The semantics of derivational morphology

# Linguistische Arbeiten



# Volume Seriesvol-ume/Seriesvolume

# The semantics of derivational morphology

Theory, methods, evidence

Edited by
Sven Kotowski & Ingo Plag

Authorinfo

ISBN 978-3-11-021808-4
e-ISBN (PDF) 978-3-11-021809-1
ISSN 0179-0986

**Library of Congress Cataloging-in-Publication Data**
A CIP catalog record for this book has been applied for at the Library of Congress.

**Bibliographic information published by the Deutsche Nationalbibliothek**
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.


Cover image: Cover-Firma
Typesetting: le-tex publishing services GmbH, Leipzig
Printing and binding: Druckerei XYZ
♾ Printed on acid-free paper
Printed in Germany

www.degruyter.com

…

…

Ingo Plag, Lea Kawaletz, Sabine Arndt-Lappe, Rochelle Lieber

# Analogical modeling of derivational semantics

## Two case studies

**Abstract:** Recent work on the semantics of deverbal nominalizations (e.g. Lieber 2016; Lieber & Andreou 2018; Lieber & Plag 2021; Schirakowski 2021) has shown that categorical generalizations in this domain are often empirically inadequate. The variability observed in the data cannot be accommodated by theoretical models that rely on rule-based behavior of particular classes of words. One alternative to such approaches is analogical modeling. Analogical models are computational algorithms that work on the basis of a lexicon in which forms are stored together with their properties. Based on the similarity of a given form with the forms stored in the lexicon, the given form is assigned a probability of a particular outcome (for instance, the choice of a particular morphological form, or the choice of a particular interpretation). So far, the semantics of complex words has played only a marginal role in the analogical modeling of morphology, which has mainly focussed on formal properties of complex words (e.g. their phonological structure) to predict morpho-phonological alternations (e.g. stress assignment to compounds, Plag et al. 2007, Arndt-Lappe 2011) or affix selection in situations of competition (e.g. Eddington 2002, Arndt-Lappe 2014).
This paper implements the analogical approach using the AML algorithm (Skousen et al., 2013) to address two fundamental problems of derivational semantics. The first one is that one meaning or concept may be expressed by more than one form ('affix competition'), the second one is that one affix may give rise to different meanings ('affix polysemy'). The paper presents one case study for each of the two problems. In the first study we model the choice between English *-ing* and conversion to derive deverbal nouns, in the second we model the choice of different interpretations for English *-ment* derivatives. The two studies show that both the influence of semantics on affix competition and affix polysemy can be successfully modeled using analogy. The models are not only quite accurate in their predictions.

**Ingo Plag, Lea Kawaletz, Sabine Arndt-Lappe, Rochelle Lieber,** IP and LK: Department of English language and linguistics, Heinrich-Heine-University Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany, email corresponding author: ingo.plag@uni-duesseldorf.de; SA-L: Universität Trier, Fachbereich II, Anglistik,D-54296 Trier, Germany; RL: English Department, Hamilton Smith Hall, University of New Hampshire, 95 Main Street, Durham, NH 03824, USA

https://doi.org/10.1515/futur-2012-0001

Closer inspection also reveals that important generalizations of different grain size fall out automatically in AML. The article demonstrates that analogical modeling is a cognitively plausible and computationally tractable mechanism also in the domain of derivational semantics.

**Keywords:** derivation, semantics, analogical modeling

# 1 Introduction

Although derivational semantics has gained more attention over the years (e.g. Bauer et al. 2015), fundamental questions remain under debate. This article deals with two of them, affix competition and affix polysemy. 'Affix competition' refers to the common phenomenon that different affixes often encode the same meaning. For instance, *-ity* and *-ness* in English are taken to derive abstract nouns with no clear difference in meaning between the two morphological categories (see Bauer et al. 2013, 257f). Another case is the formation of deverbal nominalizations by *-ing* and conversion (as in *the betting* vs. *the bet*). 'Affix polysemy' refers to the fact that the derivatives of a particular morphological category can have different interpretations. For instance, English *-ment* derivatives have been shown to exhibit change-of-state readings, stimulus readings, result-state readings and a few others (e.g. Bauer et al. 2013; Kawaletz & Plag 2015; Plag et al. 2018; Kawaletz 2021). A note on terminology is in order here. For ease of reference we have used the term 'affix' here in a loose sense, including non-affixational morphological processes like conversion or truncation. The expressions 'competition of morphological forms' and 'polysemy of a morphological category' would be more appropriate ways of speaking about these two phenomena.

Recent studies using modern empirical methods have shown that many generalizations in the theoretically-oriented literature concerning these two problems (some of them long-cherished) are wrong (e.g. Lieber 2016; Lieber & Andreou 2018; Lieber & Plag 2021; Schirakowski 2021). In particular, the received wisdom is being challenged by increasing evidence for the under-specification of derivational semantics and the importance of contextual and world knowledge (see, for instance, Alexiadou 2019).

These new findings present both theoretical and methodological challenges. At the theoretical level, we have to concede that categorical approaches are not able to cope with the variability attested in the data, and that these theories must either be seriously revised, or alternative theories must be developed and tested. At the empirical level, tractable statistical and computational models are needed to account for how complex words mean. The present article explores a particular, and very traditional, theoretical approach to morphology, analogy (see Arndt-Lappe

2015 for an overview). We use analogy to tackle the two problems introduced above, using a particular computational model, the AML algorithm (Skousen et al. 2013; Arndt-Lappe et al. 2018).

Analogical modeling is conceptualized here as an exemplar-based approach in which storage of individual occurrences of expressions plays a prominent role. More specifically, analogical models are computational algorithms that work on the basis of a lexicon in which forms are stored together with their properties, including the property in question ('outcome'), for instance the kind of morphological form or the interpretation they exhibit. Based on the similarity of a given form with the forms stored in the lexicon, the given form is assigned a probability of a particular outcome (for instance, 'conversion' as the morphological form, or 'result-state' as its interpretation). So far, the semantics of complex words has played only a marginal role in the analogical modeling of morphology, which has mainly focused on formal properties of simplex and complex words (e.g. their phonological structure) to predict, for example, stress (Eddington 2000), inflectional forms (e.g. Skousen 1989; Keuleers et al. 2007), morpho-phonological alternations (e.g. stress assignment to compounds, Plag et al. 2007; Arndt-Lappe 2011) or affix selection in situations of competition (e.g. Eddington 2002; Arndt-Lappe 2014).

We propose to implement analogical models in which we predict outcomes on the basis of semantic features, addressing the two fundamental problems above. In our first case study we tackle the form competition problem by modeling the choice between *-ing* and conversion based on, first, the meaning that is to be expressed with the derived noun and, second, the aspectual class of the base verb. The data set for this study consists of almost 1500 nominalization tokens from COCA (Davies, 2008-2014) that were used in Lieber & Plag (2021). It will be shown that the analogical model can quite accurately predict the choice between the two morphological forms. The model performs as well as regression models, but based on a plausible cognitive mechanism.

In our second case study we investigate how particular readings are selected for a given *-ment* derivative. For this, we make use of a data set of 40 *-ment* neologisms used in Kawaletz (2021). The nominalizations in the data set are polysemous, exhibiting up to seven different readings per type. This range of readings is systematically related to semantic properties of the base verb as well as to selectional restrictions of *-ment*. We will show that, by providing AML with a well-founded set of features pertaining to the base verbs' semantics, the algorithm does an excellent job of predicting the emergent patterns in the nominalizations' possible readings.

## 2 Modeling derivational semantics

There are a number of approaches that try to go beyond assigning impressionistic semantic labels to morphological categories. In those approaches an attempt is made to understand in more detail how meanings are generated based on the properties of base, affix and other available information. We will briefly lay out three prominent architectures before we turn to analogical modeling: syntactic approaches, Lieber's Lexical Semantic Framework, and distributional semantics.

A family of important approaches to the treatment of the semantics of nominalization are syntactic approaches, among them Distributed Morphology (Alexiadou 2001; Harley 2009), Borer's (2013) Exoskeletal model, and Nanosyntax (Baunaz et al. 2018). In these frameworks a particular reading for a derived word is attributed to a difference in syntactic structure, for instance, the location of particular entities in the array of functional projections above or below the affix.

Such an analysis works best if the range of readings is highly restricted, and clearly tied to particular categories that behave in clearly distinct ways. Broader empirical studies have shown that quite often a different situation holds, with an unexpected degree of variability in interpretation. The nominalizations in *-ing* and conversion are a case in point, as shown in Lieber & Andreou (2018) and Lieber & Plag (2021).

Another approach to modeling of the behavior derived words is Lieber's (2004, 2016, etc.) Lexical Semantic Framework (LSF). In LSF, interpretation works on the basis of the semantic representations of base and affix. Both representations are composed using particular mechanisms, including underspecification. In LSF, polysemous semantic representations can be easily derived and are part and parcel of the system, but any combination of semantic attributes with morphological form is as likely as any other. Predictions about possible tendencies cannot be derived based on this framework.

Yet another perspective on meaning is taken by distributional semantics. In this framework, meanings are determined by the contexts in which words occur. Vectors are used as meaning representations that record and count the contexts in which a given item occurs in the corpus. The closer to each other the vectors are, the closer in meaning are the words represented by those vectors. This approach can also be implemented for complex words (e.g. Marelli & Baroni 2015; Varvara 2017; Lapesa et al. 2018; Wauquier 2020; Huyghe & Wauquier 2020; Missud & Villoing 2021). Distributional semantics offers an interesting take on affix polysemy via the semantically nearest neighbors. Let us use a hypothetical example for illustration of the approach taken, for example, by Lapesa et al. (2018): The nearest neighbors of an eventive nominalization in *-ment*, such as *interment*, are verbs and other eventive nominalizations, whereas for referential nominalizations such as *pavement*, the nearest neighbors are words which themselves have a referential meaning. In this way, distributional semantic models can make predictions with respect to the kind of polysemy that particular types of bases and particular derived forms may exhibit. Overall, the framework is consistent with semantic patterns that are neither categorical nor random.

Finally, we turn to an approach that so far has been employed mainly to understand features of form rather than of meaning, analogical modeling. The present paper will explore the potential of this framework for derivational semantics.

# 3 Analogical Modeling

The term 'analogical modeling' is a cover term for a variety of computational approaches which all hold that linguistic generalization emerges through similarities in the lexicon. Four prominent ones are 'Analogical Modeling of Language (AML)' (e.g. Skousen 1989; Skousen et al. 2002; Skousen & Stanford 2007; Skousen et al. 2013 et seq.), 'Tilburg Memory-based Learner (TiMBL)' (e.g. Daelemans 2002; Daelemans & van den Bosch 2005; Daelemans et al. 2007), the 'Generalized Context Model (GCM)' (Nosofsky 1986) and the 'Minimal Generalization Learner (MGL)' (Albright & Hayes 2003; Albright 2009).

Appeal to analogy has generally had a long tradition especially in the fields of morphology and language change, but it has often been criticized, especially from a generative perspective, for a lack of rigor and restrictions. The computational analogical models mentioned above address this criticism, as they provide a tractable analysis and can generate falsifiable hypotheses that can be tested in various ways, e.g by experiments or corpus data.

In the present paper we will use the AML algorithm, and the rest of this section is devoted to introducing the reader to the specifics of this model. An analogical model basically performs a classification task. For each item that is given to the algorithm as a test item, the algorithm decides between different outcomes (e.g. *-ing* nominalization or conversion) on the basis of the distribution of similar items in the lexicon. The similarities between lexical items is computed over the properties of these items (so-called 'features') that are listed in the lexicon (i.e. coded for the data set in question). For instance, in Arndt-Lappe (2014) the choice between the rival suffixes *-ity* and *-ness* was modeled based on features that coded the syntactic category information and certain phonological properties of pertinent words.

This article presents two different semantic tasks. In the first task (see section 4) we model the choice between two forms, *-ing* nominalization or conversion nominalization based on semantic features. In the other task (see section 5) we model the choice between different interpretations of *-ment* derivatives based on the semantic properties of the base verb.

Figure 1 illustrates the architecture of the analogical model, using the choice between *-ing* and conversion as an example. The central component of the system is the lexicon, i.e. the set of stored exemplars. In this lexicon each item is represented as a set of features. In figure 1, for each item the four features 'base', 'aktionsart', 'quantification' and 'eventivity' are coded. On the basis of these features, the fifth feature ('morphological form') is to be predicted. 'Aktionsart' encodes the aktionsart of the base, 'quantification' encodes whether the derivative is interpreted as mass or count, and 'referentiality' encodes whether the derived noun is eventive or referential in its interpretation. The outcome feature 'morph. form' encodes the form of the nominalization (*-ing* or conversion).

The model can be set up to deal with types or with tokens. In our first case study, we have a data set with multiple tokens of the same words, and our analysis will be token-based. Our second case study, on the other hand, will be type-based. Having a lexicon with token representations means that the lexicon may already contain nominalizations for the same type, i.e. different instances in which a speaker has experienced pertinent words with a particular meaning before. In the toy example in figure 1, the new item to be classified is a new token of BURST. This word is already part of the lexicon with a particular nominalization (see the third row from the bottom in the box labeled 'Lexicon'). We assume, in accordance with the observed facts, that, in general, for each verb both morphological forms are available to form a nominalization. If a new token of BURST is to be classified for the morphological form of its nominalization, the system extracts from the lexicon a group of exemplars which are similar to the new item. This group is called 'analogical set'. The exemplars in the analogical set serve as analogues on the basis of which the morphological form of the nominalization will be assigned to the new item. Abstracting away from various further intricacies to be discussed below, the new item will be classified like the majority of items in the analogical set.

In figure 1, all exemplars in the analogical set share three features with the new item. The nominalization for BURST is then selected on the basis of the distribution of the nominalizations in the analogical set. In the toy example, *-ing* would be chosen in half of the cases, or with a probability of 50 percent. The algorithm gives the probability of each possible value being assigned, based on the distribution of these values amongst exemplars in the analogical set. In calculating these probabilities, AML takes into account the degree of similarity between an exemplar and the new item, as well as the number of exemplars with a particular

Lexicon

| item | base | aktionsart | quantification | eventivity | nominalization |
|---|---|---|---|---|---|
| ACCOUNT | account | activity | count | referential | conversion |
| ASSIST | assist | activity | count | referential | conversion |
| ASSIST | assist | activity | mass | referential | ing |
| BEND | bend | accomplishment | mass | eventive | ing |
| BLEND | blend | accomplishment | count | referential | conversion |
| BLEND | blend | accomplishment | mass | eventive | ing |
| BLEND | blend | accomplishment | count | eventive | conversion |
| BLINK | blink | semelfactive | count | eventive | conversion |
| BLINK | blink | semelfactive | count | eventive | conversion |
| BLINK | blink | semelfactive | count | eventive | conversion |
| BREAK | break | achievement | count | eventive | conversion |
| BREAK | break | achievement | count | eventive | ing |
| BREAK | break | achievement | count | eventive | ing |
| BURN | burn | accomplishment | count | referential | conversion |
| BURN | burn | accomplishment | mass | eventive | ing |
| BURST | burst | achievement | mass | eventive | conversion |
| CATCH | catch | achievement | mass | eventive | conversion |
| FINISH | finish | achievement | mass | eventive | conversion |

Analogical set

| item | base | aktionsart | quantification | eventivity | nominalization |
|---|---|---|---|---|---|
| BREAK | break | achievement | count | eventive | conversion |
| BREAK | break | achievement | count | eventive | ing |
| BREAK | break | achievement | count | eventive | ing |
| BURST | burst | achievement | mass | eventive | conversion |

New item: BURST

| item | base | aktionsart | quantification | eventivity | nominalization |
|---|---|---|---|---|---|
| BURST | burst | achievement | count | eventive | conversion |

**Fig. 1:** The architecture of an analogical model

set of features. The more similar an exemplar is to the new item, the more weight it receives, and the more exemplars that share a particular set of features, the greater the weight that is assigned to each of them. The latter procedure is particularly important in cases where a feature constellation is particularly frequent in the lexicon. In a classification task where exemplars sharing that feature constellation are part of the analogical set, they may outweigh exemplars that are more similar to the target word, but that are smaller in number.

Returning to our example in figure 1, the reader may have noticed that the set of features also includes ‘base’, which means that different exemplars with the same base may be listed in the lexicon. This is desirable since the same base may have more than one nominalization, with different quantification and eventivity readings in different contexts.

The crucial question is of course how the system determines which exemplars end up in the analogical set. As one can see in figure 1, the analogical set in our toy example contains two different types of exemplars. In the parlance of AML, the set of all exemplars sharing a particular feature constellation is called a 'gang'. The idea behind this terminology is that the words that instantiate a particular feature constellation 'gang up' in analogical sets to influence the outcome. In our example, one gang, comprising three instances of BREAK, shares the three semantic features 'aktionsart', 'quantification', and 'eventivity' with our new item BURST. The other gang consists of only one exemplar, an instance of BURST, which differs semantically from the new item in that it has a mass reading, but shares the features 'base', 'aktionsart', and 'eventivity' with the new item. What both gangs in the analogical set have in common is that they share a total of three features with the new item. This raises the question, however, why AML does not also include less similar exemplars. Consider, for example, the two items CATCH and FINISH from the lexicon, which share two features with the new item, 'aktionsart' ('achievement') and 'eventivity' ('eventive'). In AML, the degree of similarity that is relevant for exemplars to be included in the analogical set is decided for each new item individually. The rationale of this procedure is that, on the one hand, the model will always incorporate maximally similar items, but, on the other hand, items with lower degrees of similarity will be incorporated only if that incorporation does not lead to greater uncertainty with respect to the classification task. In our example, the reason why AML does not include all eventive achievement nouns is that the gangs of exemplars sharing these two features do not behave in the same way as the gangs sharing these two features plus one more feature with the new item, i.e. the items in the analogical set. As we have seen, the items in our analogical set testify to both *-ing* and conversion outcomes. By contrast, the two eventive achievement nouns CATCH and FINISH, which have a 'mass' reading, provide evidence for conversion only. They are therefore not included in the analogical set.[1]

As mentioned above, for each item to be classified, AML computes probabilities for the different outcomes. For instance, for a given item, the outcome *-ing* might have the probability of 73 percent, and the outcome *conversion* the probability of 27 percent. AML then selects randomly from analogical sets, which means that in 73 percent of the cases *-ing* will be selected. The percentages can of course also be turned into categorical decisions by way of categorizing percentages at a particular threshold, for example 50 percent for binary choices.

AML can compute probabilities for binary outcomes, but also for multinomial outcomes. In the next two sections we will first see an example of a binary outcome.

**1** A similar case are the second BLEND item and the three exemplars of BLINK. These items also share two features with the new item ('count' and 'eventive'), but they also have a uniform outcome, i.e. conversion. Readers interested in more details of how analogical sets are computed are referred e.g. to Skousen (2005).

This will be followed in section 5 with an example of a model with multiple possible outcomes.

# 4 One meaning - more than one form: *-ing* and conversion nominalizations

## 4.1 The problem

In English, verbs can be nominalized in various ways, with *-ing* suffixation and conversion both being very productive. In a broad empirical study, Lieber & Plag (2021) examine the extent to which these two types of nominalizations in English can express either eventive or referential readings, can be quantified as either count or mass, and can be based on verbs of particular aspectual classes (state, activity, accomplishment, achievement, semelfactive). The examples in (1) illustrate count and mass readings, the examples in (2) illustrate eventive and referential readings.[2]

(1) a. conversion count reading
Accent on Living 1992: And if enough people tell you you're no good, you end up believing it and become what they say you are. A CHANGE has to come.

b. conversion mass reading
Independent School 2006: Overall, then, the single greatest obstacle to implementing curricular CHANGE and, over time, establishing a culture that values continuous reflection and improvement in a school, is the general predisposition of educators to resist change itself.

c. *-ing* count reading
BNC 1992: ...tiny phosphorescent sparks around its hands, small ripples in the stone beneath its feet, a gentle breeze around its head, a sudden dampness and DRIPPING of water from the stones of the walls around it.

d. *-ing* mass reading
BNC 1978: Listen to the rain. Compare slow DRIPPING, fast gushing, trickling, etc.

**2** The examples are taken from Lieber & Plag (2021) (their examples (3), (7a) and (8a)). Example (2c) is taken directly from COCA.

**Tab. 1:** Verbs and their aspectual classes.

| | |
|---|---|
| Accomplishment/achievement | bet, cast, cut, divide, draw, exit, grant, hire, offer, order, pay, return, shift |
| Accomplishment | bend, blend, burn, change, cover, display, fall, fix, form, heat, melt, mix, repair, report, rise, spread, strip, surround, take, transfer, wash |
| Achievement | break, burst, catch, chop, find, finish, fracture, kill reach, reveal, slam, split, win |
| Activity/achievement | pass |
| Activity/state | smell, taste |
| Activity | account, assist, chat, climb, dance, design, drink, drive, embrace, float, gurgle, hold, hurt, kiss, leak, play, push, rest, ride, run, sail, spin, stay, step, stir, stretch, stumble, swim, wait |
| Semelfactive/achievement | pop, strike |
| Semelfactive/activity | beat, shake, shove |
| Semelfactive | blink, cough, drip, flash, hit, jump, kick, knock, poke, punch, spring, tap |
| State | concern, desire, doubt, fear, hate, hope, lack, love, stink, worry |

(2) a. *-ing* and conversion eventive readings
Massachusetts Review 1995: . . . a muffled BANGING began, a thumping, not rapid but steady, like the DRIP of water on a slab, a noise as though someone above was stamping one foot, heavily booted, on a bare board floor.

b. *-ing* referential reading
Mag. Inc. 1995: “Today,” says Holman, “employees in the cleaning department, for example, know not to work on the easiest-to-clean CASTING or on the one that happens to be on the top of the pile.”

c. conversion referential reading
Cosmopolitan 2006: The one downside: Skin in this category tends to have a greenish CAST.

An overview of the aspectual classes of the base verbs in Lieber and Plag (2021) is given in Table 1. For certain analyses, the aspectual classes were recoded using the three aspectual features [dynamic], [durative], and [implied endpoint] (or [endpoint] for short), as shown in table 2.

**Tab. 2:** Aspectual class and aspectual features.

| aspectual class | example | dynamic | durative | endpoint |
|---|---|---|---|---|
| state | *know, love* | - | + | - |
| activity | *push, float* | + | + | - |
| accomplishment | *cook, cover* | + | + | + |
| achievement | *arrive, find* | + | - | + |
| semelfactive | *blink, knock* | + | - | - |

Based on corpus attestations, the authors present statistical evidence that the relationship between morphological form, type of quantification, and aspectual class of base verb is neither categorical, as the syntactic models suggest, nor completely free, as Lieber's framework predicts, but rather is probabilistic.

For investigating the effect of quantification, eventivity and the three aspectual features ([dynamic], [durative], and [implied endpoint]) on the choice of morphological form in a multivariate analysis, Lieber and Plag used conditional inference trees, which is a particular type of classification and regression analysis. Conditional inference trees partition the data into subsets that share particular constellations of feature values and behave significantly differently from other subsets concerning the predicted outcome. The tree shown in figure 2 gives the result of that analysis. The nodes of the tree are numbered for reference. Each terminal node gives the total number of observations in this subset and a stacked bar chart showing the distribution of the dependent variable for the respective constellation of features.

Figure 2 shows that all variables have a say. The model is able to predict the correct morphological category for 86 percent of all cases. When the derivative has a count interpretation, conversion is the clear majority choice (see nodes 5 through 8). Among the count nominalizations, [implied endpoint], eventivity and [durative] also play a role, and do so in particular subsets. With mass nominalizations, things are more complicated. If based on state verbs ([dynamic]: no), conversion is prevalent (node 21), but if based on dynamic verbs ([dynamic]: yes), further distinctions are necessary. Eventive nominals favor *-ing* (nodes 13 through 15), but this preference depends on the presence or absence of the features [implied endpoint] and [durative]. For the referential nominals, the choice depends on the

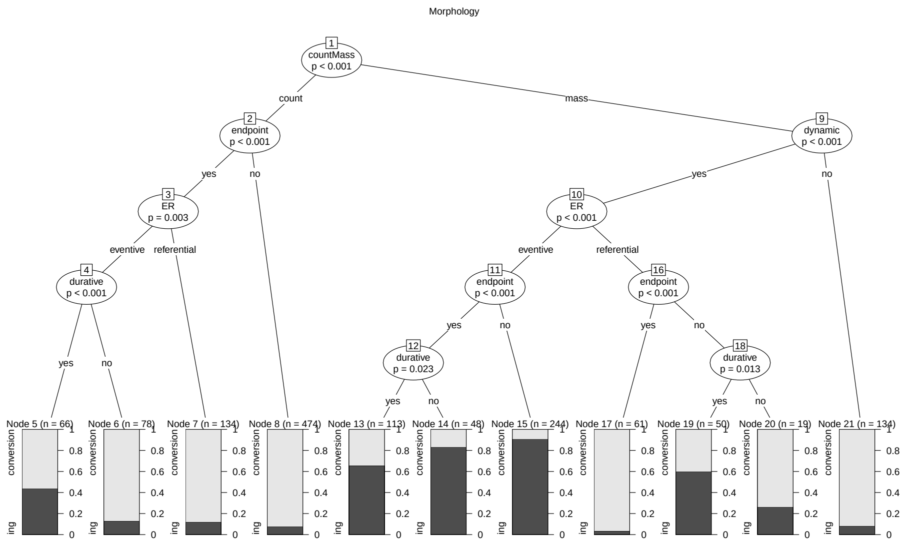


**Fig. 2:** Conditional inference tree modeling the choice of morphological form (taken from Lieber & Plag 2021).

features [implied endpoint] and [durative] (nodes 17, 19, 20). To summarize, the relationship among the variables of morphological form, eventivity, quantification and aspect is rather complex. Tendencies sometimes go in the direction suggested by past literature (e.g. *-ing* forms tend to be eventive), but sometimes contradict past predictions (conversion also tends to be eventive).

A distributional analysis showed that at least for some base verbs there is a strong tendency towards one of the two morphological forms. A mixed-effects regression analysis by Lieber and Plag supported this finding. In predicting the morphological form (in the presence of the predictors QUANTIFICATION and EVENTIVITY), the base verb plays a significant, even if quite moderate role.

The variability of the data raises the question which kind of linguistic model can account for the distribution of the two morphological forms based on the intended semantics and the aspectual properties and the identity of the base. The conditional inference tree analysis makes good predictions and can therefore serve as a reference point. However, a conditional inference tree does not seem to be a cognitively plausible model of how speakers choose a morphological form. Analogy, in contrast, is a plausible cognitive mechanism, and an analogical model, if successful, would therefore be a step forward in understanding competition between morphological forms.

## 4.2 Setting up the analogical model

For this study we used the same data set that was analyzed statistically in Lieber & Plag (2021). In particular, we employed the subset of the data for which all three semantic categories (eventivity, quantification and aspectual features) were coded unambiguously. This data set consists of 1421 tokens of nominalization, with 150 types based on 84 base verbs.[3] In addition, to test the possibility that phonological properties might also play a role in the choice between *-ing* and conversion, we also coded the segmental and syllabic structure of all items in the same way as in Arndt-Lappe's (2014) study of the competition between *-ity* and *-ness*.

The model was implemented using the 'Transparent Analogical Modeling of Language (TraML)' package (Arndt-Lappe et al. 2018), which provides rather convenient access to the model output. The algorithm was set up in such a way that the data set was used both as the lexicon file and as the test file. Specifically, for each word in the file, the morphological form is predicted on the basis of all other words in the file. In AML, this is implemented by the parameter setting 'exclude given'. In general, this kind of method is also known as 'leave-one-out' (see, for example, Daelemans & van den Bosch 2005). This simulates that a speaker decides on the form of the nominalization for a given verb on the basis of their knowledge of (potentially all) the words that are in the lexicon.[4] In our explorations of the analogical model we largely follow the rationale and procedures laid out in Arndt-Lappe (2021). We use R R Core Team (2019) for the statistical analysis.

## 4.3 Results

### 4.3.1 Accuracy

The model that includes phonological features on top of the base and the semantic features performs worse than the model without the phonological features. Only 58 percent of the nominalization forms are predicted correctly, and *-ing* is predicted correctly only in one third of the cases. It thus seems that similarities in the

**3** We refer the reader to Lieber & Plag (2021, section 3) for a detailed description of the generation of the data set and the coding procedures.

**4** Another parameter setting concerns the way the relations between gangs are computed. We used the default setting 'linear' (instead of 'squared').

phonological make-up do not help to decide whether speakers choose conversion or *-ing*.[5]

The model that includes only the base and the semantic information as features is quite successful in choosing the right morphological form. Of the 1421 forms, 84 percent are predicted correctly. This is very similar to the accuracy of the conditional inference tree analysis (86 percent), the difference is not significant (Chi-square test, $p$=0.56, $\chi^2$=0.34, $df$=1).

The output of the analogical model allows us to investigate in more detail the quality of the predictions. Figure 3 shows the certainty of predictions separately for conversion and *-ing* nominals. For each of the 945 conversion cases and 476 *-ing* nominals in the data set, the y-axis plots the probability with which the model predicts conversion (left panel) and *-ing* (right panel). We can see that the probabilities in the left panel are generally higher, and that percentages below 50 percent (which means wrong prediction if we dichotomized the probabilities at the 50 percent level), are quite rare for conversion nouns. Dichotomizing these probabilities, we end up with only 11 percent wrong predictions of conversion and 27 percent wrongly predicted *-ing* derivatives. Again this is in the same ballpark as the predictions of the tree analysis, in which 10 percent conversions and 23 percent of all *-ing* nouns were wrongly predicted. In sum, the accuracy of the analogical model is highly satisfactory.

The distribution of the percentages as shown in figure 3 also means that AML predicts speakers to be generally very certain and less variable when choosing conversion: On average, the probability of conversion within analogical sets is somewhere near 90 percent, and the variance is rather low. This is different when *-ing* is predicted. Here, the average probability is only around 70 percent, and there is greater item-by-item variance. This means that speakers will be less certain about choosing *-ing*.

### 4.3.2 Gangs: How similar forms behave

The output of the AML model also allows us to have a closer look at how similar forms behave. This is interesting from a theoretical perspective as this behavior can demonstrate how rather clear generalizations can emerge in an analogical system. For these kinds of analysis it is useful to inspect specific feature constellations. In the case of our toy example shown in figure 1, the words in the analogical set all

---

**5** This null effect for phonological features may be due to problems of feature coding or feature weighting. This issue is left for future studies to explore.

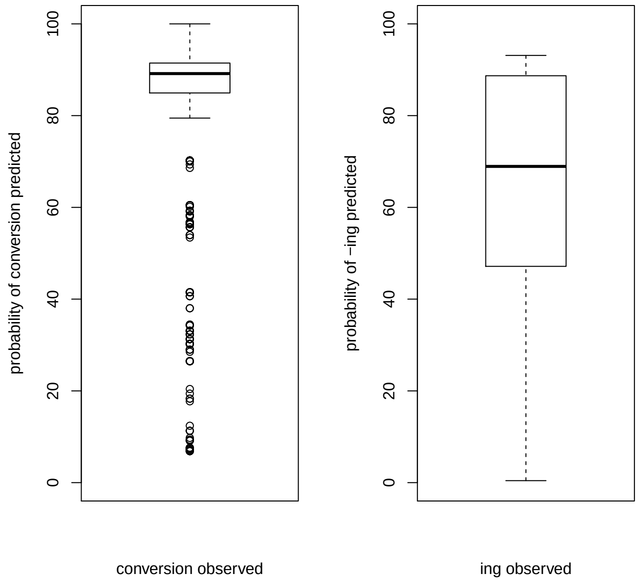


**Fig. 3:** Certainty of predictions for conversion vs. *-ing*.

share the same feature constellation (i.e. `break;achievement;count;eventive`).[6] However, this is not necessarily the case in all analogical sets. For instance, for one of the tokens of the conversion noun *bend* in our data set, the words in its analogical set instantiate three different feature constellations as shown in table 3. Nominalization tokens of the following verbs instantiate these constellations: *bend, burn, change, display, fall, fix, form, heat, melt, mix, repair, rise, spread, strip, take, transfer, wash.*

The target token has the feature constellation `bend;count;eventive;yes;yes;yes`. All three constellations are equally similar to the constellation of the target item in that they all lack one of the target item's features. Our target token thus has three gangs in its analogical set (instantiated by the nominalization tokens of the 17 verbs mentioned above). Depending on how many words instantiate a particular feature constellation, and depending on how uniformly this feature constellation behaves concerning the outcome, a gang is more or less influential,

**6** We use a different font-type and semicolons between feature values to indicate feature constellations.

**Tab. 3:** Feature constellations. Empty cells indicate features not shared with the target item.

| item | base | quantification | referentiality | dynamic | durative | endpoint |
|---|---|---|---|---|---|---|
| BEND | | count | eventive | yes | yes | yes |
| BEND | *bend* | count | | yes | yes | yes |
| BEND | *bend* | | eventive | yes | yes | yes |

and contributes more or less to the probability of a particular choice of outcome (Skousen et al. 2002; Skousen 2002).

To see emergent generalizations, we can now inspect the relation between particular kinds of gangs and particular outcomes by using the output files of TraML. In our first analysis we included all gangs that were most influential in their analogical set (i.e. that had the greatest say in the vote for the outcome) and that had outcome probabilities for either conversion or *-ing* of more than 80 percent. The threshold of 80 percent means that those gangs can be taken to behave in an almost uniform way. Table 4 shows these gangs and the outcomes with which they are associated, first for conversion nouns, then for *-ing* nominalizations.

**Tab. 4:** Highly predictive feature constellations and their outcomes.

| base | quantif. | referent. | dynamic | durative | endpoint | tokens |
|---|---|---|---|---|---|---|
| ***conversion*** | | | | | | |
| | count | eventive | yes | no | no | 122 |
| | count | eventive | yes | no | yes | 49 |
| | count | eventive | yes | yes | no | 237 |
| | count | referential | no | yes | no | 3 |
| | count | referential | yes | no | yes | 7 |
| | count | referential | yes | yes | no | 58 |
| | count | referential | yes | yes | yes | 79 |
| | mass | eventive | no | yes | no | 130 |
| | mass | referential | yes | yes | yes | 23 |
| fear | | eventive | no | yes | no | 1 |
| love | | eventive | no | yes | no | 2 |
| ***-ing*** | | | | | | |
| | mass | eventive | yes | no | no | 1 |
| | mass | eventive | yes | no | yes | 9 |
| | mass | eventive | yes | yes | no | 177 |
| drive | mass | | yes | yes | no | 2 |

The highly predictive gangs are in line with the tendencies or generalizations proposed in the literature, in that count readings are often associated with conversion, and mass readings with *-ing* nominalization. In AML these tendencies find an explanation as emerging from the similarities of target words with the words in the lexicon. The influence of individual base words is also evident, but naturally holds only for small numbers of items. The distribution of the number of tokens shows a very important property of analogical models: generalizations can be drawn on the basis of shared features, and the generalizations may thus concern very small numbers of forms or rather large numbers of forms, depending on how many forms share these features.

Finally, we can inspect the relation between feature constellations and observed versus predicted outcomes. We take the same gangs as before and plot two networks that visualize these relations. Figure 4 shows the network of the gangs that lead to a conversion prediction, figure 5 shows the network for the predicted *-ing* nouns.[7] Recall that these predictions are at a probability level of 80 percent or more. In the center of each cloud the feature constellation is given in small letters. The actual target forms sharing this feature constellation are given in capital letters. They are connected with arcs to the feature constellation that they instantiate.

**7** We used the igraph package (Csardi & Nepusz 2006) in R (R Core Team 2019) to produce the graphs in figures 4 and 5.

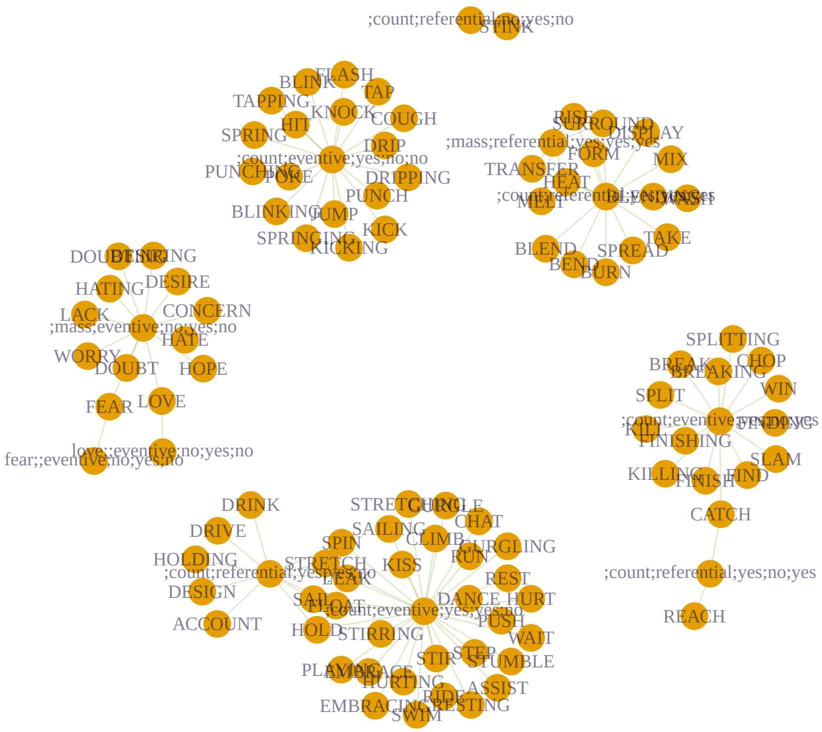


**Fig. 4:** Network of gangs for conversion items.

We can see that, as expected, in the conversion network most observed forms are conversion nouns, and in the *-ing* network most observed forms are *-ing* nominalizations. We can also see that the same forms may participate in more than one gang. For instance, LOVE (in the leftmost network in figure 4) participates in two gangs, with complementary and overlapping features (`love;-;eventive;no;yes;no` and `-;mass;eventive;no;yes;no`). Another case of the same forms participating in more than one gang is SPIN, STRETCH, LEAK, SAIL, FLOAT, HOLD (in the bottommost network in figure 4). With these items we see two feature constellations (`-;count;eventive;yes;yes;no` and `-;count;referential;yes;yes;no`) that have opposing features (`eventive` vs. `referential`), but also overlapping features (`count;yes;yes;no`). The presence of opposing semantic features for different

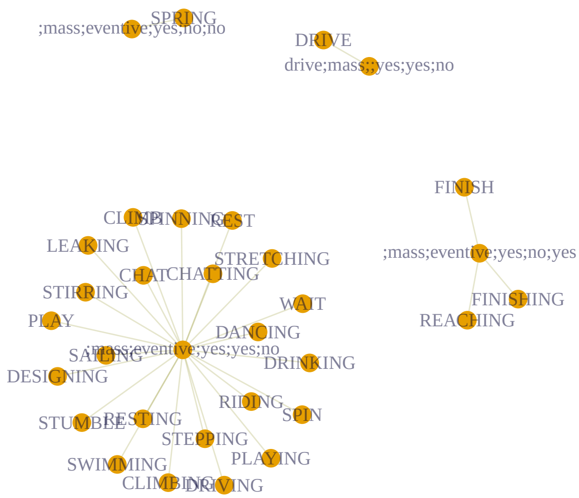


**Fig. 5:** Network of gangs for *-ing* items.

tokens of the same type means that these words are polysemous. The conversion nouns just mentioned (like many others) have both an eventive and a referential interpretation. As this kind of polysemy is wide-spread it necessarily introduces variability and uncertainty also into the construction of the analogical sets, and thus into the decisions. At the conceptual level, this variability and uncertainty is part and parcel of analogical reasoning, and not an unwelcome disturbance of an otherwise neat categorical system. Recall in this context that classification accuracy is typically measured in terms of a majority vote among exemplars in the analogical sets. However, this majority vote is a generalising simplification of what the algorithm computes as a probabilistic outcome for each individual test word.

Another important observation is that a given feature constellation may be instantiated by different target forms derived from the same verbal base. This can be most clearly seen in the rightmost network in figure 5, which includes FINISH and FINISHING, both of which share the feature constellation `-;mass;eventive;yes;no;yes`. Similar cases are PLAY and PLAYING, or CHAT and CHATTING, in the bottom left network of figure 5. Again, this kind of situation is a source of variability and uncertainty that the analogical system deals with straightforwardly.

### 4.4 Summary: One meaning – more than one form

In this case study we have modeled a data set in which the competition between two morphological forms is determined by semantics, but in a non-categorical and highly complex fashion. The traditional statistical analysis provided by Lieber and Plag had revealed the intricate interaction of different features.

We have shown that the analogical model is able to cope successfully with the challenges of this data set at the same level of accuracy as a conditional inference tree model. However, the conditional inference tree has several conceptual disadvantages. First, it does not include the influence of the individual base although there is evidence that this is a relevant predictor. Second, wrong predictions are to be treated as unexplained exceptions to some rule. Third, the interaction of predictors is taken into account, but the nature and cause of these interactions is unclear. Fourth, the statistical algorithm cannot easily map onto plausible cognitive mechanisms (but see section 5 for discussion of this point).

In contrast, the analogical model has several advantages. It is based on a cognitively plausible mechanism, it is rather accurate in its predictions, and what is conceptualized as the interaction of variables in a regression model emerges naturally on the basis of similarities in these properties between lexical items. Furthermore, the model makes satisfactory generalizations at different levels of granularity, and the notion of 'exception to the rule' becomes obsolete. Finally, the role of the base is a straightforwardly included in an analogical model alongside other properties of the words involved.

In summary, although analogy has hitherto been applied mainly on the basis of formal properties, analogy also proves to be a very promising approach when it comes to semantic properties. Semantic properties can be used equally well (and successfully) to understand the choice between two morphological forms. We may now turn to the second general problem in derivational semantics, affix polysemy.

## 5 One form – many meanings: *-ment*

### 5.1 The problem

We speak of affix polysemy when one affix is able to generate several possible and related readings. This is wide-spread in English: Lieber (2016) investigates a total of 27 nominalizing suffixes and shows that 19 of these are able to produce more than one reading (p. 60f). The suffix that we will focus on in this part of the paper, *-ment*, can be considered highly polysemous, as is illustrated in (3) with examples from the literature (Gadde 1910; Marchand 1969; Bauer et al. 2013; Lieber 2016; the actual semantic labels partly differ between authors).

(3) a. EVENT: *ceasement*
b. ACTION: *repayment*
c. STATE/CONDITION: *contentment*
d. RESULT: *improvement*
e. PRODUCT: *pavement*
f. INSTRUMENT/MEANS: *refreshment*
g. inanimate PATIENT/THEME: *investment*
h. LOCATION: *establishment*

In addition, it is often the case that several readings are possible even for one and the same derivative. For example, according to the *Oxford English Dictionary Online* (OED), *refreshment* can denote "a means of refreshment" as well as the "action of refreshing a person or thing." That is, *refreshment* can not only exhibit reading (3-f), but also reading (3-a). At the same time, a given derivative will most likely not exhibit the full range of readings that its affix can potentially produce. For example, the OED does not list *refreshment* as 'something which has been refreshed' or as 'the place of refreshing' (readings (3-g) and (3-h)).

It has been observed that the readings which are possible for a given derivative are often predictable. For example, INSTRUMENT nominalizations derive from verbs that denote actions requiring an INSTRUMENT participant (e.g. *season* > *seasoning*, *equip* > *equipment*; Bauer et al. 2013, 213-4). For deverbal *-ment* neologisms, Kawaletz (2021) has shown that their full range of readings can be predicted, given a detailed enough decomposition of their bases' semantics. The base offers an array of semantic elements, and the suffix selects from this array in a systematic way, producing a polysemous *-ment* derivative. Kawaletz (2021) models the semantics of both the verbal input and its derived output with frames, which are recursive typed feature–value structures that formalize the semantic representations.

A crucial question, however, remains unanswered in Kawaletz's work: By means of what mechanism do speakers determine for a given derivative which readings are possible, given a base verb with particular semantic properties? In this section, we test analogy as a possible mechanism, using Kawaletz's data set and a translation of her frame formalizations into an AML-readable feature matrix.

## 5.2 Setting up the analogical model

In the analogical model, we want to predict the readings that are possible for a *-ment* derivative, given the semantic properties of its base. For this, we use Kawaletz's (2021) data set of *-ment* neologisms. The data set consists of 40 types with a data base of 502 semantically annotated attestations from various corpora. For each type, a number of readings are available, and each type enters the data base for the present study, i.e. our AML lexicon, with as many entries as there are attested readings for this type. The base verbs belong to two semantic classes (see Levin 1993; Kipper Schuler 2005): change-of-state verbs (e.g. *embrittle, discolor, unfold*; henceforth COS verbs) and verbs of psychological state (e.g. *enrapture, reassure, stagger*; henceforth psych verbs). Kawaletz (2021) proposes seven distinct frames, representing semantically slightly different subgroups of verbs within these two classes.

As explained above, for AML, a feature matrix is needed that encodes the properties of the exemplars of the lexicon. This poses a challenge because the semantic properties of the words in Kawaletz's data set are coded in a very different format, i.e. with hierarchical, recursive attribute–value matrices (called 'frames', see Barsalou 1992a,b; Löbner 2013), similar to those used in other frameworks (such as HPSG or Sign-based Construction Grammar; see Pollard & Sag 1994; Sag 2012). To serve as input for AML these frames must be transformed into a two-dimensional table (i.e. feature matrix). In order to illustrate the problem, let us have a look at the frame in figure 6, which is a semantic representation of the possible readings of COS verbs in the form of a lexeme formation rule (see Plag et al. 2018). The derivative *embrittlement*, for example, would be an instantiation of this rule.

The lexeme-formation rule includes phonological information (PHON), syntactic information (CAT), semantic information (SEM), and possible readings (REF), both for the derivative (first column) and for its base (second column, introduced by M-BASE). Small numbered boxes are used for (co-)indexation. The *change-of-state causation* $\wedge$ *come-into-being causation* event denoted by the base verb is a complex event with two subevents (CAUSE and EFFECT). This complex event has four participants, ACTOR, PATIENT, INSTRUMENT and PRODUCT (nodes $\boxed{1}$ to $\boxed{4}$), and the complex event structure given in nodes $\boxed{5}$ and $\boxed{6}$. The first subevent, CAUSE, is underspecified; it can be any kind of event with any kind of participant. The second subevent, EFFECT, is specified as a *change-of-state* $\wedge$ *come-into-being*, during which the PATIENT attains a *state*, and the PRODUCT comes into existence. The labels used in the frames are hierarchically related to one another, which is formalized in a 'type hierarchy.' For example, *change-of-state causation* $\wedge$ *come-into-being*

$$
\left[\begin{array}{ll}
\multicolumn{2}{l}{\textit{lexeme}} \\
\text{PHON} & /\boxed{z}\text{-}\textit{ment}/ \\
\text{CAT} & \text{N} \\
\text{SEM} & \boxed{0} \\
\text{M-BASE} & \left[\begin{array}{lll}
\multicolumn{3}{l}{\textit{lexeme}} \\
\text{PHON} & \boxed{z} & \\
\text{CAT} & \text{V} & \\
\text{SEM} & \boxed{0} & \left[\begin{array}{lll}
\multicolumn{3}{l}{\textit{change-of-state causation} \wedge \textit{come-into-being causation}} \\
\text{ACTOR} & \boxed{1} & \\
\text{PATIENT} & \boxed{2} & \begin{bmatrix} \multicolumn{2}{l}{\textit{entity}} \\ \text{ON-REGION} & \boxed{7} \end{bmatrix} \\
\text{INSTRUMENT} & \boxed{3} & \\
\text{PRODUCT} & \boxed{4} & \begin{bmatrix} \multicolumn{2}{l}{\textit{entity}} \\ \text{REGION} & \boxed{8} \end{bmatrix} \\
\text{CAUSE} & \boxed{5} & \begin{bmatrix} \multicolumn{2}{l}{\textit{event}} \\ \text{PARTICIPANT} & \boxed{1} \end{bmatrix} \\
\text{EFFECT} & \boxed{6} & \left[\begin{array}{lll}
\multicolumn{3}{l}{\textit{change-of-state} \wedge \textit{come-into-being}} \\
\text{PATIENT} & \boxed{2} & \\
\text{PRODUCT} & \boxed{4} & \\
\text{RESULT-STATE} & \boxed{9} & \begin{bmatrix} \multicolumn{2}{l}{\textit{state}} \\ \text{PATIENT} & \boxed{2} \end{bmatrix}
\end{array}\right]
\end{array}\right] \\
 & \multicolumn{2}{l}{\textit{part-of}\,(\boxed{8}, \boxed{7})} \\
 & \multicolumn{2}{l}{\text{REF} = \left\{\boxed{0}\right\}}
\end{array}\right] \\
\multicolumn{2}{l}{\text{REF} = \left\{\boxed{0}, \boxed{1}, \boxed{3}, \boxed{4}, \boxed{5}, \boxed{6}, \boxed{9}\right\}}
\end{array}\right]
$$

**Fig. 6:** Lexical rule for the possible readings of *-ment* derivatives derived on the basis of COS verbs like *embrittle*, adapted from Kawaletz (2021, 102).

*causation* is a subtype of *event*, and a PRODUCT is a special kind of RESULT (see Kawaletz 2021, 108 for details).

In the frame-based approach, the semantics of an affix is describable as its potential to perform a referential shift on the frames of its bases (see also Plag et al. 2018). The set of possible readings for a given derivative is specified as "REF = {$\boxed{0}$, $\boxed{1}$, $\boxed{3}$, $\boxed{4}$, $\boxed{5}$, $\boxed{6}$, $\boxed{9}$}". This notation signifies that these seven nodes represent possible interpretations of the derived noun (e.g. *embrittlement*), reflecting the polysemy of the derivative. According to Kawaletz (2021), the derivative can refer to the whole, complex event ($\boxed{0}$), to either event participant except for PATIENT ($\boxed{1}$, $\boxed{3}$, $\boxed{4}$), to one of the two subevents ($\boxed{5}$ or $\boxed{6}$), or to the RESULT-STATE ($\boxed{9}$).

We translated the frames into a feature matrix using the following strategies. As the first feature we took the base. All other features encode semantic properties of the base verbs. The next set of features encoded whether a given frame element (e.g. the attribute PRODUCT or the node *state* in figure 6) occurs in the frame (features 3 to 13). We then coded at which level of embedding the element occurs (features 14 to 21). Finally, we included a feature that encoded the presence of co-indexation (feature 22). We only included features that were distinctive between verbs.

Table 5 illustrates the coding for two types, with two readings of *embrittlement* and one reading of *approvement*. We chose these two derivatives as their base verbs come from two different verb classes, COS verbs (*embrittlement*) and psych verbs (*approvement*). The information in row one ('item') functions as an identifier. For each item, 21 features are coded. On the basis of these, the feature in row 23 ('reading') is to be predicted.

With regard to the presence or absence of particular semantic elements, we see, for instance, that the presence of the two attributes INSTRUMENT and PRODUCT in the frame of *embrittlement* is encoded in features five and eight. The frame for verbs like *approve of*, on the other hand, has neither of the two, which is why the corresponding cells in rows five and eight are left empty. In accordance with the leave-one-out method, this is a meaningful bit of information for the model.

The hierarchical information, i.e. the degree of embedding of a frame element in the type hierarchy, is coded by the tags '+1' or '+2.' For example, ACTOR+1 (feature 3) is embedded one level below its parent node, coding for its daughters (here AGENT and CAUSER). ACTOR+2 (feature 4), is further embedded by one level, where we find STIMULUS – a category which is only relevant for psych verbs.

We further addressed the recursive frame structure by spelling out attribute paths, showing where and how deeply a frame element is embedded in the given frame (features 14 to 21). For example, the value *state* is deeply embedded in the *embrittle*-frame, and can be reached via the attribute path 'EFFECT:COS_RESULT-STATE:state' (feature 20; note also how the path builds up in features ten and

**Tab. 5:** Partial lexicon for the *-ment* data: bases *embrittle* and *approve of*. Attributes are given in small caps, values in regular type font. Abbreviations: att = attribute, cib = come into being, cos = change of state, eff = effect, ent = entity, evt = event, exp = experiencer, instr = instrument, part = participant, pat = patient, psy = psych, refnode = reference node, rs = result state, st = state, stim = stimulus, und = undergoer, val = value (see Kawaletz 2021 for details).

| feature | item 1 | item 2 | item 3 |
|---|---|---|---|
| 1. item | embrittlement | embrittlement | approvement |
| 2. base | embrittle | embrittle | approve_of |
| 3. ACTOR+1 | AGENT/CAUSER | AGENT/CAUSER | CAUSER |
| 4. ACTOR+2 | | | STIM |
| 5. INSTR | INSTR | INSTR | |
| 6. PAT+1 | | | EXP |
| 7. RESULT | RESULT | RESULT | |
| 8. RESULT+1 | PRODUCT | PRODUCT | |
| 9. CAUSE:val | CAUSE:evt | CAUSE:evt | |
| 10. EFF:val | EFF:cos | EFF:cos | |
| 11. EFF:val+1 | EFF:cos&cib | EFF:cos&cib | |
| 12. st+1 | st | st | psy-st |
| 13. cos | cos&cib | cos&cib | |
| 14. refnode | causation | causation | state |
| 15. . . . _ST | UND:ent_ST | UND:ent_ST | UND:ent_ST |
| 16. . . . _ST+1 | PAT:ent_ST | PAT:ent_ST | EXP:entity_PSY-ST |
| 17. CAUSE:val_ATT | CAUSE:evt_PART | CAUSE:evt_PART | |
| 18. EFF:val_ATT | EFF:cos_RS | EFF:cos_RS | |
| 19. EFF:val_ATT+1 | EFF:cos_RS | EFF:cos_RS | |
| 20. EFF:val_ATT:val | EFF:cos_RS:st | EFF:cos_RS:st | |
| 21. EFF:val_ATT:val+1 | EFF:cos_RS:st | EFF:cos_RS:st | |
| 22. INSTR/STIM | | | |
| 23. reading | transposition | instrument | transposition |

18). Verbs like *approve of*, on the other hand, denote psych states, which is why their reference node is labeled *state* (feature 14) or, on a finer level of granularity, *psych-state* (feature 12).

The last frame property encoded in the feature matrix is the presence or absence of co-indexed nodes. Feature 22 contains explicit information with regard to co-indexation, namely whether or not the values of the attributes INSTRUMENT and STIMULUS are co-indexed. This feature sets apart two subgroups of psych verb from the remaining data set. Other co-indexed nodes are not distinctive and are therefore not included in the model.

The resulting feature matrix contained one row for each reading of a given type. We see this illustrated in table 5 for the noun *embrittlement*, for which table 5 shows two entries, each with a different reading. Obviously, for a given base verb, the semantic structure of its base verb is always the same, hence the features 1 to 22 are identical. The two entries only differ in their reading. Overall, the lexicon contains 194 items, i.e. combinations of type–reading pairs.

Using the resulting feature matrix as input, the model was set up using the same parameter settings as in the study described in section 4 ('exclude given', 'linear'). The feature matrix contains empty cells, which indicate the absence of the respective feature. Given that the presence or absence of a feature is meaningful (see below), AML was instructed to treat empty cells not as missing information, but as valid feature values. This is encapsulated in the parameter setting ‘include nulls’.

## 5.3 Results

In this section we will present different aspects of the analogical model we implemented to predict possible readings of *-ment* derivatives. We will begin with an analysis of the predictions to see how successful the model is. This will be followed by an investigation of other aspects of the model in order to gain further insights into the nature of the model and its architecture.

### 5.3.1 Predictions: The polysemy of *-ment*

In this section we want to answer the question of whether the analogical model is capable of correctly determining which readings are available for a given derivative, given a base verb with particular semantic properties and a lexicon containing derivatives of the same morphological category.

Figures 7 and 8 give a visual impression of the predictions made by AML. For each derivative, a stacked bar gives the different readings that were predicted and the probability of their being selected. Readings not shown ended up with a probability of zero, i.e. they were not predicted to be available readings. For example, the bar labeled *abridgement* (to the very left in figure 7) shows that AML predicts seven different readings for this derivative. The probabilities for a given combination of noun and reading is color-coded, with shades of blue representing predicted eventive readings and shades of orange representing predicted participant readings.[8]

Overall, we see that AML predicts (with variable, non-zero percentages) four or more readings for all derivatives but one (*musement*, see below for further discussion). This means that *-ment* is predicted by the algorithm to produce highly polysemous derivatives. This corresponds to Kawaletz’s 2021 findings. Moreover, the readings that AML predicts largely correspond to those that are attested. Of the 194 combinations of derivative and readings that are attested in Kawaletz’s data base, AML predicts 189, i.e. 97 percent. The probabilities of the different

**8** Each bar subsumes all items for a given derivative. For example, *abridgement* is represented by six items in the lexicon because it is attested in six different readings. Since the coding of the features 1 to 22 is the same for these six items, the predictions for feature 23 (the reading) are also the same.

readings are mostly quite evenly distributed for a given derivative, with no clear majority decisions. With some verbs, certain readings are not very likely (e.g. 'implicit product' for *congealment* and *debauchment*), which is indicated by the low percentage of that reading.

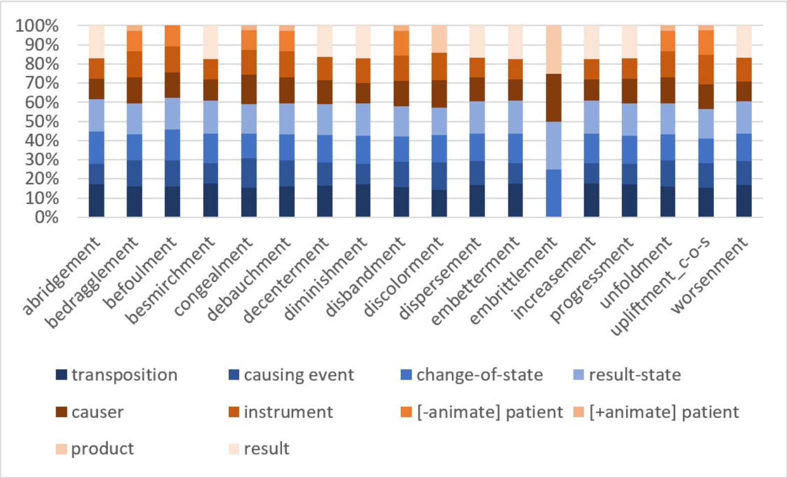


**Fig. 7:** Predictions of readings for derivatives in the COS subset of the data.

Let us examine a number of predictions more closely to see what motivated them and what this teaches us about the model. We start with two readings for which AML does very well: CHANGE-OF-PSYCH-STATE and TRANSPOSITION. CHANGE-OF-PSYCH-STATE is a reading that is specific to a subgroup of psych nouns, namely those that are based on verbs with a *change-of-psych-state* subevent (as opposed to just a *psych-state* subevent or to no complex event structure at all; see also table 5 on page 25). AML predicts this reading precisely for those nouns for which it is attested, namely *endullment, enragement, soothment* and *upliftment* (towards the right in figure 8). TRANSPOSITION is a reading that is available for all productively formed *-ment* derivatives. AML predicts this reading across the board, as shown by the very dark blue portions of the bars in figures 7 and 8).

The only derivative for which TRANSPOSITION is not predicted is *embrittlement*. In fact, Kawaletz (2021) finds the derivative attested in seven different readings, but AML only predicts four of these. *Embrittlement* is thus predicted to be less polysemous than the other COS derivatives (see figure 7). Where does this exceptional behavior originate? A look into the analogical sets reveals that these predictions are based on only one other derivative, *discolorment* (see also figure 9 below). The

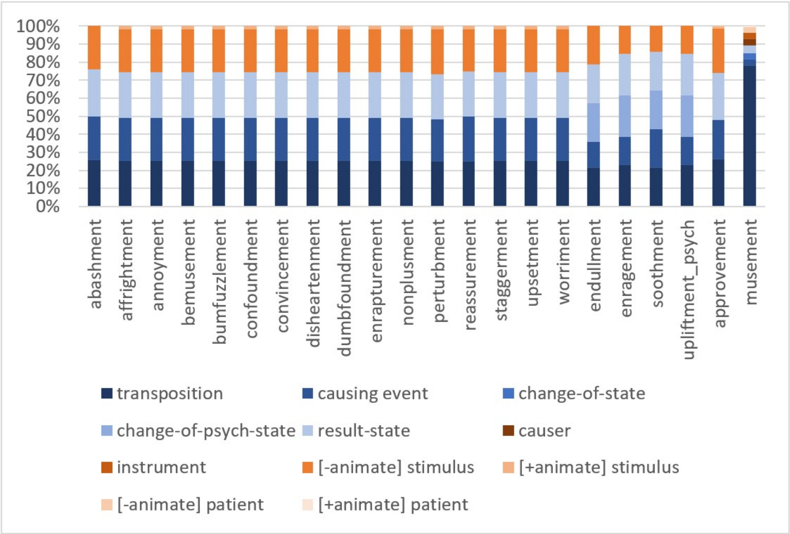


**Fig. 8:** Predictions of readings for derivatives in the psych subset of the data.

attestations for *discolorment*, in turn, are patchy: Based on the frame for the base *discolor*, the readings TRANSPOSITION, INSTRUMENT and CAUSING-EVENT should be possible, but Kawaletz (2021) does not find them attested for this derivative. Since the AML lexicon does not contain items representing *discolorment* in these readings, then, they are not predicted for *embrittlement* either. In section 5.3.3 we will come back to the issue of such gaps in the lexicon data.

Two further derivatives for which AML's predictions stand out are the psych nouns *approvement* and *musement* (see the two right-most bars in figure 8). We have already mentioned that, for some reason, *musement* is the only derivative for which AML makes a clear majority decision (0.78 probability of a TRANSPOSITION reading). *Approvement*, on the other hand, is predicted to have four different readings, two of which (CAUSING-EVENT and RESULT-STATE) should not be possible, based on its base verb frame. An additional four such unexpected readings surface with a very low probability of 0.02 percent.

These odd predictions are made because *approvement* and *musement* are odd as well: Both have unique base verbs in the data set, sharing their semantic representation with no other base verbs. What is more, their frames model simple events (psych-state and psych-action, respectively) and therefore differ substantially from the other five frames, which all model complex events. AML is thus confronted with two feature constellations that are very different from all the others in the lexicon. An inspection of the derivative's gangs tells us how the model dealt with this:

It resorted to calculating *approvement*'s and *musement*'s analogical set members on the basis of very few gang features. For example, one of *approvement*'s gangs contains only the feature value 'UNDERGOER:entity_STATE,' which it shares with all bases except for *muse over*. Because both *approvement* and *musement* are unique within the data set, AML had to resort to such an underspecified feature constellation. As we have seen above, this results in wrong predictions for these two nominalizations.

Another interesting observation is that the model predicts the readings [+animate] STIMULUS and [+animate] PATIENT. According to the literature, *-ment* does not produce [+animate] readings (see Kawaletz & Plag 2015; Lieber 2016; Kawaletz 2021). There are, however, exceptions to this rule: In Kawaletz's (2021) data set, *befoulment* is attested in an [+animate] PATIENT reading and *abashment* is attested in an [+animate] STIMULUS reading. These exceptional items in the lexicon lead to AML predicting such readings, albeit with a very low probability, for those nouns which have *befoulment* and/or *abashment* in their analogical set (for instance *bedragglement*; see figure 6 above). For these derivatives, the probabilities range from 2.56 percent to 2.7 percent for an [+animate] PATIENT reading and from 1.45 percent to 1.7 percent for an [+animate] STIMULUS reading. Note that *befoulment* and *abashment* themselves are not predicted to exhibit [+animate] readings due to the leave-one-out method. Although these predictions do not match the attested readings, they do reflect linguistic reality: [+animate] readings are not entirely impossible for speakers, but merely very unlikely.

### 5.3.2 Analogical sets: Emergent base verb classes

Kawaletz (2021) builds on the common assumption that there is a systematic relationship between the semantics of the base and the semantics of the derivatives of a particular morphological category. In her approach, verbs of a particular semantic class (e.g. psych verbs) produce derivatives that are systematically different from derivatives of other classes in terms of their possible interpretations. In an analogical approach, verb classes such as those used by Kawaletz do not play a role as analytical entities. If anything, such a class would emerge as a set of verbs that share a particular feature constellation. And if these feature constellations are really influential in choosing possible interpretations, they should emerge as analogical sets in the analogical model. If Kawaletz's analysis is on the right track, and at the same time AML can successfully model her data, the analogical sets emerging in the AML model should reflect the verb classes that feature prominently in Kawaletz's approach.

Figure 9 visualizes the analogical sets that the AML algorithm has used to classify the test items. This offers us an insight as to which sets of base verbs are similar to each other, and at the same time influential for the task. Each node represents a derivative, and the two semantic classes of base verb are indicated by color: derivatives with psych verbs as bases ('psych items' for short) are blue, while derivatives with COS verbs as bases ('COS items' for short') are yellow. An arrow between two words is drawn if a given word ('word 1') is part of the analogical set that is used for the prediction of the interpretation of the other word ('word 2'). Quite often, word 1 is in the analogical set of word 2, and word 2 is also in the analogical set of word 1. In such cases, we get arrows pointing in two directions. If many words are in each others' analogical sets, we see the emergence of rather large clouds. A cloud can thus be interpreted as a set of words that are similar to each other and mutually predictive for the decision to be taken.

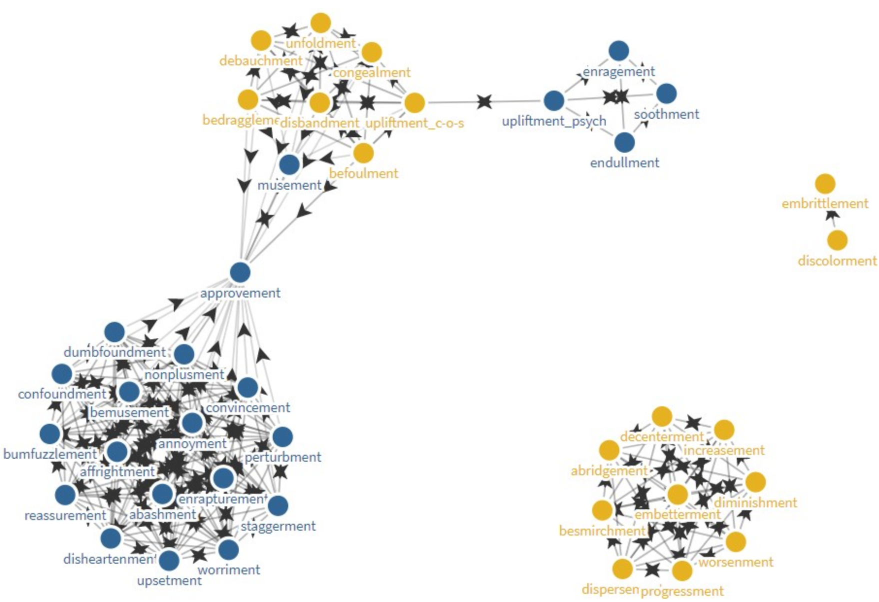


**Fig. 9:** Analogical sets of the *-ment* derivatives. The verb classes of the derivatives are indicated by color (blue = psych verbs, yellow = COS verbs). The source node of an arrow represents an item which is in the analogical set of its target.

The first observation that can be made upon inspecting figure 9 is that the items that AML deems similar form five clouds. This means that five classes of base verb emerge in the model. We can also see that the blue nodes (i.e. the psych items) and the yellow nodes (i.e. the COS items) largely keep to themselves, with few interconnections between blue and yellow clouds. This signifies that, by and large, psych items tend to be like psych items, while COS items tend to be like COS items.

Let us look at individual clouds. The very small cloud with only two yellow nodes at the right edge indicates that the algorithm has used *embrittlement* to classify only one other nominalization, i.e. *discolorment*, and vice versa. The nodes representing the psych items form two clusters, one in the bottom left and one in the top middle. The nodes representing COS items, on the other hand, form three clusters, one in the bottom right, one in the mid-right, and one in the top left. We can thus say that two subclasses of psych verbs and three subclasses of COS verbs emerge from the model.

These subclasses largely match Kawaletz's (2021) semantic classification of base verbs. Kawaletz distinguishes two types of psych causation event, namely those with a caused psych-state and those with a caused change-of-psych-state. The corresponding items can be found in the bottom left cluster and in the top middle cluster, respectively. For the COS verbs, Kawaletz distinguishes whether or not they lexicalize a RESULT. There are three corresponding subclasses of COS verb in her study: those that do not lexicalize a RESULT (here: items in the top left cluster), those that do (items in the bottom right cluster), and those that lexicalize PRODUCT – a hyponym of RESULT (mid-right cluster).

Two of Kawaletz's base verb classes, however, do not emerge in the model. The first is psych verbs that denote a simple psych-state, the second is psych verbs that denote a psych-action. This result is not surprising since these are the classes that have only one member each, *approve of* and *muse over*, respectively. As already discussed in section 5.3.1, this scarcity of information poses a problem for AML. In the figure, we see that the nodes representing *approvement* (between the top left and the bottom left cluster) and *musement* (within the top left cluster) have arrows coming in not only from other psych items, but from COS items as well. This means that, due to the underspecified gang feature constellations we observed earlier, AML has also used COS items to classify *approvement* and *musement*.

Another connection between two clouds is visible at the top of the figure, where we see a two-way arrow between the nodes representing *upliftment_ cos* and *upliftment_ psych*. An inspection of the two derivatives' gangs reveals that this mutual influence between a COS item and a psych item takes place because of the shared polysemous base verb *uplift*: One of their gangs contains the feature 'base,' which is specified for both derivatives as *uplift*. This gang is only marginally influential, but it leads to the two derivatives appearing in each other's analogical sets.

### 5.3.3 Gaps: missing readings

Generally speaking, a gap is a reading that can in principle be produced by *-ment*, but that is not attested for a given derivative. Such gaps are interesting as they can serve as a litmus test for the reliability of the algorithm. There are two kinds of gaps, which we will call 'accidental' and 'systematic' gaps. We will discuss each of these in turn.

Accidental gaps are readings that should be possible for a given derivative, but are not attested. For example, *-ment* can produce CHANGE-OF-STATE readings, and *abridgement* has a *change-of-state* node in its base verb frame. *Abridgement* should therefore be attested in a CHANGE-OF-STATE reading – but it is not (at least not in Kawaletz's data set). Accidental gaps can be attributed to scarcity of data in combination with potential partial blocking effects (see Kawaletz 2021, 168). Therefore, our analogical model would ideally fill in those gaps, predicting all readings that can be expected based on the base verb frame. That is, it should predict CHANGE-OF-STATE for *abridgement.*

Above, we saw for the example of *embrittlement* that accidental gaps in the training data can lead to unwelcome predictions if the analogical set for a given item is very small. Mostly, however, it can be observed that the gaps in the data are ironed out by AML: There are 28 accidental gaps in the lexicon, and AML predicts 26 of these readings. For example, AML does predict CHANGE-OF-STATE for *abridgement* because the reading is attested for all members of the derivative's analogical set.

Let us turn to the systematic gaps. These are readings that should not be possible for a given derivative, and that indeed are not attested. For example, *-ment* can produce STIMULUS readings, but *abridgement* does not have a STIMULUS participant in its base verb frame. *Abridgement* should therefore not be attested in a STIMULUS reading – and it indeed is not. Here, our model should *not* fill in the gaps, predicting only readings that are expected based on the base verb frame, and not predicting those that are not. That is, it should not predict STIMULUS for *abridgement.*

Here, the model also does well. There are 299 combinations of reading and derivative which should not be possible given the base verb frames, and AML correctly predicts that 258 of these should not be possible. For example, STIMULUS is not predicted for *abridgement* because none of the members of the derivative's analogical set has this reading attested.

In sum, the model does well with regard to both accidental and systematic gaps. But why is it not 100 percent successful? There are two reasons for this. First, the training data is not entirely categorical. As described in section 5.3.1, in the data set there are two exceptional combinations of derivative and reading (i.e. [+animate] PATIENT for befoulment and [+animate] STIMULUS for *abashment*). Second, scarcity of data led to two unique feature constellations (i.e. the items representing *approvement* and *musement*, respectively). When these four exceptional items figure in the analogical set of a given derivative, this results in predictions that do not match the attested readings. For example, the item representing *befoulment* in an [+animate] PATIENT reading leads to this reading being predicted for seven derivatives, albeit with very low probability (2.56 percent to 2.7 percent).

Interestingly, these predictions, which at first glance may appear wrong or at least odd, do not present unwelcome results. Rather, they reflect linguistic reality: First, the existence of [+animate] readings for *befoulment* and *abashment* shows that speakers may use other derivatives in this reading as well – this usage is just so unlikely that Kawaletz did not find it attested in the corpora, although it might appear wiht further data. Second, scarcity of data is not something that only the model is confronted with, but speakers as well. Very rare items will therefore also pose a challenge to speakers, the only (but essential) difference being that a speaker can resort to disambiguation by context.

## 5.4 Summary: One form – many meanings

In this study of the polysemy of *-ment*, we have seen that AML does an excellent job in predicting the patterns in the nominalizations' possible readings. First, the semantic classes of base verbs asserted by Kawaletz (2021) emerge in the form of analogical sets. The different sets nicely instantiate the distinction between COS verbs and psych verbs, and even capture the more fine-grained distinctions between different subclasses of COS verbs and psych verbs proposed int he literature, respectively. Second, the readings which AML predicts for the derivatives largely correspond to the predictions that can be made on the basis of the base verb frames. Finally, with regard to gaps in the training data, we have seen that systematic gaps largely remain unpredicted by the algorithm. In contrast, the accidental gaps ('accidental' in terms of the available data in Kawaletz 2021) come out as truly accidental because the algorithm predicts the possible existence of these words. This works correctly, however, only if enough data are available. In general, wrong or uninterpretable results only occur when the training data is insufficient, so that AML does not have enough data points to go by. In such cases, the presence of accidental gaps or of unique base verbs (*approve of* and *muse over* in our data set) leads to out-of-band connections between analogical set clouds, which in turn produce wrong predictions.

## 6 Discussion and conclusion

In this paper, we have investigated two fundamental problems of derivational semantics, affix competition and affix polysemy. It has been the first attempt to implement analogy as mechanism to solve essentially semantic, as against formal, selection problems. We have seen that analogy is indeed able to capture the gradient nature of the pertinent empirical data in an adequate manner. Generalizations emerge as a natural consequence of analogy at different levels of generality, depending on the degree of similarity between lexical entries.

With regard to affix competition, our test case has demonstrated that an analogical algorithm is as good at predicting the right morphological form as regression analysis (in the form of conditional inference trees). The crucial difference between the two is conceptual. While analogy works on the basis of a plausible cognitive mechanism, it is unclear what cognitive correlates might be evoked for conditional inference trees, unless one interprets regression itself as a kind of analogical model. As argued, for example, by Guzmán Naranjo (2020, 225), both models (and indeed many kinds of neural networks) are conceptually very similar because they "capture the same basic intuition: items that are similar belong to the same class". Still, the underlying statistical computations may differ quite a bit, to the effect that the predictions for individual words may also differ, and to the effect that certain factors are more, or less, important across algorithms. A case in point is the influence of the individual base verb, which cannot be meaningfully taken into account in the tree analysis.

With regard to affix polysemy, we have shown that AML can be successfully employed to model the choice between different interpretations for English *-ment* derivatives. The resulting model has many properties that are most welcome at the theoretical and empirical level. The predictions are highly accurate, lexical classes emerge as a by-product of the similarities between words (and need not be stated as indispensable separate entities), and systematic gaps are detected. Even the occasional failures of the system are an indication of the high quality of the model: Too few data lead to insecurities and wrong predictions.

Overall, the two case studies have demonstrated that derivational semantics can be fruitfully analyzed using analogical modeling.

## Bibliography

Albright, Adam. 2009. Modeling analogy as probabilistic grammar. In James P. Blevins & Juliette Blevins (eds.), *Analogy in grammar*, 185–213. Oxford: OUP.

Albright, Adam & Bruce Hayes. 2003. Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90. 119–161.

Alexiadou, Artemis. 2001. *Functional structure in nominals: Nominalization and ergativity*, vol. 42 Linguistik Aktuell/Linguistics Today. Amsterdam and Philadelphia: J. Benjamins.

Alexiadou, Artemis. 2019. Review of R. Lieber, English nouns: The ecology of nominalization. *English Language and Linguistics* 23(3). 735–739.

Arndt-Lappe, Sabine. 2011. Towards an exemplar-based model of stress in English noun–noun compounds. *Journal of Linguistics* 47(11). 549–585.

Arndt-Lappe, Sabine. 2014. Analogy in suffix rivalry: The case of English *-ity* and *-ness*. *English Language and Linguistics* 18(03). 497–548.

Arndt-Lappe, Sabine. 2015. Word-formation and analogy. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation - an international handbook of the languages of Europe*, 822–841. Berlin: de Gruyter Mouton.

Arndt-Lappe, Sabine. 2021. Different lexicons make different rivals. In *Workshop on affix rivalry*, Université de Fribourg.

Arndt-Lappe, Sabine, Ingo Plag, Kai Koch & Mikalai Krot. 2018. Transparent analogical modeling of language (TrAML). Software package.

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

Barsalou, Lawrence W. 1992b. Frames, concepts, and conceptual fields. In Adrienne Lehrer & Eva Feder Kittay (eds.), *Frames, fields and contrasts*, 21–74. Hillsdale: Lawrence Erlbaum Associates.

Bauer, Laurie, Lívia Körtvélyessy & Pavol Štekauer (eds.). 2015. *Semantics of complex words* Studies in morphology. Springer.

Bauer, Laurie, Rochelle Lieber & Ingo Plag. 2013. *The oxford reference guide to English morphology*. Oxford: Oxford University Press.

Baunaz, Lena, Eric Lander, Karen De Clercq & Liliane Haegeman. 2018. Nanosyntax: the basics. In Lena Baunaz, Liliane Haegeman, Karen De Clercq & Eric Lander (eds.), *Exploring nanosyntax*, 3–56. Oxford: Oxford University Press.

Borer, Hagit. 2013. *Structuring sense: volume iii: taking form*, vol. 3. Oxford University Press.

Csardi, Gabor & Tamas Nepusz. 2006. The igraph software package for complex network research. *InterJournal* Complex Systems. 1695. https://igraph.org.

Daelemans, Walter. 2002. A comparison of analogical modeling of language to memory-based language processing. In Royal Skousen, Deryle Lonsdale & Dilworth B. Parkinson (eds.), *Analogical modeling: An exemplar-based approach to language*, 157–179. Amsterdam: John Benjamins.

Daelemans, Walter & Antal van den Bosch. 2005. *Memory-based language processing*. Cambridge: Cambridge University Press.

Daelemans, Walter, Jakub Zavrel, Ko van der Sloot & Antal van den Bosch. 2007. Timbl: Tilburg memory based learner, version 6.0, reference guide. ilk technical report 07-03.

Davies, Mark. 2008-2014. The Corpus of Contemporary American English: 450 million words, 1990-present.

Eddington, David. 2000. Spanish stress assignment within the analogical modeling of language. *Language* 76(1). 92–109.

Eddington, David. 2002. A comparison of two analogical models: Tilburg Memory-Based Learner versus Analogical Modeling. In Royal Skousen, Deryle Lonsdale & Dilworth B.

Parkinson (eds.), *Analogical modeling*, 141–156. Amsterdam / Philadelphia: John Benjamins.

Gadde, Fredrik. 1910. *On the history and use of the suffixes -ery (-ry), -age and -ment in English*. Lund: Berlingska Boktryckeriet.

Guzmán Naranjo, Matías. 2020. Analogy, complexity and predictability in the Russian nominal inflection system. *Morphology* 30(3). 219–262. 10.1007/s11525-020-09367-1.

Harley, Heidi. 2009. Compounding and distributed morphology. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding* Oxford Handbooks in linguistics, 129–144. Oxford, New York: Oxford University Press.

Huyghe, Richard & Marine Wauquier. 2020. What's in an agent? *Morphology* 30(3). 185–218.

Kawaletz, Lea. 2021. *The semantics of English -ment nominalizations*: Heinrich-Heine-Universität Düsseldorf dissertation.

Kawaletz, Lea & Ingo Plag. 2015. Predicting the semantics of English nominalizations: A frame-based analysis of *-ment* suffixation. In Laurie Bauer, Pavol Stekauer & Livia Kortvelyessy (eds.), *Semantics of complex words*, 289–319. Dordrecht: Springer.

Keuleers, E., D. Sandra, W. Daelemans, S. Gillis, G. Durieux & E. Martens. 2007. Dutch plural inflection: The exception that proves the analogy. *Cognitive Psychology* 54. 283–318.

Kipper Schuler, Karin. 2005. *Verbnet: A broad-coverage, comprehensive verb lexicon*. Philadelphia, PA: University of Pennsylvania dissertation.

Lapesa, Gabriella, Lea Kawaletz, Ingo Plag, Marios Andreou, Max Kisselew & Sebastian Padó. 2018. Disambiguation of newly derived nominalizations in context: A distributional semantics approach. *Word Structure* 11(3).

Levin, Beth. 1993. *English verb classes and alternations: A preliminary investigation*. Chicago: University of Chicago Press.

Lieber, Rochelle. 2016. *English nouns: The ecology of nominalization*. Cambridge: Cambridge University Press.

Lieber, Rochelle & Marios Andreou. 2018. Aspect and modality in the interpretation of deverbal *-er* nominals in English. *Morphology* 28(3). 9. 10.1007/s11525-018-9321-7.

Lieber, Rochelle & Ingo Plag. 2021. The semantics of conversion nouns and *-ing* nominalizations: a quantitative and theoretical perspective. *Journal of Linguistics* .

Löbner, Sebastian. 2013. *Understanding semantics*. 2nd edn. London: Routledge.

Marchand, Hans. 1969. *The categories and types of present-day English word-formation*. Munich: Beck.

Marelli, Marco & Marco Baroni. 2015. Affixation in semantic space: Modeling morpheme meanings with compositional distributional semantics. *Psychological review* 122(3). 485.

Missud, Alice & Florence Villoing. 2021. Investigating the distributional properties of rival *-age* suffixation and verb-to-noun conversion in French. *Verbum* XLIII(1).

Nosofsky, Robert M. 1986. Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General* 115. 39–57.

*Oxford English Dictionary Online*. 2021. https://www.oed.com/ (January 21, 2021). Oxford: Oxford University Press.

Plag, I., G. Kunter & S. Lappe. 2007. Testing hypotheses about compound stress assignment in English: A corpus-based investigation. *Corpus Linguistics and Linguistic Theory* 3. 199–232.

Plag, Ingo, Marios Andreou & Lea Kawaletz. 2018. A frame-semantic approach to polysemy in affixation. In Olivier Bonami, Gilles Boyé, Georgette Dal, Hélène Giraudo & Fiammetta Namer (eds.), *The lexeme in descriptive and theoretical morphology*, 546–568. Berlin: Language Science Press.

Pollard, Carl & Ivan A. Sag. 1994. *Head-driven phrase structure grammar*. Chicago: University of Chicago Press.

R Core Team. 2019. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. https://www.R-project.org/.

Sag, Ivan A. 2012. Sign-based construction grammar: An informal synopsis. In Hans Boas & Ivan A. Sag (eds.), *Sign-based construction grammar*, 69–202. Stanford: CSLI publications.

Schirakowski, Barbara. 2021. *Nominalisierte infinitive im spanischen*. Berlin/Boston: De Gruyter.

Skousen, Royal. 1989. *Analogical modeling of language*. Dordrecht: Kluwer Academic.

Skousen, Royal. 2002. Issues in Analogical Modeling. In Royal Skousen, Deryle Lonsdale & Dilworth B. Parkinson (eds.), *Analogical Modeling*, 27–48. Amsterdam / Philadelphia: John Benjamins.

Skousen, Royal. 2005. Analogical Modeling (49). In Reinhard Köhler, Gabriel Altmann & Rajmund G. Piotrowski (eds.), *Quantitative Linguistik.Ein internationales Handbuch = Quantitative linguistics* (Handbücher Zur Sprach- Und Kommunikationswissenschaft = Handbooks of linguistics and communication science = Manuels de linguistique et des sciences de communication / mitbegr. von Gerold Ungeheuer. Hrsg. von Armin Burkhardt ... ; Bd. 27), 705–716. Berlin: de Gruyter. 10.1515/9783110155785.11.705.

Skousen, Royal, Deryle Lonsdale & Dilworth B. Parkinson (eds.). 2002. *Analogical modeling: An exemplar-based aapproach to language*. Amsterdam: John Benjamins.

Skousen, Royal & Thereon Stanford. 2007. *Am:: Parallel*. available from http://humanities.byu.edu/am/. availablefromhttp://humanities.byu.edu/am.

Skousen, Royal, Thereon Stanford & Nathan Glenn. 2013. Algorithm::AM. https://github.com/garfieldnate/Algorithm-AM.

Varvara, Rossella. 2017. *Verbs as nouns: Empirical investigations on event-denoting nominalizations*: PhD dissertation, University of Trento dissertation.

Wauquier, Marine. 2020. *Confrontation des procédés dérivationnels et des catégories sémantiques dans les modèles distributionnels*. Toulouse: Université Toulouse 2 - Jean Jaurès Phd dissertation.