The Exploitation of Spatial Information in Narrative Discourse

Blake Stephen Howald
Georgetown University
bsh25@georgetown.edu

E. Graham Katz
Georgetown University
egk7@georgetown.edu

Abstract

We present the results of several machine learning tasks that exploit explicit spatial language to classify rhetorical relations and the spatial information of narrative events. Three corpora are annotated with figure and ground (granularity) relationships, mereotopologically classified verbs and prepositions, and frames of reference. For rhetorical relations, Naïve Bayesian models achieve 84.90% and 57.87% accuracy in classifying NARRATION and BACKGROUND / ELABORATION relations respectively (16% and 23% above baseline). For the spatial information of narrative events, K* models achieve 55.68% average accuracy (12% above baseline) for all spatial information types. This result is boosted to 71.85% (28% above baseline) when inertial spatial reference and text sequence information are considered. Overall, spatial information is shown to be central to narrative discourse structure and prediction tasks.

1 Introduction

Clauses in discourse are related to one another in a number of semantic and pragmatic ways. Some of the most prominent are temporal relations that hold among the times of events and states described (Partee, 1984; Pustejovsky et al., 2003) and the rhetorical relations that hold between a pair of clauses (Mann and Thompson, 1987; Asher and Lascarides, 2003). For example, (1) illustrates the NARRATION relation which obtains between (1a-b) and between (1b-c).

(1) a. Klose was sitting with his teammates.
b. He walked to the sidelines.
c. Then he entered the game.

Because of the temporal properties of NARRATION (Asher and Lascarides 2003, p. 462), the event described in (1a) is taken to precede that described in (1b) and (1b)'s event to precede (1c)'s. As Asher and Lascarides show, there is a close tie between the rhetorical structure of a discourse and its temporal structure. In (2), for example, the fact that the clauses are related by ELABORATION entails that the temporal relation between (2a) and (2b) is inclusion.

(2) a. Klose scored a goal.
b. He headed the ball into the upper corner.

We observe that the spatial relations among the locations of the events described in these discourses are also highly determined by the rhetorical relations between the clauses used to describe them. In the NARRATION-related discourse (1), there is a spatial progression: Klose is located relative to his teammates (1a), he then moves from the bench to the sidelines (1b), and then he moves from the sidelines into the game (1c). In the ELABORATION-related discourse (2), there is no such progression.

In this paper, we investigate the degree to which the spatial structure of discourse and its rhetorical structure are co-determined. Using supervised machine learning techniques (Witten and Frank, 2002), we evaluate two hypotheses: (a) spatial information encoded in adjacent clauses is highly predictive of the rhetorical relations that hold between them and (b) spatial information is highly predictable based on associated spatial information within narrative event clauses. To do this, we build a corpus of narrative texts which are annotated both for spatial information (figure and ground (granularity) relationships, mereotopologically classified verbs and prepositions, and frames of reference) and rhetorical relations (a binary NARRATION vs. ELABORATION/BACKGROUND distinction discussed in Section 3.2). This corpus is then used to train two types of classifiers - one type that classifies the rhetorical relations holding between clauses on the basis of spatial information, and another type that classifies spatial relationships within clauses where the NARRATION relation holds. The results support both hypotheses and indicate the centrality of spatial information to narrative discourse structure and associated classification tasks.

2 Background and Related Research

2.1 Rhetorical Relations

Rhetorical relations describe the role that one clause plays with respect to another in a text and contributes to a text's coherence (Hobbs, 1985). As such, these relations are pragmatic features of a text. In NLP generally, classifying rhetorical relations has been an important area of research (Marcu, 2000; Sporleder and Lascarides, 2005) and has been shown to be useful for tasks such as text summarization (Marcu, 1998). The inventory of rhetorical relations in Segmented Discourse Representation Theory (SDRT) (Asher and Lascarides, 2003) is widely used in these applications. This inventory includes the following relations, illustrated by example: NARRATION: *Klose got up. He entered the game.* ELABORATION: *Klose pushed the Serbian midfielder. He knew him from school.* BACKGROUND: *Klose entered the game. The pitch was very wet.* EXPLANATION: *Klose received a red card. He pushed the Serbian midfielder.* CONSEQUENCE: *If Klose received a red card, then he pushed the Serbian midfielder.* RESULT: *Klose pushed the Serbian midfielder. He received a red card.* ALTERNATION: *Klose received a red card or he received a yellow card.* CONTINUATION: *Klose received a red card. Ronaldo received a yellow card.*

In previous work, rhetorical relations have been predicted based on a range of features including discourse connectives, relation location, clause length, part-of-speech, content and function words, and syntactic features (Marcu and Echihabi, 2002; Lapata and Lascarides, 2004). These systems have a wide range of average accuracies for all relations sought to be predicted - e.g. 33.96% (Marcu and Echihabi, 2002) to 70.70% (Lapata and Lascarides, 2004) - and individual relations - e.g. RESULT - 16.21% and EXPLANATION - 75.39% (Marcu and Echihabi, 2002) and CONTRAST - 43.64% and CONTINUATION - 83.35% (Sporleder and Lascarides, 2005). Our focus is on the NARRATION, BACKGROUND and ELABORATION relations, which account for over 90% of the discourses in our corpus.

2.2 Spatial Language and Discourse

Spatial language has been discussed in a number of NLP contexts. For example, linking natural language with physical locations via semantic mark-up (e.g. SpatialML (MITRE, 2009)); spatial description and wayfinding tasks (e.g. Anderson et al., 1991); and dialogue systems (e.g. Coventry et al., 2009), just to name a very few. Perspectives on spatial language are similarly varied in terms of their focus and theoretical background (e.g. cognitive, semantic and syntactic); however, common threads do emerge. First, all physical spatial references are reducible to figure and ground relationships (Talmy, 2000). In English, these are triggered by a deictic verb or adverb (e.g. *went, here*) (3a); a spatial preposition (e.g. *in, at*) (3b); a particle verb (e.g. *put on, got out*) (3c); or a motion verb (e.g. *drive, follow*) (3d).

(3) a. [Ronaldo]$_{figure}$ is [here]$_{ground}$.
b. [Ronaldo]$_{figure}$ is in [the park]$_{ground}$.
c. [Ronaldo]$_{figure}$ rolled over [Ø]$_{ground}$.
d. [Ronaldo]$_{figure}$ ran to [the park]$_{ground}$.

Second, figure and ground relationships qualitatively vary by the type of verb and preposition creating the relationship. These differences can be modeled in mereotopology, which defines spatial relationships in terms of regions and connections (e.g. RCC-8 (Randell et al., 1992)). We follow Asher and Sablayrolles (1995) who classify prepositions based on the position (Position - *at*, Initial Direction - *from*, Medial Position - *through*, Final Position - *to*) and contact (Inner - *in*, Contact - *against*, Outer

- *along*, and Outer-Most - *beyond*) of two regions (figure and ground). For verbs, Muller (2002) proposes six mereotopological classes: Reach, Leave, Internal, External, Hit, and Cross. Pustejovsky and Moszkowicz (2008) mapped Muller's classes to FrameNet and VerbNet and propose ten general classes of motion (Move, Move-External, Move-Internal, Leave, Reach, Detach, Hit, Follow, Deviate, Stay).

Third, figure and ground relationships vary by the perspective used to describe the relationship. For this discussion, perspective takes two forms, granularity of spatial description (following Montello (1993)) and frames of reference (following Levinson (1996)). Granularity refers to the level of detail in a given spatial description. Montello (1993, p. 315) indicates four spatial granularities based on the cognitive organization of spatial knowledge (summarized in (4)).

(4) a. Ronaldo jumped on the ball.
b. Ronaldo is in the corner.
c. Ronaldo is running around the field.
d. Ronaldo is in Cape Town.

(4a) is a **Figural** granularity which describes space smaller than the human body. (4b) is a **Vista** granularity which describes space from a single point of view. (4c) is an **Environmental** granularity which describes space larger than the body with multiple (scanning) point(s) of view. (4d) is a **Geographic** granularity which describes space even larger than the body and is learned by symbolic representation.

Frames of reference provide different ways of describing the same spatial relationships. For example, given a static scene of Ronaldo sitting on a bench next to his coach, each utterance in (5) would be an accurate spatial description.

(5) a. **Deictic**: *Ronaldo is there.*
b. **Contiguity**: *Ronaldo is on the bench.*
c. **Named Location**: *Ronaldo is at the sideline.*
d. **Relative**: *Ronaldo is in front of me.*
e. **Intrinsic**: *Ronaldo is behind his coach.*
f. **Absolute**: *Ronaldo is north of his coach.*

(5a-c) are non-coordinated as they relate just the figure and ground. Coordinated information, relating the figure to an additional entity within the ground, occurs in (5d-f). Frames of reference apply to both static and dynamic relationships (Levinson, 1996, p. 360).

In terms of attending to spatial information in discourse, Herman (2001) argues that spatial information patterns in narrative discourse carve out spatially defined domains that group narrative actions. In particular, the emergence and change in different types of spatial reference to physical location (discourse cues) create maps of the narrative actions. These discourse cues include figure, ground and path (motion) relationships (3); frames of reference (5); and deictic shifts - *here* vs. *there*. Herman's demonstration is based on ghost story narratives that are rich in spatial reference.

Howald (2010) showed in a corpus of serial killer first person narratives, also rich in spatial reference, that these spatial narrative domains, in the form of abstract Pre-Crime, Crime and Post-Crime events, were predicted to a 90% accuracy from three spatial features (figure, ground, and spatial verb) and discourse sequence. Overall, research by Herman (2001) and Howald (2010) demonstrates some level of dependency between spatial information and discourse structure. The present research addresses the specific question of whether there is a systematic relationship between spatial information and temporal information via rhetorical relations and the spatial architecture of narrative events.

3 Data and Annotation

3.1 Data

Three corpora of narrative discourse were annotated with rhetorical and spatial information. These corpora were then used to train and test machine learning systems. Summarized in Table 1, the three different narrative corpora selected for analysis were: (1) narratives from serial criminals (CRI) - oral and written confession statements and guilty pleas; (2) American National Corpus Charlotte Narrative and Conversation Collection (Ide and Suderman, 2007) (ANC) - oral narratives in conversations collected in a sociolinguistic interview format; and (3) The Degree Confluence Project (DEG) - this project, which seeks to map all possible latitude-longitude intersections on Earth, requires that participants who visit these intersections provide written narratives of the visit for inclusion on the project's website.

Table 1: Relation and Spatial Clause Distribution

Corpus	**ANC (n=20)**	**DEG (n=20)**	**CRI (n=20)**	**Total (N=60)**
Total Clauses	588	611	1,710	2,909
Spatial Clauses	260	354	932	1,546
Average	44.21	57.93	54.50	53.14
Total Rhetorical	568	591	1,690	2,848
Spatial Rhetorical	259	345	929	1,533
Average	45.59	58.37	55.00	53.82

20 narratives from each corpus were selected. There was a total of 2,909 (independent) clauses with 1,546 of those clauses containing spatial information - spatial clauses (53.14% on average). There was a total of 2,848 relations with 1,533 of those relations where *both* clauses contained spatial information - spatial rhetorical (53.82% on average).

3.2 Spatial Information and Rhetorical Relation Annotation

We developed a coding scheme for spatial information that consolidates the insights on spatial langauge discussed in Section 2.2.

- FIGURE is an indication of grammatical person or a non-person entity (**1** = *I, my*; **2** = *you, your*; **3** = *he, she, it, his, her*; **4** = *we, our*; **5** = *you, your*; **6** = *they, their*; **NP** = *the purse, a bench, three cars*);
- VERB is one of the four mereotopological classes - a consolidation of Pustejovsky and Moszkowicz's (2008) ten classifications (**State** = *was, stay, was sitting*; **Move** = *run, go, jump*; **Outside** = *follow, pass, track*; **Hit** = *attach, detach, strike*);
- PREPOSITION is one of four mereotopological classes based on Asher and Sablayrolles (1995) (**Positional** = *in, on*; **Initial** = *from* ; **Medial** = *through*; **Final** = *to*);
- GROUND is one of four granularities (**Figural**, **Environmental**, **Vista**, **Geographic**) (see (4) above);
- FRAME is one of six frames of reference (**Deictic**, **Contiguity**, **Named Location**, **Relative**, **Intrinsic**, **Absolute**) (see (5) above).

The three corpora were annotated by one of the authors. Annotation occurred one narrative at a time and any information from that narrative could be used to resolve rhetorical relations and spatial information. A reference sheet including several examples of each coding element was available to the annotator. The annotation happened in two phases. First, each pair of clauses was annotated with an SDRT relation. Second, each clause that contained a physical figure and ground relationship was identified. The figure, ground, preposition and verb were annotated with a **Figure**, **Verb**, **Preposition**, **Ground**, and **Frame**. We illustrate with (6) where the NARRATION relation obtains between (6a-b).

(6) a. Kaka kicked the ball into the goal.
b. Then he ran to the left side of the bench.

The spatial annotation of (6a) is: FIGURE = **NP**, *the ball*; VERB = **Hit (H)**, *kicked*; PREPOSITION = **Final (F)**, *into*; GROUND = **Environmental (E)**, *the goal*; and FRAME = **Contiguity (C)**. The spatial annotation of (6b) is: FIGURE = **3**, *he*; VERB = **Move (M)**, *ran*; PREPOSITION = **Final (F)**, *to the left side of*; GROUND = **Environmental (E)**, *the bench*; and FRAME = **Intrinsic (INT)**. The distribution of spatial rhetorical relations is summarized in Table 2.

Table 2: Spatial Rhetorical Relation Distribution per Corpus

Relation	ANC	DEG	CRI	Total
NARRATION	133	124	654	911
BACKGROUND	74	87	238	399
ELABORATION	34	63	17	114
CONTINUATION	14	27	10	51
RESULT	3	22	0	25
EXPLANATION	0	16	1	17
ALTERNATION	0	0	9	9
CONSEQUENCE	1	6	0	7
Total	259	345	929	1,533

An additional individual was queried for inter-rater reliability against the author annotation. The rater was given roughly one-third of the data (10 narratives (4 ANC, 4 DEG, 2 CRI) accounting for 510 spatial clause pairs), the same example sheet used by the author, and as much time as needed to complete the task. Average agreement and Cohen's kappa statistics (Cohen, 1960) were computed between the inter-rater and the author for the spatial annotations and NARRATION, BACKGROUND, and ELABORATION codings. Individually, BACKGROUND and ELABORATION have low interannotator agreement (κ = 32.92 and 54.20 respectively), but these two relations were often confused (26% of BACKGROUND relations coded as ELABORATION and 12% of ELABORATION relations coded as BACKGROUND). As illustrated in (7-8), both BACKGROUND and ELABORATION add information to the surrounding state of affairs.

(7) a. Klose entered the game.
b. The pitch was very wet.

(8) a. Klose pushed the Serbian midfielder.
b. He knew him from school.

As evidenced by the annotation confusions, the difference between these relations is difficult to distinguish and the distinction made by Asher and Lascarides (2003) is subtle - BACKGROUND's temporal consequence is one of *overlap* and ELABORATION, a subordinating relation, is one of *part-of*. However collapsing these relations resulted in a fairly reliably distinguished category. Average agreement and kappa statistics are summarized in Table 3.

Table 3: Agreement and Kappa Statistics for Relation and Spatial Codings

Coding	Agreement (%)	Kappa (κ)
All Rhetorical Relations	71.97	60.27
NARRATION	86.32	74.36
BACKGROUND / ELABORATION	73.40	62.20
Figure	94.91	89.92
Verb	90.90	81.80
Preposition	78.35	56.70
Granularity	87.87	75.74
Frame	69.38	38.76

For rhetorical relations, the average agreement and kappa statistic are consistent with previously reported performances (e.g. Agreement = 71.25 / κ = 61.00 (Sporleder and Lascarides, 2005)). We have not been able to find previously reported performance accuracies for NARRATION, ELABORATION and BACKGROUND relations specifically. However, κ statistics from 60.00 to 75.00 and above are considered acceptable (e.g. Landis and Koch, 1977). For the spatial codings, the average agreements are relatively high with **Preposition** and **Frame** falling lowest. There is no basis for direct comparison of these numbers to other research as the coding scheme is novel.

4 Machine Learning Experiments

We constructed two machine learning tasks to exploit the annotated spatial information to determine what contributions the information is making to narrative structure. The first task evaluates the prediction of NARRATION and BACKGROUND/ ELABORATION relations based on pairs of spatial clauses. The second task evaluates the prediction of spatial information types, based on the other spatial information types in that clause, in individual clauses where the NARRATION relation holds.

4.1 Rhetorical Relation Prediction

4.1.1 Methods and Results

Task 1 builds a 2-way classifier for the NARRATION and BACKGROUND/ ELABORATION relations. Clause pairs were coded as vectors (n = 1,424) - for example, the vector for (6) is **NP3**, **HM**, **FF**, **EE**, **CINT**. These vectors were used to train and test (10-fold cross-validation) a number of classifiers. The Naïve Bayes classifier performed the best. Results are reported in Table 4.

Table 4: Naïve Bayes Classification Accuracy and F-Measures for Task 1

NARRATION	**Accuracy (% / baseline)**	**Precision**	**Recall**	**F-Score**
ANC	63.29 / 58	.676	.633	.654
DEG	75.71 / 61	.803	.757	.779
CRI	90.12 / 73	.822	.901	.860
TOTAL	84.90 / 68	.808	.841	.824
BACK/ ELAB	**Accuracy (% / baseline)**	**Precision**	**Recall**	**F-Score**
ANC	57.89 / 41	.532	.579	.555
DEG	70.11 / 38	.642	.701	.670
CRI	45.63 / 26	.624	.456	.527
TOTAL	57.87 / 35	.622	.567	.593

For all corpora combined, the majority class ("baseline") for NARRATION is 68% and 26% for BACKGROUND / ELABORATION; the classifier performs 16% and 22% above baseline respectively. The difference between the NARRATION and BACKGROUND / ELABORATION relations and baselines is statistically significant for each corpus and all corpora combined - ANC: χ^2 = 25.64, d.f. = 1, p $\leq$ .001; DEG: χ^2 = 33.86, d.f. = 1, p $\leq$ .001; CRI: χ^2 = 22.69, d.f. = 1, p $\leq$ .001; and TOTAL:χ^2 = 34.09, d.f. = 1, p $\leq$ .001.

4.1.2 Discussion

Again, we have not been able to find reported results for a direct comparison of NARRATION and BACKGROUND/ ELABORATION. However, the 84.90% and 57.87% (at 16% and 22% over baseline) performance of our Naïve Bayesian model is consistent with results reported in similar tasks. For example, Marcu and Echihabi (2002) report an average accuracy of 33.96% (5-way classifier) and 49.70% (6-way classifier) based on training with very large data sets. Sporleder and Lascarides (2005) report a 57.55% average accuracy, based on training with large data sets, which is 20% over Marcu and Echihabi's 5-way classifier and almost 40% over a random 20% baseline. Lapata and Lascarides (2004) report an average accuracy of 70.70% for inferring temporal relations based on training.

We ran an additional set of experiments to determine the relative contribution of spatial features to predict NARRATION and BACKGROUND / ELABORATION relations. As shown in Table 5, **Figure** and **Verb** outperform **Ground**, **Preposition** and **Frame** in accuracy. **Figure** performs at a 71% average accuracy (85% for NARRATION and 40% for BACKGROUND/ ELABORATION) and **Verb** performs at a 74% average accuracy (84% for NARRATION and 54% for BACKGROUND/ ELABORATION). **Figure** and **Verb** appear to be most discriminating. Note that we are not suggesting that *subject* and *verb* generally are similarly discriminatory - **Figure** and **Verb** in this task are overtly spatial. Despite the performance of **Figure** and **Verb**, different subsets of spatial information worked better (we ran all permutations of spatial features - the top five are listed in Table 5). However, the difference in performance is negligible. For example, the best subset of **Figure**, **Verb** and **Ground** (85% and 58%) only performed 1% above NARRATION and BACKGROUND/ ELABORATION prediction based on all five features combined.

Table 5: Single and Combined Spatial Feature Performance

Feature	NARRATION	BACK/ ELAB	**Features**	NARRATION	BACK/ ELAB
Figure (F)	85.58	40.33	**FVG**	85.24	58.33
Verb (V)	84.59	54.97	**VGP**	84.34	58.33
Prepostion (P)	97.34	1.00	**FVGR**	86.33	56.45
Ground (G)	97.33	1.00	**FV**	86.56	56.90
Frame (R)	98.02	2.00	**VG**	85.37	57.33

These results tell us several things about the relationship between spatial information and rhetorical structure as it applies to narrative discourse. First, spatial information predicts rhetorical structure as good as non-spatial types of linguistic information reported in other investigations and with many fewer features. For example, Sporleder and Lascarides (2005) rely on 72 different features falling into nine classes whereas we rely on 14 features in five classes. This suggests that spatial information is not only central to rhetorical stucture, like temporal components, but central to the task of prediction. Second, while the type of spatial information that predicts rhetorical structure is based on the primary figure and ground relationship, it is the qualitative semantic variations within these elements that is providing the discrimination. It is the organization of spatial relationships - (**Verb** and **Preposition**) and the perspective provided by the narrator (**Figure**, **Ground** and **Frame**) combined - rather than any individual elements.

4.2 Spatial Information Prediction

4.2.1 Methods and Results

Task 2 is a series of five experiments. Each experiment builds a classifier for each type of spatial information: a 6-way classifier for **Frame**; a 5-way classifier for **Figure** (**Figure** types **2** and **5** did not occur in our corpus); and 4-way classifiers for **Ground**, **Preposition** and **Verb**. Single clauses that contribute to the NARRATION relation were coded as vectors (n = 911) - for example, the single vectors for (6a) and (6b) are **NP**, **H**, **F**, **E**, **C** and **3**, **M**, **F**, **E**, **INT**. These vectors were used to train and test (10-fold cross-validation) a number of classifiers to predict one of the five spatial features given the remaining four. The K* classifier performed the best. Results are reported in Table 6. For all corpora combined, the K* classifier performs above baseline for all spatial information (**Figure** = 9%, **Verb** = 17%, **Preposition** = 9%, **Ground** = 19%, **Frame** = 8%) (χ^2 = 20.95, d.f. = 4, p $\leq$ .001).

4.2.2 Discussion

Even though the accuracies of predicting spatial information are significantly above baseline, we sought ways to boost performance by considering implicit spatial information. For those clauses without explicit spatial information, we extended the annotation of the previous clause's coding based on the inertia of

Table 6: K* Classification Accuracy and F-Measures for Task 2

Spatial Information	Accuracy (% / baseline)	Precision	Recall	F-Score
Figure	47.97 / 38	.464	.480	.428
Verb	67.32 / 50	.635	.673	.640
Preposition	53.69 / 46	.492	.537	.499
Ground	53.59 / 34	.530	.536	.519
Frame	55.67 / 47	.507	.557	.511

narrative texts. Rapaport, et al. (1994) discuss the temporal inertia of narrative texts - time moves forward through narrative events. In the absence of updating, information is maintained. We suggest that inertia applies to spatial information as well. For example, given the clauses - *John entered the room. He sat down.* - we make the assumption that John sat down in the room that he entered. We illustrate with (9).

(9) a. Kaka kicked the ball into the goal.
NP, **H**, **F**, **E**, **C**, **.33**
b. The goaltender yelled in frustration.
NP, **H**, **F**, **E**, **C**, **.66**
c. Then Kaka ran to the left side of the bench.
3, **M**, **F**, **E**, **INT**, **1**

No explicit spatial information exists in (9b). We took the coding from the explicit spatial information in (9a) and maintained it for (9b). New explicit spatial information occurs in (9c) and the coding is updated. Further, we included explicit sequence information as a measure of a given clause's proportional position within the text (**.33**, **.66** and **1**). In the absence of overt temporal specification (occuring in only 10% of the clauses in our corpus), the sequence information, a textual feature, parallels the temporal progression (and inertia) of narrative events. This added 560 additional vectors (n = 1,471). The K* classifier still performed the best. The results are summarized in Table 7.

Table 7: K* Classification Accuracy and F-Measures for Task 2 Boosted Vectors

SPATIAL INERTIA	Accuracy (% / baseline)	Precision	Recall	F-Score
Figure	51.73 / 41	.509	.517	.473
Verb	70.22 / 48	.673	.700	.679
Preposition	57.30 / 47	.571	.573	.540
Ground	62.61 / 35	.636	.626	.611
Frame	59.82 / 44	.574	.598	.564
SPATIAL INERTIA + SEQUENCE	**Accuracy (% / baseline)**	**Precision**	**Recall**	**F-Score**
Figure	70.56 / 41	.702	.706	.699
Verb	79.33 / 48	.789	.793	.790
Preposition	67.91 / 47	.676	.679	.674
Ground	72.39 / 35	.721	.724	.721
Frame	69.06 / 44	.678	.691	.681

Inclusion of the spatial inertia values improves performance of the K* classifier in all cases (χ^2 = 40.59, d.f. = 4, p $\leq$ .001). Inclusion of sequence information improves performance even further (χ^2 = 102.36, d.f. = 4, p $\leq$ .001). Note that, despite the increase in performance, sequencing information alone does not do as well, indicating that spatial information still plays a discriminatory role. Using sequence information alone as a baseline (**Figure** = 47%, **Verb** = 52%, **Preposition** = 47%, **Ground** = 44%, **Frame** = 48%;), the normalized performance values above sequence baseline become **Figure** = 23%, **Verb** = 27%, **Preposition** = 28%, **Ground** = 20%, and **Frame** = 21%.

The ability to predict spatial features appears to be dependent both on a patterned distribution of the per-clause spatial information (increased by spatial inertia) and on the textual feature of sequence (temporal inertia). This seems to hold despite the specific subject matter or spatial characteristics of a given narrative. Considering the complete spatiotemporal picture for narrative clauses yields the best prediction results and suggests that the spatial information structure of narrative discourse represents some type of organization akin to what Herman (2001) and Howald (2010) have evaluated in spatially-rich narratives. Based on the tasks presented here, this organization appears to be fundamental and relative to formal temporally-informed discourse structure.

5 Conclusion

Exploration of the spatial dimension in narrative discourse provides interesting and robust possibilities for computational discourse analysis. We have described two machine learning tasks which exploit spatial linguistic features. In addition to improving on existing prediction systems, both tasks empirically demonstrate that, when available, certain types of spatial information are predictors of the rhetorical structure of narrative discourse and the spatial information of narrative event sequences. Based on these results, we indicate that spatial structure is related to temporal structure in narrative discourse.

The coding scheme proposed here models complex and interrelated properties of spatial relationships and perspectives and should be generalizeable to other non-narrative discourses. Future research will focus on different discourse corpora to determine how spatial information is related to rhetorical structure. Additional future research will also focus on automation of the annotation process. The ambiguity of spatial language makes automatic extraction of spatial features infeasible at the current state of the art. Fortunately, average agreement and kappa statistics for coding of the spatial information and rhetorical relations are within acceptable ranges. The annotated spatial features are semantically deep and useful for not only computational discourse systems, but tasks that involve the semantic modeling of spatial relations and spatial reasoning.

Acknowledgments

Thank you to David Herman and James Pustejovsky for productive comments and discussion and to Jerry Hobbs for suggesting the Degree Confluence Project as a source of spatially rich narratives. Thank you also to four anonymous reviewers for very helpful insights.

References

[1] Anne Anderson, Miles Bader, Ellen Bard, Elizabeth Boyle, Gwyneth Doherty, Simon Garrod, Stephen Isard, Jacqueline Kowtko, Jan McAllister, Jim Miller, Catherine Sotillo, Henry Thompson, and Regina Weinert. 1991. The HCRC Map Task Corpus. *Language and Speech*, 34:351–366.

[2] Nicholas Asher and Alex Lascarides. 2003. *Logics of Conversation*. Cambridge University Press, Cambridge, UK.

[3] Nicholas Asher and Pierre Sablayrolles. 1995. A Typology and Discourse Semantics for Motion Verbs and Spatial PPs in French. *Journal of Semantics*, 12(2):163–209.

[4] Jacob Cohen. 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1):37–46.

[5] Kenny Coventry, Thora Tenbrink, and John Bateman. 2009. *Spatial Language and Dialogue*. Oxford University Press, Oxford, UK.

[6] David Herman. 2001. Spatial Reference in Narrative Domains. *Text*, 21(4):515–541.

[7] Jerry R. Hobbs. 1985. On The Coherence and Structure of Discourse. CSLI Technical Report, 85-37.

[8] Blake Howald. 2010. Linguistic Spatial Classifications of Event Domains in Narratives of Crime. *Journal of Spatial Information Science*, 1.75–93.

[9] Nancy Ide and Keith Suderman. 2007. The Open American National Corpus (OANC), available at http://www.AmericanNationalCorpus.org/OANC.

[10] Richard Landis and Gary Koch. 1977. The Measurement of Observer Agreement for Categorical Data. Biometrics, 33(1):159–174.

[11] Mirella Lapata and Alex Lascarides. 2004. Inferring sentence internal temporal relations. In Proceedings of NAACL-04, 153–160.

[12] Stephen C. Levinson. 1996. Language and Space. *Annual Review of Anthropology*, 25(1):353–382.

[13] William Mann and Sandra Thompson. 1987. Rhetorical Structure Theory: A Framework for The Analysis of Texts. *International Pragmatics Association Papers in Pragmatics*, 1:79–105.

[14] Daniel Marcu. 1998. Improving Summarization Through Rhetorical Parsing Tuning. In *The 6th Workshop on Very Large Corpora*, 206–215.

[15] Daniel Marcu. 2000. The Rhetorical Parsing of Unrestricted Texts: A Surface-Based Approach. *Computational Linguistics*, 26(3):395–448.

[16] Daniel Marcu and Abdessamad Echihabi. 2002. An Unsupervised Approach to Recognizing Discourse Relations. In *Proceedings of ACL-02*, 368–375.

[17] MITRE. 2009. SpatialML: Annotation Scheme for Marking Spatial Expressions in Natural Language, Version 3.0. April 3, 2009.

[18] Daniel R. Montello. 1993. Scale and Multiple Psychologies of Space. In A. Frank and I. Campari (eds.), *Spatial Information Theory: A Theoretical Basis for GIS* (LNCS 716), 312–321. Springer-Verlag, Berlin.

[19] Philippe Muller. 2002. Topological Spatio-temporal Reasoning and Representation. *Computational Intelligence*, 18(3):420–450.

[20] Barbara Partee. 1984. Nominal and Temporal Anaphora. *Linguistics and Philosophy*, 7(3):243–286.

[21] James Pustejovsky and Jessica Moszkowicz. 2008. Integrating motion predicate classes with spatial and temporal annotations. *COLING 2008*:95–98.

[22] James Pustejovsky, José Castaño, Robert Ingria, Roser Saur, Robert Gaizauskas, Andrea Setzer, and Graham Katz. 2003. TimeML: Robust Specification of Event and Temporal Expressions in Text. *In Proceedings of the IWCS-5, Fifth International Workshop on Computational Semantics*.

[23] David Randell, Zhan Cui, and Anthony Cohn. 1992. A Spatial Logic Based on Regions and Connection. *Proceedings of KR92*, 394–398. Los Altos, CA: Morgan Kaufmann.

[24] William Rapaport, Erwin Segal, Stuart Shapiro, David Zubin, Gail Bruder, Judith Duchan, Michael Almeida, Joyce Daniels, Mary Galbraith, Janyce Wiebe and Albert Yuhan. 1994. Deictic Centers and the Cognitive Structure of Narrative Comprehension. Technical Report No. 89-01. Buffalo, NY: SUNY Buffalo Department of Computer Science.

[25] Caroline Sporleder and Alex Lascarides. 2005. Exploiting Linguistic Cues to Classify Rhetorical Relations. *Proceedings of Recent Advances in Natural Language Processing (RANLP-05)*, 532–539.

[26] Leonard Talmy. 2000. *Toward a Cognitive Semantics, Volume 2*. The MIT Press, Cambridge, MA.

[27] Ian Witten and Eibe Frank. 2002. *Data Mining Practical Machine Learning Tools and Techniques with Java Implementation*. Morgan Kaufmann.