Automatic Identification of Motion Verbs in WordNet and FrameNet

Parvin Sadat Feizabadi and **Sebastian Padó**
Institut für Computerlinguistik
Universität Heidelberg
Im Neuenheimer Feld 325
69120 Heidelberg
feizabadi,pado@cl.uni-heidelberg.de

Abstract

This paper discusses the automatic identification of motion verbs. The context is the recovery of unrealized location roles from discourse context or "locational inference", a special case of missing argument recovery. We first report on a small corpus study on verb classes for which location roles are particularly relevant. This includes motion, orientation and position verbs. Then, we discuss the automatic recognition of these verbs on the basis of WordNet and FrameNet. For FrameNet, we obtain results up to 67% F-Score.

1 Introduction

A central problem in natural language processing (NLP) is that a considerable portion of the meaning of texts is typically not expressed overtly. It must be provided by the reader through inference, as described, e.g., by Norvig (1987): *An inference is defined to be any assertion which the reader comes to believe to be true as a result of reading the text, but which was not previously believed by the reader, and was not stated explicitly in the text.* The omission of assertions is by no means a sign of erroneous or "sloppy" language use. In fact, pragmatic considerations, notably the maxim of quantity (Grice, 1975) encourage speakers not to express all information.

Omission of material is thus a fundamental fact of language. Its importance for NLP can be appreciated in the framework of textual entailment inference, a semantic meta-task which is defined as the decision, given two short texts T (Text) and H (Hypothesis), whether *readers of T would agree that H is (almost) certainly true* (Dagan et al., 2009). The reasoning requirements of many NLP tasks can be phrased as textual entailment queries. Examples include answer validation in QA (Peñas et al., 2008); IE (Romano et al., 2006); and multi-document summarization (Harabagiu et al., 2007).

Within textual entailment, methods which do not try to recover missing assertions in T often fail to correctly derive the existence of an entailment relation when H realizes an assertion that T does not. Recent work has verified that this routinely happens for sentences in discourse, as the following example shows (Mirkin et al., 2010):

(1) **Context:** *China* seeks solutions for its coal mine safety problem. [. . .]

T: A recent accident has cost more than a dozen miners their lives.

H: An accident *in China* has killed several miners.

The Hypothesis **H** in Example (1) can be inferred almost completely from the Text **T**, with the exception of the prepositional phrase "in China", which is not realized locally in the Text. It can however be inferred from the context sentence.

This paper presents a pilot study in the context of *locational inference*, i.e., the recovery of unexpressed assertions that concerns the spatial configuration of events. Determining location information is an important subpart of inference and plays an important part in Information Extraction (Leidner et al., 2003) as well as Question Answering (Greenwood, 2004) and has been accorded an important role in the analysis of narratives (Herman, 2001; Howald, 2010).

Proceedings of KONVENS 2012 (Main track: oral presentations), Vienna, September 19, 2012

Our ultimate goal is to develop computational methods that can automatically retrieve omitted locations. The present paper takes the first step by defining the task, performing a focused data analysis, and evaluating the ability of the two most widely used resources in computational linguistics, WordNet and FrameNet, for the purpose of identifying motion verbs, for which locational inference is particularly relevant.

Plan of the paper. Section 2 defines the task of locational inference and discusses its challenges. Section 3 presents a corpus annotation study which provides ground truth of motion verbs and location roles as well as an analysis. Section 4 evaluates how well WordNet and FrameNet can be used to tackle the identification of motion verbs.

2 Locational Inference and Motion Verbs

2.1 Locational Inference

Locational inference is the special case of the recovery of unrealized arguments or null instantiations for semantic roles that denote places and directions. Figure 1 shows our concept of a processing pipeline for location inference. The task consists of several subtasks. The first subtask (I.) is the recognition of verbs that actually require locational inference. This subtasks again decomposes into two parts: first we verify that the verb in question requires a location (1), and then we check that the current instance of the lemma leaves at least one location unrealized so that it must be inferred (2). The second subtask, recovery (II.), then attempts to identify the missing location from the available locations in context (3., 4.). The modeling part of this paper focuses on the first recognition step, that is, the decision of whether an event requires a location (Step 1).

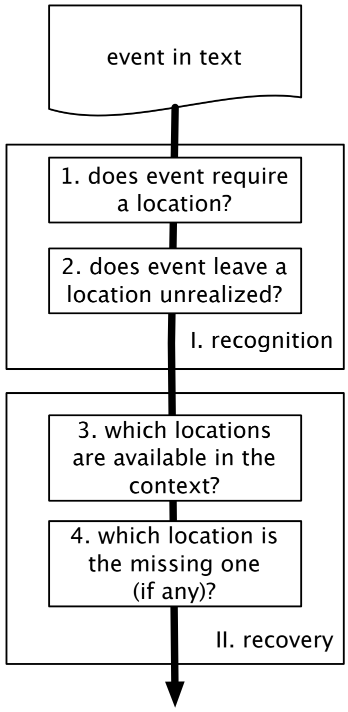

Figure 1: Processing pipeline for locational inference

2.2 Related Work

This task shows similarities to some existing NLP processing paradigms. The most notable is semantic role labeling or SRL (Gildea and Jurafsky, 2002), in the sense that locations can be seen as specific semantic roles. Semantic roles have also been employed as a basis for deciding entailment (Burchardt et al., 2009). The division of locational inference into recognition and recovery is also reminiscent of the frequent decomposition of SRL into recognition and classification steps. Traditionally, SRL concentrates on locally realized arguments, but awareness is growing that SRL needs to include non-local arguments. Gerber and Chai (2010) found many implicit arguments of nouns to be recoverable from prior context, and a recent SemEval task (Ruppenhofer et al., 2010) explicitly included unrealized arguments.

A second related task is coreference resolution (Soon et al., 2001). The analogy applies in particular to the second part of locational inference (recovery). Similar to coreference resolution, we have to construct a set of contextually available "antecedents" for missing locations and pick the correct one. However, here face "zero anaphora" (Fillmore, 1986), which again makes it difficult to identify features. Silberer and Frank (2012), address the general problem but find it to be very difficult. By focusing on location roles, we hope to simplify the problem through limiting the semantic types of possible antecedents.

2.3 Locational Inference and Motion Verbs

A fundamental question about locational inference is what verbs it applies to. While it can be argued that (almost) all events are take place somewhere, our intuition was that for many events in a text it is difficult to pin down where exactly they happen. This is true for many states as well as events. Consider this example:

(2) Although the concept **was** a simple one, Allan **thought** it had potential.

Here, the discourse context does not provide any hint as to the location of the marked verbs. At the same time, the location is not centrally important to the event either; therefore inferring it does not appear crucially important.

The situation is substantially different for some verb classes; notably verbs of motion (including self motion and caused motion), verbs of orientation, and verbs of position. For verbs in these classes, locations play an integral role in their semantics, as can be argued either on the basis of decomposition (Jackendoff, 1990) or of corpus evidence (Baker et al., 1998).

Based on these observations, we decided to focus on these three classes of verbs which require a location in this paper: motion, orientation, and position verbs (henceforth called motion verbs for the sake of simplicity). Thus, *see* and *ask* are not included in the scope of our study, but *arrive* and *sit* are. By taking this stance, we do not mean to imply that locational inference is irrelevant for non-motion verbs. We merely take the decision to focus on these comparatively "clear" cases for the purpose of this pilot study.

2.4 A FrameNet-based Characterization of the Motion Domain

There is no readily available comprehensive definition of motion, orientation, and position verbs in any existing computational resource. Part of the problem are unclear boundaries between motion verbs and change-of-state verbs. For example, in Levin's verb classification (Levin, 1993), class 10 (*Verbs of Removing*) appear to be a good candidate to contain (caused) motion verbs; however, the description of subclass 10.3 (*Clear verbs*) notes that "some of [the verbs'] uses are better characterized as verbs of change of state".

However, in a small pilot experiment we found annotators' intuitions on what constitutes a motion verb to deviate too much for reliable annotation. We decided to ground our notion of motion (plus orientation and position) verbs based on FrameNet. FrameNet (Baker et al., 1998) is a semantic dictionary of English, based on the concept of semantic frames from Fillmore's frame semantics (Fillmore, 1985). Each frame describes a prototypical situation and comes with a set of semantic roles describing the participants and objects of that situation. It also lists a set of lexical units, word senses that can introduce the frame. In this way, frames group verbs into semantic classes. FrameNet includes about 10,000 lexical units, and 800 semantic frames.

More specifically, we started out from FrameNet's MOTION frame which assumes five location roles:

- SOURCE: "If A is the source of a motion event with mover B, then B is at place A at the beginning of the event".
- GOAL: "If A is the goal of a motion event with mover B, then B is at place A at the end of the event".
- PATH: "If A is the path of a motion event with mover B, then B is at place A at sometime during the movement event, but neither at the beginning of the event nor at the end."
- PLACE: "Place identifies the setting in which the Theme's movement takes place without a specified Path. "
- DIRECTION: "If A is the direction of a motion event from source B to goal C, then A is [...] a straight line from B to C."

These five roles appear to be the maximum set of location roles that motion verbs support in English. We then identified all FrameNet frames with at least one of these roles, manually excluding about 10 frames which used the same role names for non-locations (e.g., SOURCE in the case of AUTHORITY as in *military*$_{\text{SRC}}$ *power*). This step yielded a set of 34 frames, shown in Table 1. A list of the verbs from these frames, together with the lists of rolesets FrameNet provides for the each verb (possibly more than one), formed the basis of our corpus study.

ADORNING	ARRIVING
CAUSE_FLUIDIC_MOTION	CAUSE_MOTION
COTHEME	DEPARTING
EMANATING	EMITTING
EMPTYING	EVENT
FILLING	FLUIDIC_MOTION
GIVING	IMPORT_EXPORT
LIGHT_MOVEMENT	MASS_MOTION
MOTION	MOTION_DIRECTIONAL
MOTION_NOISE	OPERATE_VEHICLE
PATH_SHAPE	PLACING
POSTURE	PRECIPITATION
QUITTING	RECEIVING
REMOVING	RESIDENCE
RIDE_VEHICLE	SELF_MOTION
SENDING	SMUGGLING
TAKING	TRAVEL

Table 1: List of Motion frames in FrameNet

3 A Corpus Study of Motion Verbs and Location Roles

3.1 Annotation

The goal of the annotation is to produce a ground truth of human's understanding of location information from text. Annotators were asked to identify all motion verbs in the corpus and annotate their location roles.

For our annotation, we chose the 4000-word short story "The Black Willow" from the "fiction" subset of MASC, the Manually Annotated Subcorpus of the American National Corpus (Ide et al., 2008). We decided on a fiction text since fiction is less stylistically standardized than newswire text and we therefore expect to find more natural narrative structures.

To perform the annotation, we used the MMAX2 Annotation Tool (Müller and Strube, 2006). It is a graphical user interface for creating, browsing, visualizing and querying linguistic annotations on multiple levels. We annotated information on raw text, with no linguistic analysis other than sentence segmentation.

Annotators were provided with the list of motion verbs according to FrameNet (cf. Section 2.4), with the understanding that this list was possibly incomplete. It was complemented by detailed annotation guidelines[1]. For verbs, annotators specified whether the verb was in the FrameNet list, and if it was, whether the instance could be covered by any of the role sets offered by FrameNet.

[1]The guidelines are available from `http://www.nlpado.de/~sebastian/data.shtml`.

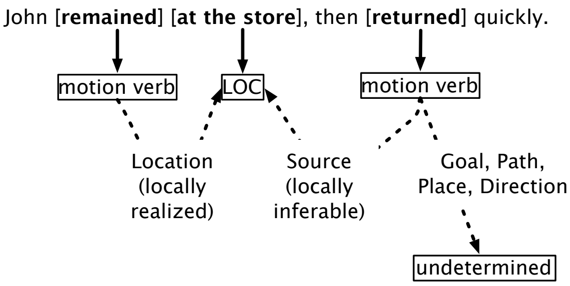

Figure 2: Annotation example

Each location was tagged with one of four possible statuses: *locally realized, locally inferable, globally inferable*, *world knowledge* and *undetermined*. The first option is for roles realized explicitly in the verb's predicate-argument structure. The second and third option describe roles which are inferable from context – "locally inferable" is for locations within a three-sentence window (previous, current, and next sentence), and "globally inferable" is for locations that can be found outside this window. "World knowledge" covers cases where an unrealized role can be completed based on world knowledge triggered by the predicate-argument structure (see below for examples). The last option, "undetermined", indicates that the related location role remains completely undetermined: it is not mentioned explicitly anywhere in the text and cannot be determined based on world knowledge either.

Figure 2 illustrates most of these categories. It shows a sentence with one position verb and one motion verb and one location. The position verb is annotated with just one location role (LOCATION) which is locally realized. The motion verb, in contrast, is tagged with five location roles none of which are locally realized. The SOURCE role is locally inferable from the first verb, and the others are all undetermined, due to the lack of discourse context in the example.

3.2 Annotation Reliability

The annotation was performed by two annotators with very good proficiency in English and graduate-level background in linguistics. The agreement of the two annotators, computed as ex-

Agreement on motion verbs	94%
Agreement on PLACE role	60%
Agreement on SOURCE role	64%
Agreement on GOAL role	73%
Agreement on PATH role	74%
Agreement on DIRECTION role	68%

Table 2: Annotation reliability (exact match)

act match, is shown in Table 2. Agreement on motion verbs, which this paper focuses on, is very good. Agreement on roles is lower. We believe that this is at least in part due to our detailed annotation scheme (cf. the status types described in Section 3.1). This interpretation calls for further investigation in the future. After computing agreement, the two annotators resolved annotation differences to produce a joint gold standard.

3.3 Analysis of the Annotation

Among the 4000 words of "The Black Willow", we found 731 verbs, of which 208 belonged to the motion domain and were thus annotated with information regarding their location(s). Of these 208, 143 (69%) were in the FrameNet-derived list, but 65 were not. Furthermore, There were 9 instances among the 143 where FrameNet did not offer the correct role set (i.e., the correct sense).

The numbers indicate that FrameNet is a solid, but not complete, resource to guide annotation of motion verbs. Annotators cannot be told to rely on the completeness of the lists that they are provided. Examples of missing coverage include cases where the motion verb is not listed in FrameNet, although the correct frame is available (e.g., *lean back*) as well as cases of missing frames. This latter case includes, for example, the frame IMPACT which lists several motion verbs but was not selected because it uses frame-specific roles (*Impactor, Impactee*) rather than directly recognizable motion roles.

We also identified a total of 190 realized location phrases. Motion verbs and locations were linked by a total of 268 roles. Thus, each verb is linked to an average of 1.3 realized location roles, either local or inferred, and many of the locations are referred to by more than one event. Given the theoretical maximum of 683 location roles that the 208 verbs could realize according to their FrameNet rolesets, some 40% of the roles were realized somewhere in the discourse.

The 268 roles decompose into 188 locally realized roles (i.e., within the same predicate-argument structure), 43 locally inferrable ones (in the direct context), and 37 globally inferrable ones (in the complete document context). These numbers indicate that there is indeed considerable need for locational inference: only 188 out of 683 roles are realized locally. Unfortunately, only another 80 can be recovered by finding their "antecendents" in the discourse using the surface-based methods to have in mind (cf. Section 2.1).

Thus, we find that even for motion verbs, a considerable number of location roles remains undetermined even taking the complete document into account. An analysis by location role found PATH to be the role with the highest ratio (74%), and PLACE the one with the lowest ratio (39%). GOAL (56%), SOURCE (54%), and DIRECTION (54%) are located between the extremes. We investigated the background of these fairly high numbers and found the main reason to be that undetermined location roles are often inferable not from the discourse context, but from the meaning of the predicate itself or from other roles. This happens in the example "they kicked up dust [$_{\text{PATH}}$ along the roadway]" where the PATH role supports the inference that the PLACE of the kicking event is the roadway, too. In "he left [$_{\text{SOURCE}}$ the room]", the predicate triggers the inference that the DESTINATION is simply "some place outside the room". Finally, in "the sun had reached [$_{\text{GOAL}}$ the trees]", the sky understood as SOURCE. This inference requires world knowledge about the sun as well as the relative position of the sky and the trees.

4 Automatic Identification of Motion Verb Instances

The first requirement for automating locational inference is an automatic means to identify motion verbs in running text (cf. Step 1 in Figure 1). Given that the definition of the motion domain is essentially a semantic one, the second contribution of this paper is a set of experiments with the goal of recognizing motion verbs in text. We follow a knowledge-based approach, using two standard CL resources, WordNet and FrameNet.

FrameNet was already introduced in Section 2.4. WordNet is an electronic lexical database, in which

English nouns, verbs, adjectives, and adverbs are organized into synonym sets (synsets) at the sense level (Miller et al., 1990). WordNet includes more than 150,000 words grouped into 117,000 synsets, among which there are over 11,000 verbs.

The main challenge in identifying motion verbs for locational inference is word sense disambiguation: Many verbs have motion as well as non-motion senses. For example, *to cross* can be a motion verb (*The ship crossed the ocean*), in which case it would be subject to locational inference, but it can also mean a gesture (*Peter crossed himself*) or a trickery event (*John was crossed by the con man*) – in the last two cases we currently assume that no locational inference takes place. A potential additional complication arises from the fact that the motion domain is a classical source domain for metaphorical mappings (Lakoff and Johnson, 1980). That is, many motion verbs are used to express metaphorical motion, orientation, or position (*Colin moves towards the Labour position / is on the Tory side*). Fortunately, a preliminary corpus analysis indicated that metaphorical usages of motion verbs behave similar to literal usages with regard to locations and locational inference. That is, even though it might be ultimately desirable to distinguish literal and metaphorical motion verbs in order to avoid unwanted inferences, such as *the Labour position* is a physical place, we avoid this distinction in the current study.

Thus, the concrete motion verb prediction task is as follows: Given the information in one of the two resources, we (a) define the set of motion verbs in terms of the resources; (b), we disambiguate the verb instances in the text with the resource; (c) we predict an instance to be a motion verb if it is included in the set from (a).

4.1 Preprocessing

Given that our manual annotation is based on FrameNet, it seems natural to use FrameNet frames for disambiguation. However, the development of standalone FrameNet-based WSD is made difficult by the incomplete coverage of word senses by FrameNet (Erk, 2005) as a consequence of which many instances should be left unassigned. The results of our own annotation confirm this observation (cf. Section 3.3). We therefore performed Word Sense Disambiguation on the basis of WordNet 3.0 using UKB, a state-of-the-art unsupervised graph-based word sense disambiguation tool (Agirre and Soroa, 2009). Since UKB requires part-of-speech information, so also performed part-of-speech tagging with TreeTagger (Schmid, 1994).

We also evaluated the quality of the preprocessing components. First, we found that TreeTagger only recognized 89% of the gold standard motion verbs as verbs. Most of the missing cases were phrasal verbs which the tagger could not handle correctly; this leads to an upper bound of 89% recall for any model building on the TreeTagger output. Then, we annotated the verbs with WordNet synsets to evaluate UKB. When compared against the gold standard, its exact match accuracy is almost exactly 50%, comparable to the verb results reported by Agirre and Soroa on the Senseval all-words datasets. However, accuracy improves to 73% if we evaluate just the coarse-grained decision motion verbs vs. non-motion verbs.

4.2 Disambiguation Strategies

The numbers reported in the previous subsection leave open the question of whether WSD is actually good enough to serve for the selection of motion verbs in our application.

To explore this question, we will compare three strategies for Word Sense Disambiguation. The first strategy is no disambiguation at all (NoWSD). It classifies a verb instance as a motion verb if any sense of the lemma is in the resource-derived list of motion verbs (cf. above). The second strategy, WSD, classifies an instance as a motion verb if its UKB-assigned synset is in the list of motion verbs. The third strategy, PredomSense, leverages the observation that WSD has a hard time beating the *predominant sense heuristic* (McCarthy et al., 2004) which assigns the predominant, or first, sense to all instances. Here, this strategy means that we treat all verbs as motion verbs whose first sense, according to WordNet, is in the list of motion verbs.

4.3 Motion verbs in FrameNet

Defining motion verbs based on frames. This approach builds directly on the intuition developed in Section 2.4: we characterize the motion domain through a set of motion frames recognizable by the

	Precision	Recall	F_1-Score
NoWSD	43	**87**	58
WSD	68	48	56
PredomSense	**73**	62	**67**

Table 3: Motion verb recognition results with FrameNet

	Precision	Recall	F_1-Score
NoWSD	38	**55**	45
WSD	83	18	30
PredomSense	**87**	32	**47**

Table 4: Motion verb recognition results with WordNet

use of location roles. The resulting list was shown in Table 1. We created an initial list of motion verbs by listing all verbal lexical units for these frames. However, as discussed in Section 3.3, this list is far from complete. A second problem is that the WSD labels assigned by UKB are WordNet synsets. Both problems can be solved together by mapping map the FrameNet lexical units onto WordNet. We used the mapping constructed by Shi and Mihalcea (2005) to determine the matching synsets for each lexical unit[2]. In order to improve coverage, we added all hyponyms of these synsets. This corresponds to the assumption that any hyponyms of a lexical unit l can evoke the same frame as l. The resulting set comprises 2838 synsets.

Results. Table 3 shows the results for the three strategies defined in Section 4.2. NoWSD (row 1) shows that without disambiguation, 87% of the motion verbs are detected, but the precision is only 43%. The 13% false negatives are either cases of missing frames, or of missing lexical units in FrameNet and WordNet (phrasal verbs). The 57% false positives are due to ambiguous words with non-motion senses. Using WSD (row 2) substantially improves precision, but hurts recall, with a small loss in F-Score. Finally, the predominant sense heuristic outperforms WSD in both precision and recall. It cannot rival NoWSD in recall, but the higher precision yields a net gain in F-Score of 9%, the overall best results.

These numbers show that when the first sense of a verb is a motion sense, then heuristic assumption that instances of this verb belong to the motion domain outperform the UKB-provided disambiguation. (Note however that UKB tries to solve a more difficult task, namely fine-grained sense assignment.) The heuristic is nevertheless far from perfect: Among the 27% false positives, there are high-frequency high-ambiguity verbs like *take*, but we also find that many motion verbs specifically have a concrete motion sense but also a more abstract non-motion sense, often in the mental or cognitive domain. Examples are *struggle, lean, confront, follow*.

[2]Newer mappings have been created by, e.g., Tonelli and Pighin (2009).

4.4 Motion Verbs in WordNet

While FrameNet is a comprehensive resource for defining motion verbs, it is available only for a small number of languages and suffers from coverage problems. We therefore also experimented with using only WordNet, which is available for a large number of languages.

Defining motion verbs in the WordNet hierarchy. WordNet uses a troponymy relation in the verbal domain which organizes verb senses into hierarchy structured by specificity. For example, *talk* is a troponym of *communicate*; in turn, *whisper* is a troponym of *talk*. Within this hierarchical organization, we can define the motion domain in WordNet as a (small) set of subtrees by identifying the root of these subtrees. This is similar to how, for example, semantic classes for selectional preferences are often represented in terms of WordNet subtrees (Resnik, 1996). The challenge is to find a set of nodes whose subtrees cover as much as possible of the motion domain while avoiding overgeneration. An inspection of WordNet led us to two nodes that meet these conditions well. The first one is "move, locomote, travel, go" (synset ID 01818343-V), which covers the motion domain. The second one is "to be (occupy a certain position or area; be somewhere)" (synset ID 02629830-V), which covers the orientation and position domains. The WordNet motion list formed by these nodes and their hyponyms comprises a total of 1090 synsets.

Results. The results for the three strategies from Section 4.2 applied to WordNet are shown in Table 4. Generally, we observe that WordNet, works considerably worse than FrameNet. This is not surprising, given the additional layer of information that FrameNet provides, and the simplistic motion domain model we adopted for WordNet.

WordNot notably suffers from low recall. Even in the NoWSD condition, many verbs that are motion verbs are not included in the WordNet-derived list of motion verbs, the recall being only 55%. The reason appears to be that a number of motion verbs in particular are scattered in WordNet outside our two chosen subtrees. Examples include *crouch*, a hyponym of the synset *to change (to undergo a change)*, and *stand*, a hyponym of the synset *to be (have the quality of being)*. However, these motion verbs are not easy to assign to complete subtrees that can also be designated as motion subtrees.

Not surprisingly, NoWSD has by far the lowest precision of the three conditions, since many instances of verbs that have motion senses but also other senses are mistagged as motion verbs. The precision improves dramatically for the WSD condition, from 38% to 83%. However, the recall takes a further major hit down to 18%, and thus the resulting F-Score is very low. In the PredomSense condition, precision increases even somewhat further, and the decline in recall compared to NoWSD is not quite as pronounced. Consequently, the PredomSense condition shows the overall best F-Score for WordNet-based models. For both FrameNet- and WordNet-based models, therefore, it seems currently preferable to employ a predominant sense heuristic over performing full-fledged word sense disambiguation. The best WordNet-based result, 47% F-Score, is however still 20% below the best FrameNet-based result of 67%. An example for an instance that is wrongly classified as a motion verb even in the high-precision PredomSense condition is *to follow*, whose first WordNet sense concerns motion, but which was used in the corpus for “obeying” (following commands).

5 Discussion and Conclusion

In this paper, we introduced the task of locational inference and sketched a processing strategy. We presented a corpus study which provided evidence for the relevance of locational inference and performed an experiment which compared WordNet and FrameNet for the recognition of the motion domain (motion, orientation, and position verbs). We found FrameNet to be a useful tool to define these semantic domains. Processing can also proceed when just WordNet is available, but unsurprisingly with lower results. Comparing different word sense disambiguation schemes, the unsupervised WSD system UKB could not beat the simple "predominant sense heuristic".

Concerning the automatic recognition of motion verbs, one avenue of future research is the refinement of the characterization of motion verbs in FrameNet and WordNet. As we have observed, the location role-based collection of motion frames misses some frames which should be added to the list of frames. As for WordNet, our current definition which limits itself to two subtrees in the WordNet verb hierarchy leads to a very high precision but a low recall. Since WordNet was not designed specifically with the motion domain in mind, many other motion verbs are scattered throughout the verb hierarchy, and their distribution is difficult to describe succinctly. We will experiment with the precision/recall trade-off that arises from adding more synsets to the definition of motion verbs.

In terms of annotation, we also made, but not yet acted upon, the observation that many locations mentioned in a discourse are related hierarchically. For example, a person can be concurrently said to be in a seat, in the carriage and on the road, all of which fill the PLACE role with varying degrees of specificity. These descriptions can be said to be related through bridging.

Finally, we plan to bring these threads together in the realization of a complete pipeline for locational inference (cf. Figure 1). Our current experiments only concern the very first step of this pipeline, and the recovery and assignment of location roles that remain locally unrealized remains the ultimate goal of our research program.

Acknowledgments. This work was partially supported by the EC-funded project EXCITEMENT (FP7 ICT-287923) and the Research Training Group "Coherence in language processing" of Heidelberg University.

References

Eneko Agirre and Aitor Soroa. 2009. Personalizing PageRank for word sense disambiguation. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, pages 33–41, Athens, Greece.

Collin F Baker, Charles J Fillmore, and John B Lowe. 1998. The Berkeley FrameNet project. In *Proceedings of the joint Annual Meeting of the Association for Computational Linguistics and International Conference on Computational Linguistics*, pages 86–90, Montreal, QC.

Aljoscha Burchardt, Marco Pennacchiotti, Stefan Thater, and Manfred Pinkal. 2009. Assessing the Impact of Frame Semantics on Textual Entailment. *Journal of Natural Language Engineering*, 15(4):527–550.

Ido Dagan, Bill Dolan, Bernardo Magnini, and Dan Roth. 2009. Recognizing textual entailment: Rational, evaluation and approaches. *Natural Language Engineering*, 15(4):i–xvii.

Katrin Erk. 2005. Frame Assignment as Word Sense Disambiguation. In *Proceedings of the 6th International Workshop on Computational Semantics*, Tilburg, The Netherlands.

Charles J Fillmore. 1985. Frames and the Semantics of Understanding. *Quaderni di Semantica*, IV(2):222–254.

Charles J Fillmore. 1986. Pragmatically controlled zero anaphora. In V Nikiforidou, M VanClay, M Niepokuj, and D Feder, editors, *Proceedings of the Berkeley Linguistics Society*, volume 12, pages 95–107.

Matthew Gerber and Joyce Y Chai. 2010. Beyond nombank : A study of implicit arguments for nominal predicates. In *Proceedings of ACL*, pages 1583–1592, Uppsala, Sweden.

Daniel Gildea and Daniel Jurafsky. 2002. Automatic labeling of semantic roles. *Computational Linguistics*, 28(3):245–288.

Mark A Greenwood. 2004. Using Pertainyms to Improve Passage Retrieval for Questions Requesting Information About a Location. In *Proceedings of the SIGIR Workshop on Information Retrieval for Question Answering (IR4QA)*, pages 17–22.

Paul Grice. 1975. Logic and conversation. In P Cole and J Morgan, editors, *Syntax And Semantics*, pages 1–13. Academic Press, New York.

Sanda Harabagiu, Andrew Hickl, and Finley Lacatusu. 2007. Satisfying information needs with multi-document summaries. *Information Processing and Management*, 43(6):1619–1642.

David Herman. 2001. Spatial reference in narrative domains. *Text*, 4:515–541.

Blake Stephen Howald. 2010. Linguistic spatial classifications of event domains in narratives of crime. *Journal of Spatial Information Science*, 1(1):75–93.

Nancy Ide, Collin F. Baker, Christiane Fellbaum, Charles Fillmore, and Rebecca Passonneau. 2008. MASC: The Manually Annotated Sub-Corpus of American English. In *Proceedings of LREC 2008*, pages 2455–2461, Marrakech, Morocco.

Ray S Jackendoff. 1990. *Semantic Structures*. The MIT Press, Cambridge, MA.

George Lakoff and Mark Johnson. 1980. *Metaphors we live by*. University of Chicago Press, Chicago, IL.

Jochen L Leidner, Gail Sinclair, and Bonnie Webber. 2003. Grounding spatial named entities for information extraction and question answering. In *Proceedings of the HLT/NAACL 2003 workshop on the analysis of geographic references*, pages 31–38, Edmonton, Canada.

Beth Levin. 1993. *English Verb Classes and Alternations*. University of Chicago Press, Chicago, IL.

Diana McCarthy, Rob Koeling, Julie Weeds, and John Carroll. 2004. Finding Predominant Word Senses in Untagged Text. In *Proceedings of the 42th Annual Meeting of the Association for Computational Linguistics*, pages 279–286, Barcelona, Spain.

George Miller, Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine Miller. 1990. Introduction to WordNet: An On-Line Lexical Database. *International Journal of Lexicography*, 3(4):235–244.

Shachar Mirkin, Ido Dagan, and Sebastian Padó. 2010. Assessing the Role of Discourse References in Entailment Inference. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 1209–1219, Uppsala, Sweden.

Christoph Müller and Michael Strube. 2006. Multi-level annotation of linguistic data with MMAX2. In Sabine Braun, Kurt Kohn, and Joybrato Mukherjee, editors, *Corpus Technology and Language Pedagogy New Resources New Tools New Methods*, volume 3 of *English Corpus Linguistics*, pages 197–214. Peter Lang.

Peter Norvig. 1987. Inference in Text Understanding. In *Proceedings of the National Conference on Artificial Intelligence*, pages 561–565, Seattle, WA.

Anselmo Peñas, Álvaro Rodrigo, Valentı̀n Sama, and Felisa Verdejo. 2008. Testing the Reasoning for Question Answering Validation. *Journal of Logic and Computation*, 18:459–474.

Philip Resnik. 1996. Selectional Constraints: An Information-Theoretic Model and its Computational Realization. *Cognition*, 61:127–159.

Lorenza Romano, Milen Kouylekov, Idan Szpektor, Ido Dagan, and Alberto Lavelli. 2006. Investigating a Generic Paraphrase-Based Approach for Relation

Extraction. In *Proceedings of the 11th Conference of the European Chapter of the ACL*, pages 401–408, Trento, Italy.

Josef Ruppenhofer, Caroline Sporleder, R. Morante, Collin Baker, and Martha Palmer. 2010. Semeval-2010 task 10: Linking events and their participants in discourse. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, pages 45–50, Uppsala, Sweden.

Helmut Schmid. 1994. Probabilistic Part-of-Speech Tagging Using Decision Trees. In *Proceedings of the International Conference on New Methods in Language Processing*, pages 44–49, Manchester, UK.

Lei Shi and Rada Mihalcea. 2005. Putting Pieces Together: Combining FrameNet, VerbNet and WordNet for Robust Semantic Parsing. In *Proceedings of the 6th CICLing*, pages 100–111. Springer.

Carina Silberer and Anette Frank. 2012. Casting implicit role linking as an anaphora resolution task. In **SEM 2012: The First Joint Conference on Lexical and Computational Semantics – Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation (SemEval 2012)*, pages 1–10, Montréal, Canada, 7-8 June. Association for Computational Linguistics.

Wee Meng Soon, Hwee Tou Ng, and Daniel Chung Yong Lim. 2001. A Machine Learning Approach to Coreference Resolution of Noun Phrases. *Computational Linguistics*, 27(4):521–544.

Sara Tonelli and Daniele Pighin. 2009. New features for FrameNet–WordNet mapping. In *Proceedings of the Thirteenth Conference on Computational Natural Language Learning*, Boulder, CO, USA.