3.1 Allgemeine Angaben zum Teilprojekt D6 neu

3.1.1 Titel: Lexical-semantic factors in event interpretation

Lexical-semantic factors in event interpretation

3.1.2 Fachgebiete und Arbeitsrichtung:

Computational Linguistics, Psycholinguistics, Lexical Semantics, Syntax-Semantics Interface, Corpus-Driven Modelling

3.1.3 Leiter/in:

Juniorprofessor Dr. Padó, Sebastian, geb. 24.07.1978

Universität Stuttgart
Institut für maschinelle Sprachverarbeitung
Azenbergstrasse 12
70174 Stuttgart

Telefon:+49-(0)711-685 81394
Telefax:+49-(0)711-81366
E-Mail: pado@ims.uni-stuttgart.de

Ist die Stelle des Leiters/der Leiterin des Projektes befristet?

☐ nein  ☑ ja, befristet bis zum 31.03.2012
☐ eine weitere Beschäftigung ist vorgesehen bis zum 31.03.2015
   (abhängig von positiver Evaluation)
3.1.4 In dem Teilprojekt sind vorgesehen:
• Untersuchungen am Menschen oder am menschlichen Material
  □ ja □ nein

  Die erforderliche Zustimmung der zuständigen Ethikkommission
  liegt dem Antrag zum Teilprojekt in Kopie bei  □ ja
  □ nein

• klinische Studien
  □ ja □ nein

• Tierversuche
  □ ja □ nein

• gentechnische Untersuchungen
  □ ja □ nein

• Untersuchungen an humanen embryonalen Stammzellen
  □ ja □ nein

  Die gesetzliche Genehmigung liegt vor  □ ja □ nein

3.1.5 Beantragte Förderung des Teilprojektes im Rahmen des Sonderforschungsbereichs (Ergänzungsausstattung)

<table>
<thead>
<tr>
<th>Haushaltsjahr</th>
<th>Personalmittel</th>
<th>Sachmittel</th>
<th>Investitionsmittel</th>
<th>Gesamt</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010/2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2014/1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Beträge in Tausend EUR)

3.2 Zusammenfassung

Short summary. The project D6 will address the interpretation process of Verb-Object pairs which are potential cases of logical metonymy and as such are able to imply events that are not overtly realized. An example of such a case is help with the dishes, which is generally understood as help wash the dishes, while other cases like help with the cleaning do not require the recovery of an event. We will use a combination of psycholinguistic and corpus-driven computational approaches to investigate the understudied role of lexical-semantic factors in the interpretation of such sentences. Our focus will be on: (a), the behaviour of borderline that may or may not involve covert events (where indicators for the two interpretations may be ambiguous or contradictory); and (b), the shape and the semantic structure of the set of events recovered for given Verb-Object pairs. Additional practical
results will include a data-driven model for the prediction of such interpretations, and
distributional criteria to identify novel candidates for constructions whose interpretation
relies on covert events.

**Extended summary.** The D6 project will investigate the topic of the SFB, *incremental
specification in context*, in the context of the interpretation of sentences involving *covert
events* which are implied, but not realized on the surface (e.g., *begin a song*, in neutral
context, implies *begin singing a song*).

The interpretation of such sentences is of considerable interest both cognitively and
computationally. Cognitively, the recovery of missing event has been linked to reading times
that are systematically longer than for similar instances without a covert event (e.g., *begin
a holiday*). Computationally, such sentences are problematic because current methods for
constructing semantic representations are, as a rule, guided by the overt material. Thus,
they cannot draw inferences involving the covert events, such as *Tom sings a song* being
entailed by *Tom begins a song*.

D6 will consider both steps of interpretation in logical metonymy, namely first the
decision between simple composition (where no event is inserted) and enriched composition
(where it is), and then the specification of this event in context. The novelty of our project
lies in the investigation of the role played by lexical-semantic factors in the interpretation
of logical metonymy. These factors have been neglected in existing work.

With regard to the simple/enriched composition split, the literature assumes that the
*ontological type* of the object (event/object) is the main determinant. Since this is arguably
an oversimplification, we will investigate (a), the behaviour of ambiguous nominalizations
and nouns; (b), the respective influence of ontological type and plausibility; and (c), the role
played by aspectual considerations in interpretation. Then, we will investigate the inter-
pretations speakers produce for covert events. Previous work has described interpretations
merely on the level of individual verbs. Our counter-hypothesis is that interpretations are
better understood in terms of *concepts*. We will investigate how intra-sentential context
drives the specification of the set of possible interpretations, and what the semantic struc-
ture within the concepts (i.e., the pairwise semantic relations between verbs expressing the
concept) is. This process will be analysed in parallel for English and Italian. Both of these
investigations will use psycholinguistic methods to ground our analyses in both online and
offline correlates of human language processing.

In parallel, we will develop corpus-driven, broad-coverage computational models that
can predict both the simple/enriched composition distinction for sentences potentially in-
volving covert events, and, if appropriate, the full range of interpretations appropriate for
the present context. These models will be situated within the framework of vector spaces.
Vector space models can be induced from corpora without recourse to other knowledge
sources, and have been successfully used to model the plausibility of predicate-argument
combinations, since the representations they contain are shaped by word *usage*.

Finally, we aim at making the insights that we will gain on the influence of context on
event interpretation useful beyond logical metonymy alone. We will do so in three ways.
First, we will phrase the recovery of covert event interpretations as an entailment task,
so to make our phenomenon accessible to a larger audience in computational lexical se-
manics. Second, we will study another phenomenon that appears to be fairly similar to
logical metonymy, namely Italian complex *da*-nominals. Third, we will develop distributio-
nal criteria that allow the corpus-driven identification of (candidate) constructions whose interpretation involves enriched composition.

Deutsche Zusammenfassung. Das Projekt D6 befaßt sich mit der Interpretation von Verb-Objekt-Paaren, zu denen Interpretation möglicherweise Ereignisse rekonstruiert werden müssen, die auf der Oberfläche fehlen (logische Metonymie). Beispielsweise wird in mit dem Geschirr helfen das Ereignis, bei dem geholfen wird, nicht genannt; in neutralem Kontext wird der Satz aber als synonym zu helfen, das Geschirr abzuwaschen verstanden. Im Gegensatz dazu muss und kann bei mit dem Abwasch helfen kein Ereignis eingefügt werden. Bisher wurde die Rolle, die bei diesem Prozess von lexikalisch-semantischen Faktoren gespielt wird, nur vereinfacht darstellt. Wir werden diese Rolle mit psycholinguistischen und komputationellen Methoden untersuchen. Wir fragen (a), nach dem Verhalten von Grenzfällen, bei denen unklar ist, ob Ereignisse rekonstruiert werden müssen, da die Indikatoren ambig oder widersprüchlich sind; und (b), nach den Grenzen und der semantischen Struktur der Menge an Ereignissen, die für ein gegebenes Verb-Objekt-Paar von Hörern rekonstruiert werden. Weitere praktische Ergebnisse des Projektes werden ein korpusbasiertes Modell zur Interpretation solcher Verb-Objekt-Paare sein sowie eine weitergehende distributionale Charakterisierung von Konstruktionen, deren Interpretation die Rekonstruktion nicht genannter Ereignisse umfasst.

3.3 Ausgangssituation des Teilprojekts

3.3.1 Stand der Forschung

Polysemy, metonymy and coercion. Polysemy is a central fact of lexical semantics. A particular case is regular polysemy, where lexemes have several readings between which a conventionalized semantic relation holds, such as the readings of newspaper as a physical object, an information object, and a company (Apresjan, 1973; Nunberg, 1979). While polysemy is often characterized simply in the form of multiple word senses, this does not appear to be possible in the case of logical metonymies (Pustejovsky, 1995), which arise for example for experiencer verbs such as enjoy that can subcategorize both for an event-denoting VP (which we will call the “long variant”) and for an object-denoting NP (the “short variant”). Assuming an uniform semantic representation for enjoy, this results in a type conflict for at least one of the cases during composition. Furthermore, the interpretation of such verbs when combined with an NP can be influenced by context. Consider the following two pairs of paraphrases

(3.1) a. Peter enjoyed the book.
    b. Peter enjoyed reading the book.

(3.2) a. The goat enjoyed the book.
    b. The goat enjoyed eating the book.

Even though the VPs of (3.1a) and (3.2a) are identical, their paraphrases, (3.1b) and (3.2b) respectively, are not. Thus, it is assumed that in such cases, a type shifting process
takes place which “reinterprets” the object-denoting NP as an event involving the object in question.

Jackendoff (1997), building on Pustejovsky’s work, has coined the terms *simple composition* and *enriched composition* for the two cases. Simple composition is the “standard case” where the meaning of an expression results directly from the meaning of its parts and that applies to the long variant. In contrast, the enriched composition of the short variant involves

other material that is not expressed lexically, but that must be present [...] either (i) in order to achieve well-formedness [...], or (ii) in order to satisfy the pragmatics of the discourse or extralinguistic content.

In theory-neutral terms, the interpretation of logical metonymies involves *non-overt events* that must be recovered if the expression is to be understood correctly.

**Mechanisms and representations of non-overt events.** Strong compositional approaches have been challenged by cases of syntax-semantics mismatch, because, in cases of logical metonymies, *non-overt events* to be recovered are not given at the syntactic level. Different semantic theories have thus provided different representations for such phenomena:

(3.3) a. The author began the book.
   b. The author began *writing* the book.

The *Generative Lexicon (GL) approach* (Pustejovsky, 1995; Jackendoff, 1997) observes that in (3.3a) the verb *begin* selects for an event-denoting argument, but the argument *the book* denotes an entity. The *type-mismatch* requires a process of *type-shifting* or *complement coercion* to occur, thus coercing the interpretation of the argument into an event-related meaning, such as (3.3b). The bases and constraints for such operation are assumed to be grounded in the lexicon and to be contained in the lexical entry for the coerced argument in the form of a *qualia structure* ¹: the non-overt event in (3.3b) is lexically determined and can be retrieved with the coercion operation.

Fodor and Lepore (1998) criticize the GL approach by pointing out that it is not clear why, if the non-overt event for *begin a car* should be taken from the *TELIC quale* for car, it does not mean *begin driving a car*. The non-overt event here is assumed to be inferred from the speakers’ knowledge of the properties of car, rather than being computed from its lexical entry, which is instead assumed to be atomistic.

de Almeida and Dwivedi (2008), following Wurmbrand (2001, 2004), accept the analysis in Fodor and Lepore (1998) (“lexical items are atomistic representations: they do not encode information beyond their own denotations”), and propose an alternative account for examples like (3.3a): the triggering factor in the interpretation of non-overt events is assumed to be not a type mismatch but rather a structurally unfulfilled gap in syntax.

¹A qualia structure is assumed to provide an exhaustive description of properties of the item, by referring to its physical makeup (*CONSTITUTIVE quale*: books are made with paper), its taxonomic structure (*FORMAL quale*: books are a kind of physical object), its associated purpose/function (*TELIC quale*: books are read) and the factors which brought it about (*AGENTIVE quale*: books are written).
and semantic structure, and the gap is assumed to be conceptually filled by pragmatic inferences.

Asher (2007) argues against the assumption that the interpretation of the non-overt event can not derive from the lexical meaning of the noun. See the reported examples:

(3.4) Sam enjoyed the zibzab.
(3.5) *The book started at 10 am.
(3.6) [Last week Julie painted her house.] She started with the kitchen, then proceeded to the living room and bedroom and finished up with the bathroom.

Asher argues that, when hearers are asked about sentences with made-up nouns like 3.4, "most of them report that he enjoyed doing something with the zibzab, but they have no idea what". This observation suggests that, if words which are not in the lexicon can trigger an enriched composition structure, then this can not be ascribed to a lexicalized qualia structure. Another argument against qualia structures is that it would overgenerate non grammatical sentences like in 3.5: "it seems difficult to deny phrases like the book an event reading if events are associated with book in the lexicon, as in the case in classic GL". On the other hand, note that the example in 3.6 can be made grammatical with a plausible background sentence: sentences showing logical metonymy are seen as underspecified and discourse context is seen as crucial to solve the underspecification (SDRT theory, Asher (1993), Asher and Lascarides (2003)): the solution of the underspecification and the construction of the discourse structure "are codependent tasks".

**Concrete approaches to computing interpretations for logical metonymy.** The problem of computing interpretations for logical metonymy has received considerable attention in the theoretical literature but little in the empirical literature; remarkable exceptions are Lapata and Lascarides (2003) and Lapata et al. (2003).

Lapata and Lascarides (2003) present a probabilistic data-driven model to predict “missing” verbs for a range of matrix verbs that allow for logical metonymy and to rank their plausibilities. The interpretation of a verbal metonymy (enjoy the book) is modelled as the joint distribution \( P(e, o, v) \) of the matrix verb \( v \) (enjoy), its object \( o \) (book) and the events \( e \) co-occurring with \( o \) (reading, writing, publishing, editing...). The model is extended to the treatment of metonymic adjectives and evaluated against human judgments.

Lapata et al. (2003) present a model based on Lapata and Lascarides (2003), which focuses on one aspect of intra-sentential context (the influence of the sentential subject) to determine if the telic or the agentive quale is accessed, as in:

(3.7) a. Peter begann das Buch. (neutral condition)
    b. Der Student begann das Buch. (assumed interpretation: reading, telic quale)
    c. Der Author begann das Buch. (assumed interpretation: writing, agentive quale)

The model’s prediction on non-overt events in logical metonymies are tested against those elicited by humans in a completion experiment.

Both approaches make use of intra-sentential context, and are therefore not suited for the interpretation of metonymies for which the statistically dispreferred interpretation is
the correct one and which require the integration of a wider discourse context; see the following examples from Lascarides and Copestake (1998):

(3.8) a. Peter enjoyed the book. (assumed interpretation: reading)

   b. Peter enjoyed the book made out of marzipan. (assumed interpretation: eating)

As noted by Lascarides and Copestake (1998), “in such cases, open-ended inference may be required about books and the nature of reading” and “it is implausible that a selectional restriction will work”.

Another issue that is not addressed by the current computational models is the different behavior of entity-denoting nouns (as in Peter enjoyed the book) and event-denoting nouns (as in Peter enjoyed the holiday): event-denoting nouns would not require the retrieval of a non-overt event. The intuition here is that an event-denoting noun should not trigger the same mechanisms which lead speakers to trigger an enriched composition interpretation and should rather constrain the interpretation of potential metonymies.

Computational semantics, entailment, and paraphrase. A different perspective on enriched composition events comes from the requirements of natural language processing (NLP), broadly construed as technologies that help humans search and manage information in free text. Efficient information access needs to abstract away from the variability of realization on all linguistic levels and to be able to characterize the underlying semantic relations between sentences and texts. Of particularly interest is the relation of entailment which allows systems to derive new true statements from known true statements. A recent successful approach to conceptualizing entailment is textual entailment (Dagan and Glickman, 2004; Dagan et al., 2009), which sidesteps the issues of semantic representation by grounding entailment in human judgments at the level of plain text. More specifically, textual entailment holds if naive speakers of a language agree that, if a first piece of text (the premise P) is true, then the second piece of text (the hypothesis H) is true as well.

The benefit of textual entailment in the current context is that it allows us to phrase enriched composition as a classification task: textual entailment is supposed to hold between a short variant like (3.9a) and a long variant that provides an appropriate event, e.g. (3.9b), but not between a short variant and a long variant with a wrong event like (3.9c)).

(3.9) a. Peter enjoyed the book.

   b. Peter enjoyed reading the book.

   c. Peter enjoyed burning the book.

In consequence, enriched composition should be amenable to the methods and analyses developed for the “Recognizing Textual Entailment” (RTE) challenges, a series of shared tasks on recognizing textual entailment on samples of real-world sentence pairs motivated by different NLP applications. The challenges have been organized yearly since 2005 and have attracted participants from a broad range of approaches, from surface overlap to dependency structure rewriting and full-blown semantics construction (Bar-Haim et al., 2005; Bos and Markert, 2006; Harmeling, 2007).

Across all approaches, one of the most prominent causes of missed positive instances of entailment was found to be lexical gaps, lexical differences between text and hypothesis
that could be bridged due to the imperfect coverage of lexical resources. While the gap can be closed at the level of individual words at least to a considerable extent with the use of ontologies like WordNet (Fellbaum, 1998) or Dekang Lin’s thesaurus (Lin, 1998), the data-driven acquisition of substitutions at the level of entire phrases (paraphrases) is still an open research question. Virtually all approaches either (a) use distributional similarity of argument headwords in monolingual corpora (Lin and Pantel, 2002; Glickman and Dagan, 2004); or (b) exploit translational relations in parallel corpora (Barzilay and Lee, 2003; Bannard and Callison-Burch, 2005; Szpektor and Dagan, 2008).

Crucially, what none of these approaches currently offers is a systematic account of possible paraphrases for certain constructions. Their output is usually a table of equivalence string pairs, which means that no predictions are possible for unseen expressions. Reliable estimates of recall are hard to obtain, but there is evidence that only a small fraction of paraphrases are represented in current resources (Pantel and Pennacchiotti, 2006; de Marneffe et al., 2009). In addition, monolingual methods can only recover pairs of expressions of the same type (typically transitive and intransitive verbs), and parallel corpora-based approaches are severely limited by the availability of such resources. Thus, it appears like no current RTE system can systematically account for enriched composition.

Cognitive studies of the interpretation of “short variants”. A third perspective on enriched composition is from a cognitive points of view. A number of psycholinguistics and neurolinguistics studies have investigated behavioral and neural evidence for enriched composition (for a review, see Pylkkänen and McElree (2006)). Both self-paced reading and eye-tracking methodologies have been used to detect higher processing costs for sentences which are assumed to require a coercion operation, like the author began the book. Control conditions were non-coerced conditions with a preferred event like the author wrote the book and with a non-preferred event like the author read the book (McElree et al., 2001; Traxler et al., 2002; McElree et al., 2006a,b), sentences with an event-denoting NP like the author began the fight (Traxler et al., 2002) and sentences with overtly expressed events like the author began writing the book (Pickering et al., 2005). Such results are usually related with Pustejovsky’s coercion model (Pustejovsky, 1995).

Nevertheless, it has been argued (de Almeida, 2004) that processing costs for the coerced condition are comparable with the ones for a non-coerced non-preferred condition (the author read the book), thus supporting an alternative account based on Fodor and Lepore (1998), and arguing that the obtained results can also reflect a post-access inferential process triggered by the underspecification of the V + NP type shifting construction. Pickering et al. (2005), on the other hand, argue that there is no ground to state that ”type-shifting constructions behave similarly to non-preferred constructions”, because the detecting of longer reading times or eye fixations for coerced conditions, though not specifically committing to qualia structure theory, is nevertheless a sign of higher processing costs.

De Almeida’s work remarkably takes into account the role of context in type-shifting operations, beyond what is lexically determined (Lascarides and Copestake, 1998; de Almeida and Dwivedi, 2008). Context effects are also taken into account by Traxler et al. (2005), following the intuition that “coerced senses are computed from a broader range of properties than the Qualia structure of the complement noun”: they reported a limited facilitation effect of context on processing costs for coerced sentences, which does not occur when the event is mentioned by previous context (the author wrote every morning in his
conservatory... that particular day, he began a chapter about the main villain), but rather when the entire event sense is (the author began the chapter about the main villain in his apartment... although she began it well before lunch, it took him a long time to get going).

More insights came from a MEG study by Pylkkänen and McElree (2007) (see also Pylkkänen et al. (2008)), showing that complement coercion does not modulate the same brain activity found in clear cases of semantic mismatch between a verb and its complement, thus suggesting that the coercion process is not triggered by a semantic mismatch, nor it shares the same brain activity with decision making processes.

3.3.2 Eigene Vorarbeiten

Sebastian Padó has worked on broad-coverage data-driven models of selectional preferences, which provide verb-specific graded predictions of the plausibility of nominal arguments in null context (Padó et al. 2007) based on vector space models of word meaning constructed from syntactically annotated corpora (Padó and Lapata 2007). One particular focus of recent work was the development of the concept of inverse preferences, that is, noun-specific knowledge about plausibility of verbs in which the nouns participate (Erk and Padó 2008, Erk et al. submitted). These selectional preference models, which make predictions at the lemma level, thus represent a starting point for models that can predict likely predicates at the instance-level. His PhD thesis was on the topic of cross-lingual induction of semantic role information in parallel corpora (Padó and Lapata 2006, 2009).

Alessandra Zarcone has worked on data-driven models and psycholinguistics investigation of event types (Vendler’s classes) in Italian. In a first Maximum Entropy model, event types identification was modelled as a supervised classification task (Lenci and Zarcone 2006). In a second model, event types were identified in an unsupervised way by Self-Organizing Maps (Zarcone and Lenci 2008). For both models, the interaction of various contextual factors (syntactic, morphological and semantic features) was crucial to determine the event type of a sentence and to model context-driven semantic shifts. A priming study further investigated category effects and their interaction with morphology (Zarcone 2009).

3.3.3 Liste der publizierten einschlägigen Vorarbeiten

I. Begutachtete Veröffentlichungen


II. Eingereichte Veröffentlichungen (mit Datum der Einreichung)


III. Nicht begutachtete Veröffentlichungen


3.4 Planung des Teilprojekts (Ziele, Methoden, Arbeitsprogramm)

3.4.1 Fragestellung

Project D6 will be concerned with cases of enriched composition in the sense of Jackendoff (1997) quoted above. We will concentrate on cases of enriched composition that involve covert events which are recovered automatically by hearers during interpretation, such as logical metonymy:

(3.10) a. Peter enjoyed the book. (short variant)

b. Peter enjoyed reading the book. (long variant)

Figure 3.1 sketches the putative interpretation process for logical metonymies. The “long variant”, in which the matrix verb has a VP complement, receives the standard compositional interpretation. In contrast, the “short variant” is structurally ambiguous, allowing for two interpretations: The first one (option (1)) is a simple composition of a transitive verb with its object; the second one (option (2)) is enriched composition involving an additional event.

The interpretation of “short variants” is of considerable interest both cognitively and computationally. Cognitively, the recovery of the missing event (as in option (1)) has been linked to systematically longer reading times when compared with option (2). Computationally, the construction of meaning representations for “short variants” requires the integration of knowledge that is not present on the surface.
While both of these points have been investigated before, all prior studies have made strong simplifying assumptions about the role of lexical-semantic factors in the interpretation of “short variants”, despite the fact that the interpretation is determined primarily by the identity of the matrix verb ($V_1$) and the head of the object NP.

The goal of project D6 is therefore to extend our understanding of lexical-semantic factors in enriched composition both in terms of depth (how does it actually work?) and breadth (where does it apply?). This goal has a cognitive, a computational, and a conceptual component:

- On the cognitive level, we want to understand (a), how semantic properties of the lexical properties determine the choice between simple and enriched composition; (b), how they determine the space of interpretations of the covert event in enriched composition; and (c), how that space is semantically structured;

- On the computational level, we want to understand how the influence of these factors can be accounted for in corpus-driven models of lexical semantics to (a), predict the influence of enriched composition on human language processing in terms of reading times and acceptability judgments; and (b), take advantage of these insights for semantic interpretation in natural language processing;

- On the conceptual level, we want to investigate the concept of “enriched composition” in its two-fold cognitive and linguistic nature and examine to which other constructions it is applicable.

The following paragraphs give more detail on these questions:

**What factors determine the choice of simple vs. enriched composition?** A null hypothesis about enriched composition that can be drawn from work on theoretical semantics posits that enriched composition takes place to resolve type conflicts. According to this view, it should be possible to draw the distinction between options (1) and (2) in Figure 3.1 on the basis of the semantic type of the argument (object-denoting vs. event-denoting). However, this raises at least three questions which we will examine cognitively and computationally.

First, this criterion does not make clear predictions for ambiguous cases. Do German nominalizations that have both an event and an object reading, like the following example, lead to longer reading times?

(3.11) Die Helfer bereiteten das Strassenfest vor. Sie begannen mit der Absperrung.
A second interesting case is formed by nouns that are not nominalizations, but still behave ambiguously with respect to the object/event distinction. An example is the word “game”, which was taken to be an event noun in previous studies, but which also supports enriched composition as in the following example, where enjoy the game can be paraphrased as enjoy watching the game. Do such cases lead to longer reading times or not?

(3.12) After the tournament, the players relaxed and enjoyed the baseball game with friends and family in premium reserved infield seats at discount prices.

Finally, it has been argued that a type conflict alone does not suffice to make enriched composition possible: Cases like the following are generally perceived as ungrammatical rather than interpreted through enriched composition.

(3.13) * He began the encyclopaedia.

Is this ungrammaticality due to the Aktionsart of the salient events associated with encyclopaedias, refer, which is a non-durative event?

How do contextual lexical factors determine the recovered event in enriched composition? In the case of enriched composition, a missing event needs to be recovered. However, combinations of matrix verbs and objects generally do not restrict interpretation to a single event; instead, data analyses that we have performed indicate that often a range of possible events exists. Consider the example for a Verb+PP construction in Table 3.1, which demonstrates three context-specific interpretations sampled from the WWW. The first one is apparently about difficulties in obtaining food and the second deals with food processing. The third one, which was found in the context of a discussion about pet health, asks for advice about food choice. Only the last one corresponds well on the small set of interpretations suggested by, e.g., Generative Lexicon approaches. We are going to ask specifically:

- How is the space of interpretations for a given matrix verb-object combination constrained without and within context?
- How is the space of interpretations for a given matrix verb-object combination structured? How can it be represented either in terms of vector space models, or in terms of ontologies?
- How do these results carry over across languages?
How can models for enriched composition make an impact on NLP? Ideally, the insights we will gain on the influence of context on event interpretation, and the modelling strategies for implementing practical model of event interpretation, should also be beneficial for natural language processing. However, the relevance of our work is clearly dependent on our ability to demonstrate the relevance of our methods beyond the specific phenomenon of logical metonymy. We will therefore ask the following questions:

- How can we phrase the recovery of covert event interpretations in terms of a “recognition of textual entailment” (RTE) task? The RTE setup has gained importance as a meta-task over the last years, which makes it possible both to apply other semantic models to our specific phenomenon, and to apply our own models to other data.
- Can we apply the models we have developed for logical metonymy to the interpretation of Italian complex da-nominals, a construction that promises to behave similar to logical metonymy?
- Reaching out further, can we identify other constructions in a corpus-based manner to which the concept of enriched composition involving covert events is applicable? To this end, how can we give a distributional characterizations of enriched composition?

### 3.4.2 Ziele

- Investigation of processing and interpretation of logical metonymy as an instance of enriched composition involving covert events. In contrast to existing work, we place a special focus on lexical-semantic factors and their role in the resolution of the single/enriched composition ambiguity as well as the semantic structure of possible event interpretations.
  - Exploration of the “borderline” between simple and enriched composition and of the factors determining the choice of simple vs. enriched composition: ontological type of object argument, plausibility, aspectual features of the object;
  - Exploration of the range of possible interpretations in enriched composition: influence of the lexical-semantic properties of the intra-sentential context; internal structure of elicited representations
- A fully implemented, data-driven model for computing the interpretation of potentially metonymic $V + N$ combinations.
  - Insights in the role of context and plausibility in shaping interpretations and inferences, and consequences for modelling strategies
- A large dataset for textual entailment consisting of possible and impossible interpretations for covert events. This dataset can be used as a testbed for other models of covert event interpretation.
- A distributional characterization for constructions that can understood as involving enriched composition with covert events, and an inventory of such constructions.
3.4.3 Methoden und Arbeitsprogramm

The Arbeitspakete of D6 fall into two “core” cognitive packages (Arbeitspakete 1 and 3), two “core” computational modelling packages (Arbeitspakete 2 and 4), and two auxiliary packages (Arbeitspakete 5 and 6). The distinction within the core packages into a cognitive and a computational (lexical semantic) part corresponds to the twin goals of our project. Before we describe the individual APs, we sketch two methodological issues: Types of experimental data, and languages.

Experimental data acquisition. In the cognitive APs, we acquire both offline data (reflecting the results of human language processing while ignoring the timecourse of processing) and online data (reflecting the timecourse of human language processing).

- Offline data: elicitation of paraphrases. Such data can be gathered using a paraphrasing task where we will ask participants to provide interpretations for “simple forms”. Such tasks have been used recently in lexical semantics with good success to obtain empirical characterizations of word senses (McCarthy and Navigli, 2009). Furthermore, the availability of “crowdsourcing” platforms like Amazon Mechanical Turk (www.mturk.com) allows us to tap into naive native speakers’ judgments and obtain large datasets in relatively short time. Recent studies have verified that the quality of judgments is high, provided that the task can be phrased in a way that is intuitive for naive native speakers (Snow et al., 2008; Callison-Burch, 2009).

- Offline data: norming studies. Norming data are gathered by asking participants to quantitatively assess some property of a linguistic stimulus, typically on a five-point or
on a seven-point scale. Typical applications of norming studies are the acquisition of
detailed plausibility ratings for paraphrases (Lapata and Lascarides, 2003; Mitchell
and Lapata, 2008), predicate-argument combinations (Brockmann and Lapata, 2003),
or quality of machine translation hypotheses (Callison-Burch et al., 2008). A recent
development is that norming studies can be sped up considerably by using web-based
crowdsourcing platforms such as Amazon’s Mechanical turk (Callison-Burch, 2009).

- Online data: reading times (RT). RT methods have traditionally been used as ex-
 perimental correlates of processing cost (Just et al., 1982). RT paradigms typically
  compare one (or more) experimental conditions (e.g. The author began the book) with
  a control condition (e.g. The author wrote the book), which should only differ with
  regard to the manipulated experimental variable. The conditions are then compared
  with regard to their reading times, and differences in RT are ascribed to differences in
  processing costs between the two conditions. In particular, established RT methods
  for the study of processing costs of logical metonymy are self-paced reading (McElree
  et al., 2001; Traxler et al., 2002; de Almeida, 2004), speed-accuracy tradeoff (McElree
  et al., 2006b) and eye-tracking (Traxler et al., 2002; Pickering et al., 2005; Traxler
  et al., 2005; McElree et al., 2006a).

Languages. Project D6 will be focussed on English, in order to retain comparability to
previous studies and to profit from existing large-scale resources. In addition, some of the
experiments and the modelling will be carried out also on Italian, a language belonging to
a different language family (Romance), and having less lexical overlap with English than
French. The comparison with Italian will serve both to verify our findings on a second
language, thus avoiding language-specific idiosyncrasies, and to create a bilingual gold
standard. This is particularly interesting for Arbeitspaket 3, the investigation of the internal
semantic structure of the range of covert events, which raises the question of how well these
spaces correspond for two different languages. Since the offline web-based experiments
planned for Arbeitspaket 3 are cost-effective enough to be carried out for both English and
Italian, we will perform these experiments in parallel for both languages. Arbeitspaket 4.4
will address the question of cross-lingual modelling of these phenomena. In Arbeitspaket 6,
we will build on these cross-lingual insights to ultimately use the English model enriched
composition in a novel Italian construction.

Arbeitspaket 1: Human choice between simple composition and enriched com-
position

The first Arbeitspaket will be concerned with gathering data about the human interpretati-
on of simple vs. enriched composition from cognitive experiments. The main questions are:
Under what circumstances does enriched composition take place? What are the semantic
properties guiding the decision between ungrammaticality, simple composition or enriched
composition?

Three main factors assumed to affect the interpretation of logical metonymies will be
investigated:

1. entity-denoting objects vs. event-denoting objects (Arbeitspaket 1.1);
2. subject of the main verb (Arbeitspaket 1.2);

3. aspectual features of the object (Arbeitspaket 1.3);

**Arbeitspaket 1.1: entity-denoting objects vs. event-denoting objects.** The examples reported in the literature show how different objects can require a non-overt event or select for an enriched composition interpretation:

(3.14)  

| a. | Peter enjoyed the book. (assumed interpretation: enriched composition, entity noun) |
| b. | Peter enjoyed the holidays. (assumed interpretation: simple composition, event noun) |
| c. | Peter enjoyed the game. (assumed interpretation: ambiguous noun) |
| d. | Peter enjoyed the translation. (assumed interpretation: polysemous, nominalization) |

Generally speaking, entity-denoting nouns are assumed to require enriched composition (3.14a), whereas event-denoting nouns (3.14b) are not. Nevertheless, some problems arise from this distinction. Nouns can be ambiguous with respect to their event/entity reading (3.14c, game session vs. game object), while nominalizations can show regular event/entity polysemy (3.14d, translation process vs. result of the translation); polysemous nominalizations can refer back to the corresponding event (*translation → translate*), but ambiguous nouns show a much less straightforward alternation between the event meaning and the entity meaning.

The main aim of Arbeitspaket 1.1 will be the analysis of the different behavior of these four categories (event nouns, entity nouns, vague nouns, nominalizations) in a potential enriched composition context (3.14) and of the role played by different objects in selecting a simple composition interpretation vs. an enriched composition interpretation.

Another issue addressed by Arbeitspaket 1.1 will be the role of the matrix verbs which allow for enriched composition (*begin, finish, enjoy*, etc.): do these verbs show uniform behavior when determining processing costs differences between simple and enriched composition?

The investigation of the role played by the object will be articulated in three steps:

1. **selection of a balanced set of stimuli** (event nouns, entity nouns, polysemous nouns, nominalizations): it will be necessary to define indicators of *entity-hood* vs. *event-hood*, in order to select both unambiguous nouns (event nouns and entity nouns) and polysemous nouns and nominalizations (the definition of selection criteria for the four categories will benefit from Arbeitspaket 2.1 and from the cooperation with project B3);

2. **web-based norming study** to evaluate the stimuli set (does *Peter enjoyed the translation* mean *Peter enjoyed performing the translation* or *Peter enjoyed doing something with the translation*?);

3. **reading-time experiment** to explore how the entity-event distinction correlates with the simple vs. enriched composition distinction; the experiment will have a two-factors design (*object, main verb*): the factor *object* will have four levels (event nouns,
entity nouns, polysemous nouns, nominalizations). The factor matrix verb will evaluate the effect of different matrix verbs in determining processing costs differences between simple and enriched composition. We expect different matrix verbs to have a similar behavior; as for the factor object, we expect results to be coherent with the literature (RT for event nouns (simple composition) < RT for entity nouns (enriched composition)), and we are interested in comparing the ambiguous conditions (polysemous nouns and nominalizations) with those bounds.

Arbeitspaket 1.2: Plausibility considerations: The subject. Another element playing a crucial role in the incremental interpretation of simple vs. enriched composition is the subject. When the object of the VP is ambiguous between an entity interpretation and an event interpretation (e.g. icing in 3.15), some subjects can select for an enriched composition interpretation, some others can suggest a simple composition interpretation:

(3.15) a. The cook finished the icing. (assumed interpretation: simple composition, he finished the process of icing the cake)
b. The child finished the icing. (assumed interpretation: enriched composition, he finished eating the icing)

Note that here it is crucial to consider ambiguous objects, because both interpretations should be available.

The investigation of the role played by the subject will be articulated in two steps:

1. web-based norming study to elicit different interpretations, and potential subjects for each of the interpretation (simple vs. enriched);
2. reading-time experiment to explore the different contribution of different types of subjects to determine simple vs. enriched composition; materials will be pairs of sentences with ambiguous objects like in 3.15. The experiment will have a one-factor design (subject) with two levels: a) subject selecting for simple composition (3.15a); b) subject selecting for enriched composition (3.15b); we predict b) to show slower reading times than a).

Arbeitspaket 1.3: Aspectual features of the object. It has been observed how Aktionsart features related to the object can play a role in determining the grammaticality of the metonymy and its interpretation:

(3.16) simple composition
a. Peter began the translation. (+durative event noun)
b. *Peter began the leap. (-durative event noun)

(3.17) enriched composition
a. Peter began the book. (+durative entity noun)
b. *Peter began the dictionary. (-durative entity noun)

2by “durative entity” we refer here to entities for which the “non overt event” performed on them is durative: e.g. “read a book” vs. “consult a dictionary”.

17
Arbeitspaket 1.3 aims at evaluating how simple and enriched composition interpretation differ when the object belongs to a dispreferred Aktionsart class. Note that the simple composition in 3.16a should show smaller processing costs than the enriched composition in 3.17a; but what happens in 3.16b and 3.17b? Does the dispreferred object lead to a reinterpretation process? Is reinterpretation in 3.16b more costly than enriched composition in 3.17a and 3.17b?

The investigation of the role played by aspectual features will be articulated in three steps:

1. **selection of a balanced set of stimuli** (durative event nouns, non-durative event nouns, durative entity nouns, non-durative entity nouns);
2. **web-based norming study** to elicitate possible paraphrases/plausibilities for each of the interpretation (simple vs. enriched);
3. **reading-times experiment** to explore the different contribution of different types of durative features to determine simple vs. enriched composition; the experiment will have a two-factors design (object, aspect) with two levels each (entity/event; durative/non-durative).

**Arbeitspaket 2: Computational classification: Simple vs. enriched composition**

The goal of Arbeitspaket 2 is the implementation of broad-coverage, unsupervised corpus-driven models that can predict the experimental data acquired in Arbeitspaket 1. These models need to be able to account for the three factors that presumably influence the interpretation: ontological type of the noun, plausibility, and aspectual properties.

While each of these factors is fairly straightforward to model on its own, our goal is to construct a model that can account for all three factors without the need of individual optimization for each dataset. This situation is ubiquitous not only in cognitive studies, where integrated models have a considerably higher cognitive validity by virtue of their generality, but also in natural language processing, where it is often the case that information from different, heterogeneous knowledge sources needs to be integrated. In this manner, we also address an important current bottleneck in work on corpus-driven models of lexical semantics. Work in recent years has concentrated mainly on deriving models for a wide range of individual lexical-semantic phenomena, but has paid comparatively little attention to the question of consolidating these individual models into general architectures (or at least parameter settings) that are able to solve multiple tasks. This is an important direction to explore (compare, e.g., the ESSLII 2008 workshop on this topic, Baroni et al. (2008)).

We will therefore investigate different architectures for integrating knowledge about our three factors: an integrated architecture, and a modularized architecture, both of which raise their own challenges.

Both of these approaches will of course depend on a good understanding of which surface features correlate with the factors of interest in order to integrate these factors into the models. While we have an account of such features for plausibility (through the clustering-based probability model for plausibility created by D4) and for aspect from Zarcone and Lenci (2008), we need to formalize such an account for the ontological type of nouns (Arbeitspaket 2.1). Subsequently, actual integration of the accounts for the three factors will happen in Arbeitspaket 2.2 and 2.3, respectively.
The focus of this Arbeitspaket is on the modelling of the data created in Arbeitspaket 1, and a detailed analysis of the resulting models. Arbeitspaket 5 will complement this perspective with an externally motivated, entailment-based evaluation.

Arbeitspaket 2.1: Predicting the ontological type for nouns. In the first phase of the SFB, Project B3 has developed a list of local context cues for the sortal disambiguation of German nominalizations (Kountz et al., 2007; Spranger and Heid, 2007). We will use these criteria as a basis for asking three questions:

- How well do these criteria carry over to nouns that are ambiguous between an event and an object reading, but which are not nominalizations (i.e., cases like game)? Are there criteria that are only applicable to nominalizations? Are there other criteria that were not previously identified, but arise from the extension of the focus?

- How well do the criteria of Kountz et al. (2007), which were developed for German, carry over to English? Again, are there criteria that do not apply to English? Are there other criteria in English that did not occur in German?

- How can these criteria be combined into a classifier that predicts event vs. object readings for arbitrary nouns, ideally without any manual annotation? For this purpose, we will look into methods developed for the unsupervised prediction of animacy vs. inanimacy (Övrelid, 2009).

Arbeitspaket 2.2: Modularized architecture. A comparatively straightforward way to arrive at a complete model is a modularized account. After the completion of Arbeitspaket 2.1, we will have three modules corresponding to the three subtasks in Arbeitspaket 1. The goal of Arbeitspaket 2.2 is to integrate them into one single model that makes the correct predictions across all three datasets.

The existence of a number of models for a given task whose potentially divergent predictions must be combined into a single prediction is usually addressed using ensemble learning methods such as weighted voting (Dietterich, 2000) and more recently non-parametric combination methods (Jiménez and Márquez, 2008).

An important contrast with the standard situation, however, is that our models will be informed by psycholinguistic datasets. That is, the individual models are not trained on one common dataset. Rather, each model is trained on a subspace of the training data where only a small number of factors is varied while all others are carefully controlled. As a result, each model will be unaware of the relevance of the factors accounted for by the other models.

In Arbeitspaket 2.2, we will investigate the impact of this situation on the performance of standard ensemble learning methods. We assume that it will be necessary to develop a meta-model for the prediction of model relevance that determines for unseen data items which ones of the available models should participate in the prediction process. The central challenge will be an accurate component of lexical similarity to determine to which data items a model’s predictions should extend, and to which ones they should not.
Arbeitspaket 2.3: Integrated architecture. Our second strategy for modeling the three phenomena will use an integrated architecture, i.e., a single data structure. For this purpose, we will use a vector space model of word meaning (Lund and Burgess, 1996; Landauer and Dumais, 1997). Vector space models yield flexible, corpus-based representations of word meaning that reflect *usage*. As such, their distributional patterns can be expected to encode some of the conceptual knowledge about plausible predicate-argument structures that is necessary for the modelling tasks (Padó et al., 2007). Vector spaces provide fundamentally graded, similarity-based predictions that can however be reified into classifications using nearest neighbour-based methods (Daelemans and van den Bos, 2005).

Vector space models of word meaning exhibit the abovementioned problem of dataset specificity to a large degree. They involve a large number of free parameters (context definitions, context weighting, and dimensionality of the space, among others). Thus, existing studies mainly perform model selection over a small portion of the parameter space using held-out datasets. In this situation, it is an attractive idea to optimize the model on multiple datasets concurrently, in order to remove parameter settings that can only account well for particular datasets. The result will be a smaller set of more generalizable parametrizations.

The exemplar models developed by project A2 represent instances of such a model that has been demonstrated to work across multiple datasets (Schütze et al., 2007). We will therefore collaborate with A2 on transferring insights, architectural ideas, and testable predictions from the A2 model onto our vector space model.

More specifically, we will adopt framework of dependency-based vector spaces models (Padó and Lapata, 2007), where target words are represented in terms of the words in their dependency context (e.g., their subjects, objects and modifiers). The central parameter in this type of model is the set of *context weights*, which quantifies the influence of each context element. These parameters have received little attention in the literature so far, but have arguably a large impact. We will phrase the acquisition of good context by weighting parameters as an optimization problem which allows us to estimate these parameters on the basis of a handful of annotated “seeds”, as has recently been proposed for optimizing random graph walk models (Minkov and Cohen, 2008).

In this setting, the challenge of accounting for three different datasets corresponds to finding a context weighting that accounts simultaneously for the informativity of three different factors. This might turn out not to be feasible. For example, process-modifying predicates are only applicable to events, but not to states (Spranger and Heid, 2007), which indicates that this dimension should be assigned a high weight for the modelling of the object/event data from Arbeitspaket 1.1. At the same time, it is unclear whether this is appropriate for modelling (e.g.) the plausibilities from Arbeitspaket 1.2. If it should turn out to be impossible to find one generic parametrization that works across datasets, we will explore graph-based models that have been proposed in Baroni and Lenci (2009) which can encode different, interacting “layers” of similarities.

Arbeitspaket 3: Human recovery of covert events

The goal of this Arbeitspaket is to investigate the actual range of interpretations that speakers/hearers produce in enriched composition, and its internal semantic structure. Previous studies have concentrated on cloze completion (where only one answer was elicited, e.g.
McElree et al. (2001); Traxler et al. (2002)) or on presenting participants with pre-chosen constructions and measuring either correlates of their processing load, or the plausibility of provided interpretations (e.g. Lapata and Lascarides (2003)).

Our general hypothesis is that the range of interpretations for an enriched composition construction is structured in a small number of verb “clusters”, which correspond to salient events and which are closed under entailment. Consider as an example the construction finish a cake, for which we expect two main clusters to be elicited:

- \{baking, preparing, \ldots\} – the “creation” event
- \{eating, devouring, munching, \ldots\} – the “consumption” event

If true, this result will have important ramifications for the representation of recovered events (i.e., their “complex” paraphrases). To test this hypothesis, we will gather a large dataset of interpretations arising in enriched composition. We will proceed in three steps, all of which rely solely on offline data and are therefore comparatively fast and cheap to set up:

1. We will analyze the conceptual structure of the set of possible interpretations (Arbeitspaket 3.1);

2. We will compare the distribution of the elicited events to the distribution of elicited subjects (Arbeitspaket 3.2);

3. We will obtain plausibility ratings for two classes of events that are not named by participants ( (a) entailed events, e.g. read → interpret; (b) related but non entailed events, e.g. read → skim) in order to analyse what semantic relations hold between possible events (Arbeitspaket 3.3);

Another issue addressed by this Arbeitspaket will be the analysis of how much the semantic structure of the space of covert events can be valid for a language other than English like Italian. The offline web-based experiments planned for Arbeitspaket 3 will be carried out for both English and Italian, in order to gather insights on the cross-linguistic variance of the space of covert events in logical metonymy.

**Arbeitspaket 3.1: Semantic structure of interpretations.** The first issue we will investigate is the range of interpretations that speakers produce in a given enriched composition context. Do the interpretations fall into one or several clusters? Which internal structure does the space of the interpretation show? We expect the interpretations to fall into small number of clusters closed under entailment (cf. above).

Arbeitspaket 3.1 will proceed in three steps:

1. **elicitation of a set of interpretations:** given a set of V + Obj stimuli (e.g. begin a poem), in a web-based experiment subjects will be asked to provide as many non-overt events as possible within a certain time frame (e.g. begin a poem → reading, writing, composing).
2. **elicitation of entailment judgments**: the elicited events will be the materials for a second web-based experiment, where subjects will be given pairs of events obtained from step 1 (e.g. $i_1$ begin reading a poem, $i_2$ begin writing a poem, $i_3$ begin composing a poem) and will be asked to judge the entailment between these interpretations. E.g., does $i_1$ imply $i_2$? does $i_1$ imply $i_3$? does $i_3$ imply $i_2$? etc.

3. **clustering**: The judgments from step 2 impose a structure onto the events elicited in step 1. We expect that for each V + Obj stimulus, the events will form a small number of meaning clusters. These clusters will form the input for Arbeitspakete 3.2 and 3.3.

**Arbeitspaket 3.2: Subjects vs. event clusters.** In Arbeitspaket 3.1, we have considered the semantic structure of recovered events, considering only the matrix verb and its object, while ignoring the potential impact of the matrix verb’s subject on interpretation (Lapata and Lascarides, 2003). Our hypothesis is that the specification of a subject is able, as a rule, to resolve the ambiguity that arises from the presence of several meaning clusters (as investigated in Arbeitspaket 3.1). For example, begin writing/composing a poem $\rightarrow$ poet; begin reading a poem $\rightarrow$ student; finish baking/preparing a cake $\rightarrow$ cook; finish eating/devouring a cake $\rightarrow$ child.

Arbeitspaket 3.2 will be articulated in two steps:

1. **elicitation of subjects**: given each V + Vcov + Obj stimulus (e.g. begin writing/reading a poem, finish eating/baking a cake), in a web-based experiment speakers will be asked to provide a plausible subject for the given stimulus (e.g. begin writing a poem $\rightarrow$ poet; begin reading a poem $\rightarrow$ student; finish baking a cake $\rightarrow$ cook; finish eating a cake $\rightarrow$ child);

2. **plausibility rating**: we will combine the subjects with the interpretation clusters that they were elicited for, and with “mismatching” interpretations from other clusters, and elicit plausibility ratings for both. Our expectation is that the mismatching subjects will receive a lower rating than the matching subjects.

**Arbeitspaket 3.3: Boundaries of acceptability: basic level effect, semantic relatedness.** The aim of Arbeitspaket 3.3 is to test how the set of interpretations that was actually provided by participants generalizes with regard to semantic relations. Assuming the example begin a poem, if speakers produced reading but not an hypernym like interpret, how plausible would they consider the hypernym? And would a semantically related but not implied term like understand show comparable plausibility?

Arbeitspaket 3.3 will proceed in two steps:

1. **enrichment of stimuli set**: stimuli set $A$ obtained from Arbeitspaket 3.1 will be enriched with stimuli sets $B$ and $C$. Stimuli set $B$ will contain a hypernym for each cluster in $A$ (e.g. read $\rightarrow$ interpret); stimuli $C$ will contain a semantically-related but yet non implied term (a sister term or a troponym, e.g. read $\rightarrow$ skim); $B$ will be divided into $B_1$, containing direct hypernyms, and $B_2$, containing more abstract events, because we expect verbs in $B_1$ to be more salient than abstract events in $B_2$ due to the basic level effect (Rosch et al., 1976);
2. **plausibility rating**: elements in A, B and C will be the materials for a web-based plausibility rating experiment. The hypothesis is that elements in A and B will show comparable plausibility ratings, and higher plausibility ratings than C; categories at the middle level of an ontology are more salient, so we expect higher plausibility ratings for B₁ than for B₂.

**Arbeitspaket 4: Computational prediction of covert events**

The goal of Arbeitspaket 4 is to develop a computational model of covert event interpretation that does not make independent predictions for individual verbs, but instead returns a concept representation for the covert event that mirrors the semantic structure of human interpretations found in the data from in Arbeitspaket 3.

In contrast to the distinction between simple and enriched composition (Arbeitspakete 1 and 2), which involves a number of factors, we assume that the determination of appropriate interpretations for covert events is primarily a question of making appropriate plausibility judgments for predicate-argument relations. These are determined by the lexical-semantic properties of the predicate and its argument(s) as well as the broader discourse context.

One of the simplest possible models of this type is the interpretation model proposed by Lapata and Lascarides (2003). It uses conditional probability as a proxy for plausibility. However, in addition to the problem of estimating conditional probabilities for low-frequency items, this model only returns predictions for individual verbs, and thus cannot make guarantees about the semantic coherence of the predicted interpretations.

**Arbeitspaket 4.1: Similarity-based modelling and analysis.** As an alternative, we will investigate the use of knowledge-lean vector space-based models of plausibility (Padó et al., 2007). This model can describe plausible fillers for argument positions of predicates as density functions in a vector space, and in this manner return a semantically coherent set of fillers. We have previously shown that this model can also be applied inversely, to predict plausible events given an argument (Erk and Padó, 2008), which is crucial for the prediction of event interpretations. A central issue for the interpretation of covert events with a vector space model is the extent to which the topology of the vector space corresponds to the set of actually produced interpretations. Since this is far from certain, the focus of subtask 4.1 will therefore be on analysing its performance on the specific task of modelling the structure of the set of interpretation from Arbeitspaket 3.1, and later on in the “boundary case” data from Arbeitspaket 3.3, always comparing the performance of our model against the frequency-based baseline. Individual steps will include:

- We will recover the individual meaning clusters (cf. Arbeitspaket 3.1) in the vector space through clustering.

- We will characterize the relations between “neighbouring verbs” in each cluster in terms of semantic relations between events. This space will presumably contain spurious verbs, therefore we will investigate ways of removing unwanted semantic relations either by finding suitable parametrizations of the space, or by filtering out unwanted candidates. Conversely, the clusters will miss verbs produced by our informants, which we will attempt to recover.
Vector space models yield graded plausibility predictions, but we require cut-off criteria to return finite lists of interpretations. We will investigate which criteria (e.g., similarity thresholds (Schütze et al., 2007), or outlier detection (Erk, 2006)) yield the best correspondence to human judgments. This task will again draw on experiences of the project A2 with their exemplar models, as well as our own experiences from work on Arbeitspaket 2.

Arbeitspaket 4.2: Combining the influence of objects and subjects. The goal of this subtask is to model the data collected in Arbeitspaket 3.2, i.e., the influence of the matrix verb’s subject on the set of interpretations for the covert event. This will involve integrating plausibility considerations from the subject and object. If our hypothesis from AP 3.2 is correct (subjects generally correspond to unique interpretation clusters), then the only role of the subject representations in vector space should be to select one cluster from among the set of object-determined clusters created by the model from 4.1, without modifying them.

However, a more general perspective on the contribution of the subject is as an instance of prototype-based concept combination (see e.g. Osherson and Smith (1981); Hampton (1991); Kamp and Partee (1995)), where the event prototype(s) evoked by the object, and the event prototype(s) evoked by the subject, are combined into a single composite event. In this sense, the strategy outlined in the previous paragraph is a type of activation-based concept combination where the subject information activates the best-matching event, but does not modify its representation. However, it has been argued that such modification (also called integration) is a necessary feature of concept combination (see e.g. Smith et al. (1988)). Such a scheme can be formalized in a vector space model as addition or multiplication of the density functions, corresponding to the intersection and union sets of the most plausible object-evoked and subject-evoked events, respectively. We will compare the results of the two combination schemes and analyse them.

Arbeitspaket 4.3: Hybrid models: Integration of external knowledge. The analysis of the data in Arbeitspaket 3 will have yielded an understanding of the semantic relations between events in each interpretation cluster. As described there, our hypothesis is that the verbs of each cluster will be related by entailment relations. This suggests that models of covert event interpretation should be able to take advantage of well-established resources encoding entailment relations, such as WordNet, as well as recent work on the large-scale acquisition of lexical entailment relations from corpora (Szpektor et al., 2004; Akhmatova and Dras, 2009). This subtask will investigate methods for combining knowledge from these resources with the vector-space model developed in 4.1 and 4.2.

Since the two approaches are largely independent, we expect their combination to surpass the quality of either model alone. As an illustration, one possible mode of combination would be to adopt only the most plausible verbs from the vector space, and then supplement this list based on the external knowledge resources. We will analyse such combination strategies with respect to the precision/recall tradeoff that they offer.

If Arbeitspaket 3 will show that entailment is not the main (or the only) semantic relation among verbs in an interpretation cluster, then the focus of 4.3 will shift from entailment relations to other semantic relations as necessary.
**Arbeitspaket 4.4: Cross-lingual modelling**  
As stated above, the experiments in Arbeitspaket 3 will be carried out for both English and Italian to verify the validity of our results across languages. Within this multi-lingual setup, the advantage of using vector space models is that construction of these models does not require any resources other than large parsed corpora, which are available for Italian as well. This allows us to perform the model analyses and experiments in 4.1 and 4.2 in parallel for English and Italian.

On the other hand, no large entailment resources (used in 4.3) are as yet available for Italian; and for other resource-poor languages, no high-accuracy parsers are available either. Subtask 4.4 will therefore use the bilingual English/Italian dataset as a test case for the cross-lingual induction of covert event interpretations. Ideally, such an approach allows the induction of models for target languages with only minimal resource requirements in the target language.

We will induce a vector space whose dimensions and vectors have both an English and an Italian label so that the space can be induced from an English parsed corpus, but make predictions about Italian verbs (Rapp, 1999; Widdows et al., 2002). We will begin by developing a method to induce a bilingual space without the need for a parallel corpus, e.g. by bootstrapping dimensions. Next, we will compare the quality of the predictions for Italian with the performance of an Italian-only model. Finally, we will experiment with methods for combining cross-lingual with monolingual information where the latter is available.

**Arbeitspaket 5: Representation and evaluation**

The purpose of Arbeitspaket 5 is to make the results of APs 1–4 available beyond our particular modelling framework, vector spaces (AP 5.1), and to establish the interpretation of covert events as a phenomenon of interest in the entailment community (AP 5.2).

**Arbeitspaket 5.1: Ontology-based representation of covert event interpretations**

Arbeitspaket 4 develops a vector space-based representation of covert event interpretation. Even though we will aim at ensuring that these models cover all interpretations produced by our human informants, corpus-driven models always inherit the limitations of the corpora they are constructed from, and it cannot be guaranteed that they yield all desirable generalizations.

Arbeitspaket 5.1 will specifically investigate another option for representing the event interpretations elicited in Arbeitspaket 3, namely by reference to the WordNet ontology. In this endeavour, we share the intuitions of project B5 that the existing ontology is insufficient in that its verbal part encodes only troponymy relations, but not other lexical relations like entailment (cf. AP 3.1.).

We will consider a sample of covert event interpretations for selected short variants. We will determine in what respects the English WordNet ontology will have to be extended (e.g., with a broader array of hyponymy relations) or modified to make it possible to specify human-produced event interpretations by reference to contingent subtrees of the ontology. For this task, we will collaborate with project B5, which will have pioneered methods for a similar purpose for French, and test to what extent these methods can be used to (semi-)automate the extension of the ontology for English as well.
Arbeitspaket 5.2: Covert event recovery as textual entailment  Up to this point, the focus of project D6 has been on obtaining data on human interpretation of covert events, and on modelling this data. For practical reasons, evaluation will proceed separately for the simple/enriched composition decision (AP 2) and for event recovery in enriched composition (AP 4). Arbeitspaket 5.2 will provide an evaluation setting that comprises both interpretation processes and will position our work in the more general context of semantic processing and inference.

We will do so by phrasing the interpretation of short forms as a textual entailment task. Textual entailment (Dagan and Glickman, 2004) has gained momentum as an evaluation setting for semantic processing components that attempts to avoid commitment to specific representations and to applications. The task that processing components are asked to solve is, given a text and a hypothesis, to determine whether the text entails the hypothesis or not. Text and hypothesis are provided in unanalysed natural language, and the notion of entailment that is assumed is “common sense” entailment (i.e., grounded in human judgment, and not by necessity indefeasible). Various NLP tasks (among them IE, QA, and MT Evaluation) can be reduced to or at least informed by the integration of entailment recognition components (Dagan et al., 2009).

Since 2004, yearly workshops have been organized as forums for the development of approaches to textual entailment, always involving a shared dataset for quantitative evaluation. Until now, these datasets have always been “general-purpose”, sampled from application-motivated datasets but involving essentially unlimited language variation. Following up on earlier discussions (Zaenen et al., 2005), and in order to become more relevant to the lexical semantics community, it was decided at this year’s ACL TextInfer workshop to extend the shared tasks at future workshops with phenomenon-specific datasets.

We will take this opportunity to bring our work to the attention of the semantic processing and inference community. We will propose to the workshop organizers to use covert event recovery as the phenomenon-specific dataset in the 2013 shared task. This dataset can be created by pairing “short variants” with their interpretations to obtain positive instances. With respect to negative instances (that is, short variants combined with interpretations that they do not support), we will build on the work from AP 1.3 and 3.3 to sample different classes of negative examples – random combinations, combinations that are ungrammatical due to aspectual constraints, and combinations that are semantically close but unacceptable.

Arbeitspaket 6: Identifying and modelling new constructions involving covert events

Arbeitspakete 1 to 5 concentrate on investigating enriched composition for one construction, namely logical polysemy. While this construction is not rare in corpora (Briscoe et al., 1990), it is not ubiquitous either. Our intuition is that the concept of covert events, and the models that we have developed for their interpretation, are applicable to a range of other phenomena and are thus of broader interest to lexical semantics and NLP than if the recovery of covert events were an isolated phenomenon.

Arbeitspaket 6 will address this question of generalizability in two steps. In 6.1, we will directly model a phenomenon where we strongly suspect the existence of covert events, namely Italian da-compounds. 6.2 is a more exploratory task, namely the mining of parallel
corpora for configurations that involve unrealized events, and the question of generalization from individual corpus examples to phenomena.

Arbeitspaket 6.1. Specific constructions: Interpretation of Italian da-complex nominals. There is a substantial body of work on the interpretation of compound nouns and complex nominals. In computational terms, the task is almost invariably phrased as a classification problem where classes correspond to the semantic relations that hold between the parts of the compound (Rosario and Hearst, 2001; Nastase and Szpakowicz, 2003; Moldovan et al., 2004). Next to the problem of agreeing on an inventory of relations – proposals vary between 5 and more than 30 relations – we note that computational work on modelling complex nominals is limited almost exclusively to English, where complex nominal construction is highly productive.

Arbeitspaket 6.1 will consider the interpretation of complex nominals in Italian, where the formation of complex nominals is considerably more restricted (Busa and Johnston, 1996). Complex nominals are constructed with post-modifying prepositional phrases, a frequent choice being N da N:

\[(3.18)\]

\[\begin{align*}
a. & \quad \text{bread knife / coltello da pane} \\
& \quad \text{shoe box / scatola da scarpe} \\
& \quad \text{parcel paper / carta da pacchi} \\
& \quad \text{women's shoes / scarpe da donna} \\
& \quad \text{hunting knife / coltello da caccia} \\
& \quad \text{race car / macchina da corsa} \\
& \quad \text{sewing machine / macchina da cucire}
\end{align*}\]

Nominals with da fall into two broad categories, (a) and (b), which appear to correspond well to the simple/enriched composition split in logical metonymy. The instances in 3.18a can all be understood as cases of covert events that are recovered in interpretation (e.g. bread knife \(\rightarrow\) knife to cut bread; shoe box \(\rightarrow\) box to store shoes); not surprisingly, Busa and Johnston (1996) have suggested a Generative Lexicon account of these cases. In contrast, the examples in 3.18b have a “purpose” interpretation that does not support paraphrasing by any specific event beyond a generic “use for” relation.\(^3\) We think that the reason for the unavailability of a covert event is the use of an event noun as modifier, analogously to the mechanism in logical metonymy. We will ask the following two central questions:

- On the cognitive level, does the interpretation of Italian da-nominals show similar properties to the interpretation of logical metonymy as investigated in Arbeitspaket 1? That is, do da-nominals with object-denoting modifiers (which presumably involve enriched composition) show increased processing times in contrast to nominals with event-denoting modifiers (which presumably do not)?

- On the computational level, what modifications have to be made to the models for the interpretation of Italian logical metonymies from AP 4 (in particular, 4.2) to be applied to the interpretation of da-nominals? One regularity of logical metonymies

\[^3\text{In logical polysemy, instances of simple composition can also support paraphrasing by a generic “use for” relation.}\]

27
that does not carry over is that the syntactic positions with respect to the matrix verb in a logical metonymy correspond to those of the recovered event. This is not true for da-nominals: compare shoe box → box contains shoes with women’s shoes → women wear shoes.

Arbeitspaket 6.2. Mining parallel corpora for covert events. Arbeitspaket 6.2 will analyse corpora in order to obtain candidates for constructions that can be understood as involving the recovery of covert events, with the aim of determining the quantitative impact of the models we develop in D6 for broad-coverage semantics interpretation, e.g. in the context of natural language processing. Alternatively, the goal of 6.2 can be understood as constructing a corpus-based characterization of the notion of covert event.

The basis for this task will be formed by large, bilingual, word-aligned parallel corpora. Such corpora can be understood as providing two potentially different views on the underlying content (Dyvik, 1998). Correspondence patterns in such corpora have therefore been used widely to uncover other types of non-overt information like word sense (Diab and Resnik, 2002), paraphrases (Bannard and Callison-Burch, 2005), or multi-word expressions (Zarrieß and Kuhn, 2009). In this context, our initial working characterization of covert events will be “verbs or deverbal nouns in one sentence that cannot be directly aligned to verbs or deverbal nouns in its translation”.

Clearly, this characterization encompasses many spurious instances of unaligned events that have nothing to do with covert events. Some of them arise from imperfections in alignment, but others are connected to the simplistic nature of the initial definition, or are merely cases of idiosyncratic translation. We will therefore refine the characterization as follows:

- We will remove spurious instances by developing shallow filters that are based on parts of speech (e.g., systematic translations of verbs by adjectives, Padó and Erk (2005)) and symmetry properties, since we would expect that “true” covert events are able to arise in both directions of translation.

- We will develop structure-based filters that incorporate our hypotheses about the nature, expressed in terms of dependency structure, of the “short variant” and “long variant” of constructions in which covert events are required. An initial, rather restricted, characterization is that the “short variant” consists of two head words in a direct dependency relation, while these words are separated by a new predicate in the “long variant”. The results of AP 3 will yield additional constraints with regard to the distribution of events in the “long variant”.

- Finally, we will develop methods to determine whether such structures are idiosyncratic, or are instances of a general construction. One test that we will use is to treat the parallel corpus as a basis for predicting unseen “short variants” and testing whether these short variants actually exist in a test set. As test set, we will use monolingual web corpora, which exist in large quantities, in order to alleviate false negatives that arise from data sparsity.
### 3.4.4 Zeitplan

<table>
<thead>
<tr>
<th></th>
<th>Jahr</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>1 Human choice in simple/enriched composition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.1 Entity- vs. Object-denoting objects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.2 The influence of subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.3 Aspectual constraints</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Classifying simple/enriched composition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1 Predicting ontological types</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.2 Integrating plausibility</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.3 Integrating aspectual well-formedness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Human recovery of covert events</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.1 Semantic structure of interpretation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.2 Interpretations and subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.3 Boundaries of acceptability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Computational prediction of covert events</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.1 Similarity-based modelling and analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.2 Combining objects and subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.3 Hybrid models</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.4 Cross-lingual modelling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Representation and evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.1 Ontology-based representation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.2 Covert event recovery as textual entailment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 New constructions with covert events</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.1 Interpretation of Italian da-compounds</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.2 Mining parallel corpora</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Work in AP 2 follows the corresponding subtasks in AP 1 with a 6 month delay for data collection, with the exception of AP 2.1 which can start on the transfer of predictors for ontological status right away.

- The data collection phase in AP 3 will take place during the first part of 2012 to provide data for APs 4 and 5. Analysis of the data will be spread out over the rest of 2012 and 2013.

- AP 4 will start in 2010/11 with a phase of creating infrastructure for the two subtasks with additional requirements (entailment resources for 4.3, bilingual vector spaces which build on some Italian infrastructure in 4.4). The bulk of the work will happen between 2012 and 2014.

- AP 5 will also take place in 2013/2014.

- In AP 6, subtask 6.1 forms one block to be worked on in 2013. 6.2 can start right away in 2010, with a second phase of activity in 2013/14 to apply the models developed in earlier APs to the identified data.
3.5 Stellung innerhalb des Sonderforschungsbereichs

3.5.1 Stellung zum Gesamtkonzept des SFBs

This project investigates logical metonymies as an example of enriched composition. This phenomenon involves underspecification on two levels: First, the syntactic surface form is twofold ambiguous between a “simple composition” reading and an “enriched composition” reading which involves a covert event. Second, the covert event, in turn, is initially completely underspecified. Note that this is not a “classical” ambiguity in the sense of a choice between a fixed number of possible readings; rather, there are (potentially infinite) sets of lexical items that make appropriate and inappropriate covert events.

We focus on the lexical-semantic parameters of the specification steps that take place in the interpretation of logical metonymy, investigating cognitive and computational questions in lockstep.

- With regard to the first step of interpretation, we aim at gaining insight how different conceptual properties of the lexical material (Aktionsart properties of events, ontological types of nouns, plausibility of predicate-argument structures) interact, and subsequently ask what an integrated architecture for a practical, data-driven model has to look like in order to explain all factors.

- With regard to the second step of interpretation, we investigate the semantic constraints on the range of interpretations, the pairwise relations between alternative realizations of recovered events, and the relationship between these relations and the entailment concept.

3.5.2 Interaktion mit anderen Teilprojekten

<table>
<thead>
<tr>
<th>TP</th>
<th>Beziehung</th>
<th>Inhalt der Austauschbeziehung aus Sicht von &lt;TP&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2</td>
<td>Collaboration</td>
<td>Application of A2’s Memory Sequence Model to simple/enriched composition decision (AP 2); Application of A2’s Unified Exemplar Model to recovery of covert events (AP 4)</td>
</tr>
<tr>
<td>B3</td>
<td>Input</td>
<td>Factors for disambiguating nominalizations</td>
</tr>
<tr>
<td></td>
<td>Output</td>
<td>Automatic acquisition of regular/inverse selectional preferences</td>
</tr>
<tr>
<td>B5</td>
<td>Collaboration</td>
<td>Common interest in certain verbs (like “begin”); Representation of covert events in terms of enriched ontology</td>
</tr>
<tr>
<td>D4</td>
<td>Collaboration</td>
<td>Application of the D4 plausibility model to logical metonymy data</td>
</tr>
</tbody>
</table>

Cooperation with other scientists. The project plans to cooperate with Prof. Dr. Katrin Erk at the Department of Linguistics, University of Texas, Austin. Katrin Erk and Sebastian Pado have worked together on vector space-based models of plausibility and concept combination, which will be continued in the context of computational Arbeitspakete, in particular AP 4.

The project plans to cooperate with Dr. Alessandro Lenci at the Department of Linguistics, Università di Pisa. Alessandro Lenci and Alessandra Zarcone have a history of collaboration on both computational and psycholinguistic models of event types, which
provides a fertile ground for future collaboration in the context of both cognitive and computational Arbeitspakete, in particular in the analysis of enriched composition in Italian and in the development of distributional models.

### 3.6 Abgrenzung gegenüber anderen geförderten Projekten der Teilprojektleiterinnen und Teilprojektleiter

None.

### 3.7 Ergänzungsausstattung für das Teilprojekt

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TV-L E13</td>
<td>TV-L E13</td>
<td>TV-L E13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TV-L E13 50%</td>
<td>TV-L E13 50%</td>
<td>TV-L E13 50%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>StHi</td>
<td>StHi</td>
<td>StHi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>zus.:</td>
<td>zus.:</td>
<td>zus.:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2013</th>
<th>2014/1</th>
</tr>
</thead>
<tbody>
<tr>
<td>TV-L E13</td>
<td>TV-L E13</td>
</tr>
<tr>
<td>TV-L E13 50%</td>
<td>TV-L E13 50%</td>
</tr>
<tr>
<td>StHi</td>
<td>StHi</td>
</tr>
<tr>
<td>zus.:</td>
<td>zus.:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Kleingeräte (bis EUR 10.000,-)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbrauchsmaterial</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Versuchstiere</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reisen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sonstiges</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Investitionen (Geräte über 10.000,- EUR (brutto))</th>
<th>Investitionsmittel insges.</th>
<th>Investitionsmittel insges.</th>
<th>Investitionsmittel insges.</th>
<th>Investitionsmittel insges.</th>
<th>Investitionsmittel insges.</th>
</tr>
</thead>
</table>

1) P = Personalmittel

Tabelle wird von der Geschäftsführung anhand der Angaben des Teilprojekts erstellt.
### 3.7.1 Personal im Teilprojekt

<table>
<thead>
<tr>
<th>Name, akad.Grad Dienststellung</th>
<th>engeres Fach des Mitarbeiters</th>
<th>Institut der Hochschule oder der außeruniv. Einrichtung</th>
<th>Beantragte Förderperiode</th>
<th>Entgeltgruppe</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mitarbeit im Teilprojekt in Wochenstunden (beratend:B)</td>
<td></td>
</tr>
</tbody>
</table>

#### Grundausstattung

3.7.1.1 wissenschaftl.
Personal
(einschl. Hilfskräfte)

3.7.1.2 nichtwissenschaftl.
Personal

#### Ergänzungsausstattung

3.7.1.3 wissenschaftl.
Personal
(einschl. Hilfskräfte)

3.7.1.4 nichtwissenschaftl.
Personal

---

1) Bitte durchnumerieren und Aufgabenbeschreibung nachfolgend erläutern

2) Bitte Verfahrensgrundsätze der DFG zur Bezahlung wissenschaftlichen Personals beachten.

Tabelle wird von der Geschäftsführung anhand der Angaben des Teilprojekts erstellt.
Aufgabenbeschreibung von Mitarbeiterinnen und Mitarbeitern der Grundausstattung für die beantragte Förderperiode
zu 1: Juniorprofessor Dr. Sebastian Padó
  Principal investigator; coordinates project D6, oversees collaboration with other projects and is responsible for D6

Aufgabenbeschreibung von Mitarbeiterinnen und Mitarbeitern der Ergänzungsausstattung für die beantragte Förderperiode
zu 1: Alessandra Zarcone (TVÖL 13/2)
  PhD student, half-time position. Linguistics.
  Design and realization of experiments (Arbeitspakete 1 and 3).
  Interpretation of compounds and generalization (Arbeitspaket 6)
zu 2: N.N. (TVÖL 13/2)
  PhD student, half-time position. Computer science or Computational Linguistics.
  Work on computational models (Arbeitspakete 2 and 4)
  Representation and evaluation of these models (Arbeitspaket 5)
zu 3: N.N. (Stud. HK)
  Student assistant, computational linguistics.
  Support for setup and data analysis for experimental work;
corpus analysis; implementation work

3.7.2 Aufgliederung und Begründung der Sachmittel (nach Haushaltsjahren)

<table>
<thead>
<tr>
<th></th>
<th>2010/2</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014/1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sächliche Aufwendungen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grundausstattung</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ergänzungsausstattung</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Für Sächliche Aufwendungen stehen als Grundausstattung voraussichtlich zur Verfügung:

Für Sächliche Aufwendungen werden als Ergänzungsausstattung beantragt (entspricht den Gesamtsummen „Sachmittel“ in Tabelle 3.7):

Tabelle wird von der Geschäftsführung anhand der Angaben des Teilprojekts erstellt.

Begründung zur Ergänzungsausstattung der Sachmittel

Gesamt: 4x4000 = 16000 EUR

Lesezeitenmesser mit Software (http://www.cedrus.com/ordering/europe.htm) 600 EUR
Honorierung der Versuchsteilnehmer in Arbeitspaket 1 (alle Experimente zusammen):
Online-Experimente: 200 Teilnehmer à 5 Euro 1000 EUR
Offline-Experimente: 120 Teilnehmer à 5 Euro 600 EUR

Honorierung der Versuchsteilnehmer in Arbeitspaket 3 (alle Experimente zusammen):
Offline-Experimente für Englisch: 200 Teilnehmer à 5 Euro 1000 EUR
Offline-Experimente für Italienisch: 200 Teilnehmer à 5 Euro 1000 EUR

Honorierung der Versuchsteilnehmer in Arbeitspaket 6.1:
Online-Experimente: 100 Teilnehmer à 5 Euro 500 EUR

Reise nach Pisa zur Erhebung von italienischen Daten für AP 3 (2 Wochen im Jahr 2012) 2000 EUR
Reise nach Pisa zur Erhebung von italienischen Daten für AP 6 (2 Wochen im Jahr 2013) 2000 EUR
Literaturverzeichnis


