An XMG account of multiplicity of meaning in derivation

Marios Andreou, Simon Petitjean
Heinrich Heine Üniversität Düsseldorf, Düsseldorf, Germany

1 Introduction

More often than not, the products of derivational processes are interpreted in more than one way. This multiplicity of meaning is particularly evident in deverbal nominalizations (Lieber, 2004; Rainer, 2014; Andreou & Petitjean, 2017; Plag *et al.*, forthcoming). Derived words that are based on the suffix *-al*, for example, may denote either situations (e.g. *removal* "the act of removing") or entities (e.g. *rental* "the thing one rents").

In this paper, we focus on deverbal nominalizations with the suffix *-al* that are based on causation events. Causation events have a rich bipartite structure which captures complex relationships between situations (events and states) and entities. This complex structure allows one to identify and test constraints that might be at effect on the types of arguments which *-al* targets.

The aim of the paper is threefold. First, to best describe the behavior of *-al* on causation events and, thus, capture the multiplicity of meaning exhibited by *-al* nominalizations. Second, to identify prominent constraints on the types of situations and entities *-al* targets. This will allow us to inform the discussion on the way one can greatly reduce overgeneration of readings. In particular, the identification of constraints will be a contribution to the literature on the way one can predict and generate those readings which are possible for a given derivative and, at the same time, rule out those readings which are not possible (Lieber, 2004; Booij, 2010; Rainer, 2014; Andreou & Petitjean, 2017; Plag *et al.*, forthcoming). Third, to best model these constraints using XMG.

Our approach is based on the framework of Frame Semantics as developed in Petersen (2007), Kallmeyer & Osswald (2013), and Löbner (2013, 2014, 2015).[1] A frame is a general format of mental representations of concepts which is also applicable to linguistic phenomena. It is a recursive attribute-value structure that provides information about the referent of the frame. Attributes are applied to a given possessor in a frame structure and assign a value to it. To adduce an example, Figure 1 gives the partial frame for *ball* in the form of an attribute-value matrix.

[1]Frames also figure in works on Lexical Functional Grammar (Bresnan, 2001), Head-Driven Phrase Structure Grammar (Pollard & Sag, 1994), and Sign-based construction grammar Sag (2012). Fillmore's frames (Fillmore, 1982) are used in the FrameNet project (Fillmore & Baker, 2010). In the present paper, we will use Frames as defined in the work of Petersen (2007), Kallmeyer & Osswald (2013), and Löbner (2013, 2014, 2015), which is inspired by the work of Barsalou (1992a,b, 1999).

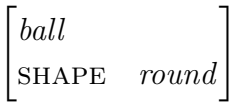

Figure 1: Partial frame for *ball*

This attribute-value matrix illustrates that the referent of the frame, *ball*, has an attribute SHAPE. This attribute assigns the value *round* to the referent of the frame.

Word formation in Frame Semantics is generally treated in terms of referential shifts (Löbner, 2013; Plag *et al.*, forthcoming). In particular, reference is shifted from the original referent to a new referent. For example, as we will see in the analysis, the suffix *-al* can target particular arguments of the base verb and shift reference from the original referent (i.e. causation event) to a new referent (e.g. theme). As recently shown by a number of studies on nominalizations (Lieber, 2004, 2016; Kawaletz & Plag, 2015; Andreou & Petitjean, 2017; Plag *et al.*, forthcoming), not all arguments of the verb can be targeted by affixation. The identification of prominent constraints on the types of arguments that can be targeted by a particular affix is still an open issue and has implications for the way we describe, model, and implement a particular derivational process in XMG.

What is XMG? XMG (eXtensible MetaGrammar, Crabbé *et al.* (2013)) is a modular and extensible tool used to generate various types of linguistic resources from an abstract and compact description. This description, the metagrammar, relies on the concepts of logic programming and constraints. XMG comes with a system of dimensions, allowing to separate the different levels of linguistic description (e.g. syntax and semantics), and providing dedicated languages adapted to the structures the user wishes to generate. In this work, the dimension we used is the <**frame**> dimension, proposed in Lichte & Petitjean (2015), where semantic frames can be described using typed feature structures descriptions.

The rest of this paper is structured as follows: In section 2, we describe and analyze the behavior of *-al* nominalizations in context. This will allow us to identify prominent constraints on the types of situations and entities that can be targeted by *-al*. In section 3, we provide an analysis of the multiplicity of meaning exhibited by *-al* nominalizations in XMG. Section 4 concludes the paper.

2 Data and analysis

In this paper, we follow the classification of VerbNet (Kipper-Schuler, 2006) that is inspired by the classification of Levin (1993) and we focus on the suffix *-al* on causation events. In particular, we examine the following verb classes: put verbs (e.g. *bury*), remove verbs (e.g. *remove*), banish verbs (e.g. *recuse*), deprive verbs (e.g. *deprive*), send verbs (e.g. *transmit*), contribute verbs (e.g. *betroth*), verbs of future having (e.g. *bequeath*), equip verbs (e.g. redress), get verbs (e.g. *procure*), obtain verbs (e.g. *retrieve*), amuse verbs (e.g. *arouse*), verbs of change of state (e.g. *disperse*), free verbs (e.g. *acquit*), addict verbs (e.g. *dispose*), and base verbs (e.g. *construe*).

We chose to work with causation events since these verbs have a rich bipartite structure which captures complex relationships between situations and entities. Thus, by using

causation events as testbed we can identify constraints on the types of situations and entities *-al* targets. In particular, we can ask the following question: Are all situations and entities targeted by *-al* affixation or are there general constraints on the types of arguments *-al* targets?

A typical causation event comes with a bipartite structure that comprises a CAUSE and an EFFECT (Kallmeyer & Osswald, 2012; Plag *et al.*, forthcoming). It involves a relationship between situations and entities in which a particular entity (e.g. an originator in the sense of Borer, 2014) causes another entity (i.e. a theme) to go from an initial situation to a result situation (Lieber, 2004; Levin, 1993; Rappaport Hovav & Levin, 2008). The following two attribute-value matrices illustrate this state of affairs. Figure 2 gives the structure of a change of state verb such as *renew* and Figure 3 illustrates the structure of a verb of change of possession such as *bequeath*.

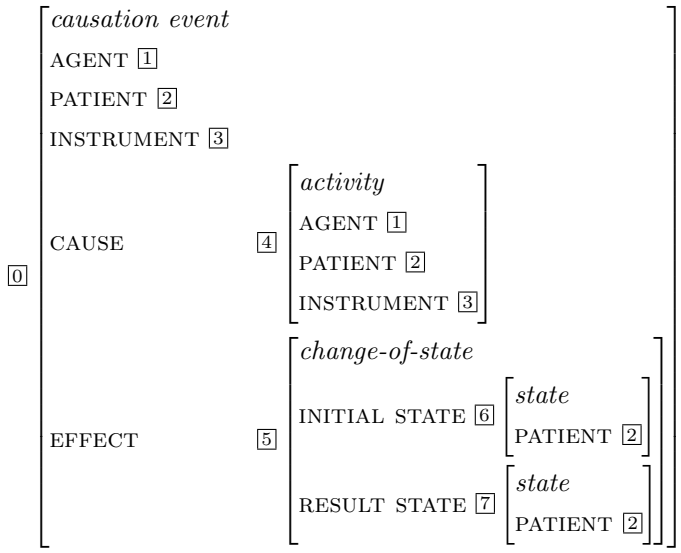

Figure 2: Change of state verbs

Figure 2 models that *renew* involves a relationship between the participants agent, patient, and instrument, in which the agent causes the patient to go from an initial state to a result state.

Another example which shows that causation events generally involve two sub-events, a cause and an effect, is given in Figure 3 which models a future having verb such as *bequeath*. This verb describes caused possession of the kind 'x causes y to have z', in which x is the agent, y is the recipient, and z is the theme (Goldberg, 1995; Jackendoff, 1990; Rappaport Hovav & Levin, 2008). Thus, Figure 3 models this state of affairs as a relationship between an agent, a theme, and a recipient, in which there is an initial situation in which the agent has possession of the theme, and a result situation in which the recipient has possession of the theme (Andreou & Petitjean, 2017).

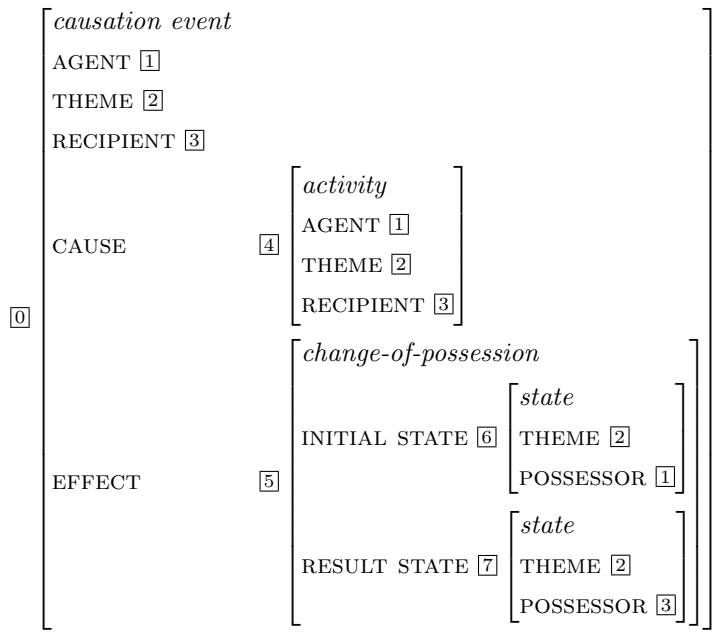

Figure 3: Verbs of change of possession

Let us now present the findings of our study with respect to possible readings of *-al* nominalizations. We use data from the Corpus of Contemporary American English (COCA, Davies, 2008). Among the readings we find in causation events, the most productive are the event and result readings. (1) includes event readings and (2) provides result readings.

(1) Event reading

a. One can perhaps gain a further glimpse of this sort of process of **construal** in a 1979 conversation of Serra, Annette Michelson, and Clara Weyergraf. Michelson began the interview by asking Serra how and when he came to filmmaking. (COCA ACAD 2015)

b. This results in delays in the **disbursal** and utilization of funds – especially at the Gram Panchayat level. (COCA ACAD 1998)

c. If it is morally unacceptable to repatriate even a convicted illegal alien criminal, then it is all the more unacceptable to repatriate someone who "merely" has crossed the border illegally. This undermining of alien **removals** is behind the constant protests demanding to "stop deportations now." (COCA MAG)

(2) Result reading

a. Introverts proved more able to focus on the task of color identification while disregarding the emotional content and had significantly better reaction times. Concludes Haas: Introverts, who exhibit a higher resting state of **arousal**, "don't need the same kind of outside entertainment." (COCA MAG 2010)

b. At the same time as it emerged that Fitzroy was terminally ill with 'a rapid consumption', Henry learned of Margaret Douglas's **betrothal** to Thomas Howard. (COCA MAG 2013)

c. Smith, 54, is the nephew of a slain American president. As a younger man, he was the defendant in a salacious Palm Beach rape trial that ended in his **acquittal**, though not before the nation devoured stories of late-night, alcohol-fueled carousing that included then-Sen. (COCA NEWS 2014)

In the examples in (1), the nominalization lexicalizes the event denoted by the verb. This type of nominalization is also referred to as 'transpositional' in that the nominalization 'transposes' (recategorizes) the word from verb to noun without altering the sense of the verbal base. Thus, *construal*, *disbursal*, and *removals* can be paraphrased as "event/process of construing", "event/process of disbursing", and "event/process of removing", respectively.

In the examples in (2), the nominalization has a result reading in that it lexicalizes "the outcome of verb-ing". Thus, *arousal*, *betrothal*, and *acquittal* can be paraphrased as "the (result) state of arousing", "the outcome of betrothing", and "the outcome of acquitting".

Observe that in both (1) and (2), contextual cues may guide us to a particular reading. For example, *the process of construal* flags a transpositional eventive reading and *a higher state of arousal* guides us towards a result state reading.

One may also find *-al* nominalizations that lexicalize the inanimate theme, that is, "the thing verb-ed, the thing affected by verb-ing". Consider the examples in (3):

(3) Inanimate theme

a. Planning for and pursuing invoices is necessary in any case. After **renewals** are paid in July or August (or the first two months), September (or the third month) is a good time to start setting up projection reviews for these resources. (COCA ACAD 2015)

b. The room was technically full of locals, people from Bianca' s life before she headed West, friends who crossed the bridge searching for more affordable **rentals** in Williamsburg or Long Island City. (COCA FIC 2015)

c. In any case, your best bet is to roll the money into a traditional IRA; otherwise, you' ll get a big tax bill. Smaller **withdrawals** from the IRA, on the other hand, will likely be taxed at a lower rate. (COCA MAG)

In (3), we observe that *renewals* are "the things one renews (e.g. subscriptions), *rentals* are "the things that someone rents (e.g. a house, an appartment)", and *withdrawals* are "the things one withdraws (i.e. money).

A closer inspection of the data in (1)-(3) reveals that the suffix *-al* can manipulate the frame of a verb and target certain arguments of it. In particular, it can target the causation event argument, the result situation argument, and the theme argument. Thus, the referent of a form derived by *-al* can be identified with some of the arguments of the verbal base, but not all of them. Observe, for instance, that the referent of *-al* derivatives is never the agent, the recipient, the cause, the effect or the initial situation.

In what follows, we undertake the nontrivial task of identifying possible constraints on the types of entities and situations *-al* targets.

As far as entities are concerned, there seems to be a constraint on the animacy of the referent of *-al* nominalizations. In particular, the referent of *-al* nominalizations cannot be [+animate]. This explains why we find inanimate theme readings but not agentive

readings.

In what follows, we test this constraint on animacy. Consider the following examples:

(4) a. Agentive reading
The path down to the sea is shaded by lemon groves. There is also an elevator to the private beach, where a saltwater pool, sun decks, a bar and seaside restaurant, along with a well-equipped gym and **boat rentals**, await. (COCA MAG 2001)

b. Instrument reading
If I hadn't read the article in your magazine, my precious dogs would be in continued danger. Enclosed is my **renewal**. Thanks for the great information. (COCA MAG 2003)

c. Asset reading
The farmer who owned the barn had asked - and received - a thousand dollars in **rental**. (COCA FIG 2004)

Although the examples in (4) are not primary readings of *-al* nominalizations, they can, nevertheless, inform the discussion on the constraint on animacy. In (4-a), *boat rentals* has an agentive reading. This seems to militate against the hypothesis that the referent of *-al* nominalizations cannot be [+animate]. On closer inspection, however, the context suggests that the referent of *boat rentals* is inanimate. It is the company that rents boats. In any case, this reading is highly lexicalized. In (4-b), *renewal* is interpreted as an instrument since it is the participant in the *renew* event that is manipulated by the agent, and with which an intentional act is performed. In our example, it is the form of renewal of subscription. Thus, the referent of *renewal* is inanimate. Finally, the argument that seems to be lexicalized in (4-c) is the asset argument, that is the value of something. In our example, *rental* lexicalizes this argument since its reading can be paraphrased as "the amount of money one has to pay for renting the barn". To sum up, the examination of secondary readings of *-al* nominalizations confirms the hypothesis that there is a constraint on animacy on the referent of *-al* forms.

Let us now turn to situations. Is there a constraint on the types of situations that can be targeted by *-al*? As mentioned above, the structure of causation events typically includes the causation event argument, a cause, an effect, an initial situation, and a result situation. In our data, there are no cases in which the cause, the effect or the initial situation are targeted by *-al*. As shown in (1) and (2), *-al* nominalizations in our data give rise only to transpositional eventive readings and result situation readings. Let us elaborate upon the latter reading, i.e. result situation. The result situations described by the various subclasses in our data are not homogeneous. In particular, verbs such as *arouse* describe a change of emotional state, verbs such as *bequeath* describe a change of possession, and verbs such as *remove* describe a change of location. Are all these situations targeted by *-al*?

Our data suggest that the only result situation that is compatible with *-al* is the result state. The only example in which we identified a different reading is given below:

(5) In a **burial** in Gyeongju, South Korea, archaeologists uncovered armor of a fifth-century A.D. warrior and his horse, as well as dozens of serving vessels used in traditional burial rituals. (COCA ACAD 2009)

This reading involves the put verb *bury* which describes a change of location. The use of *burial* with the reading of result location (e.g. tomb, grave), however, is highly lexicalized and only used in archeology. Thus we can safely conclude that the referential argument of *-al* forms is not compatible with arguments of the type location.

The identification of these constraints allows one to comment on the way one can handle multiplicity of meaning in derivation. In the relevant literature (Lieber, 2004; Booij, 2010; Rainer, 2014; Andreou & Petitjean, 2017; Plag *et al.*, forthcoming), there are two approaches to multiplicity of meaning in derivation. Under the first appoach, i.e. monosemy, more concrete meanings of affixes derive from a general highly underspecified meaning that is capable of taking into account all possible readings of an affix.

In order to apply the monosemy approach to *-al*, we have to reduce multiplicity of meaning by identifying meanings that are shared by all *-al* derivatives. As follows from the analysis of our data, *-al* derivatives denote (a) eventualities (e.g. event 'transpositional' readings), and (b) entities (e.g. inanimate theme readings). Thus, the abstract core meaning of *-al* can be characterized as 'eventuality or entity having to do with X' (with 'X' denoting the base).

Monosemy approaches to the semantics of derivation are confronted with two problems. The first problem is that it is very hard to establish a unitary meaning for an affix. In particular, the aim of monosemy approaches is to reduce multiplicity of meaning by postulating a unitary abstract meaning. Forms derived by *-al*, however, denote both eventualities and entities. Thus, the disjunction 'eventuality or entity' that is needed in order to capture the multiplicity of meaning of *-al* derivatives reveals that the desirable underspecified meaning of affixes cannot always be reduced to a single unitary meaning.

The second problem with the monosemy approach to the semantics of derivation is (massive) overgeneration. As we saw earlier, the abstract meaning for *-al* informs us that *-al* forms denote both eventualities and entities. What kind of predictions follow from the abstract meaning 'eventuality or entity having to do with X'? This particular formulation of the abstract meaning of *-al* leads one to expect that *-al* derivatives could in principle denote all entities and all eventualities. Our data, however, suggests that not all entities and not all eventualities can be denoted by *-al* derivatives. For instance, the referent of an *-al* derivative may be the inanimate theme (e.g. *money* in the case of *withdrawal*) but not the agent.

Under the second approach, i.e. polysemy, there is multiplicity of meaning in word formation patterns. Given the architecture of Frame Semantics, the multiplicity of readings exhibited by *-al* nominalizations can be captured with the use of an inheritance hierarchy of lexeme formation rules (Riehemann, 1998; Koenig, 1999; Booij, 2010; Bonami & Crysmann, 2016; Plag *et al.*, forthcoming). Inheritance hierarchies allow one to generalize over derived formations and capture shared characteristics between them as we show in Figure 4.

Figure 4 gives the inheritance hierarchy of lexeme formation rules ('lfr') for deverbal nominalizations ('v-n') in *-al*. This hierarchy involves two dimensions, namely phonology (PHON) and semantics (SEM). The first dimension, i.e. phonology, is shared by all *-al* nominalizations. In particular, all *-al* nominalizations have the phonology /$\boxed{1}$+al/. Boxed numerals such as $\boxed{1}$ are called tags and are used in feature structures to indicate structure sharing. That is, to show that the respective values are identical. In Figure 4, this means that the value of the first part of the phonology of the derived lexeme is

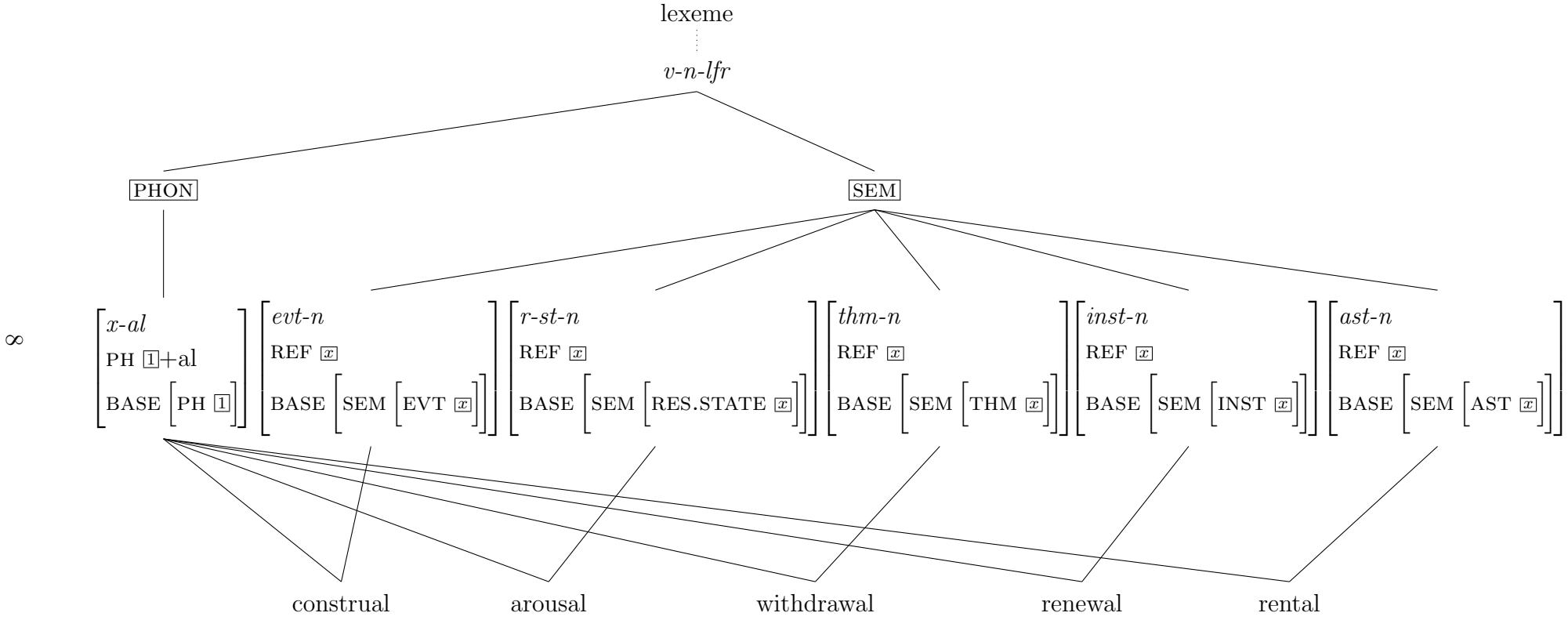

Figure 4: Inheritance hierarchy of lexeme formation rules for the suffix *-al*

identical to the value for the phonology of the base. The second part of the phonology of the derived lexeme is, of course, contributed by the affix, i.e. /al/.

Although *-al* nominalizations are based on the same phonological pattern, their semantics differs. The semantic dimension in the inheritance hierarchy in Figure 4 captures the different readings exhibited by *-al* forms. In accordance with the analysis suggested by our data, when the reference of a form in *-al* is identified with the event argument ('evt') of the base, we get an eventive 'transpositional' reading and when it is identified with the result state argument ('r-st') of the base, we get a result state reading. In a similar vein, a theme reading arises when the reference of an *-al* nominal is identified with the theme argument ('thm') of the base, an instrument reading when it is identified with the instrument argument ('inst') of the base, and finally an asset reading when it is identified with the asset argument ('ast') of the base. The lowest level of Figure 4 shows that *-al* forms inherit their characteristics from both dimensions, i.e. phonology and semantics. In particular, all *-al* forms share the same phonology, but their semantics differs.

In this section, we identified the range of readings available to *-al* forms and described the way this range could be accounted for under the monosemy and polysemy approach. In the next section, we will use the type constraints we identified in this section, in order to predict and generate those readings which are possible for an *-al* form and, at the same time, rule out those readings which are not possible.

3 XMG implementation

The XMG compiler is a tool which has already been used to generate a wide range of linguistic resources, focusing on different levels of linguistic description, such as syntax and semantics, or even interfaces between them. Syntactic resources developed with XMG are tree-based grammars such as Tree Adjoining Grammars (Crabbé (2005), Kallmeyer *et al.* (2008), Gardent (2008) for instance) or Interaction Grammars such as Perrier (2007). Other types of resources include lexicons of fully inflected forms, which were generated from morphological descriptions as in Duchier *et al.* (2012), or frame-based semantic descriptions. In this work, even though we are interested in both morphology and semantics, we will only focus on the description of the semantics. On the morphological side, the description is trivial as it only consists in combining a verb and a given affix.

An XMG implementation is a program (called metagrammar) composed of a set of classes, which are reusable abstractions. A class describes a partial linguistic structure, which is in our case the frame for a given class of verbs. Classes can be reused by other classes (imported), to add information to the partial description. This is what will be done by the classes modeling derivations: they will import the descriptions of the verb frames and augment them by defining the semantic reference corresponding to one reading of the derivation. The descriptions shown in this article mainly consist of typed feature structures. By using unification variables in their description, the feature structures are combined to describe more complex frames. An XMG program is non-deterministic: it uses underspecification and disjunction, meaning that every class can describe zero, one or more structures. When the metagrammar is processed by the XMG compiler, all the structures described in the classes are computed and written into an output file (using

the XML or JSON format).

The implementation that we present aims at generating the frames corresponding to all the attested readings for the derivations. For space limitations, below we focus on two classes, namely, verbs of change of possession and verbs of change of state. The proposed analysis can, nevertheless, be extended to additional verb classes in a similar and straightforward manner.

We first need to describe the frame given in Figure 3, by means of a XMG class which we will name rent. This abstraction describes the class of verbs of change of possession:

```
class rent
export ?X0
declare ?X0 ?X1 ?X2 ?X3 ?X4 ?X5 ?X6 ?X7
{<frame>{
  ?X0[causation,
    agent: ?X1[entity, animacy:[animate]],
    theme: ?X2,
    recipient : ?X3[entity] ,
    cause: ?X4[activity ,
               agent:?X1,
               theme:?X2,
               recipient :?X3[entity, animacy:[animate]]
               ] ,
    effect : ?X5[change_of_possession,
                 initial −state: ?X6[ initial_state ,
                                      theme:?X2[entity],
                                      possessor:?X1],
                  result −state:  ?X7[result_state ,
                                      theme:?X2[entity],
                                      possessor:?X3] ] ]
                 }
}
```

where the first lines define the set of unification variables which can be used within the class (**declare**) and outside of it (**export**). <**frame**> means that the description belongs to the Frame Semantics dimension. The structure described in the frame dimension, labeled by **?**X0, is a straightforward translation of the one in Figure 3, with the addition of information on animacy, where all variables **?**X0 ,..., **?**X7 stand for the boxed numbers from $\boxed{0}$ to $\boxed{7}$. The only variable which can be accessed outside of the class is **?**X0 (cf. **export ?**X0). In the same fashion, we define the class of verbs of change of state shown in Figure 2:

```
class renew
export ?X0
declare ?X0 ?X1 ?X2 ?X3 ?X4 ?X5 ?X6 ?X7
{<frame>{
 ?X0[causation,
       agent: ?X1[entity, animacy:[animate]],
       patient: ?X2,
```

```
        instrument: ?X3[entity],
        cause: ?X4[activity,
                   agent:?X1,
                   patient:?X2,
                   instrument:?X3[entity, animacy:[animate]]
                       ],
        effect : ?X5[change_of_state,
                   initial −state: ?X6[ initial_state , patient:?X1],
                   result −state:  ?X7[result_state , patient:?X3] ] ]

                  }

}
```

To define the scope-over relation mentioned earlier, we can use a new abstraction (a class we will name al_nominal). This class, as its name suggests, models the semantics of *-al* derivatives, which for the purposes of this first example are based on verbs of change of possession.

```
class al_nominal
import rent[]
declare ?Ref
{
  <frame>{
    [al−lexeme,
      m−base:[event,
              sem:?X0]
      ref:?Ref
    ]
    ;
    ?X0 >* ?Ref;
  }
}
```

With **import** rent[], i.e. we make the structure defined in the class rent available in the current class, together with its variables (we can refer to all variables **?**X0 ,..., **?**X7 in the current class as they are exported by rent). The operator >* means that there must be a "path" (as it would be in a graph representation[2] of the frame) from the root **?**X0 to the semantic reference **?**Ref. Concretely, the compiler will try to generate structures where the reference is identified with another label, starting with the whole frame (**?**X0), and then exploring all of its subparts, recursively. As we do not specify any other constraint here, the reference will be able to unify with every possible subpart. This means that readings such as initial state, which should be ruled out, are also generated when this first version of the metagrammar is executed.

In this first implementation we modeled an approach to multiplicity of meaning which is close to a version of the monosemy approach under which there are no constraints on

[2]An attribute-value matrix can be seen as a directed graph in which every attribute-value pair is an edge labeled by the attribute and pointing to the node representing the value.

types, and showed that it leads to massive overgeneration. In the next section we focus on the second approach to multiplicity of meaning: polysemy.

An open question is how we can model the polysemy approach in XMG and constrain possible readings. We suggest that there are two ways to tackle this issue. First, via a fully specified (and explicit) rule, which will replace the scope over relation in the previous class al_nominal:

> **{**?X0=?Ref | ?X2=?Ref | ?X7=?Ref**}**

where `|` and `=` are respectively the disjunction and the unification operators, `?X0`, `?X2` and `?X7` respectively correspond to the boxed numbers 0, 2 and 7 of Figure 3, and `?Ref` is a variable representing the semantic reference.

Under this approach, possible readings are considered as generalizations over already attested derivatives. Thus, agent, recipient, and initial state readings are ruled out since they are not part of the possible readings in the fully-specified-rule; the rule models readings that are already attested in *-al* derivatives. However, this implementation is totally specific to a given class of verbs, here the one described in the class `rent`. More XMG code would have to be written for the derivation of other verb classes, where the reference would be identified with different unification variables. In our case, we used consistent variable namings in the class `renew` (the variables corresponding to the attested readings are also `?X0`, `?X2` and `?X7`), making it easily compatible with this implementation, but it would not be as straightforward for frames with different numbers of features. For example, for a verb class where the *-al* nominalization has four different readings, a different XMG class with four alternatives of variable unifications would have to be used.

Another way to model the polysemy approach in XMG is the introduction of an underspecified rule with constraints on types. Only the types of the feature structures will determine if one reading should be valid or not, which means that we do not need to provide explicitly the set of variables that may be unified with the semantic reference. In the case of our verb classes, the referent of an *-al* nominal can have three possible types: causation, result state, or entity.

> ?X0 >∗ ?Ref**;**
> **{** ?Ref[result_state] | ?Ref[causation]
> | ?Ref[entity , animacy:[inanimate]] **}**

Here, the first line is once again the scope over relation, but of course, in this case, only the structures where no type constraint is violated will eventually be generated.

In the second line, we express the fact that the referent of an *-al* derivative can have any of the three types previously stated. In the case of entity, only the theme should be a possible referent. We, therefore, add information about animacy (here, inanimate), which makes the reference of *-al* derivatives incompatible with frames of type animate, such as the agent and the recipient. This is in accordance with findings in the literature on possible constraints on animacy (see Kawaletz & Plag, 2015 on the suffix *-ment*). When the referent of an *-al* derivative is a state, the type result_state is given to prevent unification with the initial state frame (of type initial_state). This way, agent, recipient, and initial state readings are ruled out because frame unification only succeeds if types are compatible. The type constraints (for example incompatibility of event and entity) are also specified in the metagrammar. This is done globally, meaning that the type

constraints will apply to all the structures described in the metagrammar. The constraints defining our type hierarchy are introduced by the keyword **frame**−constraints as follows:

```
frame−constraints ={
    event −> eventuality,
    state −> eventuality,
    state event −> −,
    eventuality entity −> −,
    derived−lexeme −> lexeme,
    ment−lexeme −> derived−lexeme,
    lexeme eventuality −> −,
    eventuality entity −> −,
    causation −> event,
    activity −> event,
    change_of_possession −> event,
    change_of_state −> event,
    causation activity −> −,
    causation change_of_possession −> −,
    causation change_of_state −> −,
    change_of_state change_of_possession −> −,
    experiencer −> entity,
    stimulus −> entity,
    experiencer stimulus −> −,
    initial_state result_state −> −,
    initial_state −> state,
    result_state −> state,
    animate inanimate −> −,
    animate −> animacy,
    inanimate −> animacy,
    animacy eventuality −> −,
    animacy entity −> −,
    entity −> animacy:animacy,
    animacy lexeme −> −
    }
```

Three types of constraints are used here. Subsumption constraints, such as causation −> event, mean that an atomic type (here causation) is a subtype of another type (event). An incompatibility constraint, such as causation activity −> − means that a structure cannot have both of the two given types: here, a frame cannot be of type causation and of type activity. Finally, feature constraints, such as entity −> animacy:animacy ensure that all the structures having a given type have a given feature. In our case, structures of type entity will all have an attribute animacy of type animacy.

This implementation is directly compatible with the verbs described in the class `renew`, and does not depend on the naming of the variables used in the frame description. Therefore, an XMG abstraction describing verbs from another class, even if it is written by another linguist who uses different naming conventions can be combined with the al_nominal class. Of course, for verb classes in which readings are not limited to the

same types (causation, result_state and inanimate entity), new XMG abstractions for *-al* nominalization would have to be written. In these new XMG classes, only the type constraints would differ, and they could be directly reused for all other verb classes with similar behaviors.

4 Conclusion

In the present paper, we tackled the issue of multiplicity of meaning in derivation by offering a detailed analysis of *-al* derivatives. We used corpus extracted data to identify the range of readings available to *-al* derivatives and to establish possible constraints on the types of arguments *-al* targets. Finally, we modeled these constraints using XMG.

In a nutshell, we showed that the referent of an *-al* derivative can be identified with certain types of situations and entities, and not all of them. This has implications for the way we model multiplicity of meaning in derivation, since it shows that it is not always possible to reduce the meaning of a particular affix to a single unitary meaning.

Our XMG implementation corroborates the idea that the introduction of constraints into the semantics of an affix allows one to predict and generate those readings which are possible for a given derivative and rule out other readings which are not possible. These constraints have the form of type constraints and specify which arguments in the frame of the verbal base are compatible with the referential argument of the derivative. The introduction of type constraints rules out certain readings because frame unification only succeeds if types are compatible.

In the present paper, we focused on *-al* derivatives. The next step is to apply the proposed analysis to the modeling of other affixes as well. This will allow us to identify which constraints are specific to particular classes or affixes, and which constraints are shared across classes or affixes. For example, the suffixes *-ance*, *-ment*, and *-ure* show similar characteristics to the suffix *-al*, in that the referent of forms derived by these affixes is never [+animate]. They differ, however, from one another with respect to other characteristics. For example, *-ance*, *-ment*, and *-ure* are compatible with the location argument of the verbal base, whereas *-al* is not, and *-ure* is not compatible with the instrument argument of the verbal base, whereas *-ance*, *-ment*, and *-al* are. The main advantages of the metagrammatical framework will become more obvious as the linguistic resource grows: for example, inheritance will help sharing information across classes with similar behaviors.

Acknowledgements

The authors gratefully acknowledge financial support by the Deutsche Forschungsgemeinschaft (DFG Collaborative Research Centre 991: 'The Structure of Representations in Language, Cognition, and Science', Project C08 'The semantics of derivational morphology: A frame-based approach')

References

Andreou M. & Petitjean S. (2017). Describing derivational polysemy with xmg. In I. Eshkol & J.-Y. Antoine, Eds., *Actes de TALN 2017, 24e Conférence sur le Traitement Automatique des Langues Naturelles*, volume 2, p. 94–101.

Barsalou L. W. (1992a). *Cognitive psychology: An overview for cognitive sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Barsalou L. W. (1992b). Frames, concepts, and conceptual fields. In A. Lehrer & Eva Feder Kittay, Eds., *Frames, fields and contrasts: New essays in semantic and lexical organization*, p. 21–74. Hillsdale, NJ: Lawrence Erlbaum Associates.

Barsalou L. W. (1999). Perceptual symbol systems. *The Behavioral and Brain Sciences*, **22**, 577–660.

Bonami O. & Crysmann B. (2016). Morphology in constraint-based lexical approaches to grammar. In A. Hippisley & G. T. Stump, Eds., *Cambridge Handbook of Morphology*, p. 609–656. Cambridge: Cambridge University Press.

Booij G. (2010). *Construction morphology*. Oxford: Oxford University Press.

Borer H. (2014). *Taking Form. Structuring Sense, Volume III*. Oxford: Oxford University Press.

Bresnan J. (2001). *Lexical Functional Syntax*. Malden – Oxford: Blackwell.

Crabbé B. (2005). *Représentation informatique de grammaires fortement lexicalisées : Application à la grammaire d'arbres adjoints*. PhD thesis, Université Nancy 2.

Crabbé B., Duchier D., Gardent C., Le Roux J. & Parmentier Y. (2013). XMG : eXtensible MetaGrammar. *Computational Linguistics*, **39**(3), 1–66.

Davies M. (2008). The Corpus of Contemporary American English: 400+ million words, 1990-present.

Duchier D., Magnana Ekoukou B., Parmentier Y., Petitjean S. & Schang E. (2012). Describing Morphologically-rich Languages using Metagrammars: a Look at Verbs in Ikota. In *Workshop on "Language technology for normalisation of less-resourced languages", 8th SALTMIL Workshop on Minority Languages and the 4th workshop on African Language Technology*, p.˜_, Istanbul, Turkey.

Fillmore C. J. (1982). Frame semantics. In T. linguistic society of Korea, Ed., *Linguistics in the Morning Calm*, p. 111–137. Seoul: Hanshin Publishing Co.

Fillmore C. J. & Baker C. (2010). A frames approach to semantic analysis. In B. Heine & H. Narrog, Eds., *The Oxford Handbook of Linguistic Analysis*, p. 313–340. Oxford: Oxford University Press.

Gardent C. (2008). Integrating a Unification-Based Semantics in a Large Scale Lexicalised Tree Adjoining Grammar for French. In *Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008)*, p. 249–256, Manchester, UK: Coling 2008 Organizing Committee.

Goldberg A. (1995). *Constructions: A Construction Grammar approach to argument structure*. Chicago, IL: University of Chicago Press.

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

Kallmeyer L., Lichte T., Maier W., Parmentier Y. & Dellert J. (2008). Developing a TT-MCTAG for German with an RCG-based Parser. In *The sixth international conference on Language Resources and Evaluation (LREC 08)*, p. 782–789, Marrakech, Morocco.

Kallmeyer L. & Osswald R. (2012). A frame-based semantics of the dative alternation in Lexicalized Tree Adjoining Grammars. In C. Piñón, Ed., *Empirical Issues in Syntax and Semantics 9*, p. 167–184. ISSN 1769-7158.

Kallmeyer L. & Osswald R. (2013). Syntax-driven semantic frame composition in lexicalized tree adjoining grammars. *Journal of Language Modelling*, **1**(2), 267–330.

Kawaletz L. & Plag I. (2015). *Predicting the Semantics of English Nominalizations: A Frame-Based Analysis of -ment Suffixation*, In L. Bauer, L. Körtvélyessy & P. Štekauer, Eds., *Semantics of Complex Words*, p. 289–319. Springer International Publishing: Cham.

Kipper-Schuler K. (2006). *VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon*. PhD thesis, University of Pennsylvania.

Koenig J. (1999). *Lexical Relations*. Stanford Monographs in Linguistics. Cambridge University Press.

Levin B. (1993). *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago, IL: University of Chicago Press.

Lichte T. & Petitjean S. (2015). Implementing semantic frames as typed feature structures with XMG. *Journal of Language Modelling*, **3**(1), 185–228.

Lieber R. (2004). *Morphology and lexical semantics*. Cambridge: Cambridge University Press.

Lieber R. (2016). *English nouns: The ecology of nominalization*. Cambridge: Cambridge University Press.

Löbner S. (2013). *Understanding semantics*. London: Arnold, 2nd, revised edition edition.

Löbner S. (2014). Evidence for frames from human language. In T. Gamerschlag, D. Gerland, R. Osswald & W. Petersen, Eds., *Frames and concept types: Applications in Language, Cognition, and Philosophy*, p. 23–67. Dordrecht: Springer.

Löbner S. (2015). Functional concepts and frames. In T. Gamerschlag, D. Gerland, R. Osswald & W. Petersen, Eds., *Meaning, Frames, and Conceptual Representation*, p. 35–62. Düsseldorf: dup.

Perrier G. (2007). A French Interaction Grammar. In G. Angelova, K. Bontcheva, R. Mitkov, N. Nicolov & K. Simov, Eds., *International Conference on Recent Advances in Natural Language Processing - RANLP 2007*, p. 463–467, Borovets, Bulgarie: IPP & BAS & ACL-Bulgaria INCOMA Ltd, Shoumen, Bulgaria.

Petersen W. (2007). Representation of concepts as frames. In J. Skilters, Ed., *The Baltic international yearbook of cognition, logic and communication*, volume 2, p. 151–170.

Plag I., Andreou M. & Kawaletz L. (forthcoming). A frame-semantic approach to polysemy in affixation. In O. Bonami, G. Boyè, G. Dal, H. Giraudo & F. Namer, Eds., *The lexeme in descriptive and theoretical morphology*. Berlin: Language Science Press.

Pollard C. & Sag I. (1994). *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press, Stanford : CSLI Publications.

Rainer F. (2014). Polysemy in derivation. In R. Lieber & P. Štekauer, Eds., *The Oxford handbook of derivational morphology*, Oxford Handbooks in linguistics, p. 338–353. Oxford: Oxford University Press.

Rappaport Hovav M. & Levin B. (2008). The English dative alternation: The case for verb sensitivity. *Journal of Linguistics*, **44**, 129–167.

Riehemann S. Z. (1998). Type-based derivational morphology. *Journal of Comparative Germanic Linguistics*, **2**(1), 49–77.

Sag I. A. (2012). Sign-based construction grammar: An informal synopsis. In H. C. Boas & I. A. Sag, Eds., *Sign-Based Construction Grammar*, p. 69–202. Stanford, CA: CSLI Publications.