Fabienne Martin, Dennis Spohr, Achim Stein

Representing a Resource of Formal Lexical-Semantic Descriptions in the Web Ontology Language

This paper presents an approach to disambiguating verb senses which differ wrt. the inferences they allow. It combines standard ontological tools and formalisms with in-depth formal semantic analysis and is therefore more formalised and more detailed than existing lexical semantic resources like WordNet and FrameNet. The resource presented here implements formal semantic description of verbs in the Web Ontology Language OWL and exploits its reasoning potential based on Description Logics for the disambiguation of verbs in context. After a thorough discussion of the theoretical motivation as well as the manual formal semantic analysis, we present details on the disambiguation process, which is based on a mapping from the French version of EuroWordNet to the Suggested Upper Merged Ontology. In addition to this, we focus on the selectional restrictions of verbs wrt. the ontological type of their arguments, as well as their representation as necessary and sufficient conditions in the ontology. Finally, we discuss how we make use of the Semantic Web Rule Language SWRL in order to calculate the inferences that are permitted on the selected interpretation.

1 Introduction

Verbs raise a number of challenges for computational linguistic applications, two of which will be addressed in this paper. First, a lot of them are highly polysemous, which makes a careful disambiguation a prerequisite for the application of semantic web technologies. As an example, the French verbs *encourager* and *pousser* are normally translated by different verbs in German, as illustrated for *encourager* in (1)-(3), and for *pousser* in greater detail in section 2:

(1) Un terroriste a encouragé$_1$ ma voisine à poser une bombe dans la cave.
A terrorist has encouraged my neighbour to place a bomb in the basement.
Ein Terrorist hat meine Nachbarin ermutigt, eine Bombe im Keller zu legen.

(2) La lettre a encouragé$_2$ ma voisine à poser une bombe dans la cave.
The letter has encouraged my neighbour to place a bomb in the basement.
Der Brief hat meine Nachbarin dazu bewegt, eine Bombe im Keller zu legen.

(3) Le gouvernement a encouragé$_3$ la recherche sur les armes biologiques.
The government has encouraged research on biological weapons.
Die Regierung hat die Erforschung biologischer Waffen angeregt.

Note that (1) differs from (2) only by the ontological category of the subject, which is a human entity in (1) and a non-human one in (2).[1] On the other hand, (3) differs from the previous two in that its object is a non-human entity (while the object denotes a human in (1) and (2)), and in that it does not take an infinitival complement.

The second challenge concerns the computation and the weighting of the inferences triggered by verbs. The contrast between (1) and (2) offers a striking example: while the interpreter of (2) can take for granted that a bomb was placed, she can only *guess* that it was possible in (1). Let us call "actuality entailment" (AE) the entailment triggered by (2) – and to which the interpreter assigns the probability $p = 1$ – that an event satisfying the infinitival complement took place, and "weak inference" the one triggered by (1) – and to which the interpreter assigns a probability p between 0 and 1. Furthermore, we will say that when the AE is triggered, the verb instantiates its "implicative reading", and that it receives its non implicative reading otherwise.[2]

Note that the presence of the non-human subject in (2) is only a necessary condition to trigger the AE (and consequently the implicative reading). The tense of the sentence has also to be of a certain kind (namely a perfective tense, of which the *passé composé* is an example in French) for the AE to arise. The interaction between lexical semantics and information pertaining to the textual level like tense and aspect must then be modelled appropriately to capture the facts.

A model which allows to identify and weigh appropriately the inferences triggered by verbs like *encourager*, *pousser* etc. is highly desirable, since, first of all, verbs of this class are pervasive in the lexicon and heavily present in official texts.[3] Second, rating accurately the inference that an event described by a constituent took place is central for the understanding of texts, or the recognition of textual entailment.

As it is more convenient to present the implementation through specific polysemous verbs, we will first detail the lexical semantics of two specific semantically close verbs at hand, namely *pousser* 'push' and *encourager* 'encourage'.

Plan of the paper. The paper is organised as follows. Section 2 presents the theoretical part of this work by delineating the different readings of the transitive *pousser*, as well as identifying the factors allowing their disambiguation. In addition to this, the semantic differences between *pousser* and *encourager* are discussed. In Section 3, the technical part of the work starts with an introduction to the necessary background. Section 4 discusses the model in detail, focussing on the implementation of selectional argument restrictions as well as the formalisation of inference rules. Section 5 shows how these mechanisms are applied to an example sentence in order to disambiguate

[1] In (2), *encouragé*$_2$ could also be translated figuratively with *ermutigt* instead of *bewegt*.

[2] The terminology is borrowed from Karttunen (1971). Note however that verbs like *encourager* differ from what Karttunen originally calls "implicative verbs" (e.g. *réussir à*, 'manage to'), because the latter trigger an AE with any kind of tenses.

[3] Examples of verbs displaying the alternation between (1) and (2) (and thus triggering inferences of different strength in the two places) are *autoriser à P* 'to authorise to *P*', *inviter à P*, 'to invite to *P*', *aider à P*, 'to help to *P*', *permettre de P*, 'to allow to *P*', *suggérer de P*, 'to suggest to *P*', *exiger que P*, 'to demand that *P*'.

pousser and its syntactic arguments, and to calculate the inferences on the basis of this selection. We conclude in Section 6.

2 Theoretical Background and Motivation: Forceful Verbs in Spatial, Psychological and Conceptual Domains

2.1 Transitive use of *pousser*

2.1.1 Indicators for sense disambiguation

Ontological categories. On its transitive use, *pousser* has roughly four readings, respectively illustrated by the following examples:

(4) Il a poussé Paul vers la porte.
He pushed Paul towards the door. (literal, physical meaning #1, syn. *faire bouger*)

(5) Le pianiste a poussé Fauré vers Brahms.
Lit. *The pianist pushed Fauré towards Brahms.* (figurative, conceptual)
Fig. *The pianist brought the music of Fauré closer to the music of Brahms.*

(6) Il a poussé Paul vers le crime/à tuer.
He pushed Paul towards the crime/to kill. (figurative, psychological)

(7) Il a poussé des cris perçants.
He pushed penetrating cries. (literal, physical #2, syn. *émettre, produire*)

A first obvious indicator allowing to identify the targeted reading is the ontological category of the arguments. The one in (7) is the easiest to disambiguate, since it is automatically selected with an object denoting a sound object (a song, a cry, a spoken word). Disambiguating the other readings is less trivial. First, the physical reading is only available when the subject is able to trigger a physical movement by itself, i.e. is animate. It is for instance excluded in (9), but possible in (8).[4]

(8) Pierre/le vent/la fourmi a poussé y. (physical reading available)
Pierre/the wind/the ant pushed y.

[4] Note that a sentence like *La tristesse a poussé Paul au cimetière*, lit. 'Sadness pushed Paul to the cemetery' is interpreted as describing a physical movement of Pierre, which seems to go against our claim that the physical reading is not available with a inanimate subject. However, we argue that this sentence contains an ellipsis of an à-COMP (*La tristesse a poussé Paul à aller au cimetière*, lit. 'Sadness pushed Paul to go to the cemetery'), which selects the psychological reading of *pousser*. A movement is then interpreted because the AE is triggered with an animate subject and the *passé composé* tense. Two arguments militate in favour of this view. First, the AE disappears with the *imparfait* (the corresponding imperfective sentence does not entail anymore that Paul went to the cemetery), whereas a movement is still entailed with the 'real' physical *pousser*. Second, sentences of this type strongly require the preposition *à* and become deviant with other types of PPs (cf. ??*La tristesse l'a poussé dans le cimetière*, 'Sadness pushed him within the cemetery'), which again suggests that *à* is introducing a hidden infinitive here.

(9) La tristesse/la faim/l'inflation a poussé y. (no physical reading)
The sadness/the hunger/the inflation pushed y.

Second, when the object denotes an abstract entity, the conceptual reading is automatically selected, as in (10). The physical reading is excluded because one cannot physically push an abstract entity, and the psychological one is out because an abstract entity cannot be an Experiencer.

(10) x a poussé ma faim/les prix vers/jusqu'à z. (conceptual reading)
x has pushed my hunger/the prices to the point of z.

When the object denotes a physical entity, all readings are *ceteris paribus* available. The PP then becomes the disambiguating constituent: if the noun contained in the PP denotes an abstract entity, the reading is automatically figurative (psychological or conceptual), cf. (11).

(11) x a poussé y jusqu'à la rage/vers le romantisme. (no physical reading)
x has pushed y to the point of fury/towards romanticism.

To obtain the psychological reading, the object must denote an Experiencer, and the noun contained in the PP must denote a subkind of abstract entities, namely acts or dispositions to act. As e.g. *romantisme* does not pertain to this subset of abstract entities, its presence suffices to select the conceptual reading.

Other indicators. Ontological categories however do not always suffice to isolate the targeted reading, because a lot of nouns are themselves polysemous. The syntactic frame is a further relevant disambiguating indicator. First, the presence of an *à*-COMP suffices to select the psychological reading:

(12) x a poussé y à dormir/à tuer Paul. (psychological reading mandatory)
x pushed y to sleep/to kill Paul.

More interestingly, with an animate subject, the physical readings remain the only ones available in absence of a PP. In other words, with an animate subject, the figurative reading generally makes the PP compulsory. Therefore, leaving out the parenthesised component in the examples in (13) changes the interpretation of *pousser* from figurative to physical. This is also true of other movement verbs like *tirer* 'pull', *jeter* 'throw' or *entrer* 'enter', cf. (13).[5] Note that this fact casts doubt on the idea supported e.g. by Asher and Lascarides (1995) that the syntax remains the same under the metaphorical use of verbs:

(13) a. Le pianiste a poussé Fauré #(vers Brahms). (figurative, conceptual meaning)
Lit.: *The pianist pushed Fauré (towards Brahms).*
Fig.: *The pianist brought (the music of) Fauré closer to (the music of) Brahms.*

[5] To our knowledge, data of this kind are sporadically observed (cf. e.g. Adler and Asnès, 2005), but not explained yet. See Martin (2008) for a theoretical explanation.

b. Albert a poussé Yves #(dans la dépression).(figurative, psychological meaning)
Albert pushed Yves (into depression).

c. Fourier a jeté les étudiants #(dans l'athéisme).
Fourier threw the students (into atheism). (figurative, psychological reading)

d. Pierre entre #(dans l'endettement). (figurative, social reading)
Pierre is entering (into indebtedness).

Apart from some exceptions, the absence of the PPs thus suffices to select the physical reading with polysemous nouns. Therefore, in a sentence like (14), *a poussé* can only be understood as 'physically pushed'.[6]

(14) Le pianiste a poussé le compositeur. (physical reading)
The pianist pushed the composer.

2.1.2 Types of inferences

The inferences triggered by *pousser* on its transitive use basically depend on three factors, namely (i) the ontological type of the subject, (ii) the tense of the verb (cf. Section 1) and (iii) the presence of a PP and the preposition heading it.

On its psychological reading, one of the most relevant inferences concerns the occurrence of the action described in the *à*-COMP or the *à*-object (what we call the AE in the introduction). As already observed above, under its psychological reading, *pousser* entails the occurrence of an action of the Experiencer iff the subject denotes an inanimate entity and the tense used is perfective; cf. (15) as well as (1) to (3).[7]

(15) a. Marie a poussé$_{perf.}$ Paul au suicide/à se suicider.
Marie pushed Paul to commit suicide.
$\not\rightarrow$ Paul is dead.

b. Son divorce a poussé$_{perf.}$ Paul au suicide/à se suicider.
His divorce pushed Paul to commit suicide.
$\rightarrow$ Paul is dead.

c. Son divorce poussait$_{imperf.}$ Paul au suicide/à se suicider.
His divorce was pushing Paul to commit suicide.
$\not\rightarrow$ Paul is dead.

The physical reading also has an 'implicative' and a 'non-implicative' reading, although in another way. As already observed by Jackendoff (1990), *push* does not entail a movement of the Theme ((16) is not contradictory), and the same is true of the French verb *pousser* (cf. Stein (2007) and his example (17)). A movement of the Theme is thus at most a weak inference triggered by *pousser*.

[6] Recall that when the subject is inanimate, the physical reading is not available anyway. The presence of another argument than the subject and the object is thus not necessary to select the figurative reading (cf. e.g. *L'inflation a poussé les prix*, 'the inflation pushed the prices').

[7] Bhatt (1999) and Hacquard (2006) already observed contrasts of the same type for modal verbs, but Mari and Martin (2007) were the first to observe that it expands to non modal verbs like *pousser*.

(16) I pushed the car, but it didn't move.

(17) J'ai poussé la boîte, mais elle n'a pas bougé.
I pushed the box, but it didn't move.

However, in presence of certain PPs (as e.g. *jusque/dans/à*-PPs, but not with *vers/dans la direction de*-PPs), this weak inference becomes an entailment:

(18) J'ai poussé la voiture jusque/dans le garage, #mais elle n'a pas bougé.
I pushed the car until/in the garage, but it didn't move.

(19) J'ai poussé la voiture vers le/dans la direction du garage, *OK* mais elle n'a pas bougé.
I pushed the car towards/in the direction of the garage, but it didn't move.

These contrasts show again that it is crucial (i) to distinguish between strong and weak inferences (entailment or defeasible *a priori* assumptions) and (ii) to differentiate, for a same verb on a same reading, several sets of inferences varying with the syntactic context.

2.2 Semantic representation of transitive *pousser*

The fact that geometrical notions alone do not suffice to model the meaning of all spatial prepositions is already well-documented. For instance, Vandeloise (1991) showed that 'inclusion' cannot alone define the preposition *in*, whose meaning also involves force-dynamic notions like *causation*, *control* or *interaction*. Recently, Zwarts (2007) argued that the same is true of what he calls "forceful verbs" like *push* and *pull*. As he emphasises, spatial notions like *movement*, *direction* and *location* cannot capture the fact that *push* is obviously opposite to *pull*. Indeed, one cannot say that *push* and *pull* describe a movement that goes in opposite direction, since as already noticed above (examples 16 and 17), *push* does not entail a movement of the Theme, and the same is true for *pull*. It is not even clear that these verbs entail a movement of the Agent (arguably, exerting a pressure on something without moving can still be a *pushing/pulling*). According to Zwarts, it is rather the *forces* involved which go in opposite directions: the force is pointing away from the Agent in the case of *push*, and towards him with *pull*.

2.2.1 Vectors

Following Zwarts (2007), we will model the informal notion of force with the help of the mathematical notion of vectors. A vector v is an arrow, i.e. a directed line segment. *Free* vectors have a length and direction, but no starting point. *Located* vectors have a starting point $st\text{-}p(v)$. $st\text{-}p(v)=x$ means that x defines the starting point of the vector v. On the other hand, forces have two parameters, namely a *magnitude* $m(v)$ and a *direction* $dir(v)$. $dir(v) = y$ indicates that the vector v points into the direction of

y. Here, we will slightly augment the formalisation of Zwarts in order to define more precisely the direction of v. $dir(v) = y, z$ indicates that the vector v points into the direction of y and is parallel to the line joining y to z. This allows one to define on which side of y the pressure is exerted.

The two parameters m and d define a vector (the length of the vector co-varies with the magnitude of the force it represents). The *location* $loc(v)$ of the force corresponds to the physical point where the force is exerted. After Talmy (1985) and following Zwarts (2007), we will call *Antagonist* the Agent exerting the force, and *Agonist* the Patient on which the force is exerted.

Objects often tend to move by themselves in a particular direction. In that case, the Agonist has also its own force vector, which represents its inherent tendency to move by itself (generally downwards, because of gravitation). A *resultant vector* determines the result of the interaction between the forces of the Antagonist and the Agonist. This sum can be zero, when the forces of the Antagonist and the Agonist are equal but opposite.

Let us see now how Zwarts defines the meaning of *push* and *pull* with the help of vectors. Informally, their meaning can be represented in the following way. The arrows represent the force vector.

(20) *push*: Antagonist $\Rightarrow$ Agonist

(21) *pull*: Antagonist $\Leftarrow$ Agonist

It can be modelled more formally in a vector model. Let us assume that the vector v_{sp} represents the *spatial* relation between the Antagonist and the Agonist pointing from the Agonist to the Antagonist. v_{sp} provides the spatial frame for the *force* vector v_f. What *push* and *pull* express is the way v_f is aligned with respect to v_{sp}: these two vectors are opposite for *push*, and point in the same direction for *pull*. If no other forces are interacting, the Agonist will move in the direction of the force vector, i.e. away from the Antagonist in the case of *push* (and in its direction in the case of *pull*).

We observe that the definition of Zwarts rightly predicts that *push* is not appropriate when the context indicates that the force vector v_f points into the direction of the Antagonist (cf. (22)) and that *pull* is not appropriate when it indicates v_f points into the opposite direction of the Antagonist (cf. (23)):

(22) #J'ai poussé le landau vers moi.
I pushed the buggy towards me.

(23) #J'ai tiré le landau vers le mur en face de moi.
I pulled the buggy against the wall opposite of me.

2.2.2 Lexical representations

We propose to assume that forceful verbs introduce an implicit vector argument. This argument can be targeted by modifiers like *hard* (cf. *push hard*). $\mathsf{Source}(v, e)$ means that the event e is the source of the vector force v. We can then define the two forceful movements *pousser* 'push' and *tirer* 'pull' in the following way. Recall that *st-p(0)=x* means that x is the starting point of the vector v, $loc(v){=}y$ that y corresponds to the entity on which the force is exerted, and $dir(v){=}y, z$ that v is parallel to the line joining y and z. In absence of a spatial complement, z is left underspecified and the interpretation by default given the context is chosen. In presence of such a spatial PP, z is denoted by the noun in the PP.

(24) $pousser_{tr} \rightsquigarrow \lambda z \lambda y \lambda x \lambda e \lambda v[\text{Antagonist}(e,x) \wedge \text{Agonist}(e,y) \wedge \text{Source}(v,e) \wedge \text{st-p}(v) = x \wedge \text{loc}(v) = y \wedge \text{dir}(v) = y, z \wedge m(v) > 0]$

(25) $tirer \rightsquigarrow \lambda z \lambda y \lambda x \lambda e \lambda v[\text{Antagonist}(e,x) \wedge \text{Agonist}(e,y) \wedge \text{Source}(v,e) \wedge \text{st-p}(v) = x \wedge \text{loc}(v) = y \wedge \text{dir}(v) = x, z \wedge m(v) > 0]$

It is specified that the magnitude of the force vector must be superior to zero, because a *pushing/pulling* of zero magnitude is not a *pushing/pulling*. Note that this representation is underspecified wrt. the domain (e.g. spatial, psychological, social or conceptual) on which *pousser* or *tirer* are used.

2.2.3 Presupposition

This representation is still too weak though. To see it, let us compare *pousser y* and *glisser y* ('slide y'), which is arguably another forceful verb on its transitive use. A first difference, irrelevant here, is that *glisser y* entails a change of state of y, which explains the contradiction in (26):

(26) Il a glissé la lettre, #mais elle n'a pas bougé.
He slid the letter, but it didn't move.

The important point is that we do not want to say that *glisser y* is equivalent to *pousser y* modulo this entailed change of state. Indeed, while *glisser y* indicates that no other forces interact with the one initiated by the Antagonist to make y move, *pousser y* presupposes such an interacting resistant force. For instance, *pousser la lettre dans la boîte* ('push the letter in the box') will be chosen in a context where the opening of the box is stuffed with something which somehow blocks the insertion of the letter in it. In a context where resistance is totally absent, the use of *pousser* is not appropriate, cf. (27):

(27) Le vent a poussé OK la fumée épaisse qui se dégageait de l'incendie/ #l'air pur du matin.
The wind has pushed the thick smoke which spread from the fire / the clear morning air.

(28) J'ai tiré le livre de la bibliothèque.
I've pulled the book from the bookshelf.

Similarly, (28) is fully appropriate only if the book is somehow blocked in its position on the bookshelf (because it is between two books for instance).[8] On the contrary, *glisser* in *glisser la lettre dans la boîte* ('slide the letter in the box') will be preferred when nothing makes obstacle to the movement initiated by x. The presence or absence of a resistant force also differentiates *pousser/tirer* from *glisser* in the non-spatial domain:

(29) L'interprète a fait glisser Cage vers Satie.
The interpreter let Cage slide towards Satie.

(30) L'interprète a tiré/poussé Cage vers Satie.
The interpreter pulled/pushed Cage towards Satie.

For instance, (30) suggests that some force crosses the one initiated by the interpreter who wants to narrow the conceptual distance between the music of Cage and Satie, while (29) precisely suggests the absence of such a force (*la voie est libre*, 'the coast is clear').[9]

It is pointless to define the interacting vector by its position with regard to the spatial/conceptual vector linking x to y, because it can vary considerably from one context to another. But we can simply state that if v takes place, a vector v' interacting with v and of a non-null magnitude is also present.

Importantly, the entering into play of the interacting vector is not *asserted* by the forceful verbs. It rather seems to be presupposed. Look for instance to the following sentences:

(31) Il est possible qu'il ait poussé la lettre dans la boîte.
It is possible that he had pushed the letter in the box.

(32) L'interprète n'a pas poussé Brahms vers Fauré.
The interpreter hasn't pushed Brahms towards Fauré.

The speaker of (31) clearly assumes that the box entrance is stuffed with some matter which *would* originate in an interacting vector if somebody tried to push something in it, and the speaker of (32) assumes the existence of a conceptual obstacle between Brahms and Fauré which *would* trigger an interacting vector if somebody tried to narrow the two. It confirms the presuppositional nature of the inference, since (31) and (32) are classical presuppositional environments. Note that what is presupposed is not the interacting vector itself, but the obstacle between y and z that could trigger it.

Interestingly, the presupposition is kept under the intransitive use of *pousser*. For instance, when it means *aller* 'to go', the intransitive *pousser* presupposes that the

[8] Intriguingly, no interacting force is implied by *retirer*. For instance, while *J'ai retiré la plume du bureau* ('I withdrew the feather from the desk') is fine, *#j'ai tiré la plume du bureau* ('#I pulled the feather off the desk.') is odd precisely because it is difficult to imagine a force resisting to the one of the Antagonist. Another difference between the two is that *retirer* but not *tirer*, is a hyponym of *enlever*, which explains its incompatibility with directional PPs.

[9] Note that the periphrastic causative in *faire*+inf. is almost compulsory in (29) to trigger the non-physical reading of *glisser*. The lexical causative *L'interprète a glissé Cage vers Satie* oddly suggests that the musician physically pushed Cage towards Satie. We will ignore these data here.

displacement was made difficult by an obstacle manifesting a certain resistance to the Performer of the movement. This explains why (33) is acceptable whereas (34) is strange when uttered out of the blue, since going as far as the kitchen generally does not consist in a big achievement:

(33) Je suis allée jusqu'à la cuisine.
I went until the kitchen.

(34) #J'ai poussé jusqu'à la cuisine.
Lit.:*I pushed until the kitchen.*
Int.:*I went until the kitchen.*

2.3 Comparison between *pousser* and *encourager*

Obviously, it is on the psychological reading that *pousser* resembles more *encourager*. The analysis presented above allows us to easily capture the differences between the two near-synonyms. First, *encourager* is not classified as a movement verb, and as such is not acceptable with strictly directional PPs:

(35) *Pierre m'a encouragé vers le crime.
Pierre encouraged me towards the crime.

Second, *encourager* does not trigger the presupposition described above. This difference is responsible for the following contrast[10]:

(36) L'entraîneur a encouragé/?poussé l'équipe à gagner.
The trainer has encouraged/pushed the team to win.

(37) L'entraîneur a encouragé/?poussé l'équipe à bien jouer.
The trainer has encouraged/pushed the team to play well.

(38) Il a bien voulu m'encourager/?me pousser à le faire.
He has been so kind as to encourage/push me to it do.

In the psychological domain, the occurrence of an interacting vector presupposed by *pousser* translates in the following way: it indicates that x and y entertain inverse preferences; x wants y to do P (P corresponding to the proposition denoted by the infinitive), while the speaker or y prefers $\neg P$ (the 'conflict' between the two preferences can be conceived as the psychological translation of the 'resistance' between the Pusher and another entity). This explains why *pousser* becomes deviant when x and y normally entertain the same preferences.

[10] Of course, the contrast is only present if we do not assume that the team wants to lose. In any case, it is safe to say that the resistance of the team seems to be greater with *pousser* than with *encourager*.

3 Technical Background

Since the theoretical background has been established, we will now turn to the more technical part of this work. In this section, we will briefly outline the technical background, i.e. the main features of the formalisms as well as the lexical-semantic and ontological resources used in this study.

3.1 Formalisms

The formalisms that we use for building our resource have been developed in the field of the *Semantic Web*, a research area devoted among others to providing tools and formalisms for assigning meaning to web content (Berners-Lee et al., 2001). In particular, we make use of the Web Ontology Language OWL (Bechhofer et al., 2004) and the Semantic Web Rule Language SWRL (Horrocks et al., 2004).

3.1.1 OWL

The Web Ontology Language (Bechhofer et al., 2004) is a formalism based on the Resource Description Framework RDF[11] and can be expressed in XML syntax. Its main building blocks are classes (corresponding to one-place predicates in first-order logic) and properties (two-place predicates), which are both structured hierarchically and inherit restrictions (such as axiomatic definitions of classes or, in the case of properties, formal characteristics like being functional or transitive) to their subclasses and subproperties respectively. In addition to these entities, there are also individuals, which are simply instances of classes. OWL comes in three sublanguages, which differ wrt. their expressivity: OWL Lite is the least expressive sublanguage and allows for simple class definitions; OWL DL is based on Description Logics (DL), a decidable fragment of first-order logic (see e.g. Baader et al., 2003), which restricts the use of some OWL constructs in order to maintain decidability of reasoning; OWL Full is the most expressive sublanguage and imposes no restrictions on the language constructs, however at the cost of decidability. For example, in OWL Full it is possible to express that a class is an instance of another class, which is disallowed in OWL DL.

Resources defined in OWL Lite and OWL DL can be interfaced with a Description Logic reasoner (e.g. Pellet; Sirin et al., 2007) on the one hand to check the consistency of the resource, and on the other hand to infer further statements on the basis of explicit statements in the resource. Such reasoning tasks are e.g. the classification of the taxonomy in order to infer logical subclass relationships, or the classification of individuals for inferring the classes they implicitly instantiate, based on necessary and sufficient conditions in the axiomatic definitions of classes.

[11] http://www.w3.org/RDF/

3.1.2 SWRL

The Semantic Web Rule Language (Horrocks et al., 2004) adds expressivity to OWL in that it allows for the expression of Horn-like rules, i.e. disjunctive rules with at most one positive literal, as in the two equivalent formulae in 39 and 40.

(39) $\neg hasFather(x, y) \vee \neg hasBrother(y, z) \vee hasUncle(x, z)$ (Horn clause)

(40) $hasFather(x, y) \wedge hasBrother(y, z) \rightarrow hasUncle(x, z)$ (equivalent SWRL rule)

SWRL can be expressed directly in OWL syntax – so the resulting documents are still OWL compliant – and the rules can be interpreted and executed by tools such as the Jess® rule engine[12].

3.2 Lexical-semantic and ontological resources

3.2.1 EuroWordNet

The EuroWordNet project (Vossen, 1998) aimed at providing resources similar to Princeton WordNet (Fellbaum, 1998) for eight European languages, all of which are connected through an interlingual index (ILI) that contains a set of language-independent concepts. The ILI is linked to the so-called EuroWordNet Top Ontology, an upper-ontology-like collection of features that have been designed to describe the lexical-semantic relations in the wordnet. The French version of EuroWordNet contains roughly 8,300 verb senses and 24,500 noun senses, which are organised into 22,745 synonym sets and linked using lexical-semantic relations like hyponymy and meronymy.

In contrast to the scale of the resource in terms of covered senses, the detail of description is generally limited to taxonomic relations between synonym sets and does not include information on argument structure. However, the probably biggest drawback of the French EuroWordNet lies in its inaccuracy and even partial incorrectness, mainly wrt. the verbal descriptions. Therefore, only the noun hierarchy can be considered as a useful starting point for building other lexical resources, whereas the verb hierarchy can only provide a rough sketch as to the interpretation and organisation of the senses.

Other resources. Apart from EuroWordNet, there is no large-scale lexical resource of French that provides qualitatively adequate lexical-semantic analyses. While resources such as FrameNet and VerbNet (Baker et al., 1998; Kipper-Schuler, 2006) exist for English, none of these have been extended to French in a comparable way so far.

3.2.2 SUMO

Together with DOLCE (see below), the Suggested Upper Merged Ontology (Niles and Pease, 2001) is one of the most widely used ones in the NLP community, among others due to the fact that mappings have been created to Princeton WordNet (Niles and

[12] http://www.jessrules.com/

Pease, 2003) and the EuroWordNet ILI (Spohr, 2008a). SUMO comes with MILO, a mid-level ontology, as well as domain ontology extensions, which in total contain 20,000 terms and 70,000 axioms. While originally implemented in SUO-KIF – a formalism intended as first-order language – SUMO has also been translated to OWL Full, with the attempt to preserve as much as possible of the original axiomatisation.

Despite its quantitative size and degree of formalisation, SUMO has been criticised primarily wrt. the usability of its axiomatisations, since they are questionable from a modelling perspective (e.g. instances being concepts at the same time and relations being modelled as concepts). Moreover, SUMO seems to lack a clear theoretical basis, as it adopts ideas from different ontological theories (Sonntag et al., 2007).

3.2.3 DOLCE

The Descriptive Ontology for Linguistic and Cognitive Engineering is an upper-level ontology that has been designed with a strong cognitive bias (Gangemi et al., 2003a). Its classes and the relations among them have been implemented with the OntoClean methodology (Guarino and Welty, 2002), which gives the resource a formally and theoretically more solid basis than e.g. SUMO. As was mentioned above, DOLCE has also been mapped to Princeton WordNet (Gangemi et al., 2003b).

DOLCE is the first reference module of the WonderWeb library of foundational ontologies, and it has a number of extensions (e.g. an ontology of information objects). In total, DOLCE and its extensions comprise roughly 200 classes and 300 properties, and they are available as OWL versions.

4 Implementation of the Model

In the following, we will describe how we model the inference triggers mentioned in Section 2 as well as formal semantic representations like the one depicted in (24) using OWL DL and the more expressive SWRL.

4.1 Encoding of selectional argument restrictions

As was mentioned in the introduction, the primary triggers for selecting one particular meaning over another is the presence (or absence) of syntactic arguments as well as their ontological type. For example in (2), the fact that (i) *encourager* subcategorises an infinitive, (ii) the subject is inanimate, and (iii) the direct object is animate, determine the sense of *encourager* in this sentence. In order to make this information available and processable, we use a straightforward encoding of these triggers as conditions on class definitions (see also the representations in Franconi (2003)), based on conceptual classes of the DOLCE-Lite-Plus ontology and its extensions (in the examples below prefixed by *dol:* for DOLCE and *edns:* for “Extended Descriptions and Situation”).

In particular, the different senses of a verb are modelled as subclasses of a general class that denotes an underspecified representation of the verb. However, the different verb senses are only subclasses of this generic representation in the lexicon, not in our concept hierarchy, since verb senses very frequently denote different concepts that are not subsumed by a common concept. (41) below shows the class definition of the conceptual sense of *pousser* that corresponds to the one used e.g. in sentence (5) above.

(41) $pousser\_conceptual \equiv pousser$
$$\exists\ subj\ (\exists\ canDenote\ edns{:}agentive\text{-}social\text{-}object)$$
$$\exists\ obj\ (\exists\ canDenote\ dol{:}abstract)$$
$$\geq\ 3\ arg\ owl{:}Thing$$

The formalisation is to be interpreted as follows: in order to be classified as an instance of *pousser_conceptual*, it is both necessary and sufficient to be an instance of *pousser*, with a subject that can denote an agentive social object, with a direct object that can denote something abstract, and with at least one more argument (i.e. the number of values for *arg* – which is the superproperty of *subj*, *obj* and further argument properties – is at least 3; *owl:Thing* just refers to "any kind of entity"). The predicate *canDenote* used in the formalisation captures the polysemy of the nominal argument, since the classes that represent nouns contain as axioms the ontological concepts they can denote, such as e.g. the class *compositeur* with the axiom ∃ *canDenote edns:agentive-social-object*. So in other words, the subject part of the example above states that the value of the *subj* property of pousser has to be an instance of a class that can denote an agentive social object.

The properties *subj* and *obj* have been defined as *functional properties*, i.e. they can only have one value. Thus it suffices to use the existential quantifier in the axiomatisation here. However, in other cases it is necessary to use a combination of the existential and the universal quantifier, e.g. in order to express that for *pousser_physical* – irrespective of the particular syntactic configuration – it is necessary that all arguments denote something that is not abstract. Conversely, it would not suffice to use just the universal quantifier here, since that would be trivially satisfied in a case where *pousser* is used without any arguments. Thus, a definition of *pousser_physical* has to contain at least the following statements.

(42) $pousser\_physical \equiv pousser$
$$\exists\ arg\ (\exists\ canDenote\ \neg dol{:}abstract)$$
$$\forall\ arg\ (\exists\ canDenote\ \neg dol{:}abstract)$$
$$\ldots$$

The general motivation for the encoding shown in (41) and (42), which views the contextual triggers discussed above as necessary and sufficient conditions, is that a DL reasoner can infer – on the basis of a particular setting of contextual parameters (i.e. property values) – the specific type of an instance of the generic *pousser*. In the following, we will discuss the inference rules that are attached to each sense class, and which are evoked once a specific sense has been determined.

4.2 Inference rules

As was mentioned above, the different senses of *pousser* do not only differ wrt. necessary and sufficient conditions that are used to classify them, but also wrt. the inferences that may be drawn from them. In our resources, such inferences are encoded in the form of SWRL rules (see e.g. O'Connor et al., 2005), as they require inference capacities which go beyond the scope of the inventory provided by OWL DL. In order to keep the following discussion as simple as possible, we will restrict ourselves to explaining the inference rule that corresponds to the semantic description of transitive *pousser* given in (24). The SWRL rule is shown in Table 1 below, with the rule body in lines 1 to 6 and the rule head in lines 7 to 19.

The first line represents the configuration in which the rule is applicable, i.e. an instance of *pousser* with grammatical subject and object. Lines 2 to 6 make use of the SWRL extensions built-ins[13] defined within the Protégé ontology editor (Knublauch et al., 2004) in order to create the instances that are to be inserted into the representation, based on the description in (24). In line 7, a *PUSHING* event is asserted. In lines 8 and 9, the grammatical subject and object are asserted as the antagonist and agonist of the event denoted by *pousser*. Lines 10 and 11 assert a vector which has as its source the *PUSHING* event. Lines 12 to 15 assert locations which correspond to the location of the grammatical subject, the grammatical object and the underspecified entity z respectively (cf. page 7). In addition to this, the location of the subject is further the starting point of the vector (16), the location of the object is the location of the vector (17), i.e. the location where the force is exerted, and line 18 specifies that the direction of the force of the vector is parallel to the line that joins the grammatical object and the underspecified entity z, i.e. parallel to a line that has the location of y and z as points. Finally, line 19 states that the magnitude of the vector is 1.

5 Disambiguation and Calculation of Inferences

In order to select the correct reading of a verbal predicate in a sentence like (14) and, moreover, to generate the appropriate semantic representation on the basis of this choice, our system passes a number of distinct analysis steps. Basically, the system receives input from a syntactic parser and tries to determine the correct senses of both the verbal predicate and its syntactic arguments, before calculating the inferences permitted on this interpretation. The whole analysis process is summarised in Figure 1 below.

For the scope of this paper, we will ignore details on the syntactic analysis that precedes the semantic processing steps, and instead assume a syntactic parser which returns output like the one depicted in Figure 1, providing information on the predicate (*pousser*), its syntactic arguments (*pianiste* and *compositeur*), its modal context (e.g. embedding under *pouvoir*, 'can, be able to'), and the tense in which the predicate is

[13] See `http://protege.cim3.net/cgi-bin/wiki.pl?SWRLExtensionsBuiltIns`; the built-in function `createOWLThing` has been replaced with `cOT` in the table. One could say that `createOWLThing` represents the existential quantifier, although its interpretation is somewhat stronger since actual instances are asserted and created in the resource.

1	syntactic configuration required for application of rule	`pousser(?e) ∧ subj(?e,?x) ∧ obj(?e,?y) ∧`
2	create vector for representing the force	`swrlx:cOT(?v,?e) ∧`
3	create spatial location of grammatical subject	`swrlx:cOT(?locx,?x) ∧`
4	create spatial location of grammatical object	`swrlx:cOT(?locy,?y) ∧`
5	create underspecified entity	`swrlx:cOT(?z,?e) ∧`
6	create spatial location of underspecified entity	`swrlx:cOT(?locz,?z) ∧`
		→
7	assert pushing event	`PUSHING(?e) ∧`
8	assert grammatical subject as antagonist of the pushing	`antagonist(?e,?x) ∧`
9	assert grammatical object as agonist of the pushing	`agonist(?e,?y) ∧`
10	assert a vector	`VECTOR(?v) ∧`
11	assert the pushing event as source of the vector	`source(?v,?e) ∧`
12	assert locations	`LOCATION(?locx) ∧ LOCATION(?locy) ∧ LOCATION(?locz) ∧`
13	assert location of the grammatical subject	`hasLocation(?x,?locx) ∧`
14	assert location of the grammatical object	`hasLocation(?y,?locy) ∧`
15	assert location of the grammatical object	`hasLocation(?z,?locz) ∧`
16	assert the location of the grammatical subject as the starting point of the vector	`hasStartingPoint(?v,?locx) ∧`
17	assert the location of the grammatical object as the location of the vector	`hasLocation(?v,?locy) ∧`
18	assert the direction of the force of the vector to be parallel to the line joining the antagonist and the agonist (cf. page 7)	`LINE(?l1) ∧ LINE(?l2) ∧ hasPoint(?l1,?locy) ∧ hasPoint(?l1,?locz) ∧ parallel(?l1,?l2) ∧ hasDirection(?v,?l2) ∧`
19	assert the magnitude of the vector	`hasMagnitude(?v,1) ∧`

Table 1: SWRL rule corresponding to the semantic description of *pousser* in (24)

used[14]. These context features are crucial for determining the inferences that may be drawn, and thus play an important role in the semantic processing steps which build on the syntactic analysis (see below).

Disambiguation of the predicate and its syntactic arguments. In order to select the correct sense of the verbal predicate, we first disambiguate its syntactic arguments. For this, we apply a similar methodology to that presented in Spohr (2008b). In essence, this approach makes use of a mapping between the French EuroWordNet – a lexical semantic resource for French (EWN; see e.g. Vossen (1998)) – and SUMO (Niles and Pease, 2003), and that has recently been extended to provide a mapping to DOLCE.

[14] For a more detailed discussion of the syntactic analysis, the reader is referred to Boullier and Sagot (2005).

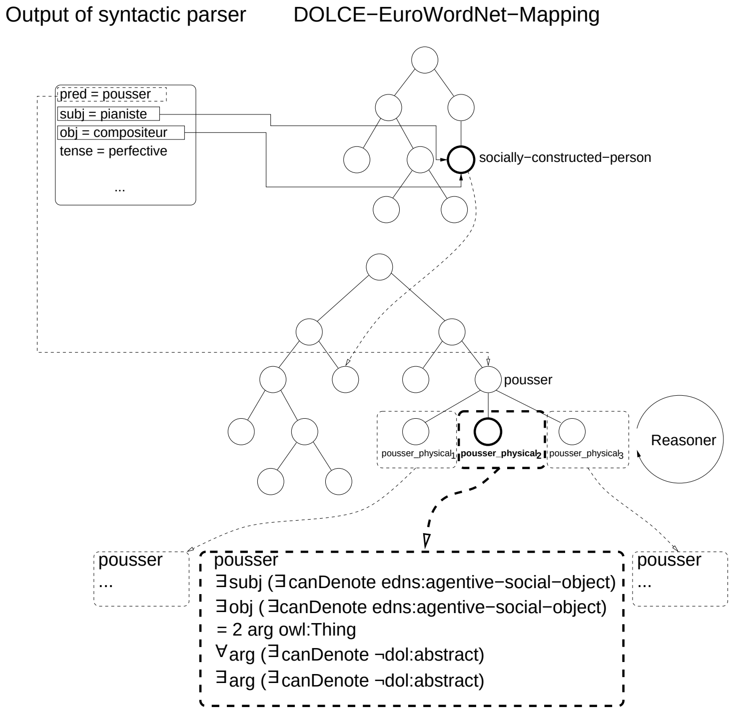

Figure 1: Schema of the process of determining the intended sense of *pousser* in (14) from syntactically parsed input

On the basis of this mapping, selectional preferences are calculated and expressed in terms of ontological concepts, rather than EuroWordNet synsets. Thus, by applying this methodology to a verb like *pousser*, we obtain lists of selectional preferences wrt. the ontological types of its subject and object (see top righthand corner of Figure 1). For the actual disambiguation, the different senses of the subject (*pianiste* in the present case) are looked up in the DOLCE-EWN mapping, and the sense scoring highest in the corresponding selectional preference list is selected. The words are then asserted as instances of the respective EWN classes (in this case *pianiste_1* and *compositeur_2*), which in this case have the necessary condition ∃ *canDenote soc:socially-constructed-person*, a subclass of *edns:agentive-social-object* (cf. (41) above).

The output of the process of disambiguating the arguments is, of course, not entirely deterministic. However, when viewed from the highly abstract level of ontological concepts, the senses distinguished in EWN are very often still closely related so that their sense distinctions have no impact on the interpretation of the verbal predicate and thus the selection of the appropriate sense. For example, although there are two senses of *compositeur* distinguished in EWN – the "non-musical" *compositeur* being that in the sense of a typographer –, they are still subsumed under the common DOLCE class *socially-constructed-person*, which suffices to select the correct sense of *pousser* irrespective of the particular interpretation of *compositeur*. Therefore, even though some of the arguments may be disambiguated towards the wrong sense, the interpretation of the verb sense stays the same and thus the inferences drawn on the basis of this selection remain unaffected. Therefore, even though some of the arguments may be disambiguated towards the wrong sense, the interpretation of the verb sense stays the same and thus the inferences drawn on the basis of this selection remain unaffected.

Once the syntactic arguments have been disambiguated, they are linked to the instance representing *pousser*. The intermediate representation obtained from the operations so far looks as follows.

(43) $pousser(e) \wedge subj(e, pianiste) \wedge obj(e, compositeur) \wedge pianiste\_1(pianiste) \wedge compositeur\_2(compositeur)$

The next step consists in determining the correct sense of *pousser*. As was mentioned in Section 4 above, selectional restrictions have been implemented as necessary and sufficient conditions on class definitions, which allows a reasoner to infer the type of the instance on the basis of these conditions. With the configuration shown in (43), the reasoner[15] can infer the instance of *pousser* as being of the more specific type $pousser\_physical_2$, as this is the only class which satisfies the condition of having agentive social objects as grammatical subject and object without any further arguments (cf. bottom of Figure 1).

Calculation of inferences. The assertion of a transitive *pousser* in combination with $subj(e, pianiste) \wedge obj(e, compositeur)$ causes the SWRL rule in Table 1 above to fire, so that the relevant inferences can be calculated and inserted into the resource. For this task we used version 7 of the Jess® rule engine[16] (see e.g. Golbreich and Imai (2004)). The result of the rule application is given below.

(44) $PUSHING(e) \wedge antagonist(e, pianiste) \wedge agonist(e, compositeur) \wedge VECTOR(v) \wedge source(v, e) \wedge LOCATION(locx) \wedge LOCATION(locy) \wedge LOCATION(locz) \wedge hasLocation(pianiste, locx) \wedge hasLocation(compositeur, locy) \wedge hasLocation(z, locz) \wedge hasStartingPoint(v, locx) \wedge hasLocation(v, locy) \wedge LINE(l1) \wedge LINE(l2) \wedge$

[15] We have used version 1.5.1 of the Pellet OWL DL reasoner (Sirin et al., 2007).
[16] http://www.jessrules.com/

$hasPoint(l1, locy) \wedge hasPoint(l1, locz) \wedge parallel(l1, l2) \wedge hasDirection(v, l2) \wedge hasMagnitude(v, 1)$

6 Conclusion

In this paper, we have presented an approach to modelling polysemous verbs, using standard formalisms such as OWL (Bechhofer et al., 2004) and SWRL (Horrocks et al., 2004). We have shown how the disambiguation of these verbs and their arguments can be performed in this model, and how inferences can be calculated and inserted into a representation that is capable of being interpreted by tools developed for the Semantic Web, such as the ontology editor Protégé (Knublauch et al., 2004).

The approach we propose has a number of advantages. One of these is that a very fine-grained distinction of senses based on contextual features enables accurate annotation of particular senses, and with it the calculation of inferences allowed by the respective sense. In addition to this, our approach combines an implementation of formal semantics with up-to-date technology for semantic processing and is therefore more formalised and more detailed than existing lexical semantic resources. The major drawback of our approach is, of course, the large amount of manual work required for the in-depth lexical-semantic analysis.

Although the system is – due to lack of broad coverage – not yet in a state of being applied to sophisticated reasoning tasks such as the RTE challenge (Recognising Textual Entailment; Dagan et al. (2005)), the inclusion of the contained knowledge into existing systems designed for such tasks seems very promising nonetheless. The RTE challenge consists in determining, given two text fragments, whether one text fragment is entailed by the other. In our examples of *pousser*, it is necessary to encode e.g. whether movement of the theme *boîte* in *"J'ai poussé la boîte"* is entailed or not, or whether the hypothesis *"A bomb has been placed in the basement"* can be inferred from a sentence like *"La lettre a poussé ma voisine à poser une bombe dans la cave"*. This shows that a high level of detail in the formal semantic description is a definite asset, and represents an important step beyond the information contained in existing lexical semantic resources.

References

Adler, S. and Asnès, M. (2005). Les compléments de degré en jusqu'à. *Travaux de linguistique*, 49:131–157.

Asher, N. and Lascarides, A. (1995). Metaphor in discourse. In *Proceedings of the AAAI Spring Symposium Series. Representation and Acquisition of Lexical Knowledge: Polysemy, Ambiguity and Generativity*, pages 3–7, Stanford, CA.

Baader, F., Calvanese, D., McGuinness, D. L., Nardi, D., and Patel-Schneider, P. F., editors (2003). *The Description Logic Handbook: Theory, Implementation and Applications.* CUP.

Baker, C. F., Fillmore, C. J., and Lowe, J. B. (1998). The Berkeley FrameNet project. In *Proceedings of the joint COLING/ACL 1998*, Montreal, Canada.

Bechhofer, S., van Harmelen, F., Hendler, J., Horrocks, I., McGuinness, D. L., eider, P. F. P.-S., and Stein, L. A. (2004). OWL Web Ontology Language Reference. W3C Recommendation. `http://www.w3.org/TR/owl-ref/`.

Berners-Lee, T., Hendler, J., and Lassila, O. (2001). The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibiliti es. *Scientific American*, 284(5):34–43.

Bhatt, R. (1999). *Covert Modality in Non-Finite Contexts*. PhD thesis, University of Pennsylvania.

Boullier, P. and Sagot, B. (2005). Efficient and robust LFG parsing: SxLFG. In *Proceedings of IWPT '05*, Vancouver, BC.

Dagan, I., Glickman, O., and Magnini, B. (2005). The PASCAL Recognising Textual Entailment Challenge. In *Proceedings of the PASCAL Challenges Workshop on Recognising Textual Entailment*, Southampton, UK.

Fellbaum, C., editor (1998). *WordNet – An Electronic Lexical Database*. MIT Press, Cambridge, MA.

Franconi, E. (2003). Natural Language Processing. In Baader, F., Calvanese, D., McGuinness, D. L., Nardi, D., and Patel-Schneider, P. F., editors, *The Description Logic Handbook: Theory, Implementation and Applications*, pages 460–471. CUP.

Gangemi, A., Guarino, N., Masolo, C., and Oltramari, A. (2003a). Sweetening WordNet with Dolce. *AI Magazine*, 24(3):13–24.

Gangemi, A., Navigli, R., and Velardi, P. (2003b). The OntoWordNet Project: extension and axiomatization of conceptual relations in WordNet. In *Proceedings of ODBASE*, Catania, Italy. Springer.

Golbreich, C. and Imai, A. (2004). Combining SWRL rules and OWL ontologies with Protégé OWL Plugin, Jess, and Racer. In *Proceedings of the 7th Protégé Conference*, Bethesda, MD.

Guarino, N. and Welty, C. (2002). Evaluating Ontological Decisions with OntoClean. *Communications of the ACM*, 45(2):61–65.

Hacquard, V. (2006). *Aspects of Modality*. PhD thesis, MIT.

Horrocks, I., Patel-Schneider, P. F., Boley, H., Tabet, S., Grosof, B., and Dean, M. (2004). SWRL: A Semantic Web Rule Language Combining OWL and RuleML. W3C Member Submission. `http://www.w3.org/Submission/SWRL/`.

Jackendoff, R. S. (1990). *Semantic structures.* MIT Press, Cambridge, MA.

Karttunen, L. (1971). Implicative verbs. *Language*, 47:340–358.

Kipper-Schuler, K. (2006). *VerbNet: A broad-coverage, comprehensive verb lexicon.* PhD thesis, Computer and Information Science Dept., University of Pennsylvania, Philadelphia, PA.

Knublauch, H., Musen, M. A., and Rector, A. L. (2004). Editing description logic ontologies with the Protégé OWL plugin. In *Proceedings of DL 2004*, Whistler, BC.

Mari, A. and Martin, F. (2007). Tense, abilities and actuality entailment. In Aloni, M., Dekker, P., and Roelofsen, F., editors, *Proceedings of the XVI Amsterdam Colloquium*, pages 151–156, Universiteit Amsterdam.

Martin, F. (2008). Forceful verbs in Spatial, Psychological and Conceptual Domains. Semantic Analysis of *push*-verbs in French. Manuscript, University of Stuttgart.

Niles, I. and Pease, A. (2001). Towards a Standard Upper Ontology. In *Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS 2001)*, Ogunquit, ME.

Niles, I. and Pease, A. (2003). Linking Lexicons and Ontologies: Mapping WordNet to the Suggested Upper Merged Ontology. In *Proceedings of the 2003 International Conference on Information and Knowledge Engineering (IKE '03)*, Las Vegas, NV.

O'Connor, M., Knublauch, H., Tu, S., Grosof, B., Dean, M., Grosso, W., and Musen, M. (2005). Supporting Rule System Interoperability on the Semantic Web with SWRL. In *Proceedings of the 4th International Semantic Web Conference*, Galway, Ireland.

Sirin, E., Parsia, B., Grau, B. C., Kalyanpur, A., and Katz, Y. (2007). Pellet: A practical OWL-DL reasoner. *Journal of Web Semantics*, 5(2).

Sonntag, D., Engel, R., Herzog, G., Pfalzgraf, A., Pfleger, N., Romanelli, M., and Reithinger, N. (2007). SmartWeb Handheld – Multimodal Interaction with Ontological Knowledge Bases and Semantic Web Services. In Huang, T. S., Nijholt, A., Pantic, M., and Pentland, A., editors, *Artificial Intelligence for Human Computing*, volume 4451 of *Lecture Notes in Artificial Intelligence*, pages 272–295. Springer, Heidelberg.

Spohr, D. (2008a). A General Methodology for Mapping EuroWordNets to the Suggested Upper Merged Ontology. In *Proceedings of the 6th Language Resources and Evaluation Conference (LREC 2008)*, Marrakech, Morocco.

Spohr, D. (2008b). Extraction of Selectional Preferences for French using a Mapping from EuroWordNet to the Suggested Upper Merged Ontolog y. In *Proceedings of the 4th Global WordNet Conference*, Szeged, Hungary.

Stein, A. (2007). Motion events in concept hierarchies: Identity criteria and french examples. In Schalley, A. and Zaefferer, D., editors, *Ontolinguistics. How Ontological Status Shapes the Linguistic Coding of Concepts*, pages 379–394. Mouton de Gruyter, Berlin.

Talmy, L. (1985). Lexicalization patterns: semantic structure in lexical form. In Shopen, T., editor, *Language typology and syntactic description III*, pages 51–149. Cambridge University Press, New York, NY.

Vandeloise, C. (1991). *Spatial Prepositions: a Case Study from French*. University of Chicago Press, Chicago.

Vossen, P., editor (1998). *EuroWordNet: A Multilingual Database with Lexical Semantic Networks*. Kluwer Academic Publishers.

Zwarts, J. (2007). Forceful prepositions. In Vyvyan, E. and Chilton, P., editors, *Language, Cognition and Space: The State of the Art and New Directions*. Equinox Publishing, London. To appear.