The semantics of derivational morphology:
A frame-based approach

Marios Andreou, Lea Kawaletz & Ingo Plag
The problem

- Affixes (or morphological processes) are frequently semantically underspecified
- Polysemy and meaning extensions of various sorts (Bauer, Lieber & Plag 2013)

Table 1: Readings of English nominalizations (Kawaletz and Plag, 2015)

<table>
<thead>
<tr>
<th>Semantic category</th>
<th>paraphrase</th>
<th>examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Result</td>
<td>‘the outcome of V-ing’</td>
<td>acceptance, alteration</td>
</tr>
<tr>
<td>Product</td>
<td>‘the thing that is created by V-ing’</td>
<td>pavement, growth</td>
</tr>
<tr>
<td>Instrument</td>
<td>‘the thing that V-s’</td>
<td>seasoning, advertisement</td>
</tr>
<tr>
<td>Location</td>
<td>‘the place of V-ing’</td>
<td>dump, residence</td>
</tr>
<tr>
<td>Agent</td>
<td>‘people or person who V-s’</td>
<td>administration, cook</td>
</tr>
<tr>
<td>Measure</td>
<td>‘how much is V-ed’</td>
<td>pinch, deceleration</td>
</tr>
<tr>
<td>Path</td>
<td>‘the direction of V-ing’</td>
<td>decline, direction</td>
</tr>
<tr>
<td>Patient</td>
<td>‘the thing affected or moved by V-ing’</td>
<td>catch, acquisition</td>
</tr>
<tr>
<td>State</td>
<td>‘the state of V-ing or being V-ed’</td>
<td>alienation, disappointment</td>
</tr>
<tr>
<td>Instance</td>
<td>‘an instance of V-ing’</td>
<td>belch, cuddle</td>
</tr>
</tbody>
</table>

- *er nominalizations:
  - fryer Agent, Instrument, Patient (Anderson & Andreou 2018)
Research questions

- Which kinds of readings or meaning extensions are possible and which ones should be impossible for a given derivative?
- What is the role of the bases?
- What is the role of encyclopaedic knowledge?
- How can the semantics of derivational morphology be formally modelled?
State of the art

Lieber’s theory of word-formation semantics (2004)

- Highly restricted set of semantic features ('skeleton')
- Conceptual knowledge representations ('body')

Problems

- Not clear how polysemy of affixes (or derivatives) comes about
Derivation as unification

\[ \text{bounded-event} \]
NOUN-DIM 3
VERB-DIM 3
M-DIM 3
INIT
FIN

\[ \text{change-of-state} \]
MANNER [heat]
ACTOR 1
THEME 2
VERB-DIM 3
M-DIM 3

\[ Na- \]
\[ \text{Gret} \text{‘to heat’} \]

(Zinova 2016, 232-3)
Derivation as unification

\[
\text{bounded-event} \land \text{change-of-state}
\]

MANNER [heat]

ACTOR 1

THEME 2

NOUN-DIM 3

VERB-DIM 3

M-DIM 3

MIN 4

THRESHOLD 5

INIT

DEG 4

FIN

DEG 6

\[
\text{scale} \land \text{temperature}
\]

\[
\text{stage}
\]

\[
5 < 6
\]
Lexical rules

- Model the interaction of the semantics of the morphological process and the semantics of the base.
- Reduce redundancy
- Capture generalizations in the lexicon
  (see among others, Bresnan 1982; Pollard and Sag 1994; Briscoe and Copestake 1999; Sag 2012; Bonami and Crysmann 2016).
Lexical rule for -ment (Plag et al. 2018)

(cf. Pollard and Sag 1994; Sag 2012; Bonami and Crysmann 2016)
Lexical rule for \textit{-ment} (Plag et al. 2018)

(cf. Pollard and Sag 1994; Sag 2012; Bonami and Crysmann 2016)
Lexical rule for -ment (Plag et al. 2018)

\[
\begin{align*}
\text{lexeme} & \\
\text{PHON} & \quad \text{-ment} \\
\text{PHON} & \quad \circ \\
\text{causation event} & \\
\text{ACTOR} & \quad 1 \\
\text{UNDERGOER} & \quad 2 \\
\text{INSTRUMENT} & \quad 3 \\
\text{M-BASE} & \\
\text{SEM} & \quad 0 \\
\text{CAUSE} & \\
\text{EFFECT} & \\
\text{REF} & = \{ 0, 3, 1, 5, 7, 2-7 \}
\end{align*}
\]

(c.f. Pollard and Sag 1994; Sag 2012; Bonami and Crysmann 2016)
Inheritance hierarchy

(c.f. Riehemann 1998; Koenig 1999; Desmets & Villoing 2009; Booij 2010; Tribout 2010; Bonami & Crysmann 2016)
This talk: Overview of our work

- **Introduction**

**Mechanisms**
- Shift of reference through derivation:
  Polysemy of *ment*
- Manipulating attribute values through derivation:
  Stereotype negation

**Computational implementation**
- Implementing derivational polysemy:
  XMG
- Distributional semantic approach to disambiguation
Polysemy in derivation: *-ment*

- Affix polysemy in deverbal nominalization with *-ment*
- Input semantics $\rightarrow$ Output semantics
- Neologisms taken from corpora
- Four input classes based on Levin 1993/VerbNet
  - Change of state verbs
  - Psych verbs
  - Putting verbs
  - Force verbs
The data

- **PATIENT**
  - I set down the scrap of doll’s dress, a **bedragglement** of loose lace hem *(COCA FIC Bk:MournersBench 1999)*

- **PRODUCT**
  - There is an obvious **embrittance** and cracking on the nonwoven fabric (Figure 6.5b). *(GoogleBooks ACAD Cellulose Based Composites 2014)*

- **EVENTUALITY**
  - Hydrides then form and can limit the fuel lifetime due to their **embrittance** of the cladding. *(Google WEB imperial.ac.uk 2014)*
Change-of-state verbs as bases
(causative/inchoative)

- Underspecified first subevent
- Reference: complex event or change of state
- Constraints

There exists at least one participant \( P \) in \( 4 \).
Every \( P \) in \( 4 \) may or may not be co-indexed with \( 1, 2, \) or \( 3 \).
Derived nouns (e.g. *embrittlement*)

There exists at least one participant $p$ in 4.
Every $p$ in 4 may or may not be co-indexed with 1, 2, or 3.

- Reference: range of possible readings
Event type and participants reflect verb class

Reference: only complex event

There exists at least one participant P in 3.
Every P in 3 may or may not be co-indexed with 1 or 2.
Derived nouns (e.g. annoyance)

- Reference: different range of possible readings
  - No EXPERIENCER-IN-RESULT-STATE
  - No clear CHANGE-OF-STATE

There exists at least one participant \( P \) in \( 3 \).
Every \( P \) in \( 3 \) may or may not be co-indexed with \( 1 \) or \( 2 \).
Patterns of possible readings can be detected and explained

... By differences between base verb classes

- **UNDERGOER**: Attested **PATIENT** vs. unattested **EXPERIENCER**
- Differences in **CHANGE-OF-STATE** attestations (not alternating? No c-o-s subevent in the first place?)

... By preferences of –ment

- Animacy constraint (no **AGENT**, no **EXPERIENCER**, no animate **PATIENT**)

... By frame-theoretical considerations

- No **INITIAL-STATE** readings (bidirectional functionality, see e.g. Löbner 2013)
Mechanisms

- Shift of reference
- Manipulation of attribute values
  - Stereotype negation as a test case
COCA SPOK 1994: Dawn Upshaw has been called the “un-diva” of the opera world, often preferring to perform innovative, relatively obscure works that emphasize words over music in an informal style, often—imagine this—even chatting with an audience at recitals.

COCA ACAD 2010: In my writing workshops I often meet the equivalent writing hobbyists. They are people who are writing what I term “coffee-break books,” simpleminded nonbooks that they turn out in short order.
Lexical rule for stereotype negation

```
lexeme
PHON /prefix-[1]/
CAT N

SEM

IND i

S-FRAME [2]!

REF

ATTRIBUTE_j

\[ -\alpha \]
Lexical rule for *non-*

```
[lexeme
  PHON /nanbük/
  CAT N
  SEM
    INDI
    S-FRAME
      CONTENT complex
      REF i
      ...
]
```

```
[M-BASE
  lexeme
  PHON /bük/
  CAT N
  SEM
    INDI
    S-FRAME
      CONTENT complex
      REF i
      ...
]
```
- Highlights the importance of structured information of perceptual aspects of meaning (i.e. functional attributes that assign properties to referents and the values these attributes take).

- Shows that the “absence” of a characteristic of the base lexeme can be treated as a change in the value of an attribute of the base lexeme and not as absence of the attribute (i.e. the general characteristic) itself.
Computational implementations

- XMG (in collaboration with Simon Petitjean)
- Distributional semantics (in collaboration with CRC Stuttgart)
Andreou and Petitjean (2017, forthcoming) used corpus extracted data to

- Identify the range of readings of -al derivatives (e.g. *rental*) and
- Identify prominent constraints on the types of situations and entities -al targets (e.g. animacy).

Constraints are given in 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.
Inheritance hierarchy
Frame to XMG

causation event
AGENT 1
PATIENT 2
INSTRUMENT 3

cause
AGENT 1
PATIENT 2
INSTRUMENT 3

activity
AGENT 1
PATIENT 2
INSTRUMENT 3

change-of-state
INITIAL STATE 6
PATIENT 2

RESULT STATE 7
PATIENT 2

class renew
export ?X0

{<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: ?X2],
       result-state: ?X7[result_state, patient: ?X2]]
}

}
class al_nominal
import rent[

declare ?Ref
{
  <frame>{
    [al-lexeme,
     m-base:[event,
        sem:?X0]
     ref:?Ref
    ]
  }
  ;
  ?X0 >* ?Ref;
}

?X0 >* ?Ref;
{ ?Ref[result_state] | ?Ref[causation]
  | ?Ref[entity, animacy:[inanimate]] }
The introduction of constraints into the semantics of an affix allows one to **predict** and **generate**

- Readings which are possible for a given derivative and,
- **Rule out** those readings which are not possible.

As a **proof of concept**, the output resulting of the XMG description was consistent with the range of readings observed in the corpus.
Distributional Semantics

- Problem: Disambiguating newly derived words in context
- Can a Distributional Semantics model do the job?
- Co-occurrence vectors: a toy example

<table>
<thead>
<tr>
<th>Context</th>
<th>Target</th>
<th>t-shirt</th>
<th>tie</th>
<th>lawyer</th>
<th>judge</th>
</tr>
</thead>
<tbody>
<tr>
<td>wear</td>
<td></td>
<td>9</td>
<td>7</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>law</td>
<td></td>
<td>1</td>
<td>3</td>
<td>7</td>
<td>9</td>
</tr>
</tbody>
</table>
Distributional Semantics

Disambiguation:

suit$_1$ ‘process in a law court’
suit$_2$ ‘ensemble of matching garments’
Method

- Low frequency deverbal -ment nominalizations (55 types, 406 tokens)
- Manually annotated: eventive, non-eventive, ambiguous
- Disambiguation by comparing
  - vectors of training nouns
  - with vectors of nominalizations in their context

TRAINING NOUNS

- *accident* → EVENT
- *bike* → NOT EVENT

ANOYMENT.1 : ?
“Such an annoymnt already happened”

ANOYMENT.2 : ?
“Fix this annoymnt!”
### Results

<table>
<thead>
<tr>
<th>Predictions</th>
<th>Annotations</th>
<th>eventive</th>
<th>non-eventive</th>
<th>ambiguous</th>
</tr>
</thead>
<tbody>
<tr>
<td>eventive</td>
<td></td>
<td>0.78</td>
<td>0.47</td>
<td>0.73</td>
</tr>
<tr>
<td>non-eventive</td>
<td></td>
<td>0.22</td>
<td>0.53</td>
<td>0.27</td>
</tr>
</tbody>
</table>
Results

![Graph showing probability distributions for eventive, ambiguous, and non-eventive interpretations.]

- eventive
- ambiguous
- non-eventive
Non-eventive nominalizations

![Box plot showing probability of event interpretation for object-abstract and object-concrete categories.]
Summary disambiguation

- It is possible to use distributional semantics to disambiguate the meaning even of newly formed words.
- This demonstrates the usefulness of the context in disambiguation.
- A window with two content words on each side suffices to make good predictions.
- Non-eventive derivatives are hard to classify as such.
- Non-eventive **abstract nouns** and **eventive nouns** are not only similar in their semantic properties, they may also occur in the same contexts. Both facts make disambiguation of such nouns a hard task.
- There are quite a few cases (15%) in which the interpretation of new words remains unclear, even for humans.
A framework for the analysis of derivational semantics
Analyses of individual morphological categories
Computational implementations

The future

Analyses of more categories
Eventive interpretations without verbal bases
Scalar interpretations
Computational modeling (XMG and Analogy)
Thank you very much for your attention!
a) Peer-reviewed publications and books


b) Other publications


Andreou, Marios. 2018. A frame-based approach to evaluative morphology. Submitted to Glossa
