Improving the Representation of Word-Formation in Multilingual Lexicographic Tools: the MuLeXFoR Database

Bruno Cartoni, LIMSI/CNRS
Marie-Aude Lefer, Centre for English Corpus Linguistics – Université catholique de Louvain

This paper introduces a new lexicographic resource, MuLeXFoR, which aims to present word-formation processes in a multilingual database designed for both language specialists (e.g. linguists, terminologists, lexicographers, NLP specialists) as well as second-language (L2) learners and trainee translators. Morphological items (e.g. affixes, compound parts, combining forms) and processes (prefixation, suffixation, compounding, conversion, etc.) pose major challenges for lexicographic work, especially with respect to the design of bilingual and multilingual resources. It is well-known that derivational affixes can take part in several word-formation rules and that, conversely, rules can be realised by means of a variety of affixes. In view of this complexity, it is often difficult to (1) provide enough information to help users understand the meaning(s) of an affix and the (near-)synonymy relations between affixes and (2) become familiar with the most frequent strategies used to translate the meaning(s) conveyed by these affixes. In fact, traditional dictionaries often fail to achieve this goal. The MuLeXFoR database tries to take advantage of recent advances in morphological description and the development of electronic multi-access database systems. The database relies on the lexematic approach to word-formation, which is especially helpful to represent morphological processes cross-linguistically. In addition, it has been entirely implemented in a multi-access database interface. The prototype described in this paper so far centres around prefixation in English, French and Italian. Two interfaces are currently available: a comprehensive interface aimed at morphological and lexicographic investigations by language specialists (MuLeXFoR-Linguists) and a second interface designed for second-language learners or trainee translators (MuLeXFoR-Learners).
Section 1 first briefly discusses the ways in which word-formation processes are currently presented in dictionaries and brings to light some of the shortcomings involved in these procedures. Section 2 then describes the theoretical approach that has been adopted to formalise word-formation processes from a multilingual perspective, viz. the lexematic approach. Section 3 introduces the MuLeXFoR database and provides a detailed description of the ways users can browse the tool. Section 4 presents the ways in which the database was adapted to second-language learners, mainly by simplifying labels and menu names and by adding information specific to the production of new words in L2. Finally, Sections 5 and 6 deal with the various implementation and data collection issues raised by our approach. The paper ends with some concluding remarks in Section 7.

1. Word-formation in dictionaries

Many bilingual dictionaries include morphological items in their lists of entries, usually with the purpose of providing information about how to interpret, translate or coin derivatives and compounds (among other things). Because many complex and compound words are not given individual entries in dictionaries, it is crucial that the word-parts which make them up be listed as headwords. In addition, the inclusion of morphological elements in dictionaries increases the users' 'morphological awareness', i.e. improves their ability to understand the factors at play in the coining of new words.

This said, the representation of morphological items and processes in monolingual and bilingual dictionaries has often been criticised (see e.g. Prcic, 1999; Dardano et al., 2006; ten Hacken et al., 2006; Cartoni, 2008a; Lefer, 2009). Prcic (1999), for example, has shown that affix descriptions in learners' dictionaries are inadequate and has argued that affix entries should be more exhaustive, i.e. they should include information on pronunciation, sense distinctions, productivity, and be based on a coherent terminology (in the sense that

morphological labels should be used more consistently throughout the dictionary). Additionally, these studies have put forward the inadequacy of relying solely on affix representation in dictionaries, which is how morphological items have been included in dictionaries so far. Derivational affixes are often polysemous in the sense that they usually display a range of possible meanings. To put it in lexematic terms (see Section 2), affixes often take part in several word-formation rules. In addition, a given meaning can often be conveyed by several affixes (e.g. *multi* and *pluri* to express 'unspecified plurality'). The selection of one affix instead of another to coin new words largely depends on a number of factors such as register, genre and domain, analogy, and productivity, all of which are crucially important in multilingual lexicographic contexts. Bilingual dictionaries to date have not yet adequately addressed these two semantic issues.

Second-language learners and trainee translators could greatly benefit from systematic comparisons of L1 and L2 word-formation systems to interpret, translate or coin (new) complex words. Such descriptions could also lead to better knowledge of word-formation rules and constraints. However, descriptions of word-formation systems can be complex and are sorely missing from traditional dictionary entries. In this respect, we concur with Prcic's (1999: 274) claim that 'description of affixes in […] dictionaries […] urgently needs an overhaul – both theoretical and methodological'. The MuLeXFoR database aims to fill this gap in lexicographic practice. Before presenting the database in greater detail, Section 2 provides a short description of the theoretical framework adopted here.

2. The lexematic approach to morphology

2.1. Word-formation representation with lexeme-formation rules

The lexematic approach to morphology (Fradin, 2003) considers derivational affixes as morphological items that take part in much more complex word-formation processes. More precisely, affixes are seen as the formal components of lexeme-formation rules (hereafter LFRs) which entail other constructional operations (such as word category changes) and which, most importantly, are semantically-driven. As such, the lexematic approach constitutes a useful starting point to deal with morphological processes as semantic unified wholes (the macro-approach) rather than dealing with individual affixes one by one (the micro-approach). Figure 1, which is inspired by Fradin's (2003) formalism, presents the French LFR of verbal reiterativity.

	INPUT			OUTPUT
(G)	X		G	reX
(F)	/X/		F	/ʀə/⊕/X/
(SX)	cat :v <SN, SN>	→	SX	cat :v , <SN, SN>
(S)	X'		S	REITER X'

Figure 1. French LRF of reiterativity (v>v)

The rule in Figure 1 schematically represents the coinage of a derived verb ('output') from a base lexeme ('input'). It is made up of the different operations that are applied to the base: formal (graphical (G) and phonological (F), by adding a prefix in the above example), syntactic (SX) and semantic (S) (by modifying the meaning of the base verb). In such representations, rules can display several formal components, whether of the same kind (e.g.

prefixes) or not (e.g. affixes and conversion, or prefixes and suffixes). The 'removal/reversal' rule that produces verbs from nouns, for instance, is realised in English by means of the prefix *dis* and by means of conversion, as shown in Figure 2.

	INPUT
(G)	X
(F)	/X/
(SX)	cat :n
(S)	X'

→

	OUTPUT	
G	disX	X
F	/dIs/⊕/X/	/X/
SX	cat :v	
S	REMOVE (X')	

Figure 2. English LFR of removal/reversal (n>v)

As appears from Figure 2, two different formal components are included in the rule (prefixation with *dis* and conversion). The syntactic and semantic parts of the rule, however, are the same for *dis* (*arm* → *disarm*) and for conversion (*milk* → *to milk*) as they both produce verbs from nouns (syntactic operation) with the meaning 'removal/reversal' (semantic operation).

2.2. Multilingual lexeme-formation rules

In this project, the cross-linguistic representation of morphological processes is based on the lexematic approach. Only a few monolingual lexicographic tools rely on this approach, such as the Database of Catalan Affixes (Bernal and DeCesaris, 2008). The approach has not yet been applied to multilingual tools, however. In the case of MuLeXFoR, lexematic morphology proved to be extremely useful to formalise multilingual LFRs that match equivalent word-formation processes in different languages. For example, one can formalise a reiterativity LFR that creates verbs from verbs (LFR_reiter(v→v)) and represents the various affixes that are used cross-linguistically to convey this meaning (*ri* in Italian, *ré* in French, *re* in English). Figure 3 illustrates the formalisation of this trilingual reiterativity LFR.

	Italian		*French*		*English*
	INPUT		INPUT		INPUT
(G)	X_{IT}		X_{FR}		X_{EN}
(F)	/X_{IT}/		/X_{FR}/		/X_{EN}/
(SX)	cat :v		Cat :v		cat :v
(S)	X_{IT}'		X_{FR}'		X_{EN}'
	↕	↔	↕	↔	↕
	OUTPUT		OUTPUT		OUTPUT
(G)	riX_{IT}		reX_{FR}		reX_{EN}
(F)	/ri/⊕/X_{IT}/		/ʀə/⊕/X_{FR}/		/ʀə/⊕/X_{EN}/
(SX)	cat :v		cat :v		cat :v
(S)	REITER. (X_{IT}')		REITER. (X_{FR}')		REITER. (X_{EN}')

where $X_{IT}' = X_{FR}' = X_{EN}$, translation equivalents

Figure 3. Trilingual LFR of reiterativity (v>v)

Figure 3 formalises the translatability of the reiterativity LFR. When the rule is applied to translationally equivalent bases (here in Italian, French and English), it produces

translationally equivalent derivatives. The formal operations (in sections G and F) are language-specific, while the other operations are the same cross-linguistically.

A further advantage of this approach applies to cases of synonymy where one rule represents several affixes. The unspecified plurality LFR is made up of three prefixes in Italian and French and two prefixes in English (It. *multi, pluri, poli;* Fr. *multi, pluri, poly*; Engl. *multi, poly*). In French, the three prefixes are described as interchangeable (Amiot, 2005), which is also probably true for the two other languages. Some pragmatic constrains are affix-specific (e.g. *poli*/*poly* tends to be restricted to the technical domain), in which cases the lexematic rule specifies information on actual use. In our multilingual framework, these constraints can be either language-specific or cross-linguistic.

	Italian		*French*		*English*
	INPUT		INPUT		INPUT
(G)	$Xsfx_{IT}$		$Xsfx_{FR}$		$Xsfx_{EN}$
(F)	/$Xsfx_{IT}$/		/$Xsfx_{FR}$/		/$Xsfx_{EN}$/
(SX)	cat :a		cat :a		cat :a
(S)	X_{IT}'		X_{FR}'		X_{EN}'
	↕	↔	↕	↔	↕
	OUTPUT		OUTPUT		OUTPUT
(G)	multi/pluri/poli$Xsfx_{IT}$		multi/pluri/poly$Xsfx_{FR}$		multi/poly$Xsfx_{EN}$
(F)	/multi/|/pluri/|/pɔli/⊕/$Xsfx_{IT}$/		/mylti/|/plyri/|/pɔli/⊕/$Xsfx_{FR}$/		/mVltI/|/pQlI/⊕/$Xsfx_{EN}$/
(M)	res (poli): (X_{IT})=techn.		res (poly): (X_{FR})=techn.		res (poly): (X_{EN})=techn.
(SX)	cat :a		cat :a		cat :a
(S)	UNSP. PLUR. (X_{IT}')		UNSP. PLUR. (X_{FR}')		UNSP. PLUR. (X_{EN}')

where $X_{IT}' = X_{FR}' = X_{EN}$, translation equivalents

Figure 4. Trilingual LFR of unspecified plurality (a>a)

The rule in Figure 4 shows the possible interchangeability of the *multi*, *pluri* and *poly*/*poli* prefixes which all convey the meaning of unspecified plurality 'many'. Pragmatic constraints can also be specified (here for *poly*) in the section (M)[1], which stresses that the base usually belongs to the technical or scientific domain. This rule can account for the translation equivalence between Fr. *multidimensionnel* and It. *pluridimensionnale* which are both based on the noun *dimension*/*dimenzione* (Engl. *dimension*).

Importantly, the coinage of prefixed relational adjectives is very peculiar in the sense that these adjectives are formed from nouns. Their formal representation in the lexematic approach shows that they are made up of a suffixed nominal base ('sfx' in 'Xsfx', where X is the nominal base). The semantic operation of the prefixation rule applies to the base noun, as represented in (UNSP.PLUR. (X)). The above example, i.e. *multidimensional,* can thus be

[1] The section (M) – for Morphology – is an optional section aimed at specifying possible pragmatic constraints.

understood as 'with many dimensions' (see Fradin, 2007 for a complete description of this phenomenon). This is the reason why the rule is represented as *n>a*.

3. The MuLeXFoR-Linguists database: general architecture

The MuLeXFoR project aims to present multilingual LFRs (as described in Section 2) in a user-friendly interface. The lexeme-formation rules are core to the database. Their surface representations (i.e. affixes and other morphological processes such as conversion) are listed for each language, together with language-specific notes. Figure 5 graphically represents this two-level architecture.

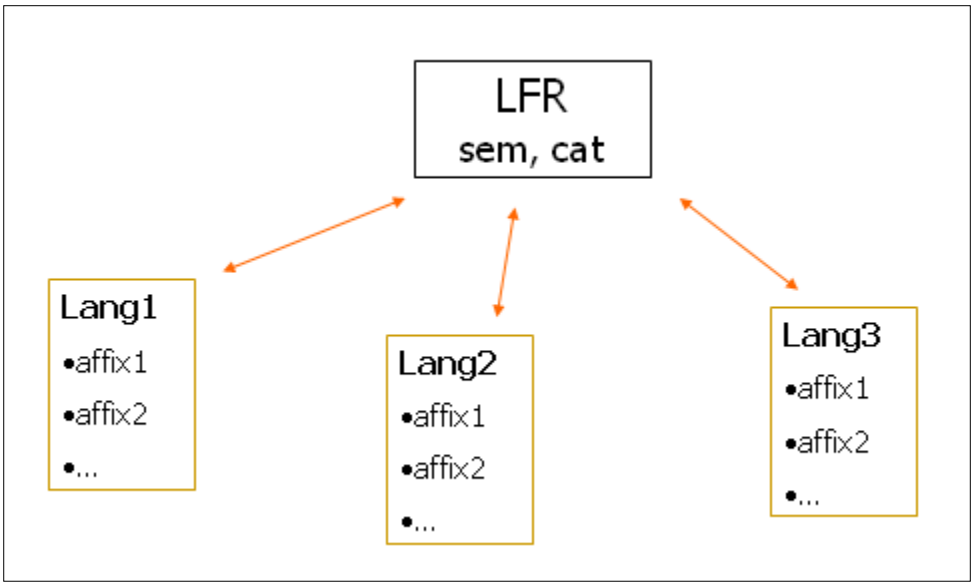

Figure 5. Two-level architecture of the database

As regards the implementation of the tool[2], the use of a multi-access and dynamic database such as this one enables users to access morphological information through different modes and languages. First, users can browse the database via the affix index for each implemented language (English, French and Italian), as is the case in any standard dictionary. Direct access is also provided to the corresponding multilingual LFR (e.g. the selection of Fr. *multi* gives access to the multilingual LFR of unspecified plurality). Users can also browse the LFRs via semantic labels (e.g. 'reiterativity', 'unspecified plurality', 'reversal and removal', 'inchoativity'), thus accessing multilingual LFRs and their respective affixes and constraints. A third access mode is also offered via the lexeme index (where the morphologically complex lexemes included in the database are listed alphabetically as headwords). The three types of access modes are described in Sections 3.1 to 3.3.

[2] The MuLeXFoR database is implemented in PHP and will soon be available on the web. Please contact the first author (Bruno Cartoni) for access information.

3.1. Affix browsing

Users can select the affix they wish to look up in the affix index and thereby get access to (1) the rules that the affix takes part in, (2) a complete description of each rule and (3) the corresponding equivalent affixes in the target languages. For example, users who wish to know how to express Engl. *multi* in French can first select the English prefix *multi* in the affix index. MuLeXFoR then provides the rule(s) that involve(s) this English prefix (in this case, 'unspecified plurality' to form adjectives from nouns (n>a) and nouns from nouns (n>n)). When clicking on one of these rules, users get a comprehensive description of the multilingual rule, including the equivalent affixes in Italian and French (*multi*, *pluri*, *poli*/*poly*), usage restrictions, and examples. This is illustrated in Figure 6.

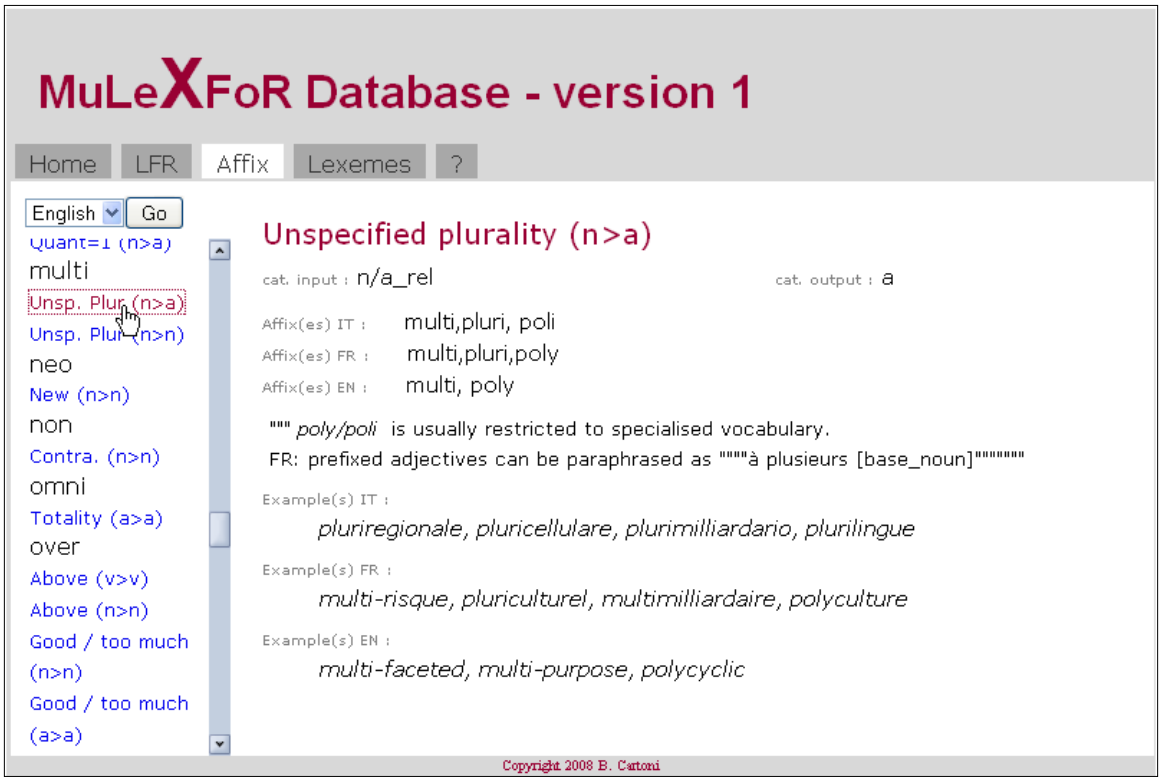

Figure 6. Browsing via the affix index

As can be seen in Figure 6, two usage notes are provided under the prefix inventories in each language. The first one, which is common to the three languages currently included in the database, states that *poly/poli* tends to be restricted to scientific vocabulary. The second usage note is specific to French and identifies non-morphological ways of conveying the same meaning as the prefixes (here by means of a prepositional phrase).

3.2. Rule browsing

The database can also be browsed via specific rules. Figure 7 illustrates the selection of the rule 'above' which produces prefixed relational adjectives from base nouns. Once we click on the rule name in the menu panel, the selected rule subsequently appears in the main panel. This provides various types of information (affixes, morphographemic information, etc.).

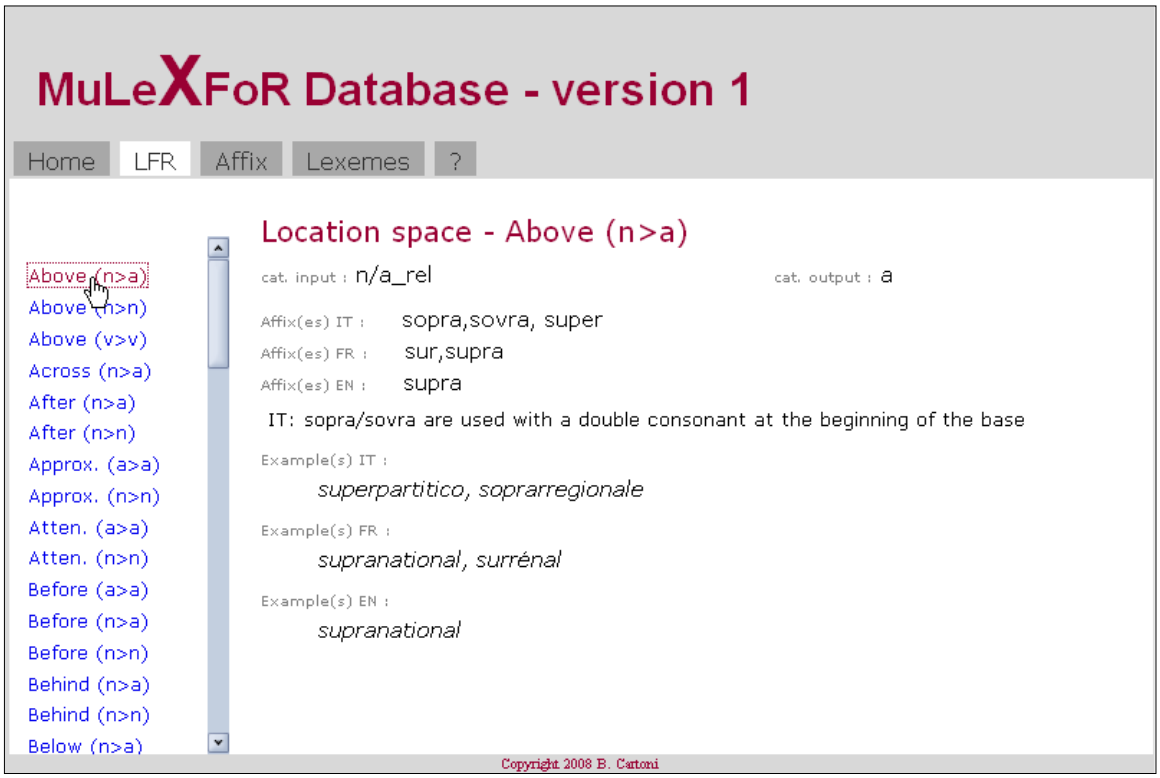

Figure 7. Browsing via the LFR index

3.3. Lexeme browsing

Any lexeme that is formed by means of a specific rule can provide direct access to the rule in question. This feature was implemented in the database by building an index with all the examples provided in the LFRs for the three languages (Engl. *anti-abortion*, *auto-suggestion*, *bi-directional*, *ex-model*, etc.). Thanks to this index, the integration of MuLeXFoR in an existing multilingual or bilingual dictionary could be envisaged. MuLeXFoR is not meant as a stand-alone application, but is rather conceived of as an add-in for multilingual and bilingual dictionaries.

There is also room for improvement in the form of a fourth access mode: words which are not included in the dictionary (e.g. neologisms) could be automatically analysed and subsequently matched to the corresponding rule (e.g. *re-look up* would be matched to the reiterativity rule). Needless to say, this feature would depend heavily on the efficiency of the morphological analyser used. However, tools such as Derif (Namer, 2009) are growing increasingly efficient and tests could be conducted to assess the reliability of morphological analysers in performing this task.

4. Adapting the database to learners' needs: MuLeXFoR-Learners

The MuLeXFoR database was originally designed for a wide range of users, i.e. it did not target an audience in particular. However, we soon realised that the labels used in MuLeXFoR-Linguists were too opaque for L2 learners or trainee translators who might struggle with terms such as 'reiterativity', 'unspecified plurality', 'inchoativity', etc.

To address this issue, another version of the database (based on the same resources) was created: MuLeXFoR-Learners. Starting from the assumption that L2 users' knowledge of the terminological descriptors used to label morphological processes and items is relatively limited, the labelling used in MuLeXFoR was considerably simplified. For example, labels such as 'again' and 'many' were inserted instead of 'reiterativity' and 'unspecified plurality'. In addition, the names of the different indexes were adapted so as to be easily understood by non-specialists (i.e. 'meaning', 'affix', 'word'), as illustrated in Figure 8.

Figure 8. MuLeXFoR – student version

Learner-specific information was also added in the database entries. Learners who use MuLeXFoR to learn how to create new words in the L2 need to identify which affixes to use in order to coin a given meaning. Consequently, whenever the LFR provides several affixes for a single rule, different notes are provided with a view to helping learners make informed choices. Providing this kind of information, however, is not an easy task. Despite some studies on morphological productivity and variation and affix alternation (e.g. Amiot, 2005; Cartoni, 2008c), affix selection remains an under-researched area. A strength of MuLeXFoR is that we benefited from insights gained from a number of extensive corpus-based studies on productivity (e.g. Lefer, 2009), which made it possible to add productivity notes. For example, in the 'before' LFR, the comment *EN: pre is much more productive than fore* helps users select *pre* rather than *fore* to express temporal anteriority.

We also provided lexical or syntactic equivalents of the morphological items described in the database as these represent interesting alternatives for users who lack confidence coining complex words in their L2. For example, the 'approximation' LFR (Engl. *quasi*) provides the note *EN: near-X compounds are also frequently used in English to convey the idea of approximation, as in 'near-dark*'. Other features could be added in future stages of the database development such as common translation errors for instance.

5. Feeding the database: the contribution of corpus data

As in any lexicographic work, implementation (i.e. acquiring the data to feed into the database) is a thorny issue. In morphological resources such as MuLeXFoR, two main types of knowledge needed to be acquired and formalised. On the one hand, the multilingual semantic rules had to be singled out and formalised. On the other, productive affixes (or other productive morphological processes) corresponding to these rules needed to be identified cross-linguistically.

The first implementation step was largely inspired by the linguistic literature that provides abstract – and hence cross-linguistically valid – semantic descriptions of morphological processes. As argued in Szymanek's (1988) study, morphological processes are closely related to basic cognitive notions, such as movement, modality, evaluation, etc. By examining various semantic descriptions of prefixation in different languages (e.g. Montermini, 2002 and Iacobini, 2004 for Italian and Amiot and Montermini, 2009 for French), six major semantic categories were identified to formalise prefixation cross-linguistically: location, evaluation, negation, quantity, modality, and inchoativity. These categories have been further divided into subcategories. For example, location is divided into space and time, and within spatial location, a distinction is further made between different positions (in front of, behind, beside, etc.). Gathering these descriptions allowed us to obtain a rather exhaustive and fine-grained set of prefixation rules (see Cartoni, 2008b for further details). The semantic categories implemented in MuLeXFoR currently focus on prefixation and, to a lesser extent, conversion. Even though suffixation is usually said to be more abstract and semantically less specified than prefixation (as its main role is to change the category of the base), a similar approach could be applied to suffixation.

Corpus-based methods and tools were used in the second stage where we aimed to determine which prefixes contribute to which rule(s) in the three languages investigated. We drew from the results of a detailed study on word-formation which focussed on machine translation from Italian into French. This study heavily relied on corpus data (*La Repubblica Corpus*; Baroni et al., 2004) (see Cartoni, 2008b). The English data along with additional French data were collected within the framework of a corpus-based contrastive study of English and French prefixation across genres (press editorials, novels and scientific articles) and disciplines (medicine, linguistics and economics) (c. 100 prefixes in each language were investigated; see Lefer, 2009). Both corpus-based studies made it possible to single out productive prefixes in the three languages investigated, together with authentic examples of neologisms formed with these prefixes.

MuLeXFoR currently contains more than 60 multilingual LFRs and c. 50 productive prefixes in French, Italian and English. Further data acquisition methods are presently under investigation to increase the coverage of the database.

6. Future developments

We first wish to extend the resource to other languages (e.g. German) and to suffixation processes. As regards other morphological items, such as combining forms, their inclusion in the database raises many more issues as they do not fit into the semantic categories currently included in the database. Semantically, these elements of Latin and Greek origin are still very closely related to the lexeme they come from and are therefore characterised by a greater lexical content than affixes (e.g. *bio*, *eco*, *geo*, *hydro*). The localisation of the interface and of the content (comments, meta-information, etc.) in other languages than English is also planned.

In addition to the obvious extension of the tool to other affixes and word-formation processes and to other languages, an assessment of the database is also planned. Two aspects will be examined: the usefulness of the database and its interoperability with other tools. The

usefulness of the database will be assessed in terms of users' expectations and needs, with special emphasis on L2 learners and trainee translators. To our knowledge, the evaluation of affix representation in dictionaries has mainly been carried out by linguists and lexicographers, while assessments involving actual users are still sorely lacking. In view of the innovative aspect of our approach, it is essential to evaluate the database in terms of human-computer interaction. Particular attention will be paid to the comprehensibility of the labels and meta-information used in the database. Indeed, even in the student version, it is often difficult to present morphological information for non-expert users. We will therefore focus on help files and pop-ups. Second, the interoperability with existing multilingual lexicographic databases and tools will be assessed. As mentioned above, MuLeXFoR is not meant as a stand-alone application but is intended for inclusion into a larger dictionary. This stage raises many issues of data organisation and browsability.

7. Concluding remarks

MuLeXFoR is a first attempt at representing morphological information in a multilingual lexicographic environment. It is well-known that word-formation is an essential component of the lexicon. However, bilingual dictionaries fail to adequately describe word-formation processes and items. The database presented in this paper relies on lexematic morphology, which makes it possible to formalise word-formation processes cross-linguistically. The use of a semantic categorisation to represent affixes in different languages seems to be a promising starting point to improve and systematise the morphological information presented in dictionaries. The database also heavily relies on corpus data extracted from multilingual corpora, making it possible to include usage notes (e.g. about register and genre variation) and propose non-morphological (near-) synonymous alternatives (lexical or syntactic). Although MuLeXFoR is still under development, we hope that the framework presented here will contribute to the improvement of the representation of morphological items in multilingual lexicographic tools.

References

Amiot, D. (2005). 'Plusieurs vs poly-, pluri- et multi-'. In Flaux, N.; Amiot, D. (eds.). *La quantification côté déterminants et côté préfixes*. Verbum 27 (4). 403-417.

Amiot, D. and Montermini F. (2009) 'Affixes et mots grammaticaux'. In Fradin B., Kerleroux F. and Plénat M.(eds.) *Aperçus de morphologie du français*. Saint-Denis, Puv. 127-141

Baroni, M.; Bernardini, S.; Comastri, F.; Piccioni, L.; Volpi, A.; Aston, G.; Mazzoleni, M. (2004). 'Introducing the 'la Repubblica' corpus: A large, annotated, TEI(XML)-compliant corpus of newspaper Italian'. In *Proceedings of LREC 2004*, Lisbon. 1771-1774.

Bernal, E.; DeCesaris, J. (2008). 'A Digital Dictionary of Catalan Derivational Affixes'. In Bernal, E.; DeCesaris, J. (eds.). *Proceedings of the XIII EURALEX International Congress* (Barcelona, 15-19 July). Barcelona: Institut Universitari de Lingüistica Aplicada, Universitat Pompeu Fabra.

Cartoni, B. (2008a). 'La place de la morphologie constructionnelle dans les dictionnaires bilingues: étude de cas'. In Bernal, E.; DeCesaris, J. (eds.). *Proceedings of the XIII EURALEX International Congress* (Barcelona, 15-19 July). Barcelona: Institut Universitari de Lingüistica Aplicada, Universitat Pompeu Fabra. 813-820.

Cartoni, B. (2008b). *De l'incomplétude lexicale en traduction automatique : vers une approche morphosémantique multilingue*. PhD thesis. Geneva: Université de Genève.

Cartoni, B. (2008c). 'Mesure de l'alternance entre préfixes pour la génération en traduction automatique'. In *Proceedings of TALN 2008*, Avignon.

Dardano, M.; Frenguelli, G.; Colella, G. (2006). 'What Lexicographers Do with Word Formation'. In Corino, E.; Marello, C.; Onesti, C. (eds.). *Proceedings XII Euralex International Congress*. Torino, Italia, September 6th-9th, 2006. Allessandria: Edizioni dell'Orso. 1115-1127.

Fradin, B. (2003). *Nouvelles approches en morphologie*. Paris: Presses Universitaires de France.

Fradin, B. (2007). 'On the semantics of Denominal Adjectives'. Ralli A., Booij G. & Scalise S. (eds) In *Online Proceedings of the 6th Mediterranean Morphology Meeting, Ithaca, Greece*.

Hacken, P. ten; Abel, A.; Knapp, J. (2006). 'Word formation in an electronic learners' dictionary: ELDIT'. *International Journal of Lexicography* 19 (3). 243-256.

Iacobini, C. (2004). 'I prefissi'. In Grossmann, M.; Rainer, F. (eds.). La formazione delle parole in italiano. Tübingen: Niemeyer. 99-163.

Lefer, M.-A. (2009). Exploring lexical morphology across languages: a corpus-based study of prefixation in English and French writing. Unpublished PhD thesis. Louvain-la-Neuve: Université catholique de Louvain.

Montermini, F. (2002). Le système préfixal en italien contemporain. Unpublished PhD thesis. Université de Paris X-Nanterre – Università degli Studi di Bologna.

Namer, F. (2009). Morphologie, lexique et traitement automatique des langues, l'analyseur DeriF. Paris: Hermes-Lavoisier.

Prcic, T. (1999). 'The treatment of affixes in the 'big four' EFL dictionaries'. International Journal of Lexicography 12(4). 263-279.

Szymanek, B. (1988). Categories and Categorization in Morphology. Lublin: RW-KUL.