Abducing Term Variant Translations in Aligned Texts

Michael Carl, Ecaterina Rascu, Johann Haller
and Philippe Langlais

Abstract. Term variation is an important issue in various applications of natural language processing (NLP) such as machine translation, information retrieval and text indexing. In this paper, we describe an 'Abductive Terminological Database' (ATDB) aiming to detect translations of terms and their variants in bilingual texts. We describe abduction as the process to infer specific term translation templates from multiple resources which have been induced from a bilingual text. We show that precision and recall of the ATDB increase when using more resources and when the resources interfere in a less restricted way. We discuss a way to feed back evaluation values into the induced resources thus allowing for weighted abduction which further enhances the precision of the tool.

Keywords: recognition of terminological variation, automatic generation of variation templates, weighted term variant abduction

1 Introduction

Terminologies play a crucial role in technical communication and the consistent use of terms is a prerequisite to avoid misunderstandings. However, terminological variation is a frequent phenomenon even in established technical domains (Daille et al., 1996; Macklovitch, 1995; Royauté, 1999). Recognizing terms and their variants plays an important role in many NLP applications such as information retrieval, machine translation (MT), validation of terminological accuracy or in the process of updating existing terminological resources.

Up to date, research on computational terminology has mainly focused on term extraction and acquisition (see (Cabré Castellví et al., 2001) for a recent overview), although some research projects investigate methods to complete existing terminologies. In order to enrich existing terminological resources Jacquemin (Jacquemin, 1996) derives candidate terms from variants of reference terms. The author exploits the fact that some variation patterns, for example coordination, include more than one term in their structure. New candidate terms can be detected through structural analysis of such variants.

On the other hand, mechanisms and classification schemas for terminological variation have been studied, amongst others, for French (Jacquemin, 2001), English (Daille et al., 1996) and German (Carl et al., 2002; Schmidt-Wigger, 1999). This knowledge has been used in a number of applications. For instance, variants of terms are used for query expansion in information retrieval thereby considerably increasing retrieval efficiency (Jacquemin and Tzoukermann, 1999). In a controlled language application, technical authors are encouraged to use the preferred base forms of terms instead of their unauthorized variants. In many cases this is, for economical, legal, or just for presentational reasons, of prime interest for the company (Schmidt-Wigger, 1999).

In this paper we investigate a method to detect terms and their variants in bilingual texts. We assume that it is unrealistic and undesirable to list all variants of a term together with all their possible variant translations in a bilingual terminological lexicon. This is particularly true for terms that are very productive. Such terms may occur in different morphological and/or syntactic forms and be rewritten as synonyms. Therefore we suggest a method to detect terminological variants and their translations in bilingual texts using a dictionary which contains translations of authorized base terms and language specific knowledge of terminological variation.

We develop a semi-automatic method to detect term variant translations using processing steps taken from abductive reasoning. The process consists of three steps which will be outlined in this paper. First a bilingual terminology and a set of general but language specific term variation patterns are extracted from a bilingual text. Then we extend the terminology database with specific variation templates by applying the general variation patterns and synonyms to each entry. The resulting database contains sufficient knowledge to map terms and their variants onto their base forms and it contains also the required information to trace back a detected variant translations to their underlying general variation mechanisms. We refer to this database as the Abductive Terminology Database (ATDB) as it is the core component of the system. We feed back values of precision into the general variation mechanism thus to weight their applicability.

In section 2, we outline the architecture of the ATDB and underpin some basic assumptions of the approach. We classify and discuss in depth some general term variation mechanisms as observed in a bilingual English–French text and we give an example of how the ATDB is produced from these resources. We describe three different databases ATDB_0,

$ATDB_1$ and $ATDB_2$ which are generated through different, increasingly larger, sets of the general variation patterns.

In section 3 we evaluate the three ATDBs by detecting terms (and their variants) in two bilingual English–French texts. We compute the precision of the tool for each of the ATDBs and for each variation template. We discuss some of the problems that have occurred and we conclude that using more resources leads to better recall and precision. However, permitting more variation patterns to co-occur simultaneously leads to more noise and, even more important, to more ambiguous term alignments. In order to rank the strength of two or more ambiguous term alignments, we feed back evaluation values into the abduction process thus enhancing the precision of the tool.

2 Abduction of Terminological Variants

The word "abduction" was introduced by Peirce (Peirce, 1955) who defined it in the following terms:

> "The surprising fact C is observed; But if A were true, C would be a matter of course; Hence, there is reason to suspect that A is true." (Peirce, 1955, p.151)

Unlike deduction, abduction starts from an observed phenomenon and formulates hypotheses that could explain it. Streiter (Streiter, 2003) points out that abductive reasoning creates hypotheses which are not logically implied by the premises. In contrast to deductive reasoning, abduction is not always correct in all reasoning steps. However, abductive reasoning should be "plausible" in a context and yield correct results in the vast majority of the cases.

Magnani (Magnani, 2001) explores the relationships between abduction, deduction, and induction in order to point out the significance of abduction in the problem solving process. He distinguishes between selective or creative abduction, which both generate plausible explanatory hypotheses. Abduction as "inference to the best explanation" additionally evaluates the generated hypotheses. This type of abduction requires a ranking of the hypotheses by evaluating their explanatory power. Magnani points out that various evaluation criteria can be used to establish this ranking. The relevance of such criteria depends on the context and on the nature of the application in which abduction is performed. He lists several criteria proposed in the literature such as correctness, empirical

adequacy, simplicity, consistency, practical usability, or even mathematical probability.

Mooney (Mooney, 2000) examines the relation between abduction and induction. He states that

> "in abduction, the hypothesis is a specific set of assumptions that explain the observations of a particular case, while in induction, the hypothesis is a general theory that explains the observations across a number of cases." (Mooney, 2000, p.183)

Mooney applies abductive learning to theory refinement. Theory refinement is the task to make an existing imperfect domain theory consistent with a set of data. For him, abduction is primarily useful for generalizing a theory to cover more positive examples. For each individual positive example that is not derivable from the current theory, abduction is applied to determine a set of assumptions that would allow it to be proven.

In a similar way the ATDB detects terminological variants in translated and aligned texts. The underlying assumption is that each term in the source text segment has also a translation in the corresponding target text segment and vice versa. In case a term translation cannot be detected, the ATDB tries to prove the presence of a variant translation. We describe an approach where most of the resources for this prove are generated off-line and are ready at hand during runtime of the tool.

In section 2.1 and 2.2, we outline the architecture and the resources of the system. Section 2.3 illustrates how an ATDB is generated from these resources and shows a sample segment of three different ATDBs. In section 3.4 we feed back translation errors to weight the resources and solve some ambiguities.

2.1 Architecture of the ATDB

In order to recognize variants of terms and their translations in bilingual aligned texts, we have adapted a monolingual terminology tool described in (Carl et al., 2002). The modified architecture is shown in figure 1. It consists of two symmetrical language sides, a left-hand side (i.e. English) and a right-hand side (i.e. French).

To detect translations of terms and their variants in an aligned English–French text, the tool requires two types of resources. The first resource is a bilingual terminology containing authorized base terms and their translations (bottom part of figure 1). This terminology may either be

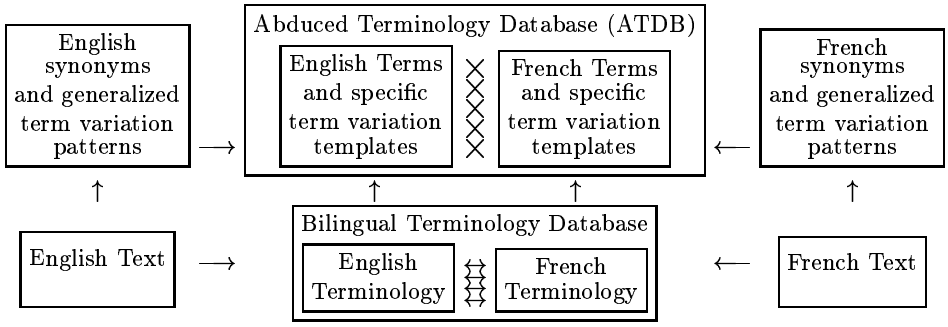

Fig. 1. Abducing a Term Database (ATDB)

provided by a terminologist or it can be automatically acquired. For the experiment reported in section 3.2 we have manually extracted a small bilingual terminology of unambiguous term translations from a bilingual text.

The second resource consists in generalized variation patterns and synonymy relations, obtained through inspection of an English–French translated text. This is illustrated in the left and right sides in figure 1. Based on these general variation patterns, a number of variants and specific variation templates are produced for every term in the bilingual terminology.

These variants and specific variation templates are stored in a database — the so-called Abduced Terminology Database (ATDB) — together with the original terms so that each variant is linked to its authorized form. While the ATDB contains sufficient knowledge to detect term variants in bilingual texts, it now also contains $m - to - n$ translation equivalences as indicated by the symbol $\times$ in figure 1. The architecture distinguishes two basic steps: (i) an induction step where a bilingual terminology and general term variation patterns are extracted from a (bilingual) text and (ii) a deduction step, where specific term variation templates are inferred from the induced resources. In section 3.4 we will introduce a third step to weigh the resources obtained in step (i). In the remainder of this section we shall discuss the first two steps.

2.2 Induction of General Variation Patterns

Generalized variation patterns were obtained through examination of an English–French aligned bilingual text. The patterns coincide largely with those described in (Daille et al., 1996; Jacquemin, 2001) who distinguish

EO_1	$(N_1;A_1)+\ N_2 \rightarrow N_2$
	inclined groove $\rightarrow$ *groove*
	lock knob $\rightarrow$ *knob*
EO_2	$(N_1;A_1)\ N\ N_2 \rightarrow N_1\ N_2$
	match grade ammunition $\rightarrow$ *match ammunition*
	c3a1 sniper rifle $\rightarrow$ *c3a1 rifle*
EO_3	$(N;A)+\ (N_1;A_1)\ N_2 \rightarrow (N_1;A_1)\ N_2$
	scope spanner wrench $\rightarrow$ *spanner wrench*
	supervised safety precautions $\rightarrow$ *safety precautions*
FO_1	$N_1\ * \rightarrow N_1$
	fusil de le tireur d'élite $\rightarrow$ fusil
FO_2	$N_1\ (N_2;A_2)\ * \rightarrow N_1\ (N_2;A_2)$
	huile polyvalente pour armes $\rightarrow$ huile polyvalente
FO_3	$N_1\ p_2\ d_3\hat{}\ N_4\ * \rightarrow N_1\ p_2\ d_3\hat{}\ N_4$
	tenon de verrouillage de sécurité arrière $\rightarrow$ tenon de verrouillage
	mesures de sécurité supervisées $\rightarrow$ mesures de sécurité
FO_4	$N_1\ (p\hat{}\ d\hat{}\ N)+\ N_2 \rightarrow N_1\ N_2$
	fusil de tireur d'élite c3a1 $\rightarrow$ fusil c3a1
	munition du type match $\rightarrow$ munition match
FO_5	$N_1\ A\ d_2\ N_3\ p_4\hat{}\ d_5\hat{}\ N_6\hat{} \rightarrow N_1\ d_2\ N_3\ p_4\hat{}\ d_5\hat{}\ N_6\hat{}$
	sens contraire des aiguilles d'une montre $\rightarrow$ sens des aiguilles d'une montre
FO_6	$N_1\ p_2\ d\ N_3 \rightarrow N_1\ p_2\ N_3$
	échelle de les millièmes $\rightarrow$ échelle en millièmes
	ressort de le percuteur $\rightarrow$ ressort de percuteur

Table 1. Omission Variation for English and French Terms

between variation patterns of derivation, insertion, permutation, and coordination. The text — which we refer to as SNIPER2 (see section 3.1 for a more detailed presentation) — is a manual on sniper training and deployment that was also used in a previous study (Carl and Langlais, 2002; Macklovitch, 1995). Even though Macklovitch (Macklovitch, 1995) reports that among the French translations of the term *sniper* only 8.5% are variants in this text, we will see that a large number of variation patterns occur. We found that omission, insertion, permutation, coordination, synonymy, derivation, and typographical variation are the most important means of terminological variation in the text. In the following sections we describe the major classes of term variation and the corresponding induced patterns. Some of the patterns presented in the following section (i.e. EI_1, EP_3, EP_4, FI_4) were induced at a later stage of the evaluation (see Section 3.2) from another text referred to as SNIPER3 in order to cover further variation phenomena.

Variation by Omission Variation by omission occurs when one or more components are deleted from multi-word terms. Table 1 presents such variation patterns for English and for French terms. For instance, the variation patterns EO_1 and EO_3 are generalizations of the observed variants *inclined groove* → *groove* and *supervised safety precautions* → *safety precautions*. Omission variants are under-specified compared to their original base form. Ommission variants are also observed for French as, for instance, in FO_4: fusil de tireur d'élite c3a1 → fusil c3a1. Following Jacquemin (Jacquemin, 1996), omission variants can be considered as being generic variants of their original, more specific, base forms.

The induced generalized patterns describe the variation mechanism using part-of-speech tags, `N`, `A`, `d`, and `p` for noun, adjective, determiner, and preposition respectively. While the indexed tags — for instance N_1 in FO_1 — map a particular word (tireur) from the left hand side to the right hand side, tags with no indexes map any word of that category. Asterisks "*" map any sequence of words irrespective of their category. The sign "+" indicates that the preceding bracketed sequence may occur more than once and the hat "^" expresses optionality of the preceeding tag. .

EI_1	$(N_1;A_1)\ N_2 \rightarrow (N_1;A_1)\ A\ N_2$ *prone position* → *prone supported position*
FI_1	$N_1\ (N_2;A_2) \rightarrow N_1\ p\ d\ (N_2;A_2)$ mouvement arrière → mouvement vers l'arrière
FI_2	$N_1\ p_2\ N_3 \rightarrow N_1\ p_2\ d\ N_3\ A$ manchon de culasse → manchon de la culasse mobile
FI_3	$N_1\ p_2\ N_3\ p_4\hat{}\ d_5\hat{}\ N_6 \rightarrow N_1\ (num)+\ p_2\ d\ N_3\ p_4\hat{}\ d_5\hat{}\ N_6$ fusil de tireur d'élite → fusil 7.62 mm x 51 mm de le tireur d'élite
FI_4	$N_1\ A_2 \rightarrow N_1\ p\ d\ N\ A_2\ p\ N$ position couchée → position de le tireur couché avec appui

Table 2. Insertion Variation for English and French Terms

Variation by Insertion In variants by insertion the structure of a term is changed through the insertion of one or more components. In variants of type FI_1 a function word is inserted in the terms while the variants of type FI_2 and FI_3 contain an additional modifier. In the case of pattern FI_2 the inserted modifier is an adjective while in case of pattern FI_3 the tag `num` matches numbers, dates and units of measurement. Similarly, English variants of type EI_1 insert an adjective (e.g. *supported*) to build

more specific variants of the base terms. Note that the inserted tags in the right hand sides of the generalized variation patterns are not indexed and thus are not bound to any particular word of the left hand side. The sequence `p d` in pattern FI_1, for instance, maps any sequence of prepositions followed by a determiner.

EP_1	$N_1\ N_2\ N_3 \rightarrow N_2\ N_3$ num N_1
	c3a1 sniper rifle → *sniper rifle 7.66 mm x 52 mm c3a1*
EP_2	$N_1\ N_2\ N_3 \rightarrow N_2\ N_3\ N_1$
	c3a1 sniper rifle → *sniper rifle c3a1*
EP_3	$N_1\ N_2 \rightarrow N_2$ p d N_1
	rifle butt → *butt of a rifle*
EP_4	N_1 p d $N_2 \rightarrow N_2\ N_1$
	pocket of the shoulder → *shoulder pocket*
FP_1 $N_1\ p_2\ N_3\ p_4\hat{}\ d_5\hat{}\ N_6\ N_7 \rightarrow N_1\ N_7\ p_2$ d $N_3\ p_4\hat{}\ d_5\hat{}\ N_6$	
	fusil de tireur d'élite c3a1 → fusil c3a1 de le tireur d'élite

Table 3. Permutation Variation for English and French Terms

Variation by Permutation Variation by permutation changes the linear order of the term constituents (Daille et al., 1996). In many cases permutation occurs together with the insertion of function or modifying words. The order of the constituents in the English term is changed in the pattern EP_2 while in patterns EP_3, EP_4, and FP_1 permutation of nouns takes place and function words are inserted. In pattern EP_1 a modifying phrase consisting of figures and/or units of measurement is inserted into the English term.

Variation by Coordination As Jacquemin (Jacquemin, 1996) points out, two or more conceptually related terms can be linked by means of coordination into a more compact structure. Common constituents are eliminated and a coordinative conjunction, or in case of enumeration a comma ",", is inserted between the basic constituents. Coordination can occur between head words or between modifiers. The variation patterns in table 4 show coordination patterns for English modifiers — ECO_1 — and English head words — ECO_2 — and for French modifiers — FCO_1 and FCO_2. In the French pattern FCO_2 coordination is combined with inser-

ECO_1	$(N_1;A_1)\ N_2 \rightarrow (N_1;A_1)$ Conj $(N;A)\ N_2$
	elevation adjustment $\rightarrow$ *elevation and windage adjustment*
ECO_2	$A_1\ N_2 \rightarrow A_1$ N Conj N_2
	visual acuity $\rightarrow$ *visual ability and acuity*
FCO_1	$N_1\ p_2\ N_3 \rightarrow N_1\ p_2$ N Conj $p_2\ N_3$
	tambours de dérive $\rightarrow$ tambours de hausse et de dérive
FCO_2	$N_1\ p_2\ N_3 \rightarrow N_1\ p_2$ d N Conj p_2 d N_3
	réglage de dérive $\rightarrow$ réglage de la hausse et de la dérive
FCO_3	$N_1\ p_2\ N_3 \rightarrow N_1\ p_2\ N_3$ Conj p_2 N
	tambours de dérive $\rightarrow$ tambours de dérive et de hausse
FCO_4	$N_1\ p_2\ N_3 \rightarrow N_1\ p_2$ d N_3 Conj p_2 d N
	réglage de dérive $\rightarrow$ réglage de la dérive et de la hausse

Table 4. Coordination Variation for English and French Terms

tion. Since in our current analysis we do not have semantic features which would allow for recognition of conceptual relatedness of the coordinated terms, we allow any noun to be inserted in position N.

ES_1	*spotting telescope* $\rightarrow$ *spotting scope*
FS_1	lunette de visée $\rightarrow$ lunette de tir

Table 5. English and French Synonyms

Synonyms In addition to the morpho-syntactical variations presented above, terms also vary in the choice of their lexemes. In order to detect such variants we allow synonyms of terms. A problem related to synonyms is to determine a sufficient context, so that unnecessary noise is avoided and a maximum number of relevant terms are found. In the context of a previous study (Carl et al., 2002), we found that synonyms are highly domain specific. This implies that for restricted terminologies we can shorten the context of a synonym without risking to produce too much noise. For the text at hand, the synonyms ES_1 and FS_1 in the left hand sides can equally be obtained by the substitution of the French modifier visée $\rightarrow$ tir and of the English head word *telescope* $\rightarrow$ *scope*.

Hamond and Nazarenko (Hamon and Nazarenko, 2001) allow synonyms to appear simultaneously in the head and in the modifier of a compound term. While this is justified in their application, in our expe-

rience synonyms in two or more positions of a term produce much more noise. Particularly when different variation patterns and synonyms overlap, the variants detected in the text and the task of finding corresponding target language equivalents become increasingly ambiguous(cf. (Carl and Langlais, 2002)).

Writing and Derivational Variants Writing variants are typographical variants including different use of hyphenation or of blanks between term constituents or the use of punctuation marks around a term constituent. In the aligned text we found writing variants such as shown in table 6; EWV_1 and EWV_2 for English and FWV_1 for French.

EWV_1	*hand stop* → *handstop*
EWV_2	*re-insert* → *reinsert*
FWV_1	munition de le type match → munition de le type "match"
FDV_1	dégagement de l'oeil → dégagement oculaire
FDV_2	couché → coucher

Table 6. English and French Writing Variants

Derivational variants make use of different word classes such as adjectives, adverbs or nouns which are derived from the same word stem. Thus, the nominal construction de l'oeil in FDV_1 is equivalent to the adjective construction oculaire. Since the standard statistical tagger and the lemmatizer used in this study (Foster, 1991) do not recognize derivational and writing variants, both variation forms are treated like synonyms.

We also (ab)use synonyms to correct wrong analyses of the morphological analyzer. In some cases, when the part-of-speech of a term or variant component is not correctly detected, an incompatible lemma is assigned. Since recognition of the words is based on the lemmas rather than on their surface forms, we re-adjust some of the analyses through synonymy relations, as for instance, in FDV_2.

2.3 Abduction of Specific Term Variation Templates

Starting from an initial database of unambiguous term translations, English and French variants are abduced using synonym lists and general variation patterns. We describe this process by means of an example. Assume, for instance, the term translations (1) *spotting telescope* ↔ lunette

d' observation and (2) *telescopic sight* ↔ lunette de visée as in table 7 are in the base terminology. A number of variants and variant templates can be abduced from these terms using the general variation patterns as indicated in the first column for English and the last column for French in table 7. From term (1) the variant *telescope* is obtained by applying variation pattern EO_1 while the variant *scope* is the result of successively applying the synonymy pattern ES_1 and the variation pattern EO_1.

By taking into account various combinations of resources, we generate three different databases: $ATDB_0$, $ATDB_1$, and $ATDB_2$. The database $ATDB_0$ is identical to the original term database, with no additional variants. From the terms in table 7, $ATDB_0$ contains only the base terms. $ATDB_1$ includes $ATDB_0$ as well as all variants that can be generated through a single application of one variation or synonymy pattern. Applying the general variation patterns described in section 2.2, there are 5 additional variants of *spotting telescope* in $ATDB_1$. In addition to the entries in $ATDB_1$, $ATDB_2$ contains variants in which synonyms co-occur with variation patterns. This produces another four variants.

Each entry in table 7 contains three sections, separated through horizontal lines. These sections represent the variants and specific variation templates for the databases $ATDB_0$, $ATDB_1$, and $ATDB_2$ respectively.

While the original terminology (and correspondingly $ATDB_0$) contains only $1 - to - 1$ term translation correspondences, ADTBs contain $m - to - n$ translation relations. Due to variation pattern FO_1, French lunette is recognized as an omission variant of the base terms lunette de visée and lunette d'observation. Accordingly, the English base term translation is either *spotting telescope* or *telescopic sight*.

Therefore, the ATDBs encode the capacity of one French term to be translated into n different English terms and vice versa. We expect that adding further terms and variation patterns to the resources increases the ambiguity of the terminology. Ambiguity is also likely to increase for higher level $ATDB_i$, $i > 2$. As we show in the next section, coverage and precision also increase with higher level ATDBs.

3 Experiments and Evaluation

In this section we evaluate the abductive terminology databases presented in section 2. We use two bilingual aligned test texts, SNIPER2 and SNIPER3 which we present in more detail in section 3.1. Terms and their translations in the test texts are manually annotated following an annotation format presented in (Carl and Fissaha, 2003). The test texts are passed

(1) *spotting telescope* ↔ lunette d' observation

Pattern	English ATDB		French ATDB	Pattern
base	*spotting telescope*			
EO_1	*telescope*			
EI_1	*spotting* `A` *telescope*			
EP_3	*telescope* `p d` *spotting*			
ECO_1	*spotting* `Conj (N;A)` *telescope*	↔	lunette d' observation	*base*
ES_1	*spotting scope*		lunette d' `d` observation `A`	FI_2
ES_1, EO_1	*scope*		lunette `d N Conj` d'observation	FCO_1
ES_1, EI_1	*spotting* `A` *scope*		lunette	FO_1
ES_1, EP_3	*scope* `p d` *spotting*			
ES_1, ECO_1	*spotting* `Conj (N;A)` *scope*			

(2) *telescopic sight* ↔ lunette de visée

Pattern	English ATDB		French ATDB	Pattern
			lunette de visée	*base*
base	*telescopic sight*		lunette de `d` visée `A`	FI_2
EI_1	*telescopic* `A` *sight*	↔	lunette `d N Conj` de visée	FCO_1
ECO_1	*telescopic* `Conj (N;A)` *sight*		lunette	FO_1
EO_1	*sight*		lunette de tir	FS_1
			lunette `d N Conj` de tir	FS_1, FCO_1
			lunette	FS_1, FO_1

Table 7. Abduction of Term Variants

through ATDB_0, ATDB_1, and ATDB_2 where terms and their variants are marked automatically. These results are compared with the manual annotation as shown in figure 2. We compute precision and recall for the three databases in section 3.2. In section 3.3 we evaluate the general variation patterns and synonyms from which the variants were produced. In section 3.4 we automatically weight these variation patterns by feeding the alignment errors back into the evaluation process. We show that precision of term alignment prediction enhances when ranking term variant translations by these weights. An overall picture of the evaluation architecture is shown in figure 2.

3.1 Resources

In this subsection we introduce two test texts which are used to evaluate the architecture in figure 2. First we present the terminology databases ATDB_0, ATDB_1 and ATDB_2.

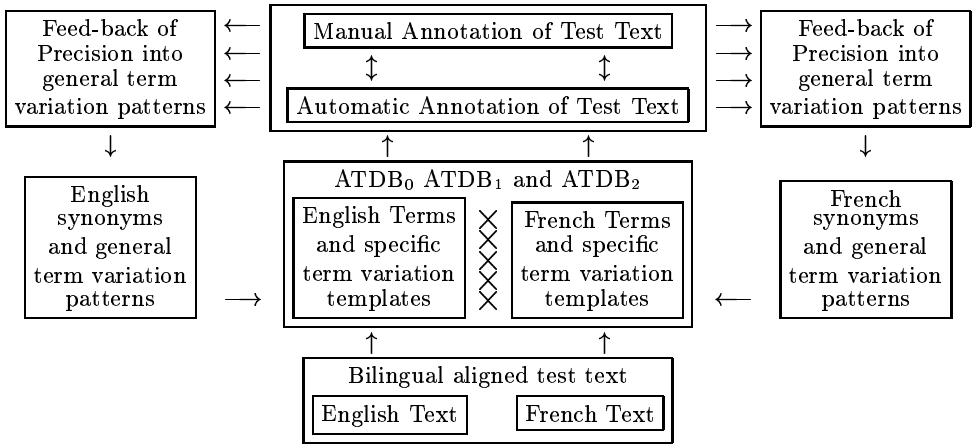

Fig. 2. Architecture of the ATDB

The Abduced Terminology Databases As outlined in section 2.1, the abduction of term variants in the ATDB requires a bilingual terminology database. The bilingual terminology used for the evaluation of the ATDB was manually extracted from SNIPER2. It contains 137 non-ambiguous term translations where each English and French term occurs exactly once in the terminology database. A small set of synonyms of the content words in the term base was also collected manually. The synonym list for English contains a total of 131 synonyms for 54 content words while the French list consists of 92 synonyms for 50 content words.

We generate three databases $ATDB_0$, $ATDB_1$ and $ATDB_2$ from these resources and the general variation patterns discussed in section 2.3. While $ATDB_0$ contains only the 137 base terms, $ATDB_1$ contains in addition 420 specific term variation templates for English and 595 variation templates for French. $ATDB_2$ has 542 English and 759 French variation templates. Table 8 plots the productivity of the general variation patterns for the tree databases. English Insertion patterns were introduced only later, as will be described in section 3.3.

According to the table, omission patterns are most productive. Thus, $ATDB_2$ contains for each base term more than one omission variant. Note that some of the term variation templates in table 8 are partially underspecified and they may contain disjunctive items. Such templates map a large number of surface strings. For instance, even though there are only 90 synonym templates, multiplying out the disjunctions they contain yields 274 different "non-ambiguous" (though still partially underspecified) templates. An estimative evaluation of the synonym templates

	$ATDB_0$		$ATDB_1$		$ATDB_2$	
	English	French	English	French	English	French
base terms	137	137	137	137	137	137
Synonyms	—	—	90	64	90	64
Omission	—	—	110	141	174	207
Permutation	—	—	24	6	42	10
Insertion	—	—	0	135	0	173
Coordination	—	—	59	112	99	168
Total	137	137	420	595	542	759

Table 8. Number of Abduced Term Variation Templates

shows that for English they contain 76% real synonyms, 7% writing variants, and 17% derivational variants. The abduced synonym templates for French contain 77% real synonyms, 6% writing variants, and 17% derivational variants.

The Test Texts The ATDB was tested on two texts, SNIPER2 and SNIPER3. The texts are excerpts from an army manual on sniper training and deployment (Macklovitch, 1995). SNIPER2 and SNIPER3 have 391 and 400 English–French aligned segments, respectively, with an average length of 19 and 22 words in the English and the French segment.

For the evaluation of the ATDBs we established "gold standards" by manually annotating term translations in SNIPER2 and SNIPER3. We follow a format proposed by (Carl and Fissaha, 2003) which consists of a three-tuple containing the number of the alignment, the position of the English (or French) term or variant within the alignment, and the corresponding French (or English) translation. This format presumes a fixed word tokenization. Punctuation marks (.,:;), the French preposition d' and the article l' as well as the English genitive marker *'s* are counted as distinct tokens. Moreover, concatenated function words like French des and du are expanded into two tokens de les and de le respectively.

(3) However , each *sniper* 's *visual ability and acuity* is slightly different .

Toutefois , l' acuité visuelle de chaque tireur d' élite varie légèrement .

Example (3) contains the term translations *sniper* ↔ tireur d'élite and *visual ability and acuity* ↔ acuité visuelle. Given this fixed tokenization, translations of terms are unambiguously marked and stored in a separate file by registering the number of the example as well as the beginning and the end of each term (or variant) in one line.

In case an English (or a French) term is not translated, we face a NIL term alignment. The example (4) illustrates such a NIL term alignment from SNIPER3.

(4) Use the nonfiring hand to support the *butt of the weapon* ...

Le tireur se sert de sa main libre pour soutenir la crosse de l' arme ...
... and ensure a firm contact between the *butt* and the shoulder .

... et la maintenir appuyée contre son épaule .

The English sentence is in the imperative mood while the French one makes use of a declarative construction. Accordingly, the French subject tireur is not translated in the English sentence. The example (4) also contains the permutation variant *butt of the weapon* (base term: *weapon's butt*) translated as base term crosse de l'arme and the ommission variant *butt* translated as the French pronoun la. However, the manual annotation of test texts does not specify whether the terms are in their base forms or whether they are derived variants or pronouns.

For SNIPER2 we annotated 1033 term correspondences; for SNIPER3 we found 872 correspondences. In the experiments in section 3.2 we initially used resources that were extracted and induced from SNIPER2. During these experiments further variation patterns were found in SNIPER3. We added some of these patterns (EI_1, EP_3, EP_4, FI_4), included 17 additional base terms to the terminology and repeated the experiments in order to study the effect on the performance of the ATDB.

In the following sections we present an overall evaluation of the ATDBs. Subsequently we evaluate the different components.

3.2 Precision and Recall

The goal of this experiment is to see how reliable the abduced variants are. Both language sides of SNIPER2 and SNIPER3 were passed through the ATDBs. First, only the base terms were marked in both language sides of the aligned texts ($ATDB_0$) and retrieved from the output (cf. figure 2). The same procedure was repeated for the terms and variants obtained by separately applying the synonym lists and variation patterns ($ATDB_1$). In a third pass we combined the synonym lists and the variation patterns when abducing variants in $ATDB_2$.

We counted the number of correct term translations recovered by the ATDB, the noise it produced, and the misses, that is the number of term translations which were not recovered. Furthermore, we computed

precision as the ratio of the *correct* recovered term translations over all annotated term translations (*correct* / *correct* + *noise*), and recall as the ratio of the *correct* recovered term translations over all recovered term translations (*correct* / *correct* + *misses*). Table 9 summarizes the results.

	SNIPER2			SNIPER3		
	$ATDB_0$	$ATDB_1$	$ATDB_2$	$ATDB_0$	$ATDB_1$	$ATDB_2$
precision	0.54	0.61	0.62	0.57	0.60	0.62
recall	0.45	0.78	0.89	0.39	0.78	0.86
correct	467	802	921	338	683	754
noise	402	508	559	250	446	468
misses	566	231	112	534	189	118

Table 9. Coverage and Precision of the ATDB

Table 9 shows that when using more resources ($ATDB_1$ and $ATDB_2$), precision *and* recall increase. The increase in recall is, however, much more significant than the increase in precision.

In case of $ATDB_0$ noise is mainly due to the fact that for terms that were detected in one side of the alignment no corresponding translation could be found in the other, i.e. these terms are translated through a variant. Thus, a NIL term alignment was wrongly detected by the tool. Moreover, the high number of noisy alignments in all ATDBs is due to a great extent to the fact that the ATDB establishes all possible connections between term pairs (and their variants) in both language sides. In example 5 the manual annotation contains two term alignments, while the ATDBs produce four term alignments, which leads to a precision of 50 %.

(5) ... to judge the distance to *targets* and as an aid in the indication of a *target* .

... pour évaluer la distance à l' objectif et pour indiquer un objectif .

Performing a deeper syntactical analysis would allow for a proper treatment of the multiple connections between terms. Our estimations show that under such circumstances, precision could increase by ca. 20%.

Example (6), on the other hand, illustrates one of the positive effects of multiple connections between term pairs and/or their variants. The two occurrences of the English synonym *weapon* are correctly linked to

one occurrence of the French arme since the preceding verbs are bound through coordination.

(6) ... to adjust the *weapon* and zero the *weapon* to the firer .

... pour ajuster et zéroter l' arme selon ses besoins .

In some cases we are unable to detect correct term translations when they involve pronouns — as, for instance, in example (4). In example (7a), too, the French term tireur is translated as the personal pronoun *he*. As Macklovich (Macklovitch, 1995) points out, translation of French subjects into an English pronoun occurs more often than the other way round. In this example, the French term crosse is the translation of the first occurrence of English *weapon 's butt* while the second occurrence of the same English phrase is translated as the French pronoun l'.

(7a) *He* balls his hand into a fist to raise the *weapon 's butt* or ...

Pour soulever la crosse , le tireur n' a qu' a serrer la main - ...
... loosens the fist to lower the *weapon 's butt* .

... pour l' abaisser, il suffit de relâcher la prise .

Example (7a) represents a manual alignment while (7b) shows a part of the output for $ATDB_1$ and $ATDB_2$ of the same pair of sentences. The ATDBs align both instances of *weapon 's butt* with the French crosse and ignore the French pronoun. We count this as one correct, one noisy and one missing term alignment.

(7b) ... to raise the *weapon 's butt* or ... to lower the *weapon 's butt*

Pour soulever la crosse , le ...

The manual annotation contains also NIL-alignments as, for instance, shown in (8). English *shot* is such a NIL-alignment in the manual annotation since it has no translation in the French segment.

(8) The *sniper* must be a near perfect *shot* .

Le tireur d' élite doit être pratiquement infaillible .

However, *shot* is a synonym of *sniper* and hence considered a possible translation of French tireur d' élite. The ATDB erroneously aligns both terms and this fact is equally counted as noise.

We noticed that for SNIPER3 the implementation of further variation patterns would allow for a more precise variant recognition. For example, the variation pattern (EP_3) covers *butt of the weapon* as a variant of the term *rifle butt*.

We repeated the experiment after having implemented some more patterns (EI_1, EP_3, EP_4, FI_4) and included 17 additional terms in the terminology. The "gold standard" was completed accordingly. The results of this second run are shown in table 10. They indicate increase in precision and recall for SNIPER3 and a very slight decrease for SNIPER2.

	SNIPER2			SNIPER3		
	$ATDB_0$	$ATDB_1$	$ATDB_2$	$ATDB_0$	$ATDB_1$	$ATDB_2$
precision	0.53	0.61	0.62	0.59	0.64	0.66
recall	0.45	0.78	0.89	0.40	0.79	0.86
correct	467	802	916	373	732	804
noise	407	510	562	259	410	422
misses	566	231	117	557	198	126

Table 10. Coverage and Precision of the ATDB

3.3 Evaluation of the General Variation Patterns

We consider the results of the latter experiment for a more detailed analysis of the general variation patterns. We classify the detected term translation correspondence according to the variation patterns and we examine their precision. Since the test text annotation does not specify the type of the term variation, we cannot compute the number of misses for each type of variation, and hence cannot compute their recall.

Base Terms Just as the results reported in tables 9 and 10 precision of term detection increases when using higher order ATDBs. The results in table 11 show that when using abduced variants the number of correct discovered term alignments increases for English and French.

Note that two parameters influence the value of precision: precision can be increased by augmenting the number of correct term translations

		SNIPER2			SNIPER3		
		precision	correct	noise	precision	correct	noise
E	$ATDB_0$	0.55	455	369	0.63	362	213
	$ATDB_1$	0.61	538	343	0.71	411	164
	$ATDB_2$	0.62	566	352	0.71	412	166
F	$ATDB_0$	0.64	456	254	0.69	346	157
	$ATDB_1$	0.71	492	201	0.75	390	127
	$ATDB_2$	0.71	492	203	0.75	390	128

Table 11. Evaluation of Variant Abduction for English and French Terms

and/or by decreasing the noise produced. As shown in table 11 when using $ATDB_1$ we increase the number of correct recovered terms and decrease noise. However, when allowing synonyms to apply simultaneously with term variation patterns, as in $ATDB_2$, we slightly increase noise as compared to $ATDB_1$.

Table 12 shows the distribution of correct base term alignments. The majority of base terms are translated as a base term ($T \leftrightarrow T$). Synonym–base term correspondences ($T \leftrightarrow S$, $S \leftrightarrow T$) also represent a considerable category for English and French. Translation of base terms by insertion (I), permutation (P), coordination (CO), or omission variants (O) seems to play a minor role for English as compared to French.

		SNIPER2			SNIPER3		
E	$\leftrightarrow F$	$ATDB_0$	$ATDB_1$	$ATDB_2$	$ATDB_0$	$ATDB_1$	$ATDB_2$
T	$\leftrightarrow T$	444	444	444	335	334	333
T	$\leftrightarrow NIL$	11	10	10	27	9	9
T	$\leftrightarrow S$	0	43	41	0	52	51
T	$\leftrightarrow I$	0	5	6	0	0	1
T	$\leftrightarrow O$	0	36	65	0	16	18
NIL	$\leftrightarrow T$	12	9	9	11	8	8
S	$\leftrightarrow T$	0	35	35	0	40	40
CO	$\leftrightarrow T$	0	1	1	0	0	0
O	$\leftrightarrow T$	0	3	3	0	6	7
P	$\leftrightarrow T$	0	0	0	0	2	2

Table 12. Translation Correspondences for English and French Terms

In $ATDB_1$ and $ATDB_2$ the detection of a wrong variant may prevent the tool from retrieving a valid term pair. One such example is the

detection of the wrong coordination variant battants de bretelles et d'un cale-main of the base term battants de bretelles which shadows the detection of the term pair *sling swivels* ↔ battants de bretelles.

		SNIPER2			SNIPER3		
		precision	correct	noise	precision	correct	noise
E	$ATDB_1$	0.42	141	194	0.39	165	263
	$ATDB_2$	0.48	157	172	0.46	196	234
F	$ATDB_1$	0.36	84	150	0.39	126	197
	$ATDB_2$	0.34	79	150	0.39	127	200

Table 13. Evaluation of Synonyms

Synonyms The number of detected French synonyms shows no important difference for $ATDB_1$ or $ATDB_2$. For English noise decreases and the number of correct alignments increases as plotted in table 13. Low precision for alignment of synonyms is due to their ambiguity. For example the word *shot* is a synonym of *sniper*, *sniping*, *fire*, and *firing*. In addition to this, *shot* is also the synonym of some omission variants in $ATDB_2$ such as *field firing*. This leads to a large number of erroneous or ambiguous term alignments which cannot always be resolved.

Nevertheless, a considerable number of correct translation pairs are detected by synonyms. As table 14 shows, synonyms are frequently translated as synonyms, but also as base terms. However, English synonyms are most often translated as French omission variants.

			SNIPER2		SNIPER3	
E	↔	*F*	$ATDB_1$	$ATDB_2$	$ATDB_1$	$ATDB_2$
S	↔	*S*	30	30	60	60
S	↔	*NIL*	1	0	9	2
S	↔	*T*	35	35	40	40
S	↔	*I*	0	1	0	0
S	↔	*O*	75	91	56	94
NIL	↔	*S*	10	7	12	9
T	↔	*S*	43	41	52	51
O	↔	*S*	1	1	1	6
P	↔	*S*	0	0	1	1

Table 14. Translation Correspondences for English and French Synonyms

Best results are achieved, when synonyms also occur in combination with other variation patterns as in $ATDB_2$.

Variation by Omission Variation by omission is by far the most productive means of variation for both English and French terms (cf. table 15). The results show that the number of correct alignments almost doubles when the variation patterns are combined with synonyms, as is the case in $ATDB_2$. However, precision remains constant or even decreases for most of the general variation patterns except for FO_2 and FO_4 where noise decreases.

		SNIPER2			SNIPER3		
		precision	correct	noise	precision	correct	noise
EO_1	$ATDB_1$	0.52	215	201	0.55	265	216
	$ATDB_2$	0.46	396	473	0.43	437	591
EO_2	$ATDB_1$	0.50	1	1	1.00	2	0
	$ATDB_2$	0.50	1	1	1.00	2	0
EO_3	$ATDB_1$	0.59	10	7	0.93	13	1
	$ATDB_2$	0.59	10	7	0.93	13	1
$\sum$	$ATDB_1$	0.52	226	209	0.56	280	217
	$ATDB_2$	0.46	407	481	0.43	452	592
FO_1	$ATDB_1$	0.58	317	226	0.53	360	324
	$ATDB_2$	0.56	548	429	0.59	556	379
FO_2	$ATDB_1$	0.56	5	4	0.00	0	3
	$ATDB_2$	0.67	6	3	1.00	3	0
FO_3	$ATDB_1$	0.41	12	17	0.80	12	3
	$ATDB_2$	0.41	12	17	0.80	12	3
FO_4	$ATDB_1$	0.25	1	3	1.00	2	0
	$ATDB_2$	1.00	5	0	1.00	2	0
FO_5	$ATDB_1$	1.00	1	0	0.00	0	0
	$ATDB_2$	1.00	1	0	0.00	0	0
FO_6	$ATDB_1$	1.00	2	0	0.00	0	0
	$ATDB_2$	1.00	2	0	0.00	0	0
$\sum$	$ATDB_1$	0.57	338	250	0.53	374	330
	$ATDB_2$	0.56	574	449	0.60	573	382

Table 15. Evaluation of Omission Variants

The most productive patterns (EO_1 and $F0_1$) generate also most of the noise. The noise generated by EO_1 has different causes illustrated by the examples (9), (10) and (11).

An important source of noise are omission variants which have very general meanings such as English *position* or the French tête. For instance,

the word *position* is wrongly recognized as a variant of *firing position*, *hawkins position* or *prone position* which leads to noise in contexts like (9 a) and (9b)

(9a) when the bolt reaches this *position*
(9b) to ensure that the scales do not move they should be taped in *position*

Similarly, French tête is an omission variant produced by pattern FO_1 from the base terms tête de la gâchette, tête de bipied en acier and tête du percuteur. Each occurrence of the word tête in the French side of an alignment produces noise if no variant of *nose of the sear*, *steel bipod spigot*, or *top of the trigger* is found in the English side of the alignment as shown in (10a-c):

(10a) *the sniper moves his head* ↔ le tireur déplace la tête
(10b) *bolt face* ↔ tête de culasse
(10c) *the nose of the cocking piece* ↔ tête du talon du percuteur

In many instances, this leads to ambiguous alignments as shown in example (11). In the context of example (11), the base terms *general purpose weapons oil* and the omission variant *oil* are recovered as translations for huile polyvalente pour armes. However, only the former is a valid translation of the French term in that context, while the latter produces noise.

(11) *oil* bottle for *general purpose weapons oil*
bouteille d' huile polyvalente pour armes

			SNIPER2		SNIPER3	
E	↔	*F*	$ATDB_1$	$ATDB_2$	$ATDB_1$	$ATDB_2$
O	↔	*O*	213	394	253	406
O	↔	*NIL*	9	9	20	33
O	↔	*T*	3	3	6	7
O	↔	*S*	1	1	1	6
NIL	↔	*O*	14	23	49	55
T	↔	*O*	36	65	16	18
S	↔	*O*	75	91	56	94
P	↔	*O*	0	1	0	0

Table 16. Translation Correspondences for English and French Omission Variants

An analysis of the correct translation correspondences in table 16 shows that omission variants translate most frequently into omission vari-

ants. Note that the symbol O includes synonyms of omission variants for $ATDB_2$.

Variation by Insertion Only a very few insertion, permutation, or coordination variants could be found in SNIPER2 and SNIPER3.

No correct English insertion variant was detected in SNIPER2. In SNIPER3 17 English insertion variants were recognized with ($ATDB_2$) with a precision of 0.53. As pointed out before, the low precision is due to multiple connections between source language and target language variants.

In the French side of SNIPER2 and SNIPER3, respectively 6 and 10, correct insertion variants were recognized by $ATDB_2$. The four variants recognized on the basis of pattern FI_2 in SNIPER2 are actually different instances of the same French variant, manchon de la culasse mobile. The pattern FI_1 allowed for the recognition of the variant dégagement de l'oeil of the term dégagement oculaire and of different variants of the term mouvement arrière: déplacement vers l'arrière et mouvement vers l'arrière. The variants dégagement oculaire and déplacement vers l'arrière use synonyms and derivational variation. In SNIPER3 the pattern FI_4 accounts for 9 of the 10 correct insertion variants. Different variants of the term position couchée could thus be recognized: position du tireur couché avec appui, position du tireur couché sans appui and position du tireur couché avec bipied.

			SNIPER2		SNIPER3	
E	$\leftrightarrow$	F	$ATDB_1$	$ATDB_2$	$ATDB_1$	$ATDB_2$
I	$\leftrightarrow$	I	0	0	5	9
T	$\leftrightarrow$	I	5	6	0	1
S	$\leftrightarrow$	I	0	1	0	0

Table 17. Translation Correspondences for English and French Insertion Variants

As table 17 shows most of the insertion variants in SNIPER2 are translated as base terms while corresponding target language insertion variants are preferred in SNIPER3.

Variation by Permutation In SNIPER2 only 3 variants by permutation were found for English and 1 for French with $ATDB_2$. In SNIPER3 2 correct variants were recognized by applying pattern (EP_3) and 1 by applying pattern (EP_4).

The correspondences between English and French patterns when variation by permutation is involved is shown in table 18. Accordingly, English variants by permutation may be translated as French base terms, permutation variants or synonyms. However, the number of permutation variants detected in the text is too low to attempt any kind of generalization.

			SNIPER2		SNIPER3	
E	$\leftrightarrow$	F	$ATDB_1$	$ATDB_2$	$ATDB_1$	$ATDB_2$
P	$\leftrightarrow$	P	1	1	0	0
P	$\leftrightarrow$	I	1	1	0	0
P	$\leftrightarrow$	O	0	1	0	0
P	$\leftrightarrow$	S	0	0	1	0
P	$\leftrightarrow$	T	0	0	2	2

Table 18. Translation Correspondences for English and French Permutation Variants

Variation by Coordination We discovered very few coordination variants in SNIPER2 and SNIPER3. Among these there were four coordinating modifiers and one head coordination for English in SNIPER2. For French only modifier coordination was found. As table 19 shows, coordination variants in both texts are translated by corresponding target language coordination variants, except for one term translation of an English variant.

Further refinement of the patterns for insertion, permutation, and coordination as well as experiments on larger corpora would allow for a more conclusive evaluation of these types of variation. Future work will be in line with the methodology of iterative refinement as outlined by Meyer (Meyer, 2001).

			SNIPER2		SNIPER3	
E	$\leftrightarrow$	F	$ATDB_1$	$ATDB_2$	$ATDB_1$	$ATDB_2$
CO	$\leftrightarrow$	CO	3	4	1	1
CO	$\leftrightarrow$	T	1	1	0	0

Table 19. Translation Correspondences for English and French Coordination Variants

3.4 Weighing Term Variant Alignments

The number of ambiguous term alignments increases for $ATDB_1$ and $ATDB_2$ compared to $ATDB_0$. While with $ATDB_1$ less than 0.6 of the automatically generated term alignments are ambiguous, with $ATDB_2$ every term alignment has on average 1.3 ambiguous links. Increased ambiguity is due to simultaneous application of synonyms and variation patterns. We have shown that this leads, in most instances, to higher values for precision and recall of term alignments.

Ambiguous alignments, as produced by the ATDBs, can be of different types. One type of ambiguity occurs when aligning underspecified variants. As discussed above (see example 9), in case the English segment contains the word *position*, the $ATDB_1$ and $ATDB_2$ do not know to which of the three base terms 12a-c the detected variant refers.

12a	*firing position*	↔	position de tir
12b	*hawkins position*	↔	position de hawkins
12c	*prone position*	↔	position couché

Another type of ambiguity occurs when aligning two different text segments in one language with a single text segment in the other language. Examples (6) and (7a) show such instances. While in example (6) both occurrences of *weapon* ↔ arme are correct, in (7a) one of the two connections *weapon 's butt* ↔ crosse is erroneous. In the current state of the tool we cannot propose a solution to automatically recognize and distinguish these cases.

We will rather examine a third type of alignment ambiguity which is illustrated in example (11). Here, the English term *general purpose weapons oil* and the variant *oil* are aligned with the French term huile polyvalente pour armes. Based on the precision of the variation patterns we shall weight these ambiguous links and show that this can considerably reduce noise.

As more terms and variation patterns become available it is likely that all types of term alignment ambiguities increase. It is therefore desirable to rank term alignments and predict which of the possible links is most reliable. For instance, given a choice of two or more alignment links, if we had an oracle which would always predict the correct link we would reach an estimated precision of 0.85 or more. The remaining 15% noise are due to wrong detection of variants (e.g. see examples 9 and 10) and translation into pronouns, which is not tackled in the ATDBs. Thus, in cases like example (11) the oracle would predict the correct alignment

general purpose weapons oil ↔ huile polyvalente pour armes and eliminate the erroneous link to *oil*.

In this section we suggest to rank (ambiguous) term alignments through the weight of the resources (i.e. general variation patterns and synonyms) which were involved in predicting this link. First, we compute for every three-tuple of a term translation term_t an English variation pattern_e and a French variation pattern_f a co-occurrence_precision according to the following equation:

$$\text{co-occurrence\_precision}(\text{term}_t, \text{pattern}_e, \text{pattern}_f) = \frac{\text{correct}(\text{term}_t, \text{pattern}_e, \text{pattern}_f)}{\text{correct}(\text{term}_t, \text{pattern}_e, \text{pattern}_f) + \text{noise}(\text{term}_t, \text{pattern}_e, \text{pattern}_f)}$$

This leads to a three-dimensional table which allows us to rank ambiguous alignments: given an English string e and a French string f we weight their alignment link as the precision of e being derived from the English term_t by means of variation pattern_e and f being derived from the same French term_t and the French variation pattern_f. We thus obtain stronger links for more precise alignments and weaker links for less precise alignments. As a general conclusion from the investigation in section 3.3, we find that when more variation patterns are involved to derive a term alignment, the precision of that alignment decreases. Thus, precision of base term alignment (13a) is 1.0 while the omission term alignment (13b) has a precision of 0.01. Both alignments can thus be ranked and a clear preference can be attributed.

(13a)	*general purpose weapons oil*	↔	huile polyvalente pour armes	1.0
(13b)	*oil*	↔	huile polyvalente pour armes	0.01

We repeated the experiment described in section 3.2 and counted as noise and correct the weighted alignment links instead of adding up 1. Table 20 compares the non-weighted precision values as described in section 3.2 and the precision values as obtained by weighting the alignment links. The evaluations in rows EV_2 only take into consideration resources extracted from SNIPER2 while those in EV_3 also include resources from SNIPER3.

From the higher precision obtained through weighted term alignments we conclude what we have already observed in the previous section: that certain variation types tend to co-occur more frequently than other variation types. We believe that this insight can be of great value and should be exploited more thoroughly in the further development of the tool.

		SNIPER2		SNIPER3	
		$ATDB_1$	$ATDB_2$	$ATDB_1$	$ATDB_2$
precision	EV_2	0.61	0.62	0.60	0.62
(not weighted)	EV_3	0.61	0.62	0.64	0.66
precision	EV_2	0.79	0.77	0.76	0.75
weighted	EV_3	0.79	0.77	0.81	0.79

Table 20. Precision of Weighted Term Translation Alignment

4 Conclusion and Outlook

This paper presents Abductive Terminological Database (ATDB), a tool designed to detect term translations and their variants in aligned bilingual texts. The tool integrates different resources, a terminology database, lists of synonyms and sets of general variation patterns, which can be combined in various ways. We have implemented and discussed a method to weigh and rank term alignments. The weights are calculated on the basis of term alignment precision. The weights are fed back into the general variation patterns and synonym lists. In our opinion this provides a powerful means not only to enhance alignment results but also to investigate on an empirical basis the success and applicability of variation patterns in different contexts.

We describe a series of experiments based on an English–French aligned text from a military domain. The assumption underlying the ATDB is that each English term has also a translation in the French text segment and vice versa. In case a term translation cannot be detected, the ATDB tries to prove the presence of a variant through abductive reasoning. As the number of resources increases and the way they interact multiplies, alignment of term translations becomes more ambiguous. However, we find at the same time that recall and precision also increase in most cases.

The tool is conceived as an independent software module that can be used in a number of NLP applications. As a pre-processor to an MT-engine it ensures a greater coverage and enables the production of more consistent translations. In (Carl and Langlais, 2002) we used a previous version of this tool to detect various term variants in the source text while the retrieved target term is generated in its authorized base form. As part of a bilingual terminology tool — for instance in a controlled translation environment — the tool is capable of detecting terminological inconsistency of draft translations and to suggest using authorized translations

instead. A similar monolingual application has already been described elsewhere (Carl et al., 2002).

Still, there are a number of open problems in our implementation of the ATDB. Anaphora and ellipsis are unsolved problems. We have also observed translation mismatches where no term translation occurs in the text or where multiple occurrences of one term are translated as a single term in the target language. To tackle some of these problems it will be useful to integrate a syntactic analysis into the tool which could still enhance precision.

References

Cabré Castellví, M. T., R. Estopà Bagot, and J. Vivaldi Palatresi. 2001. "Automatic term detection: A review of current systems." In Bourigault, D., C. Jacquemin, and M.-C. L'Homme. 2001. *Recent Advances in Computational Terminology.* John Benjamins Publishing Company, Amsterdam/Philadelphia, 53–89.

Carl, M., and S. Fissaha. 2003. "Phrase-based evaluation of word-to-word alignments." In *HLT-NAACL 2003 Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond*, Edmonton, Canada, 31–35.

Carl, M. and P. Langlais. 2002. "An intelligent terminology database as a pre-processor for statistical machine translation." In *Proceedings of COMPUTERM 02*, Taipei, 128–138.

Carl, M. J. Haller, Ch. Horschmann, D. Maas, and J. Schütz. 2002. "The TETRIS Terminology Tool". In *TAL, Structuration de terminologie*, 43(1), 73–103.

Daille, B., B. Habert, C. Jacquemin, and J. Royauté. 1996. "Empirical Observation of Term Variations and Principles for their Description". In *Terminology*, 3(2), 77–93.

Foster, G. F. 1991. *Statistical lexical disambiguation.* Master's thesis, McGill University, School of Computer Science.

Hamon, T. and A. Nazarenko. 2001. "Detection of synonymy links between terms: Experiment and results." In Bourigault, D., C. Jacquemin, and M.-C. L'Homme. 2001. *Recent Advances in Computational Terminology.* John Benjamins Publishing Company, Amsterdam/Philadelphia, 185–208.

Jacquemin, C. and E. Tzoukermann. 1999. "NLP for term variant extraction: A synergy of morphology, lexicon and syntax." In Strzalkowski, *Natural Language Information Retrieval*, Boston, MA. Kluwer, 25–74.

Jacquemin, C. 1996. "A symbolic and surgical acquisition of terms through variation." In *Proceedings on the workshop on Connectionist, Statistical and Symbolic Approaches to Learning for Natural Language Processing*, 425–438.

Jacquemin, C. 2001. *Spotting and Discovering Terms through Natural Language Processing.* The MIT Press, Cambridge, Massachusetts.

Macklovitch, E. 1995. *Can terminological consistency be validated automatically ?.* Technical report, CITI/RALI, Montréal, Canada.

Magnani, L. 2001. *Abduction, Reason and Science. Processes of Discovery and Explanation.* Kluwer Academic, Doordrecht.

Meyer, I. 2001. "Extracting knowledge-rich contexts for terminography." In Bourigault, D., C. Jacquemin, and M.-C. L'Homme. 2001. *Recent Advances in Computational Terminology.* John Benjamins Publishing Company, Amsterdam/Philadelphia, 279–302.

Mooney, R. J. 2000. "Integrating abduction and induction in machine learning." In P. Flach and A. Kakas, editors, *Abduction and Induction*, Kluwer Academic Publishers, 181–191.

Peirce, C.S. 1955. *Abduction and Induction.* Dover Books, New York.

Royauté, J. 1999. *Les groupes nominaux complexes et leurs propriétés : application à l'analyse de l'information.* Ph.D. thesis, Université Henri Poincaré-Nancy 1 College, Nancy.

Schmidt-Wigger, A. 1999. "Term checking through term variation." In *Proceedings of Terminology and Knowledge Engeneering (TKE)*, 570–581.

Streiter, O. 2003. "Abductive Explanation-based Learning Improves Parsing Accuracy and Efficiency." In *2nd SIGHAN Workshop at ACL 2003*, 7–12.

Author's adresses

Michael Carl, Johann Haller and Ecaterina Rascu
Institut für Angewandte Informationsforschung
Saarbrücken, Germany
{carl;hans}@iai.uni-sb.de
kati@rascu.de

Philippe Langlais
RALI/DIRO
Université de Montréal, Quebec, Canada
felipe@iro.umontreal.ca

About the authors

Michael Carl studied computer sciences and computational linguistics in Berlin, Paris and Hong Kong. He obtained his PhD from the University of the Saarland, Saarbrücken in 2001 on Example-based Machine Translation. He joined the IAI in 1994 and since then was involved in many projects and products in NLP and Machine Translation. In 2002 he was on a post-doctoral position at RALI/University of Montreal.

Johann Haller studied Romance languages and Linguistics in Germany and Spain. He has been involved in research and teaching on computational linguistics at SIEMENS, Systran and at Universities in Germany, France, Spain and Latin America. Since 1990 he is full professor for Machine Translation at the University of the Saarland and Director of IAI in Saarbrücken. His areas of expertise include multilingual NLP, machine translation, technical documentation, controlled language, and information retrieval.

Philippe Langlais is assistant professor in the computer science department (DIRO) at Université of Montreal. He obtained his PhD from Université d'Avignon in 1995, working on speech recognition after having worked for three years in the speech technology group at IDIAP, in Switzerland. In 1997, he was an invited researcher in CTT, in Stockholm. He joined the RALI laboratory in 1998 where he now works on bilingual applications.

Ecaterina Rascu studied English Language and Literature and French Language and Literature at the Babes-Bolyai University in Cluj-Napoca (Romania). Since 2003 she is a PhD Student at the University of the Saarland. Her main research interests concern automatic paraphrasing and machine translation.