figure 4: A standard game event only structure. (CLICK FOR GRAPH)

This figure also illustrates the point from the previous subsection, namely that the game has a rich structure of its own.

Once dialogues became more clearly delineated, it was obvious that sometimes, chat moves that annotators thought were related to one another were actually relevant to different dialogues. In (12), 564 was incorrectly attached to 561 as a Comment in the linguistic corpus. Annotators no doubt judged that Euan and Cardlinger were expressing positive attitudes towards the same event, which is natural given Euan's continuation in 565—judging by linguistic cues only, 565 seems to be both a follow-up to Cardlinger's comment in 564 and a continuation of Euan's own comment in 561 (CLICK FOR GRAPH).

(12)

561	Euan	Ooh! Clay :D
562	Server	jon rolled a 4 and a 4.
563	Server	Cardlinger gets 3 wood. Joel gets 2 ore. Euan gets 1 ore.
564	Cardlinger	that was an easy turn for me :D
565	Euan	I like this "getting resources" business.

Another way that the nonlinguistic context frequently influenced judgments about the structure of chats in our corpora is illustrated by (13). (CLICK FOR GRAPH)

(13)

71	T.K.	anyone can offer any wood?
72	william	sry no
73	GWFS	sorry - more 6s and I can oblige then :)
74	LJAY	move the robber and sure :p
75	Server	T.K. traded 3 sheep for 1 wood from a port.
76	Server	T.K. built a road.

In turn 71, T.K. expresses an interest in trading with other players, but in turns 72-74, all of the other players reject his offer. In turn 75, T.K. goes to a port to get the wood that 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. In other words, they judged that the trade was a result of the entire failed negotiation in 71-74. Consequently, while 71-74 were not grouped in a CDU in the linguistic corpus, they had to be grouped together in the situated corpus. 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.

Because our linguistic corpus was completed entirely before annotation on the situated corpus began, we can quantify the ways in which game events influenced judgments about the discursive function and structure of chat moves. Table 3 shows how many relations from the linguistic corpus persisted in the situated corpus. While these statistics are specific to our corpus, they nevertheless serve to underscore the extent to which the nonlinguistic context can influence judgements about the linguistic context.

Table 3: Table of persisting relations

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

As indicated in the table, around 20% of the relations in the linguistic corpus changed or disappeared in the situated annotation. In addition, relations were added between chat moves in the situated corpus so that the persisting linguistic annotations made up 72% of the relations between chat moves in the situated corpus. The situated annotations thus constitute a non-negligeable revision of the annotations in the linguistic corpus. Note that occasionally we kept the same arguments for a discourse relation but changed its type, which means that the nonlinguistic context sometimes affected our view of what discourse relation held between two linguistically given arguments.

4. Preservation of linguistic structure

In Section 3.3, we counted the number of relation instances from the linguistic corpus that were preserved in the situated corpus, and this provided an indication of how much the linguistic message is influenced by the game events in our chats. To really understand the impact of the game events, however, we need to look at how discourse *structure* is preserved, not only relation instances. Figure 5 illustrates the difference. The graph on the left is from the linguistic corpus and the graph on the right, from the situated corpus. As we can see, 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 corpus is the Continuation, which is represented in dark green. This small change, however, correlates with a significant increase in information about how the EDUs from the left graph figure in the larger game. We see that the top node from the left graph actually relates to a CDU of game events or states (represented by the second red node) in the incoming context, and not directly to the third node of the left graph (to which it was previously related by Continuation). 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. The difference in these two graphs shows why we should consider not only how many relations are preserved in the move to the situated corpus, but also how the larger discourse structures change in the process.

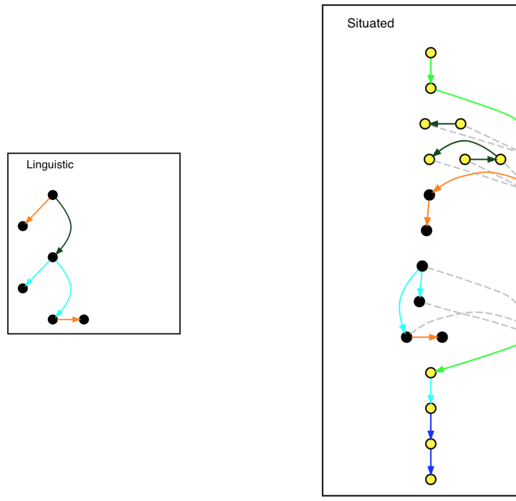

Figure 5: All but one of the relations in the linguistic structure (left) are preserved in the situated structure (right); however the linguistic structure is broken into two pieces once we move to the situated structure, and we see that the two sub-structures figure in significantly different parts of the overall game structure. (CLICK FOR GRAPH)

Since our discourse graphs are first-order structures with one designated object (a pointed model in other words), we can check for elementary embeddings of dialogue structures from the linguistic corpus into the structures of the corresponding dialogues in the situated corpus. An elementary embedding preserves all relations, functions, and designated objects in $\mathfrak{A}$ (Chang and Keisler, 1973). When such an embedding $\mathfrak{f}$ exists, $\mathfrak{f}(\mathfrak{A})$ is an *elementary substructure* of $\mathfrak{B}$.[18] In fact, in our situation, there is a canonical embedding $\mathfrak{f}$ that is the identity function on discourse units in the linguistic structure; so to say that $\mathfrak{f}(\mathfrak{A})$ is an *elementary substructure* of $\mathfrak{B}$ is just to say that $\mathfrak{A}$ is an *elementary substructure* of $\mathfrak{B}$.

To check for the relevant elementary embeddings from our linguistic dialogues into situated dialogues, however, we ignore the element *Last* from Definition 1 so that a discourse graph G is a tuple (V, E_1, E_2, ℓ). *Last* is a designated object, as in a pointed model, that is used to define constraints on how a discourse can dynamically evolve, but in checking for elementary embeddings, we want to restrict our attention to properties of static structures.

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 *linguistic annotations* occur in the situated corpus, $\mathcal{L}$ will not survive as an elementary substructure of $\mathcal{S}$. Substructure preservation is a much stronger condition

18. Here is the technical definition for our discourse structures. An elementary embedding is a one-to-one function $\mathfrak{f}$ 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 $\mathfrak{a}$ in the signature of $\mathfrak{A}$, $\mathfrak{A} \models R(b,c)$ iff $\mathfrak{B} \models R(\mathfrak{f}(b), \mathfrak{f}(c))$; $\mathfrak{A} \models g(b) = c$ iff $\mathfrak{B} \models g(\mathfrak{f}(b)) = \mathfrak{f}(c)$, and $\mathfrak{f}(\mathfrak{a}) = \mathfrak{b}$, where $\mathfrak{b}$ is the designated object of $\mathfrak{B}$.

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 linguistic corpus, or a little over 43%, are preserved as elementary substructures of the corresponding $\mathcal{S}$ structures. While preservation of relation instances shows that our individual decisions about relation instances is relatively robust even when limited to information present only in the linguistic turns, the substructure information shows that the nonlinguistic context in general affects the discourse structure of the linguistic turns.

We can delve even deeper into the effects of the game events and structure by exploring different types of discourse structures and how these different types are preserved. In what follows, we first introduce two types of discourse structures—*asymmetric* and *interleaved* structures—that are manifested in our corpora and described in Hunter et al. (2018). We then go beyond Hunter et al. (2018) by adding reflections on the relation between these structures and discourse goals in Section 4.2, and by measuring the preservation of asymmetric structures in Section 4.3.

4.1 Asymmetric and interleaved structures

We saw in Section 3 that the game moves provide a sort of backbone that guides the development and interpretation of the chat moves in our corpora. In these interactions, the game events do more than fix the reference or interpretation of chat moves—the bulk of the structural development of the exchanges actually takes place at the nonlinguistic level and the chat moves can be seen as parasitic on this structure.

The central role of game development often leads to an asymmetric semantic dependence of chat moves on the game moves in the situated corpus. Consider the example in Figure 6. 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). The 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 other words, the chat exchange depends on the game moves for its interpretation, but the game moves are independent of the chat.[19]

The asymmetric nature of the interaction between Tomm and Dave's exchange on the one hand and the game moves on the other is reflected in the graph in Figure 5. The black nodes in the graph represent EDUs; the yellow nodes, EEUs. The two red nodes represent CDUs, and the dashed edges extending from them indicate their members. The top CDU (red node) contains the four EEUs (yellow nodes) that result from Dave's rolling a 7 (so 94.0.1 and 95.0.1-95.0.3). The bottom CDU contains the EDUs (black nodes) 101 and 103-105 that constitute a failed bargaining attempt. Note that we can trace a continuous series of edges from the top yellow node of the graph to the bottom yellow node that runs through the remaining yellow nodes plus the four black nodes contained in the bottom CDU. The resulting subgraph, shown in blue, is the backbone or what we will call the *core* of the structure in Figure 5. Clearly, ignoring the nodes that are not a part of the core has no

19. Note that we count the chat moves involved in the trading session, i.e., 101 and 103–105, as game moves. See the discussion of interleaved structures below.

93.0.2	Server	It's Dave's turn to roll the dice.
94	Server	Dave rolled a 3 and a 4.
94.0.1	Server	rennoc1 needs to discard.
95	Dave	:D
95.0.1	Server	rennoc1 discarded 4 resources.
95.0.2	Server	Dave will move the robber.
95.0.3	Server	Dave stole a resource from Tomm
96	Tomm	[Oh. . .] [well that's unkind]
97	Server	Dave built a road.
98	Dave	sorry
99	Tomm	[It's okay,] [name of the game and all that.]
100	Tomm	Do you get a random one of my resources then?
101	Dave	anyone need wheat?
102	Dave	I did, yeah
103	Tomm	no thanks
104	Dave	rennoc?
105	rennoc1	i'm ok for now
105.0.0.1	Server	Dave bought a development card.
105.0.1	UI	Dave ended their turn.

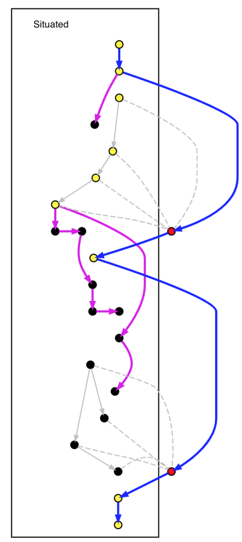

Figure 6: A dialogue and its asymmetric structure. The bold blue line traces the core, which runs through the EEU nodes (yellow nodes) contributed by Server and UI messages, some of which are grouped into a CDU (red node) for the robber sequence in 94.0.1 and 95.0.1–95.0.3, and a CDU (second red node) that groups the chat moves 101 and 103-105 involved in a trade negotiation. The periphery, shown in magenta, consists of all nodes which do not belong to the core and their incoming edges. Note that turns 96 and 99 were each divided into two EDUs, marked by square brackets in the dialogue and connected with horizontal arrows in the graph. (CLICK FOR GRAPH)

affect on the connectedness or the interpretation of the core, as the outlying nodes are all connected via outgoing edges.

The set of outlying nodes determine what we call the *periphery* of the graph in Figure 5. The periphery contains two structures, shown in magenta. The first is a tree of depth one that links the EDU 95 to the EEU 94 in the core. The second structure, which represents Dave's exchange with Tomm, is a more complex tree hanging off EEU 95.0.3, the last EEU figuring in the top CDU. The line of shorter arcs extending from 95.0.3 signifies a series of commentaries made in the two EDUs (indicated by square brackets) in turn 96, turn 98, and the two EDUs in 99. A longer arc links Tomm's clarification question in 100, followed by the answer in 102, to the core move 95.0.3. While taking the periphery away from the structure would leave the core intact, removing the core would lead to unconnectedness, and we would not represent the contribution of Dave's exchange with Tomm to the overall interaction.

Figure 5 also contains an example of 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. The core of the graph in Figure 5 is an interleaved structure: the bottom CDU is composed entirely of EDUs (101 and 103–105), but has an incoming link from EEU 97 and an outgoing link to EEU 105.0.0.1. In this case, EDUs and EEUs function together to further the game state; EEUs feed into EDUs, but unlike in the asymmetric structures in our corpus, EDUs also feed into EEUs.

Formal definitions for cores, peripheries, interleaved and asymmetric structures from Hunter et al. (2018) 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 11 in the appendix gives statistical details concerning asymmetric and interleaved structures. Around 38% (978 out of 2595) are asymmetric; on average, these structures consist of around two peripheral structures attached to the core, with the maximum being 23 peripheral structures for one dialogue. The remaining dialogues have what we call a *core-only* structure. Thus 1617 or over 62% have core-only structures—a significant increase over the 296 core-only $\mathcal{L}$ structures, which account for 26% of all the dialogues in the linguistic corpus. Around 35% of the $\mathcal{S}$ structures for dialogues in the situated corpus (902 out of 2595) are interleaved. Because almost all of the EEUs in our corpus are directly related to game development,[20] the interleaved structures in our corpus are also core structures and can be attributed to the presence of chat moves that directly impact the game, such as those involved in trade negotiations.

4.2 Asymmetric structures and discourse goals

Asymmetric structures arise anytime a discourse splits into two threads that branch off of the same node.[21] This is a commonplace occurrence at dinner parties, for instance, when a group is chatting

20. There are very limited exceptions, mostly of the sort discussed in Section 3.3, in which, say, a player accidentally joins a game and then leaves. We treated these sequences as separate threads that do not feed back into the core, so as peripheral structures.

21. 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

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 the conversation as defining the core of the discourse structure, yielding two possible asymmetric structures; the notion of a core is in principle a thematic and functional one. Figure 7 illustrates the two possibilities; in each graph, the solid line represents the core, and the dotted line, the periphery.

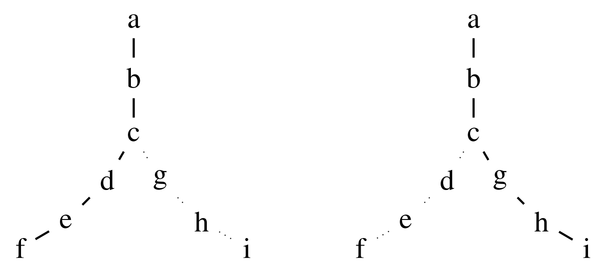

Figure 7: Two ways of distinguishing the core and periphery in a conversation that splits into two threads; the core, represented with a thick line, can be composed of either $\{a, b, c, d, e, f\}$, as in the left graph, or $\{a, b, c, g, h, i\}$, as in the right graph.

The graphs in our situated corpus 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 linguistic moves. Second, the threads that make up the periphery contain typically 2.28 nodes on average, with a range of 50 nodes (see Table 11 for more details), whereas a split in a regular conversation can can lead to an extended discussion.[22] 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 corpus 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;

generally need to be marked explicitly, so we do not think that our simplified definition is problematic for our present purposes.

22. 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.

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 EEUs via relations such as Comment (see Table 4 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).

Table 4: Breakdown of relations that connect asymmetric structures to a core in the situated corpus.

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

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 want to 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 linguistic 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 5 shows that GWFS initiated more peripheral structures than other players. (3), repeated below as (14), provides an illustration of GWFS's behavior: LJAY 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.

(14)

165	LJAY	anyone want sheep for clay?
166	GWFS	got none, sorry :(
167	GWFS	**so how do people know about the league?**
168	william	no
170	LJAY	**i did the trials**
174	TK	**i know about it from my gf**
175	GWFS	[**yeah me too,**]$_a$ [*are you an Informatics student then,* LJAY*?*]$_b$
176	TK	**did not do the trials**
177	william	has anyone got wood for me?
178	GWFS	[**I did them**]$_a$ [**because a friend did**]$_b$
179	GWFS	lol william, you cad
180	GWFS	afraid not :(
181	LJAY	[*no, I'm about to start math.*]$_a$ [*I just hang around appleton a lot*]$_b$
182	TK	sry no
183	GWFS	my single wood is precious
184	william	what's a cad?

In the linguistic and situated corpora, 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).

4.3 Measuring the preservation of structure

If we are to better understand the forces that drive discourse development, which would allow us to improve predictions about the attachment of discourse units and the labelling of these attachments, it is crucial that we examine not only instances of discourse relations and features of individual discourse units, but also the overall shapes that a discourse can take. With this in mind, we now turn to a quantitative comparison of the discourse structures in our two corpora to see how the game information in the situated corpus influenced the overall shape of the structures in the linguistic corpus. In particular, given our distinction between cores and peripheries, we can now ask how different *types* of substructures were preserved when moving from the linguistic corpus to the situated corpus. Is, for instance, the core of a preserved $\mathcal{L}$ structure a substructure of the core of the corresponding $\mathcal{S}$ structure?

For the purposes of this discussion, we restrict our study to discourse structures of either core-type or periphery-type. We say that an $\mathcal{L}$ substructure $\mathfrak{A}$ of type τ is τ *preserved* under the canonical $\mathcal{S}$ embedding $\mathfrak{f}$ into an $\mathcal{S}$ structure $\mathfrak{B}$ just in case $\mathfrak{A}$ is a substructure of a structure of type τ in $\mathfrak{B}$. For example, if we consider the structure $\mathfrak{A}$ for a core of a dialogue d in the linguistic corpus, then we say that $\mathfrak{A}$ is core-preserved under an $\mathcal{S}$ embedding $\mathfrak{f}$ 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

Table 5: Number of peripheral structures started per speaker (cutoff at 50)

emitter	#
gotwood4sheep	351
inca	164
ljaybrad123	116
ztime	102
skinnylinny	98
sabercat	98
william	81
Cardlinger	80
gramos	71
raefbrisbin	67
dmm	62
tomas.kostan	61
zorburt	57
somdechn	51
Shawnus	51
nelsen	51
Tomm	51

periphery-type structures, such that all substructures of $\mathfrak{A}$ maintain their type under the embedding, we call this *perfect type preservation*. Given our characterization, core-type, periphery-type and perfect-type preservation are more restrictive notions than elementary preservation, as defined in the introduction to this section. Figures 8-10 show examples of core-type preservation, perfect type preservation (with both core-type and periphery-type structures), and elementary preservation, respectively.

When a core structure is preserved by an $\mathcal{S}$ embedding, this indicates that the linguistic information alone was sufficient to determine the discourse function of the individual moves and also how these moves contributed to overall game play. Linguistic core structures that remain core structures in the situated corpus 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 was sufficient for determining the discursive functions of the individual moves and that these moves were not integral to the main point. Type preservation implies that whether a move belongs to the core or periphery of a conversation can be signalled by local linguistic means.

In general, perfect type preservation is rare in our corpora, meaning that information about the game state was often crucial for determining either the discourse function of individual chat moves or the type of role that they played in furthering game-play and achieving the goal of winning the game. Tables 12 and 13 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, there was an $\mathcal{S}$ embedding that preserved 70 cores, or a total of 23%, and hence were perfect type preserving. Out of 832 asymmetric $\mathcal{L}$ structures (so structures including both core and periphery-type substructures),

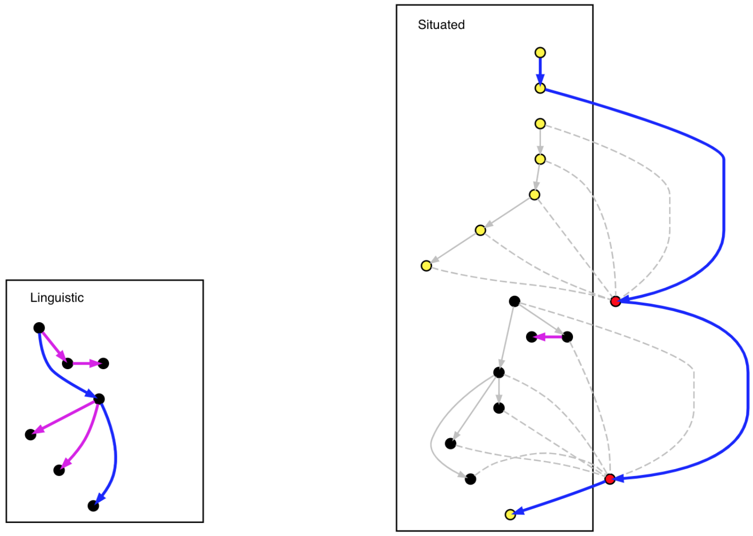

Figure 8: Core preservation. An asymmetric structure from the linguistic corpus (left image) that is preserved in the situated corpus (right image). All but one node in the linguistic structure appear as members of the same CDU (represented by black nodes) in the situated structure. (Our graph program flipped the representation of the structure when moving to the situated corpus in order to make room for EEUs, but the information content remains the same. The three bottom nodes of the CDU in the right image correspond to the three bottom nodes of the graph in the left image.) While the core of the linguistic structure is preserved as a core-type structure in the situated graph, some substructures from the periphery change type, so we have core-type preservation without perfect-type preservation. This change shows how adding information about the game state gives us a better idea of which moves are central to discourse goals and which are not. (CLICK FOR GRAPH)

only 33 (4%) exhibited perfect type preservation; an additional 65 $\mathcal{L}$ dialogue structures had their periphery-type substructures preserved with our canonical $\mathcal{S}$ embedding, though not their cores (7% of all $\mathcal{L}$ structures). On the other hand, 95 core substructures were preserved (11,5% among $\mathcal{L}$ structures with an $\mathcal{S}$ embedding).

We also looked at $\mathcal{L}$ structures, for which there was at least a canonical embedding either from the $\mathcal{L}$ periphery into the periphery of the corresponding $\mathcal{S}$ structure or from the $\mathcal{L}$ core into the $\mathcal{S}$ core. The figures for periphery type preservation in this case rise to 23% and those for core type preservation rise to 13%.

While linguistic information was more reliable in identifying peripheral elements than core elements, the point remains that linguistic information alone was insufficient for identifying core and peripheral substructures in our corpora. In fact, in a few cases, while first-order elementary preservation occurred, the types of the $\mathcal{L}$ substructures were switched when they were embedded into the $\mathcal{S}$ structures. There are some superficial reasons for this: in some dialogues, no game events are linguistically expressed and so the core of such dialogue structures would not be expected to figure in the core of the corresponding $\mathcal{S}$ structure, for example. In our corpora, it is the game events that determine the flow of the game and the game central events carry the overriding purpose of the whole conversation, and so linguistic moves of that structure might very well not reflect that

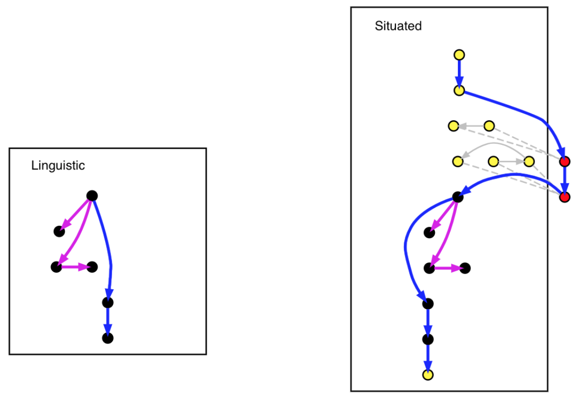

Figure 9: Perfect preservation involving both core and periphery type structures. Both the core (blue edges) and the periphery (magenta edges) in the left image are preserved in the situated corpus, as shown in the right image. (As in Figure 8, our graph program changed the placement of the first blue edge to make room for EEU nodes in the right image, but the content of the representation is the same; CLICK FOR GRAPH)

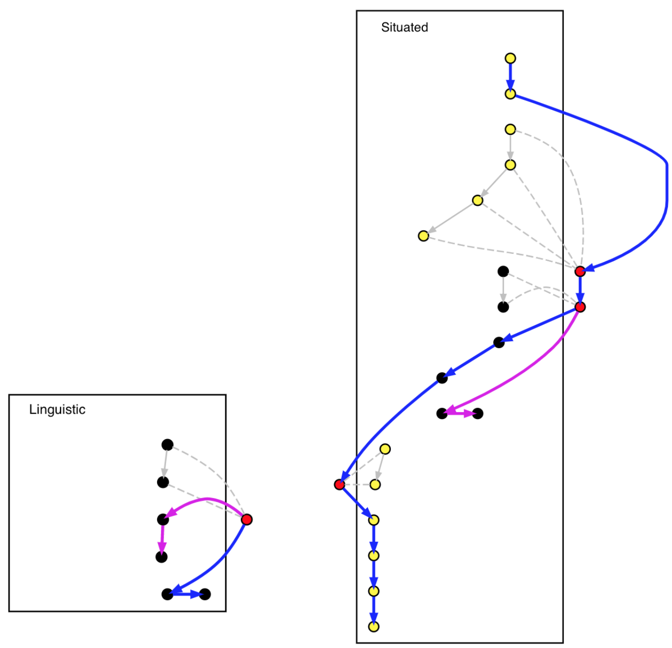

Figure 10: Elementary preservation. There is an $\mathcal{S}$ embedding of the $\mathcal{L}$ structure in the left image into the $\mathcal{S}$ structure on the right, but the types of discourse structures are reversed: what is of periphery-type in the $\mathcal{L}$ structure is of core-type in the $\mathcal{S}$ structure, and vice versa. (CLICK FOR GRAPH)

core. This means on the one hand that the results for type preservation from $\mathcal{S}$ embeddings are somewhat unsurprising for our corpora. Still, they expose the extent to which the nonlinguistic context can influence not only the interpretation of discourse moves, but its very structure, at both a local (i.e., the level of individual relations) and global level.

5. Conclusions

In this paper, we have surveyed and compared two corpora culled from an online version of the game *The Settlers of Catan*: a linguistic corpus, which consists of annotations over a set of chats from the game, and a situated corpus, which consists of the same set of chats together with descriptions of events from the games during which the players were chatting. Our aim was to illuminate and measure a variety of ways in which information from the nonlinguistic context can influence the content and structure of a discourse by contributing full propositional contents that function like independent discourse moves (as opposed to making sub-clausal contributions to otherwise linguistically-specified content, as generally happens when a deictic expression is used to refer to something in the nonlinguistic context).

After motivating the need to treat certain nonlinguistic eventualities as contributing propositional contents, we argued in Section 3 that the interpretation of situated discourse can require attributing a rich, independently evolving structure to the nonlinguistic context and that the interpretation of the nonlinguistic context can significantly impact the interpretation of even linguistic-only exchanges. The results presented in this discussion confirm many of what were largely theoretical claims in Hunter et al. (2018), but we also go beyond this work by showing more data and providing statistics on exactly how adding information about the game state in our situated corpus influenced judgments about the linguistic content and overall discourse structure of our corpora. In Section 4, we pushed the investigation to a new level by arguing that the nonlinguistic context influences not only local features of discourse structure—the set of relation instances, for example—but also global features of discourse structure. It is at the global level of discourse that reveals information about, among other things, the types of threads present in a multiparty discourse, the kind of events that fuel discourse development, and the nature of discourse goals. With this in mind, we measured global differences between discourse structures in the linguistic and situated corpora, focusing particular on the preservation (or lack thereof) of what Hunter et al. (2018) called *asymmetric structures*. We found that adding the game events to our annotations considerably impacted the overall shape of discourse structures from the linguistic corpus.

That nonlinguistic events would influence the interpretation and inferred function of linguistic moves should perhaps be unsurprising; it should be expected that a fuller understanding of a conversational context leads to a better understanding of the conversation itself. Our goal has not been to show *that* they make a difference, but *how* they make a difference. The nature and extent of nonlinguistic influence evidenced in the comparison of our linguistic and situated corpora exposes the fact 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. Moreover, contents contributed by nonlinguistic events can play the same role in the reasoning process about the linguistic content of a discourse as speech acts do; that is, they play an active role in building discourse structure, not only in interpreting it. In addition, we showed that we cannot model the nonlinguistic context as a mere collection of nonlinguistic entities: the structural and semantic relations that hold between nonlinguistic events can influence a linguistic message too. Finally, our corpora provide examples of discourses in which nonlinguistic events do more than add some arguments to discourse relations; they play the central role in discourse development and the achievement of discourse goals.

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 allowed us to build rules to compute relations between EEUs automatically with a high rate of precision and recall and also 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 Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Jan Odijk, and Stelios Piperidis, editors, *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, Istanbul, Turkey, may 2012a. European Language Resources Association (ELRA). ISBN 978-2-9517408-7-7. URL https://hal.archives-ouvertes.fr/hal-00976087/document.

Stergos Afantenos, Nicholas Asher, Farah Benamara, Anaïs Cadilhac, Cedric Dégremont, Pascal Denis, Markus Guhe, Simon Keizer, Alex Lascarides, Oliver Lemon, Philippe Muller, Soumya Paul, Vincent Popescu, Verena Rieser, and Laure Vieu. Modelling strategic conversation: Model, annotation design and corpus. In *The 16th Workshop on the Semantics and Pragmatics of Dialogue (SeineDial)*, pages 145–146, Paris, 2012b. URL https://hal.archives-ouvertes.fr/hal-01138035/document.

Stergos Afantenos, Eric Kow, Nicholas Asher, and Jérémy Perret. Discourse parsing for multi-party chat dialogues. In *Proceedings of 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 Tim Fernando. Nonincrementality and revision in dialogue. In *Proceedings of MunDial Workshop. University of Muenchen*. Citeseer, 1997. URL http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.3.2985&rep=rep1&type=pdf.

Nicholas Asher and Alex Lascarides. *Logics of Conversation*. Cambridge University Press, 2003.

Nicholas Asher, Julie Hunter, Mathieu Morey, Benamara Farah, and Stergos Afantenos. Discourse structure and dialogue acts in multiparty dialogue: the stac corpus. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Thierry Declerck, Sara Goggi, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Helene Mazo, Asuncion Moreno, Jan Odijk, and Stelios Piperidis, editors, *The Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, Paris, France, may 2016. European Language Resources Association (ELRA). ISBN 978-2-9517408-9-1. URL http://www.juliejhunter.com/uploads/3/9/6/1/39617901/lrec2016_final.pdf.

J. Baldridge, N. Asher, and J. Hunter. Annotation for and robust parsing of discourse structure on unrestricted texts. *Zeitschrift fur Sprachwissenschaft*, 26:213–239, 2007.

Lynn Carlson and Daniel Marcu. Discourse tagging reference manual. *ISI Technical Report ISI-TR-545*, 54:56, 2001. URL ftp://128.9.176.20/isi-pubs/tr-545.pdf.

Lynn Carlson, Mary Ellen Okurowski, and Daniel Marcu. *RST discourse treebank LDC2002T07*. Linguistic Data Consortium, University of Pennsylvania, 2002. URL https://catalog.ldc.upenn.edu/LDC2002T07.

C.C. Chang and H. J. Keisler. *Model Theory*. North Holland Publishing, 1973.

K. Forbes, E. Miltsakaki, R. Prasad, A. Sarkar, A. K. Joshi, and B. L. Webber. D-ltag system: Discourse parsing with a lexicalized tree-adjoining grammar. *Journal of Logic, Language and Information*, 12(3):261–279, 2003. URL https://link.springer.com/article/10.1023%2FA%3A1024137719751.

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 *SEMDIAL 2015 goDIAL*, pages 95–103, 2015a. URL https://flov.gu.se/digitalAssets/1537/1537599_semdial2015_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, 2015b. URL http://www.aclweb.org/anthology/W15-0123.

Julie Hunter, Nicholas Asher, and Alex Lascarides. A formal semantics for situated conversation. forthcoming in *Semantics and Pragmatics*, 2018. URL http://semprag.org/article/view/sp.11.10.

David Kaplan. Demonstratives. In J. Almog, J. Perry, and H. Wettstein, editors, *Themes from Kaplan*. Oxford, 1989.

A. Kranstedt, P. Kühnlein, and I. Wachsmith. 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 Henrik Iskov Christensen, Geert-Jan M. Kruijff, and Jeremy L. Wyatt, editors, *Cognitive Systems*, pages 311–364. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. ISBN

978-3-642-11694-0. doi: 10.1007/978-3-642-11694-0_8. URL 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. URL https://academic.oup.com/jos/article-abstract/26/2/109/1695463.

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. URL https://journals.linguisticsociety.org/elanguage/pip/article/view/144.html.

Y. Mathet and A. Widlöcher. La plate–forme GLOZZ : environnement d'annotation et d'exploration de corpus. In *Actes deTALN 2009 (Traitement automatique des langues naturelles)*, Senlis, 2009. ATALA, LIPN. URL http://www.glozz.org/.

Philippe Muller, Marianne Vergez-Couret, Laurent Prévot, Nicholas Asher, Farah Benamara, Myriam Bras, Anne Le Draoulec, and Laure Vieu. Manuel d'annotation en relations de discours du projet ANNODIS. *Carnets de Grammaire*, 21:1–32, 2012.

Jérémy Perret, Stergos Afantenos, Nicholas Asher, and Mathieu Morey. Integer linear programming for discourse parsing. In *Proceedings of 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.

L. Polanyi, C. Culy, M. van den Berg, G. L. Thione, and D. Ahn. A rule based approach to discourse parsing. In Michael Strube and Candy Sidner, editors, *Proceedings of 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.

Livia Polanyi. A theory of discourse structure and discourse coherence. In P. D. Kroeber W. H. Eilfort and K. L. Peterson, editors, *Papers from the General Session at the 21st Regional Meeting of the Chicago Linguistics Society*. Chicago Linguistics Society, 1985.

R. Prasad, N. Dinesh, A. Lee, E. Miltsakaki, L. Robaldo, A. Joshi, and B. Webber. The penn discourse treebank 2.0. In N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odjik, S. Piperidis, and D. Tapias, editors, *Proceedings of LREC'08*, Marrakech, Morocco, may 2008. ELRA. URL http://www.lrec-conf.org/proceedings/lrec2008/pdf/754_paper.pdf. http://www.lrec-conf.org/proceedings/lrec2008/.

Gert Rickheit and Ipke Wachsmuth. *Situated communication*, volume 166. 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. doi: 10.3765/sp.5.6. URL http://semprag.org/article/view/sp.5.6.

Antoine Venant and Nicholas Asher. Dynamics of public commitments in dialogue. In *The 11th International Conference on Computational Semantics*, pages 272–282, London, UK, 2015.

Association for Computational Linguistics. URL http://www.aclweb.org/anthology/W/W15/W15-0131.

Antoine Venant, Nicholas Asher, Philippe Muller, and Pascal Denis Stergos D. Afantenos. Expressivity and comparison of models of discourse structure. In *Proceedings of Sigdial 2013*, Metz, France, 2013. URL http://researchers.lille.inria.fr/pdenis/papers/sigdial13.pdf.

Antoine Venant, Nicholas Asher, and Cedric Degremont. Credibility and its attacks. In Verena Rieser and Philippe Muller, editors, *The 18th Workshop on the Semantics and Pragmatics of Dialogue*, pages 154–162, 2014. URL http://www.macs.hw.ac.uk/InteractionLab/Semdial/semdial14.pdf.

F. Wolf and E. Gibson. Representing discourse coherence: A corpus based study. *Computational Linguistics*, 31(2):249–287, 2005. URL http://www.aclweb.org/anthology/J05-2005.

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 2 *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 $\mathsf{P}(\mathsf{A}, \mathsf{C}) \neq \emptyset$ and A is also a core of G.

Definition 3 *Let G be a discourse graph, and G' a subgraph of G. Let $End(e)$ be the endpoints of an edge e and $End(E) = \{x \colon \exists e \in E . x \in End(e)\}$. 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 $\mathsf{G} - \mathsf{G}'$ may share nodes but form a partition over the set of relation instances or arcs in E_1.

Definition 4 $\mathsf{P}(\mathsf{G}, \mathsf{C})$*, the* **periphery** *of a structure* G *with respect to a core* C*, is such that* $\mathsf{P}(\mathsf{S}, \mathsf{C}) = \mathsf{G} - \mathsf{C}$.

Definition 5 *An* **asymmetric structure** G *is a graph with a core* C *such that* $\mathsf{C} \neq \mathsf{G}$.

6.2 Additional Tables

Table 6: Table of Basic Comparisons

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

Table 7: Table of truly non-treelike structures

	Linguistic	Situated
Truly non-treelike structures	928	1207
# truly non-treelike structures per dialogue		
Mean	0.82	0.47
Max	13	14
Min	0	0

Table 8: 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 9: Table of CDUs

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

Table 10: The distribution of relation labels for links in the linguistic corpus and the situated corpus, with a breakdown for the situated corpus 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

Table 11: Table of Asymmetric and Interleaved structures

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

Table 12: Table for Persisting core-only $\mathcal{L}$ Structures

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

Table 13: Table for Persisting Asymmetric $\mathcal{L}$ Structures

Total persisting asymmetric $\mathcal{L}$ structures	335
Perfect type preservation	95
Core type preservation	
$\mathcal{L}$ 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
Periphery type preservation	
$\mathcal{L}$ 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
Elementary preservation	
$\mathcal{L}$ 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
Elementary Preservation with inverted types	
$\mathcal{L}$ core in $\mathcal{S}$ periph // $\mathcal{L}$ periph in $\mathcal{S}$ core	5