

Lexical Functions vs. Inflectional Functions

Maarten Janssen

IULA / Barcelona

maarten.janssen@upf.edu

Abstract

Lexical Functions define structural relations between word-senses, based on semantic criteria. Inflectional functions, on the other hand, define derivational relations between words and their derived forms from a functional or semantic perspective. Inflectional functions are inspired upon lexical functions, and in many cases, the two types of functions define the same type of relations between the same pairs of words, but at a different level: the inflectional functions operate at the level of the lexical entry, the lexical functions at the level of the word-sense. This paper provides an overview of the similarities and differences between inflectional and lexical functions, and discuss the advantages and limitations of the use of inflectional functions in large-scale morphological databases.

1 Introduction

Lexical Functions (henceforth LFs) provide a way, amongst other things, to link several types of derived words to their morphological base in a structural, semantically oriented fashion (Žolkovskij and Mel'čuk, 1965). For instance, the lexical function S_1 in (1) links the deverbal noun *walker* (the value of the LF) to its verbal base *walk* (the argument of the LF), while at the same time specifying the fact that the derived noun is in fact the agentive noun expressing *someone who walks*.

(1) $S_1(\text{walk}) = \text{walker}$

That is not to say that the lexical function is intended to model morphological relations: the same relation of S_1 can hold in cases where there are two words with the right semantic relation between them, but where no morphological relation exists. An example is given in (2), where *vendor* is the word for someone who sells, but it is not a word that is morphologically related to the verb *sell*. However, in the majority of S_1 cases the noun will be morphologically derived from the verb.

(2) $S_1(\text{sell}) = \text{vendor}$

Given that lexical functions are semantically-oriented relations, they relate word-senses rather than word-forms or words. For instance in the case of (2), the word *vendor* only relates to the verb *sell* in its basic meaning of 'giving or passing something in exchange for money'. Although you can paraphrase *John sells books* to *John is a vendor of books*, you can hardly say the same for the verb *sell* in the other meanings it is, as used in such sentences as *I sold the idea to my boss*, or *This CD sold five million copies*. Therefore, the argument in (2) should strictly speaking be interpreted as the first meaning (or lexie) of the word *sell*.

The fact that lexical functions operate on word-senses rather than words makes them even more fit for the representation of derivational relations: in general, derivations take a word as their input, but the meaning

of the derived word is (often) dependent on a specific meaning of the base word. For instance, the verb *problematize* means ‘to turn something into a problem’, but the word *problem* in this definition should be interpreted specifically as ‘a difficulty that needs attention and thought’ and not a problem in its meanings of a mathematical problem or someone who causes difficulty.

However appropriate lexical functions might be with respect to strictly derivational relations, they fare less well on a group of (derivational) forms that are partially inflectional in nature. These are the forms referred to as for instance *inherent inflections* (Booij, 1995) or *transpositional inflection* (Bauer, 2004). They are the forms that you do not find in the dictionary (or find in the dictionary as a *run-on* without definition), because they are assumed to be implicitly defined by the main article. Examples are (regular) diminutives like the Dutch *kindje* (small *kind* = child), deadjectival adverbs such as *roughly*, superlatives like the Spanish *bonísimo* (very *bueno* = good), etc. In this article, this class of forms will be loosely referred to as *inherent inflections*, though not necessarily with the same meaning as intended by Booij (1995; 2004).

One of the characteristics of inherent inflections is that, like inflection, they apply to entire words independent of the word-sense. Some of these forms are reflected by lexical functions, such as for instance the relation between a verb and its event nominal as in (3).

(3) $S_0(\text{produce}) = \text{production}$

The noun *production* is the noun for the event related to the verb *produce*, independently of which of the several meanings of the words *produce* is taken into consideration. It is *the* nominal form of *produce* and not just a noun derived from it.

This means that in order to get the correct relation between each meaning of the verb *produce* and its respective event noun, there have to be as many instances of lexical functions between *produce* and *production* as there are meanings of the verb. Or in other words, the use of lexical functions for the relation between *produce* and *production* misses an important generalization, namely the fact that the relation is not restricted to a specific word meaning.

This lack of generalization was one of the motivations for the introduction of a notion of *inflectional functions* (henceforth IFs), which is a more morphologically oriented counterpart of lexical functions, operating at the level of words rather than at the level of word-senses (Janssen, 2005a). Inflectional functions have been subsequently implemented in a set of large-scale, full-form lexica called OSLIN (Open Source Lexical Information Network), starting with a lexicon for Portuguese, and currently with lexica for English, Spanish, and Catalan under development.

Given that inflectional functions are not merely a theoretical idea, but a practical solution implemented on complete lexica, the use of IFs unavoidably ran into conceptual and pragmatic problems. This article will first explain the notion of inflectional functions and its implementation in the OSLIN databases, and then discuss the advantages and limitations of inflectional functions in detail.

2 Inflectional Functions

Inflectional functions define relations between lexical entries and their derived forms. They therefore in principle model derivations, but are called inflectional because they deal with the forms that are midway between derivation and inflection, the forms of what you might call the *extended paradigm*. IFs do not model the process of word-formation: they do not indicate the type of process that led to the creation of the one from the other, but they define a semantic or functional relation between words independently of the actual process that was involved in the morphological derivation. An example of an inflectional function is given in (4), where **s0v** defines a relation between a verb and its related event nouns.

(4) $s0v(\text{descend}) = \text{descent}$

Deverbal event nouns are not the only inherent inflections modeled by IFs: other IFs are the female forms of animate nouns (not very productive in English, but *actor* - *actress* would be an example), diminutive forms of nouns, comparative adjectives, etc.

Inflectional functions are mid-way between morphological relations and semantic relations, in the sense that they define relations with both morphological and semantic characteristics. For instance, in order for there to be an **s0v** relation between two words *X* and *Y*, the following conditions have to be met:

- (a) *X* has to be a verb, and *Y* has to be a noun
- (b) *Y* has to be morphologically related to *X*
- (c) *Y* has to be an event noun referring to the abstract event expressed by the verb *X*
- (d) *Y* has to apply to all meaning of the verb *X*

Because of the semantic requirement (c), inflectional functions do not provide purely morphological information. Although the noun *ignorance* is a deverbal noun, derived in the same way from the verb *ignore* in the same way as the word *acceptance* is derived from the verb *accept*, there is a **s0v** relation between the latter two, but not between the former. Synchronically there is no direct semantic relation between *ignorance* and *ignore*, although probably there originally was. Because of the morphological requirement (b), IFs are not really semantic in the sense that LFs are (they are not used for cases like *vendor*). And because of the lack of an indication of the morphological process, IFs do not provide full morphological information since they do not specify *how* the noun is derived from the verb, but merely require that it is.

It is mainly requirement (d) that makes inflectional functions usable mainly for inherent inflections, and not for just any type of derivational relation. Inherent inflections are derivational relations that are largely inflectional in nature in the sense that the derived word can be seen as a context-dependent word-form of the word it is derived from. For instance, in the case of the relation between *descend* and *descent* in (4), the noun can be seen as the nominal form of the verb, that is to say, the form of the verb that should be used in denominal constructions. For instance, the two sentences in (5) express the same content, except that the second sentence uses a light-verb construction. Because of the light-verb construction, the use of the verb *descend* in its nominal form is required, and therefore the adverb *rapidly* has to be used in its (base) adjectival form. You could therefore say that the shift for a verb to a noun is a case of (transpositional) agreement.

- (5) He descended rapidly. ⇒ He made a rapid descent.

Inherent inflections also behave like regular inflections in the sense that the form the event noun takes is not dependent on the meaning that the verb is used in: the form *sold* simply is the past tense of the verb *sell*, and not the past tense of the verb *sell* in a specific meaning. It cannot be the case that when we are talking about selling of ideas for instance, the past tense takes a different form, say *selled* (with some exceptions which will be discussed later). And neither can the nominal form be anything different from *descent* when we use the verb in for instance its meaning of ‘the arrival of many people at the same time’.

Although the inherent inflection has to be the same for each meaning of the base word, it does not have to be unique. There can be various alternative deverbal event nouns for the same verb. For instance, the event noun for the Portuguese verb *manifestar* (to manifest) is *manifestação*, but also *manifesto*. This is comparable to the situation for verbal inflection: the past tense of *blow* is either the irregular *blew* or the modern day regularization *blowed*. But both past tenses are correct for each meaning of the verb *blow*, and so are both *manifestação* and *manifesto* correct deverbal nouns for each use of the verb *manifestar*. There can be differences in register between the variants, and there can even be a preference for a specific form related to a specific meaning, in the sense that *blowed* is a colloquial form, mostly used for colloquial meanings of the verb, but all variants should be correct for each meaning.

2.1 Words and Inflectional Functions

Inflectional Functions operate on the level of *words*. What exactly a word is in this context is best understood by looking at inflection proper. The past-tense *rung* does not belong to the sequence of letters *ring*, but to the verb (*to*) *ring*, since the noun *ring* does not have a past tense. It does not even belong to “any” verb *ring*, but to one of the homonyms of *ring* (to call; to sound a bell), since there is another verb *ring* (to put a ring around the foot of a bird) which has the past tense *ringed*. And the past tense also does not really belong to the abstract notion of a word but is more anchored in orthography: the Portuguese word for *wet* can be written as either *húmido* or *úmido* and the two forms are typically said to be different ways of writing the same word. But the female form *húmida* only belongs to the first form, and not to the second. Inflectional paradigms belong to something that is somewhere in the middle: the lexical entry, the headword as it appears in the dictionary, or to what in MTT is called the *vocable* (Mel’cuk et al., 1995).

The same holds for the extended paradigms defined by the inflectional functions. When there are different orthographic realizations of a verb, there can be different deverbal nouns for each of them: the Portuguese noun *doiramento* is the **s0v** of *doirar* (to gild) and not of its orthographic variant *dourar*. And when a noun is homonymous, it can have different female forms for each homonym: the Catalan noun *croat* has a female form *croata* when it indicates someone from Croatia, but a *croada* when it is a female crusader. But all the different word-senses (or lexies) related to a polysemous adjective have to share the same superlative form(s). So the argument of an inflectional function is a lexical entry, just like a in the case of inflection proper¹.

Although (extended) inflectional paradigms belong to lexical entries, not all the inflectional forms have to be realizable in each word-sense. There is only one noun *water* in the English language, and its plural is *waters*. But most commonly, *water* is used as a mass noun and does not have a plural. To avoid having to resort to defining homonymous entries for all words that can be used as mass nouns or count nouns (which given the *universal grinding* mechanism would be most nouns), or having to define inflection at the level of word-senses, one has to conclude that inflectional paradigms can be defective in only some of the acceptations of a lexical entry: *waters* is the plural for *water* in all senses, but it is always realized. The same holds for inherent inflections: the Portuguese word *amarela* is the female form of the noun *amarelo* (yellow), but can be used only when the word *amarelo* denotes an animate object (a pale person). We call this phenomenon *partial defectivity*.

Since the argument of an IF is not a word-form, neither is its value: the Catalan word *peixatera* (fishwife) is the female form of the word *peixater* (fishmonger). But *peixatera* is a lexical entry in its citation form, and has two inflected forms *peixatera* and *peixateres*. The value of an IF is not a lexical entry either: only in its basic meaning is the word *construction* the **s0v** of *construe*; when used for a building, it is not a deverbal event noun. Therefore, the value of an IF is in principle a word-sense, as indicated in figure 1.

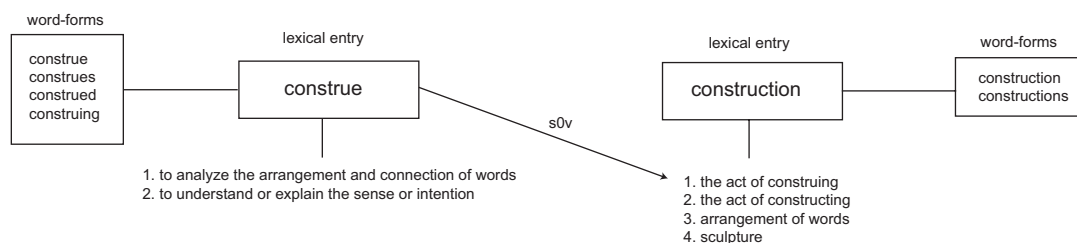


Figure 1: Inflectional Function for *construction*

Inflectional functions are explicitly intended to model forms that are mid-way between derivation and

¹The fact that the boundary between polysemy and homonymy is not absolutely clear-cut can lead to problems, as will be discussed in 3.2.

inflection, and model them in such a way that they can be interpreted either way depending on the theory/application they are used with. When considering the relation **female**, it can be taken to define morphologically derived forms, or it can be taken to define (additional) inflected forms for the base noun, as is often done for the Romance languages. But since the value of the IF is not a word-form, interpreting them as inflectional has to be done indirectly: if we want to take **s0v** as an inflectional relation (see figure 1), the word-forms (*construction* and *constructions*) of the lexical entry (*construction*) that the value of the s0v (‘the act of construing’) belongs to should be taken as inflectional forms of the lexical entry (*construe*) that is the argument of the lexical function. So indirectly, *construction* and *constructions* would be (inherently inflectional) forms of the lexical entry *construe* if **s0v** is taken to define an inflectional relation. Along the same lines, *peixatera* and *peixateres* would end up as additional inflectional functions of *peixater*.

2.2 The Open Source Lexical Information Network

The Open Sources Lexical Information Network (OSLIN) is a framework for relation-based, large-scale, full-form lexica (Janssen, 2005b). The first lexicon developed in the framework was a lexicon for Portuguese called *MorDebe*, built at the ILTEC institute in Lisbon. *MorDebe* contains around 135.000 lemmas and just under 1,5 million word-forms. Currently, lexica for Catalan, Spanish, and English are under development at the IULA institute in Barcelona, each of which currently contains over 100.000 lemmas. All these lexica are in a first instance meta-lexicographic, providing a more structured representation of the nomenclature of existing dictionaries. But the lexica have a mechanism for maintenance, and are integrated with a semi-automatic neologism detection program called NeoTrack (Janssen, 2008), meaning that they are gradually being expanded with additional lemmas. All the OSLIN lexica can be consulted online via *oslin.org*.

OSLIN is a full-form lexicon model that not merely specifies all inflectional forms explicitly, but also attempts to list all the inherent inflections of the extended paradigms. For the most extensively developed lexicon, the Portuguese *MorDebe*, all dictionarized deverbal event nouns were collected, as were all the dictionarized nominal gender cases (*gato* - *gata*) and all the deadjectival quality nouns. Apart from these extensively explored relations, a number of other lexicalized relations was collected as well, including diminutive forms and adjectival superlatives. All these relations are modeled in terms of inflectional functions. The relations used, as well as the number of relations stored, are listed in table 1. As a special case, also the *gentiles* were collected extensively, which are not relations between lexical entries and derivatives, but between toponyms and the nominal or adjectival form associated with it.

s0v	7347	event noun	acertar	-	acerto
female	8124	nominal gender	actor	-	atriz
adv0	1974	adverb of mood	abjecto	-	abjectamente
a2v	3716	participial adjective	abalar	-	abalado
able2	239	-able adjective	absorver	-	absorvível
dim	265	diminutive	animal	-	animalzinho
aum	79	augmentative	monte	-	montão
max	265	adjectival superlative	magro	-	macérrimo
genta	1561	gentile	Aachen	-	aacheniano

Table 1: Inherent inflections in *MorDebe*

Inherent inflections are in general rather regularly formed, although there are some irregular forms. But for many of the inflectional functional relations, there are various ways for deriving the same form. For instance, deverbal event nouns can be regularly formed in Portuguese with the suffixes *-ção*, *-mento*, *-dura*, *-dela*, and *-o*. For some verbs, all of these are used. For instance, for the verb *abafar* (to choke), there are the forms *abafação*, *abafamento*, *abafadura*, *abafadela*, *abafo*, and *abafeira*. But in most cases, only one

of the potential forms is actually used, and the others are considered incorrect. The correct event noun for *abolar* (to do away with) is *aboladura*, for *abdicar* (to abdicate) it is *abdicação*, for *abocar* (to put into your mouth) *abocamento*, whereas other forms such as **abocação* and **abolamento* are not correct. Which form(s) is the correct one is something that has to be learned/defined for each individual verb. Therefore, using derivational rules for these kind of relations helps to predict potential forms, but does not specify the actual lexical items used. The use of inflectional functions to model these relations systematically has proven to be a powerful way of structurally treating this strongly lexicalized type of information.

2.3 Comparison between LF and IF

Lexical functions and inflectional functions are fundamentally different animals: the arguments of IFs are lexical entries, the arguments of LFs are word-senses (lexies). IFs are morphological in nature, LFs are not. And LFs can be used to specify collocational relations, whereas IFs cannot. However, for many of the IFs, there is a corresponding LF, as shown in table 2. The first items on the list are standard LF, the ones with a star are the LF proposed in the extension of the DECIDE project (Grefenstette et al., 1996).

s0v	event noun	S ₀	0-th role noun
s0a	quality noun	S ₀	0-th role noun
a2v	participial adjective	A ₂	2nd role adjective
able2	-able adjective	Able*	Adjective of possibility
female	gender noun	Female*	Animate nominal
max	adjectival superlative	Max*	Maximal degree

Table 2: Inherent inflections - Lexical Functions

Because of the difference in nature, there is no direct mapping between LFs and IFs: not every S₀ will have a corresponding s0v. Not only because the S₀ relation is also used for deadjectival nouns, but also because IFs require a morphological link between the two elements of the relation whereas the LFs do not. Therefore, there is a Female LF relation between the Dutch words *kat* (cat) and *poes* (she-cat), but there is no **female** IF relation between them because *poes* is not morphologically derived from *kat*. And similarly, there is a S₀ relation between *steal* and *theft*, but there is no corresponding **s0v** because there is no morphological relation between the two words.

But although S₀ relations do not automatically lead to **s0v** relations, for every IF for which there is a corresponding LF the first implies the second. So the fact that there is a **s0v** relation between the verb *coincide* and *coincidence*, implies that there will be a S₀ relation for each of the word-senses of the verb *coincide* (although there are more word-senses for the noun *coincidence* that are not related to *coincide*, and there might be cases of partial defectivity). In that sense, the OSLIN lexica provide large-scale repositories of lexical functions, encoded in an efficient way.

3 Practical and Theoretical Issues

The advantages of using Inflectional Functions over Lexical Functions in a full-fledged lexicon system like OSLIN are easy to point out. First of all, OSLIN at this moment only contains word-forms and lexical entries (and proper names). Although it is the idea that in the future, a level of word-senses will be added, at this point in time that is still not the case. In the absence of word-senses, it is impossible to define LFs, since LFs by definition require word-senses.

But even if there were word-senses in OSLIN at this time, modeling the types of relations currently defined by IFs in terms of LFs would not only be inefficient, but also theoretically inaccurate. Consider a relation which could be expressed by lexical functions: diminutives. In Catalan, there are various highly polysemous nouns, such as for instance the word *mà* (hand), for which the academy dictionary (DIEC2)

lists 17 main senses and a whole range of sub-senses. In all of these meaning, the word *mà* can be expressed in diminutive form, and in each of the senses, the diminutive form is *maneta*. Listing an extensive list of relations between *mà* and *maneta* in every sense would not only create a redundant amount of relations, it would also incorrectly suggest that it is something specific for each individual meaning of *mà* that its diminutive form is *maneta*, whereas in fact it is a entry-based property of the word *mà*.

However, despite the obvious advantages, the application of inflectional functions in OSLIN raised several complications. Some of the most important ones will be discussed in this chapter. The first problem, that of the failure to distinguish between alternative forms, is not specific for inflectional functions, but equally affects lexical functions. The other problems however, are more particular for inflectional functions.

3.1 Distinguishing Alternatives

As mentioned in the previous chapter, there are often various alternative forms for the same inherent inflection. For instance, for the verb *lavar* (wash) there are 6 different event nouns mentioned in the dictionary: *lavação*, *lavadura*, *lava*, *lavadela*, *lavagem*, and *lavamento*. In principle, all these words express the same concept: the act of washing (or cleaning). But not all these words express that meaning in exactly the same way, and not all words are usable in the same context. What exactly is the difference between them is difficult to make explicit, but the word *lava* is uncommon, and is unlikely to be used in every-day expressions. The word *lavadela* has a diminutive ring to it, and is mostly applied to light washings. Where *lavagem* is used for things that have more to do with the effect of washing, *lavamento* is more associated with the result of washing (although the difference between those is far from clear).

Inflectional functions do not provide any means to distinguish between the different ways of expressing the different event nouns for *lavar*, or distinguish between any of the other alternative forms of inherent inflections. This failure to distinguish between different alternatives, or even indicate which is the most frequently used or common one, is a serious drawback in the current presentation of the inherent inflections in the (Portuguese) lexicon. And alternative inherent inflections are not rare: for instance for the deverbal nouns, there is more than one alternative for about 30% of all the verbs.

Notice that this is not a problem that is specific for the inflectional functions: since all of the event nouns are usable for each of the meanings of *lavar*, and definitely all of them with the primary meaning of *lavar*, the relation S_0 between the word-senses of *lavar* and the various event nouns would equally fail to distinguish the subtle differences between them. In fact, the problem is even slightly bigger for the lexical functions, since in the case of LFs, also the non-morphologically related event nouns are taken into account, and the lexical functions are not capable of distinguishing between *stealing* and *theft* as the S_0 for *steal*. It is an indication that although relational models like IFs and LFs partially define the meaning of the words they relate to, they do not define the meaning and use in full.

3.2 Stacking Failures

Inherent inflections can in principle be stacked. As shown in example (6), the female form and the diminutive form can be stacked to derive the female diminutive form *gatinha* (little she-cat). And as shown in (7), there is even more than one way to obtain the same result by stacking the diminutive and the female in the reverse order.

(6) **female** (*gato*) = *gata*, **dim** (*gata*) = *gatinha* $\Rightarrow$ **dim(female** (*gato*)) = *gatinha*

(7) **dim** (*gato*) = *gatinho*, **female** (*gatinho*) = *gatinha* $\Rightarrow$ **female(dim** (*gato*)) = *gatinha*

In this particular case, the meaning of the two is identical, although this is not always the case: there are such cases as the negation of the -able form of *do* and the -able form of the negation of *do*. In both cases, the correct result is *undoable*, but depending on the order of application, it is either something that cannot be done, or something that can be undone, which is hardly the same thing.

Although in principle stacking is a positive feature of inflectional functions, in certain cases it leads to undesirable results, because of two features of inflectional functions. The first feature involved is the fact that with respect to their inherent inflections, paradigms can be partially defective as explained in 2. The second feature involved is the fact that although inflectional functions take lexical entries as their arguments, the *value* of an inflectional function is in principle a word-sense and not a word: the word *llibret* in Catalan is the diminutive of *llibre* (book), but the diminutive form has obtained the lexicalized meaning of the little booklets used in the opera (libretto). As a pragmatic solution to this problem on the OSLIN website(s), in those cases, the entry for *llibret* does not display that it is the “diminutive form of *llibre*”, but that it is “*also* the diminutive of *llibre*”, indicating that it has other meanings as well.

When these two features apply to the same word, the situation becomes rather complicated. The word *ilha* (island) in Portuguese has several diminutives forms, one of them being *ilheu*. But apart from being the word for a little island, the word *ilheu* also means somebody living on the island, or an islander. Now the female form of *ilheu* in its meaning of an islander is *ilhoa*. Therefore, the definition for the word *ilheu* says that it is (also) the diminutive of *ilha*, and that it has a female form *ilhoa*. But given that there are no female little islands, the female forms does not apply in the case of the primary meaning of *ilheu*. Although this technically speaking does not say anything incorrect, it really stretches the limits of what can be done with inflectional functions. For those cases, a more word-sense driven approach like lexical functions would be much less problematic.

It should be noted that however problematic this stacking problem is, it is a practical, and not a theoretical issue: since the argument of an inflectional function is a lexical entry, and the value a word-meaning, the value a one lexical function cannot be taken as the argument of a next one. That means that strictly speaking, inflectional functions cannot be stacked as in (6). But in the practical use of inflectional function, cases such as *ilha-ilheu-ilhoa* present a serious problem for a coherent treatment.

3.3 Sense-Specific Exceptions

One of the crucial features of inherent inflections is that it is a lexical entry-level phenomenon. This implies amongst other things that all event nouns that are specific to a given word-sense should not be considered inherent inflections. Up to a point, this situation is comparable with the situation of regular inflection. The aforementioned meaning dependent past tense *ringed* versus *rang/rung* can only exist because the word *ring* is homonymous. Within the OSLIN framework, the existence of multiple inflectional paradigms is even taken as a criterion for homonymy. Since inflection is taken as a definitional criterion for the identity of words, even cases like *hang* are considered homonymous: it has either *hung* as its past tense, or *hanged* in its meaning of *killed by hanging*. This despite the fact that the latter is clearly both etymologically and semantically related to the general meaning of *hang*. Modeling the inflection correctly without assuming there to be two lexical entries for *hang* is extremely complicated, both from a practical and from a theoretical perspective.

For inherent inflection, basically the same holds: if a word has two different female gender nouns, as in the case of the *croat* in section 2.1, then the word has to be considered homonymous. However, given that inherent inflections are not really inflectional, there are several possible solutions when there are two meaning-specific inherent inflections for a given word. Consider the French word *fille* (girl; daughter)². Most French dictionaries list a single entry for both meanings of the word, considering it to be a case of polysemy rather than homonymy. However, the diminutive form *fillette* can only be a small girl, and not a small daughter. This means that *fillette* can not be considered a meaning-independent inherent inflection of *fille*. There are three different ways to solve this problem.

The first option is to say that partially because of the different diminutive forms, the word *fille* should be considered homonymous, contrary to how dictionaries treat them. There are independent reasons for doing

²Example provided by an anonymous reviewer

so (Polguère, 2008), but the different extended paradigms could be taken as a strong indication that the word is synchronically really homonymous rather than polysemous. In that case, the duplication of the lexical entry will solve the problem since the diminutive will apply to all meanings of one of the entries.

The second option is to say that *fille* is partially defective: it is not that there is another diminutive for the meaning of ‘daughter’ that is correct, it is that in its meaning of a daughter, the word does not have a diminutive. Since there is no diminutive for the French words *fil*s (son), *père* (father), or *mère* (mother) either, this could simply indicate that daughter is not of the right type to take a diminutive. That would mean that *fille* does not have a diminutive in its meaning of ‘daughter’ in the same way as *water* does not have a plural as a mass noun.

The third option is to say that the word *fillette* in this case is not an inherent inflection, but rather a meaning-dependent derivational form that happens to be formed as a regular diminutive. It is not transparently *the* diminutive form of *fille*, just a word for a small girl. In that case, there will not be an inflectional function between the two words, and the meaning-specific relation will have to be treated with, for instance, lexical functions.

Which of these three solutions is the most correct one is something that has to be considered for each individual case. In the case of *fillette*, the first of these options seems to be the most reasonable, but all three solutions will resolve the problem of seemingly meaning-specific inflectional functions. But it should be noted that for inflection proper and the clear cases of inherent inflections, cases of potential meaning-specific inherent inflections are rare.

3.4 Marginally Inflectional Forms

In the application of inflectional functions to the entire lexicon, it is obvious that there are many marginal cases. For instance, when creating a full list of all **s0v** cases, there are many examples in which it is debatable whether the noun in question is really an **s0v** deverbal noun. The first major reason for doubts is that it is sometimes not clear whether there is a morphological link between the verb and the noun. In Portuguese, many deverbal nouns are direct adaptations from Latin, and not (or not only) synchronically derived: the word *confusão* (confusion) comes from Latin, and there is no synchronic operation to derive it from the verb *confundir*. The second major reason for doubts is that there are many cases in which it is unclear if the noun (still) expresses the abstract event related to the verb, or whether it can do so for all meanings of the verb. These problems can only be solved on a case-by-case basis, using well-defined criteria to render a consistent database of deverbal nouns.

Note, however, that there are also classes of derivational forms that structurally have a marginal character. Deadjectival quality nouns in many senses behave like deverbal event nouns. The nominal form applies generally to all meaning of the adjective - a polysemous adjective like *happy* has the nominal form *happiness* in every meaning of the word. And the nominal form hardly expresses anything else than the abstract quality related to the adjective, expressed in a nominalized way. This is exemplified by the two sentences in (8) which are largely synonymous.

- (8) He is very happy. $\Rightarrow$ He has a lot of happiness.

However, different from event nouns, the quality nouns are quite regularly related to a specific meaning of the adjective, or a range of meanings of the adjective. For instance, the Portuguese adjective *fino* (fine) has a number of meanings, and a number of related quality nouns. But which quality noun is the correct one often depends on the meaning of the adjective. The noun *finura* relates to *fino* in its meaning of ‘delicate’, whereas the word *fineza* means ‘thin-ness’. And similarly, with respect to the adjective *bravo* (angry; brave), *bravura* means ‘braveness’, and *braveza* is the act of being angry (Correia, 1999).

A possible conclusion, and the current pragmatic solution in OSLIN, is that deadjectival quality nouns are not inherently inflectional, but really derivational like agentive nouns are, and should therefore not be modeled by means of inherent inflections. However, for most adjectives, the derived quality noun is not

meaning specific, and the noun reflects the abstract property related to any of the meanings of the adjective. Therefore, treating quality nouns for polysemous adjectives with lexical functions fails to take the general character of the relation between argument and value into account, in much the same way as it does for deverbal event nouns. This means that quality nouns are on the one hand too much meaning independent to be treated with purely word-sense based lexical functions, but on the other hand too meaning specific to be treated by the purely lexical entry oriented inflectional functions. And quality nouns are not unique in this respect: the same holds for instance for the agentive nouns as well.

4 Conclusion

As shown in this paper, inflectional functions provide an efficient way of modeling inflection-like derivational relations such as deverbal event nouns, female nouns, diminutives, and superlatives. Several of these relations are currently modeled by lexical functions, and inflectional functions provide an entry-level treatment of these relations rather than a word-sense based treatment, which is not only more general, but also matches the linguistic reality closer. Inflectional functions imply lexical functions, which means that the OSLIN databases in which inflectional functions are applied at large scale, indirectly provide large repositories of lexical functions.

But as also shown in this paper, there are several cases in which the application of inflectional functions is problematic, running into the boundaries of what can feasibly be expressed in terms of such entry-wide relations. However, despite these limitations, inflectional functions still provide a more efficient and manageable way of representing lexical relations in large-scale full-form lexica providing information on the extended paradigm of words.

References

- Bauer, Laurie. 2004. The function of word-formation and the inflection-derivation distinction. In Aertsen, Hannay, and Lyall, editors, *Words and their Places. A Festschrift for J. Lachlan Mackenzie*. Vrije Universiteit, Amsterdam.
- Booij, Geert. 1995. Inherent versus contextual inflection and the split morphology hypothesis. In Booij and van Marle, editors, *Yearbook of Morphology 1995*. Kluwer, Dordrecht.
- Booij, Geert. 2004. *The Grammar of Words*. Oxford University Press, Oxford.
- Correia, Margarita. 1999. *A denominação das qualidades: contributo para a compreensão da estrutura do léxico português*. Ph.D. thesis, Universidade de Lisboa.
- Grefenstette, G., U. Heid, B. M. Schulze, T. Fontenelle, and C. Gerardy. 1996. The DECIDE project: Multilingual collocation extraction. In *Papers submitted to the Seventh EURALEX International Congress on Lexicography*, Göteborg, Sweden.
- Janssen, Maarten. 2005a. Between inflection and derivation. In *East-West Encounter: Second International Conference on Meaning ⇔ Text Theory*, Moscow, Russia.
- Janssen, Maarten. 2005b. Open source lexical information network. In *Third International Workshop on Generative Approaches to the Lexicon*, Geneva, Switzerland.
- Janssen, Maarten. 2008. Neotrack: Une analyseur de neologismes en ligne. In *Actes de CINEO 2008: Actes del I Congrés Internacional de Neologia de les Llengües Romàniques*, Barcelona, Spain.
- Mel'čuk, Igor, André Clas, and Alain Polguère. 1995. *Introduction à la lexicologie explicative et combinatoire*. Duculot, Louvain-la-Neuve.
- Polguère, Alain. 2008. *Lexicologie et sémantique lexicale. Notions fondamentales*. Les Presses de l'Université de Montréal, Montréal.
- Žolkovskij, A. and I. Mel'čuk. 1965. O vozmožnom metode i instrumentax semantičeskogo sinteza. *Naučno-techničeskaja informacija*, 5:23–28.