

# Desperately Seeking Implicit Arguments in Text

**Sara Tonelli**

Fondazione Bruno Kessler / Trento, Italy  
satonelli@fbk.eu

**Rodolfo Delmonte**

Universit Ca' Foscari / Venezia, Italy  
delmont@unive.it

## Abstract

In this paper, we address the issue of automatically identifying null instantiated arguments in text. We refer to Fillmore's theory of pragmatically controlled zero anaphora (Fillmore, 1986), which accounts for the phenomenon of omissible arguments using a *lexically-based* approach, and we propose a strategy for identifying implicit arguments in a text and finding their antecedents, given the overtly expressed semantic roles in the form of frame elements. To this purpose, we primarily rely on linguistic knowledge enriched with role frequency information collected from a training corpus. We evaluate our approach using the test set developed for the SemEval task 10 and we highlight some issues of our approach. Besides, we also point out some open problems related to the task definition and to the general phenomenon of null instantiated arguments, which needs to be better investigated and described in order to be captured from a computational point of view.

## 1 Introduction

In natural language, lexically unexpressed linguistic items are very frequent and indirectly weaken any attempt at computing the meaning of a text or discourse. However, the need to address semantic interpretation is strongly felt in current advanced NLP tasks, in particular, the issue of transforming a text or discourse into a set of explicitly interconnected predicate-argument/adjunct structures (hence PAS). The aim of this task would be to unambiguously identify events and participants and their associa-

tion to spatio-temporal locations. However, in order to do that, symbolic and statistical approaches should be based on the output representation of a deep parser, which is currently almost never the case. Current NLP technologies usually address the surface level linguistic information with good approximation in dependency or constituency structures, but miss implicit entities (IEs) altogether. The difficulties to deal with lexically unexpressed items or implicit entities are related on the one hand to recall problems, i.e. the problem of deciding whether an item is implicit or not, and on the other hand to precision problems, i.e. if an implicit entity is accessible to the reader from the discourse or its context, an appropriate antecedent has to be found. However, a system able to derive the presence of IEs may be a determining factor in improving performance of QA systems and, in general, in Informations Retrieval and Extraction systems.

The current computational scene has witnessed an increased interest in the creation and use of semantically annotated computational lexica and their associated annotated corpora, like PropBank (Palmer et al., 2005), FrameNet (Baker et al., 1998) and NomBank (Meyers, 2007), where the proposed annotation scheme has been applied in real contexts. In all these cases, what has been addressed is a basic semantic issue, i.e. labeling PAS associated to semantic predicates like adjectives, verbs and nouns. However, what these corpora have not made available is information related to IEs. For example, in the case of eventive deverbal nominals, information about the subject/object of the nominal predicate is often implicit and has to be understood from the previous

discourse or text, e.g. “*the development of a prototype [→ implicit subject]*”. As reported by Gerber and Chai (2010), introducing implicit arguments to nominal predicates in NomBank would increase the resource coverage of 65%.

Other IEs can be found in agentless passive constructions ( e.g. “*Our little problem will soon be solved ∅ [→ unexpressed Agent ]*”<sup>1</sup>) or as unexpressed arguments such as addressee with verbs of commitment, for example “*I can promise ∅ that one of you will be troubled [→ unexpressed Addressee]*” and “*I dare swear ∅ that before tomorrow night he will be fluttering in our net [→ unexpressed Addressee]*”.

In this paper we discuss the issues related to the identification of implicit entities in text, focussing in particular on omissions of core arguments of predicates. We investigate the topic from the perspective proposed by (Fillmore, 1986) and base our observations on null instantiated arguments annotated for the SemEval 2010 Task 10, ‘Linking Events and Their Participants in Discourse’ (Ruppenhofer et al., 2010)<sup>2</sup>. The paper is structured as follows: in Section 2 we detail the task of identifying null instantiated arguments from a theoretical perspective and describe related work. In Section 3 we briefly introduce the SemEval task 10 for identifying implicit arguments in text, while in Section 4 we detail our proposal for NI identification and binding. In Section 5 we give a thorough description of the types of null instantiations annotated in the SemEval dataset and we explain the behaviour of our algorithm w.r.t. such cases. We also compare our results with the output of the systems participating to the SemEval task. Finally, we draw some conclusions in Section 6.

## 2 Related work

In this work, we focus on *null complements*, also called pragmatically controlled zero anaphora (Fillmore, 1986), understood arguments or linguistically

<sup>1</sup>Unless otherwise specified, the following examples are taken from the datasets made available in the SemEval 2010 task ‘Linking Events and Their Participants in Discourse’. Some of them have been slightly simplified for purposes of exposition.

<sup>2</sup><http://semeval12.fbk.eu/semeval12.php?location=tasks&taskid=9>

unrealized arguments. We focus on Fillmore’s theory because his approach represents the backbone of the FrameNet project, which in turn inspired the SemEval task we will describe below. Fillmore (1986) shows that in English and many other languages some verbs allow null complements and some others don’t. The latter require that, when they appear in a sentence, all core semantic roles related to the predicate are expressed. For example, sentences like “*Mary locked \*\**” or “*John guaranteed \*\**” are not grammatically well-formed, because they both require two mandatory linguistically inherent participants. Fillmore tries to explain why semantic roles can sometimes be left unspoken and what constraints help the interpreter recover the missing roles. He introduces different factors that can influence the licensing of null complements. These can be *lexically-based*, (semantically close predicates like ‘promise’ and ‘guarantee’ can license the omission of the theme argument in different cases), motivated by the *interpretation of the predicate* (“*I was eating ∅*” licenses a null object because it has an existential interpretation) and depending on the context (see for example the use of *impress* in an episodic context like “*She impressed the audience*”, where the null complement is not allowed, compared to “*She impresses ∅ every time*” in habitual interpretation; examples from Ruppenhofer and Michaelis (2009)).

The fact that Fillmore explains the phenomenon of omissible arguments with a lexically-based approach implies that from his perspective neither a purely pragmatic nor a purely semantic approach can account for the behaviour of omissible arguments. For example, he argues that some verbs, such as *to lock* will never license a null complement, no matter in which pragmatic context they are used. Besides, there are synonymous verbs which behave differently as regards null complementation, which Fillmore sees as evidence against a purely semantic explanation of implicit arguments.

Another relevant distinction drawn in Fillmore (1986) is the typology of omitted arguments, which depends on the type of licensor and on the interpretation of the null complement. Fillmore claims that with some verbs the missing complement can be retrieved from the context, i.e. it is possible to find a referent previously mentioned in the text / discourse

and bearing a definite, precise meaning. These cases are labeled as *definite null complements* or *instantiations* (DNI) and are *lexically specific* in that they apply only to some predicates. We report an example of DNI in (1), taken from the SemEval task 10 dataset (see Section 3). The predicate ‘*waiting*’ has an omitted object, which we understand from the discourse context to refer to ‘*I*’.

- (1) I saw him rejoin his guest, and *I* crept quietly back to where my companions were *waiting*  $\emptyset$  to tell them what I had seen.

DNIs can also occur with nominal predicates, as reported in (2), where the person having a *thought*, *the baronet*, is mentioned in the preceding sentence:

- (2) Stapleton was talking with animation, but *the baronet* looked pale and distraught. Perhaps the *thought* of that lonely walk across the ill-omened moor was weighing heavily upon his mind.

In contrast to DNIs, Fillmore claims that with some verbs and in some interpretations, a core argument can be omitted without having a referent expressing the meaning of the null argument. The identity of the missing argument can be left unknown or indefinite. These cases are labeled as *indefinite null complements* or *instantiations* (INI) and are *constructionally licensed* in that they apply to any predicate in a particular grammatical construction. See for example the following cases, where the omission of the agent is licensed by the passive construction:

- (3) One of them was suddenly *shut off*  $\emptyset$ .
- (4) I am *reckoned* fleet of foot  $\emptyset$ .

Cases of INI were annotated by the organizers of the SemEval task 10 also with nominal predicates, as shown in the example below, where the perceiver of the *odour* is left unspecified:

- (5) Rank reeds and lush, slimy water-plants sent an *odour*  $\emptyset$  of decay and a heavy miasmatic vapour.

Few attempts have been done so far to automatically deal with the recovery of implicit information

in text. One of the earliest systems for identifying extra-sentential arguments is PUNDIT by Palmer et al. (1986). This Prolog-based system comprises a syntactic component for parsing, a semantic component, which decomposes predicates into component meanings and fills their semantic roles with syntactic constituents based on a domain-specific model, and a reference resolution component, which is called both for explicit constituents and for obligatory implicit constituents. The reference resolution process is based on a focus list with all potential pronominal referents identified by the semantic component. The approach, however, has not been evaluated on a dataset, so we cannot directly compare its performance with other approaches. Furthermore, it is strongly domain-dependent.

Burchardt et al. (2005) propose in a case study to identify implicit arguments exploiting contextual relations from deep-parsing and lexico-semantic frame relations encoded in FrameNet. In particular, they suggest converting a text into a network of lexico-semantic predicate-argument relations connected through frame-to-frame relations and recurrent anaphoric linking patterns. However, the authors do not implement and evaluate this approach.

Most recently, Gerber and Chai (2010) have presented a supervised classification model for the recovery of implicit arguments of nominal predicates in NomBank. The model features are quite different from those usually considered in standard SRL tasks and include among others information from VerbNet classes, pointwise mutual information between semantic arguments, collocation and frequency information about the predicates, information about parent nodes and siblings of the predicates and discourse information. The authors show the feasibility of their approach, which however relies on a selected group of nominal predicates with a large number of annotated instances.

The first attempt to evaluate implicit argument identification over a common test set and considering different kinds of predicates was made by Ruppenhofer et al. (2010). Further details are given in the following section.

Data set	Sentences	Frame inst.	Frame types	Overt FEs	DNIs (resolved)	INIs
Train	438	1,370	317	2,526	303 (245)	277
Test	525	1,703	452	3,141	349 (259)	361

Table 1: Data set statistics from SemEval task 10

### 3 SemEval 2010 task 10

The SemEval-2010 task for linking events and their participants in discourse (Ruppenhofer et al., 2010) introduced a new issue w.r.t. the SemEval-2007 task ‘Frame Semantic Structure Extraction’ (Baker et al., 2007), in that it focused on linking local semantic argument structures across sentence boundaries. Specifically, the task included first the identification of frames and frame elements in a text following the FrameNet paradigm (Baker et al., 1998), then the identification of locally uninstantiated roles (NIs). If these roles are indefinite (INI), they have to be marked as such and no antecedent has to be found. On the contrary, if they are definite (DNI), their coreferents have to be found in the wider discourse context. The challenge comprised two tasks, namely the full task (semantic role recognition and labelling + NI linking) and the NIs only task, i.e. the identification of null instantiations and their referents given a test set with gold standard local semantic argument structure. In this work, we focus on the latter task.

The data provided to the participants included a training and a test set. The training data comprised 438 sentences from Arthur Conan Doyle’s novel ‘The Adventure of Wisteria Lodge’, manually annotated with frame and INI/DNI information. The test set included 2 chapters of the Sherlock Holmes story ‘The Hound of Baskervilles’ with a total of 525 sentences, provided with gold standard frame information. The participants had to *i*) assess if a local argument is implicit; *ii*) decide whether it is an INI or a DNI and *iii*) in the second case, find the antecedent of the implicit argument. We report in Table 1 some statistics about the provided data sets from Ruppenhofer et al. (2010). Note that overt FEs are the explicit frame elements annotated in the data set.

Although 26 teams downloaded the data sets, there were only two submissions, probably depending on the intrinsic difficulties of the task (see dis-

cussion in Section 5). The best performing system (Chen et al., 2010) is based on a supervised learning approach using, among others, distributional semantic similarity between the heads of candidate referents and role fillers in the training data, but its performance is strongly affected by data sparseness. Indeed, only 438 sentences with annotated NIs were made available in the training set, which is clearly insufficient to capture such a multifaceted phenomenon with a supervised approach. The second system participating in the task (Tonelli and Delmonte, 2010) was an adaptation of an existing LFG-based system for deep semantic analysis, whose output was mapped to FrameNet-style annotation. In this case, the major challenge was to cope with the classification of some NI phenomena which are very much dependent on frame specific information, and can hardly be generalized in the LFG framework.

### 4 A linguistically motivated proposal for NI identification and binding

In this section, we describe our proposal for dealing with INI/DNI identification and evaluate our output against SemEval gold standard data. As discussed in the previous section, existing systems dealing with this task suffer on the one hand from a lack of training data and on the other hand from the dependence of the task on frame annotation, which makes it difficult to adapt existing unsupervised approaches. We show that, given this state of the art, better results can be achieved in the task by simply developing an algorithm that reflects as much as possible the linguistic motivations behind NI identification in the FrameNet paradigm. Our approach is divided into two subtasks: *i*) identify INIs/DNIs and *ii*) for each DNI, find the corresponding referent in text.

We develop an algorithm that incorporates the following linguistic information:

**FE coreness status** Null instantiated arguments as defined in FrameNet are always *core* arguments, i.e.

they are central to the meaning of a frame. Since the coreness status of the arguments is encoded in FrameNet, we limit our search for an NI only if a core frame element is not overtly expressed in the text.

**Incorporated FEs** Although all lexical units belonging to the same frame in the FrameNet database are characterised by the same set of core FEs, a further distinction should be introduced when dealing with NIs identification. For example, in PERCEPTION\_ACTIVE, several predicates are listed, which however have a different behaviour w.r.t. the core *Body\_part* FE. ‘Feel.v’, for instance, is underspecified as regards the body part perceiving the sensation, so we can assume that when it is not overtly expressed, we have a case of null instantiation. For other verbs in the same frame, such as ‘glance.v’ or ‘listen.v’, the coreness status of *Body\_part* seems to be more questionable, because the perceiving organ is already implied by the verb meaning. For this reason, we argue that if *Body\_part* is not expressed with ‘glance.v’ or ‘listen.v’, it is not a case of null instantiation. Such FEs are defined as *incorporated* in the lexical unit and are encoded as such in FrameNet.

**Excludes and Includes relation** In FrameNet, some information about the relationship between certain FEs is encoded. In particular, some FEs are connected by the *Excludes* relation, which means that they cannot occur together, and others by the *Requires* relation, which means that if a given FE is present, then also the other must be overtly or implicitly present. An example of *Excludes* is the relationship between the FE *Entity\_1 / Entity\_2* and *Entities*, because if *Entity\_1* and *Entity\_2* are both present in a sentence, then *Entities* cannot be co-present. Conversely, *Entity\_1* and *Entity\_2* stand in a *Requires* relationship, because the first cannot occur without the second. This kind of information can clearly be helpful in case we have to automatically decide whether an argument is implicit or is just not present because it is not required.

**INI/DNI preference** Ruppenhofer and Michaelis (2009) suggest that omissibile arguments in particular frames tend to be always interpreted as definite or indefinite. For example, they report that in a sample from the British National Corpus, the interpretation

for a null instantiated *Goal* argument is definite in 97.5% of the observed cases. We take this feature into account by considering the frequency of an implicit argument being annotated as definite/indefinite in the training set.

The algorithm incorporating all this linguistic information is detailed in the following subsection.

#### 4.1 INI/DNI identification

In a preliminary step, we collect for each frame the list of arguments being annotated as DNI/INI with the corresponding frequency in the training set. For example, in the CALENDRIC\_UNIT frame, the *Whole* argument has been annotated 11 times as INI and 5 times as DNI. Some implicit frame elements occur only as INI or DNI, for example *Goal*, which is annotated 14 times as DNI and never as INI in the ARRIVING frame. This frequency list (*FreqList*) is collected in order to decide if candidate null instantiations have to be classified as DNI or INI.

We consider each sentence in the test data provided with FrameNet annotation, and for each predicate  $p$  annotated with a set of overt frame elements *FEs*, we run the first module for DNI/INI identification. The steps followed are reported in Algorithm 1. We first check if the annotated *FEs* contain all core frame elements  $C$  listed in FrameNet for  $p$ . If the two sets are identical, we conclude that no core frame element can be implicit and we return an empty set both for *DNI* and *INI*. For example, in the test sentence (6), the BODY\_MOVEMENT frame appears in the sentence with its two core frame elements, i.e. *Body\_part* and *Agent*. Therefore, no implicit argument can be postulated.

(6) Finally [she]<sub>Agent</sub> opened<sub>BODY\_MOVEMENT</sub> [her eyes]<sub>Body\_part</sub> again.

If the core FEs in  $C$  are not all overtly expressed in *FEs*, we run two routines to check if the missing FEs *CandNIs* are likely to be null instantiated elements. First, we discard all candidate NIs that appear as incorporated FEs for the given  $p$ . Second, we discard as well candidate NIs if they are excluded by the overtly annotated FEs.

The last steps of the algorithm are devoted to deciding if the candidate null instantiation is definite or indefinite. For this step, we rely on the observations collected in *FreqList*. In particular, for each

candidate  $c$  we check if it was already present as INI or DNI in the training set. If yes, we label  $c$  accordingly. In case  $c$  was observed both as INI and as DNI, the most probable label is assigned based on its frequency in the training set.

**Input:**  $TestSet$  with annotated core  $FEs$ ;  
 $FreqList$   
**Output:**  $INI$  and  $DNI$  for  $p$   
**foreach**  $p \in TestSet$  **do**  
  extract annotated core  $FEs$ ;  
  extract set  $C$  of core  $FEs$  for  $p$  in FrameNet;  
  **if**  $C \subseteq FEs$  **then**  
     $DNI = \emptyset$ ;  
     $INI = \emptyset$ ;  
  **else**  
     $C \setminus FEs = CandNIs$ ;  
    **foreach**  $c \in CandNIs$  **do**  
      **if**  $c$  is incorporated  $FE$  of  $p$  **then**  
        delete  $c$   
      **foreach**  $fe \in FEs$  **do**  
        **if**  $fe$  excludes  $c$  **then**  
          delete  $c$   
      **end**  
      **foreach**  $ni_p \in FreqList_p$  **do**  
        **if**  $c = ni_p$  **then**  
          **if**  $ni_p$  is only  $dni_p$  **then**  
             $c \in DNI$   
          **if**  $ni_p$  is only  $ini_p$  **then**  
             $c \in INI$   
          **if**  $ni_p$  is  $ini_p$  and  $ni_p$  is  $dni_p$   
          **then**  
            **if**  $Freq(ini_p) >$   
             $Freq(dni_p)$  **then**  
               $c \in INI$   
            **else**  
               $c \in DNI$   
          **end**  
      **end**  
    **end**  
  **end**  
   $return(INI)$ ;  
   $return(DNI)$ ;  
**end**  
**Algorithm 1:** DNI/INI identification

## 4.2 DNI binding

Given that both the supervised approach exploited by Chen et al. (2010) and the methodology proposed in Tonelli and Delmonte (2010) based on

deep-semantic parsing achieved quite poor results in the DNI-binding task, we devise a third approach that relies on the observed heads of each  $FE$  in the training set and assigns a relevance score to each candidate antecedent.

We first collect for each  $FE$  the list of heads  $H_{train}$  assigned to  $FE$  in the training set, and we extract for each head  $h_{train} \in H_{train}$  the corresponding frequency  $f_{h_{train}}$ . Then, for each  $dni \in DNI$  identified with Algorithm 1 in the test set, we collect all nominal heads  $H_{test}$  occurring in a window of (plus/minus) 5 sentences and we assign to each candidate head  $h_{test} \in H_{test}$  a relevance score  $rel_{h_{test}}$  w.r.t.  $dni$  computed as follows:

$$rel_{h_{test}} = \frac{f_{h_{train}}}{dist(sent_{dni}, sent_{h_{test}})} \quad (7)$$

where  $f_{h_{train}}$  is the number of times  $h$  has been observed in the training set with a  $FE$  label, and  $dist(sent_{dni}, sent_{h_{test}})$  is the distance between the sentence where the  $dni$  has been detected and the sentence where the candidate head  $h_{test}$  occurs ( $0 \leq dist(sent_{dni}, sent_{h_{test}}) \leq 5$ ).

The best candidate head for  $dni$  is the one with the highest  $rel_{h_{test}}$ , given that it is (higher) than 0. The way we compute the relevance score is based on the intuition that, if a head was frequently observed for  $FE$  in the training set, it is likely that it is a good candidate. However, the more distant it occurs from  $dni$ , then less probable it is as antecedent.

## 5 Evaluation and error analysis

We present here an evaluation of the system output on test data. We further comment on some difficult aspects of the task and suggest some solutions.

### 5.1 Results

Evaluation consists of different layers, which we consider separately. The first task was to decide whether an argument is implicit or not. We were able to identify 53.8% of all null instantiated arguments in text, which is lower than the recall of 63.4% achieved bySEMAFOR (Chen et al., 2010), the best performing system in the challenge. However, in the following subtask of deciding whether an implicit argument is an INI or a DNI, we achieved an accuracy of 74.6% (vs. 54.7% of SEMAFOR,

even if our result is based on fewer proposed classifications). Note that the majority-class accuracy reported by Ruppenhofer et al. (2010) is 50.8%.

In Table 2 we further report precision, recall and F1 scores computed separately on all DNIs and all INIs automatically detected. Precision corresponds to the percentage of null instantiations found (either INI or DNI) that are correctly labelled as such, while recall indicates the amount of INI or DNI that were correctly identified compared to the gold standard ones. Our approach does not show significant differences between the result obtained with INIs and DNIs, while the evaluation of SEMAFOR (between parenthesis) shows that its performance suffers from low recall as regards DNIs and low precision as regards INIs.

	P	R	F1
DNI	0.39 (0.57)	0.43 (0.03)	0.41 (0.06)
INI	0.46 (0.20)	0.38 (0.61)	0.42 (0.30)

Table 2: Evaluation of INI/DNI identification. SEMAFOR performance between parenthesis.

Another evaluation step concerns the binding of DNIs with the corresponding antecedents by applying the equation reported in Section 4.2. Results are shown in Table 3:

	P	R	F1
DNI	0.13 (0.25)	0.06 (0.01)	0.08 (0.02)

Table 3: Evaluation of DNI resolution. SEMAFOR performance between parenthesis.

Although the binding quality still needs to be improved, two main factors have a negative impact on our performance, which do not depend on our algorithm: first, 9% of the DNIs we bound to an antecedent don't have a referent in the gold standard. Second, 26% of the wrong assignments concern antecedents found for the *Topic* frame element in test sentences where the *STATEMENT* frame has been annotated together with the overtly expressed core FE *Message*. In all these gold cases, *Topic* is not considered null instantiated if the *Message* FE is explicit in the clause. Therefore, we can conclude that the mistake done by our algorithm depends on the missing *Excludes* relation between *Topic* and *Mes-*

*sage*, i.e. a rule should be introduced saying that one of the two roles is redundant (and not null instantiated) if the other is overtly expressed.

## 5.2 Open issues related to our approach

Even if with a small set of rules our approach achieved state-of-the-art results in the SemEval task, our performance clearly requires further improvements. Indeed, we currently rely only on the background knowledge about core FEs from FrameNet, combined with statistical observations about role fillers acquired from the training set. Additional morphological, syntactic, semantic and discourse information could be exploited in different ways. For example, since the passive voice of a verb can constructionally license INIs, this kind of information would greatly improve our performance with verbal predicates (i.e. 46% of all annotated predicates in the test set).

As for nominal predicates, consider for example sentence (8) extracted from the test set:

- (8) ‘Excuse the admiration<sub>JUDGMENT</sub> [of a connoisseur]<sub>Evaluee</sub>,’ said [he]<sub>Cognizer</sub>.

In this case, ‘admiration’ is a nominal predicate with two explicit FEs, namely *Evaluee* and *Cognizer*. The *JUDGMENT* frame includes also the *Reason* core FE, which can be a candidate for a null instantiation. In fact, it is annotated as INI in the gold standard data, because in the previous sentences a reason for such admiration is not mentioned. However, this could have been annotated as DNI as well, if only some specific quality of the person had been previously introduced. This shows that the current sentence does not present any inherent characteristic motivating the presence of a definite instantiation. In this case, a strategy based on some kind of history list may be very helpful. This could contain, for example, all subjects and direct objects previously mentioned in text and selected according to some relevance criteria, as in (Tonelli and Delmonte, 2010). A further improvement may derive from the integration of an anaphora resolution step, as first proposed by Palmer et al. (1986) and more recently by Gerber and Chai (2010).

### 5.3 Open issues related to the task

Other open issues are related to the specification of the task and to the nature of implicit entities, which make it difficult to account for this phenomenon from a computational point of view. We report below the main issues that need to be tackled:

**INI Linking:** Table 1 shows that 28% of DNIs in the test set are not linked to any referent. This puts into question one of the main assumptions of the task, that is the connection between a definite instantiation and a referent. In the test set, there are also 14 cases of indefinite null instantiations (out of 361) that are provided with a referent. Consider for example the following sentence with gold standard annotation, in which the INI label *Path* is actually instantiated and refers to ‘we’:

- (9) (We)<sub>Path</sub> allowed [him]<sub>Theme</sub> to pass<sub>TRAVERSING</sub> before we had recovered our nerve.

This again may be a controversial annotation choice, since the annotation guidelines of the task reported that ‘*in cases of indefinite omission, there need not be any overt mention of an indefinite NP in the linguistic context nor does there have to be a referent of the kind denoted by the omitted argument in the physical discourse setting*’ (Ruppenhofer, 2010).

**Position of referent:** Although we suggested that the History List may represent a good starting point for finding antecedents to DNIs, searching only in the context *preceding* the current predicate is not enough because the referent can occur *after* such predicate. Also, the predicate with a DNI and the referent can be divided by a very large text span. In the test data, 38% of the DNIs referent occur in the same sentence of the predicate, while 14% are mentioned after that (in a text span of max. 4 sentences). Another 38% of DNIs are resolved in a text span preceding the current predicate of max. 5 sentences, while the rest has a very far antecedent (up to 116 sentences before the current predicate). The notion of context where the antecedent should be searched for is clearly lacking an appropriate definition.

**Diversity of lexical fillers:** In general, it is possible to successfully obtain information about the likely fillers of a missing FE from annotated data sets only in case all FE labels are semantically well identifiable: in fact many FE labels are devoid of

any specific associated meaning. Furthermore, lexical fillers of a given semantic role in the FrameNet data sets can be as diverse as possible. For example, a complete search in the FrameNet database for the FE Charges will reveal heads like ‘possession, innocent, actions’, where the significant portion of text addressed by the FE would be in the specification - i.e. ‘possession of a gun’ etc. Only in case of highly specialized FEs there will be some help in the semantic characterization of a possible antecedent.

## 6 Conclusions

In this paper, we have described the phenomenon of null instantiated arguments according to the FrameNet paradigm and we have proposed a strategy for identifying implicit arguments and finding their antecedents, if any, using a linguistically-motivated approach. We have evaluated our system using the test set developed for the SemEval task 10 and we have discussed some problems in our approach affecting its performance. Besides, we have also pointed out some issues related to the task definition and to the general phenomenon of null instantiated arguments that make the identification task challenging from a computational point of view. We have shed some light on the syntactic, semantic and discourse information that we believe are necessary to successfully handle the task.

In the future, we plan to improve on our binding approach by making our model more flexible. More specifically, we currently treat DNI referents occurring before and after the sentence containing the predicate as equally probable. Instead, we should penalize less those *preceding* the predicate because they are more frequent in the training set. For this reason, the number of observations for the candidate head and the distance should be represented as different weighted features. Another direction to explore is to extend the training set to the whole FrameNet resource and not just to the SemEval data set. However, our approach based on the observations of lexical fillers is very much domain-dependent, and a larger training set may introduce too much variability in the heads. An approach exploiting some kind of generalization, for example by linking the fillers to WordNet synsets as proposed by (Gerber and Chai, 2010), may be more appropriate.



## References

- Collin F. Baker, Charles J. Fillmore, and J. B. Lowe. 1998. The Berkeley FrameNet Project. In *Proceedings of Coling/ACL*, Montreal, Quebec, Canada.
- C. F. Baker, M. Ellsworth, and K. Erk. 2007. Semeval-2007 task 10: Frame semantic structure extraction. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 99–104, Prague, CZ, June.
- Aljoscha Burchardt, Annette Frank, and Manfred Pinkal. 2005. Building text meaning representations from contextually related frames - a case study. In *Proceedings of the Sixth International Workshop on Computational Semantics*, Tilburg, NL.
- Desai Chen, Nathan Schneider, Dipanjan Das, and Noah A. Smith. 2010. SEMAFOR: Frame Argument Resolution with Log-Linear Models. In *Proceedings of SemEval-2010: 5th International Workshop on Semantic Evaluations*, pages 264–267, Uppsala, Sweden. Association for Computational Linguistics.
- Charles J. Fillmore. 1986. Pragmatically Controlled Zero Anaphora. In V. Nikiforidou, M. Vanllay, M. Niepokuj, and D. Felder, editors, *Proceedings of the XII Annual Meeting of the Berkeley Linguistics Society*, Berkeley, California. BLS.
- Matthew Gerber and Joyce Y. Chai. 2010. Beyond NomBank: A Study of Implicit Arguments for Nominal Predicates. In *Proceedings of the 48<sup>th</sup> annual meeting of the Association for Computational Linguistics (ACL-10)*, pages 1583–1592, Uppsala, Sweden. Association for Computational Linguistics.
- Adam Meyers. 2007. Annotation guidelines for NomBank - noun argument structure for PropBank. Technical report, New York University.
- M. Palmer, D. Dahl, R. Passonneau, L. Hirschman, M. Linebarger, and J. Dowding. 1986. Recovering implicit information. In *Proceedings of ACL 1986*, pages 96–113.
- Martha Palmer, Daniel Gildea, and Paul Kingsbury. 2005. The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics*, 31.
- Josef Ruppenhofer and Laura A. Michaelis. 2009. Frames predict the interpretation of lexical omissions. Submitted.
- Josef Ruppenhofer, Caroline Sporleder, Roser Morante, Collin F. Baker, and Martha Palmer. 2010. SemEval-2010 Task 10: Linking Events and Their Participants in Discourse. In *Proceedings of SemEval-2010: 5th International Workshop on Semantic Evaluations*.
- Josef Ruppenhofer, 2010. *Annotation guidelines used for Semeval task 10 - Linking Events and Their Participants in Discourse*. (manuscript).
- Sara Tonelli and Rodolfo Delmonte. 2010. VENSES++: Adapting a deep semantic processing system to the identification of null instantiations. In *Proceedings of SemEval-2010: 5th International Workshop on Semantic Evaluations*, Uppsala, Sweden.