Same syntax, different semantics: A compositional approach to idiomaticity in multi-word expressions

Timm Lichte & Laura Kallmeyer
University of Düsseldorf, Germany

CSSP, Paris, October 8–10, 2015
Multi-word expressions (MWEs) with literal and idiomatic meanings:

(1) \textit{John spilled the beans}.
    literal meaning: ‘John spilled the beans.’
    idiomatic meaning: ‘John revealed one or more secrets.’

(2) \textit{John kicked the bucket}.
    literal meaning: ‘John kicked the bucket.’
    idiomatic meaning: ‘John died.’
Introduction

- literal vs. idiomatic readings
- syntactic ambiguity: non-compositional
- semantic ambiguity:
  - lexicon-/disjunction-based: compositional
  - inference-based: non-compositional

⇒ How to model them with precision grammars?
⇒ What sort of ambiguity should be preferred?
⇒ One approach for all types of MWEs?

Target framework: LTAG + frame semantics

Preceding this work:
Lichte & Kallmeyer (2014; 2015)
Lichte & Kallmeyer (Düsseldorf)
introduction

- literal vs. idiomatic readings
- syntactic ambiguity (non-compositional)
- semantic ambiguity
  - lexicon-/disjunction-based (non-compositional)
  - inference-based (compositional)

⇒ How to model them with precision grammars?
⇒ What sort of ambiguity should be preferred?
⇒ One approach for all types of MWEs?
Introduction

literal vs. idiomatic readings

syntactic ambiguity

non-compositional

semantic ambiguity

lexicon-/disjunction-based

compositional

inference-based

non-compositional

⇒ How to model them with precision grammars?
⇒ What sort of ambiguity should be preferred?
⇒ One approach for all types of MWEs?

target framework: LTAG + frame semantics

preceding this work: Lichte & Kallmeyer (2014; 2015)
Outline

1. Tree-Adjoining Grammar + frame semantics
2. Former work
   - Syntactic ambiguity approaches with TAG
   - Semantic ambiguity approaches
3. New: Semantic ambiguity approach with TAG
4. Summary
Outline

1. Tree-Adjoining Grammar + frame semantics

2. Former work
   - Syntactic ambiguity approaches with TAG
   - Semantic ambiguity approaches

3. New: Semantic ambiguity approach with TAG

4. Summary
Tree-Adjoining Grammar (TAG)\textsuperscript{[2,16,17]}

- A grammar consists of elementary trees.
- Elementary trees can be combined by two operations:
  - **substitution**: replace a non-terminal leaf with an initial tree

\[
\begin{array}{c}
\text{NP} \\
\text{NP} \\
\text{N} \\
\text{Peter} \\
S \\
\text{NP} \\
\text{VP} \\
\text{V} \\
\text{repaired} \\
\Rightarrow \\
\text{NP} \\
\text{NP} \\
\text{V} \\
\text{NP} \\
\text{N} \\
\text{N} \\
\text{V} \\
\text{NP} \\
	ext{repaired} \\
\text{Peter} \\
\end{array}
\]

TAG is more powerful than CFG, but still less powerful than LFG, HPSG, TG.

Elementary trees cover an extended domain of locality.

The head immediately combines with its arguments.

no predetermined derivational order ⇒ constructionist framework!\textsuperscript{[14]}

Lexical generalizations are expressed in the metagrammar.
Tree-Adjoining Grammar (TAG)\textsuperscript{[2,16,17]}

- A grammar consists of elementary trees.
- Elementary trees can be combined by two operations:
  - substitution: replace a non-terminal leaf with an initial tree
  - adjunction: replace an inner node with an auxiliary tree

```
<table>
<thead>
<tr>
<th>VP</th>
<th>VP*</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td></td>
</tr>
<tr>
<td>easily</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S</th>
<th>VP</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>VP</td>
</tr>
<tr>
<td>V</td>
<td>NP</td>
</tr>
<tr>
<td>repaired</td>
<td></td>
</tr>
</tbody>
</table>

⇒

<table>
<thead>
<tr>
<th>S</th>
<th>VP</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>VP</td>
</tr>
<tr>
<td>AP</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td></td>
</tr>
<tr>
<td>easily</td>
<td>repaired</td>
</tr>
</tbody>
</table>
```

TAG is more powerful than CFG, but still less powerful than LFG, HPSG, TG.

Elementary trees cover an extended domain of locality.

The head immediately combines with its arguments.

no predetermined derivational order

⇒

constructionist framework!

[14]

Lexical generalizations are expressed in the metagrammar.
Tree-Adjoining Grammar (TAG)\cite{2,16,17}

- A grammar consists of elementary trees.
- Elementary trees can be combined by two operations:
  - substitution: replace a non-terminal leaf with an initial tree
  - adjunction: replace an inner node with an auxiliary tree
- TAG is more powerful than CFG, but still less powerful than LFG, HPSG, TG.
Tree-Adjoining Grammar (TAG)\textsuperscript{[2,16,17]}

- A grammar consists of elementary trees.
- Elementary trees can be combined by two operations:
  - substitution: replace a non-terminal leaf with an initial tree
  - adjunction: replace an inner node with an auxiliary tree
- TAG is more powerful than CFG, but still less powerful than LFG, HPSC, TG.
- Elementary trees cover an extended domain of locality.
  - The head immediately combines with its arguments.
  - no predetermined derivational order
  ⇒ constructionist framework!\textsuperscript{[14]}

$$S \rightarrow NP \rightarrow V \rightarrow NP \rightarrow \text{repaired}$$
Tree-Adjoining Grammar (TAG)\textsuperscript{[2,16,17]}

- A grammar consists of elementary trees.
- Elementary trees can be combined by two operations:
  - substitution: replace a non-terminal leaf with an initial tree
  - adjunction: replace an inner node with an auxiliary tree
- TAG is more powerful than CFG, but still less powerful than LFG, HPSG, TG.
- Elementary trees cover an extended domain of locality.
  - The head immediately combines with its arguments.
  - no predetermined derivational order
    \Rightarrow constructionist framework!\textsuperscript{[14]}
- Lexical generalizations are expressed in the metagrammar.
Frames emerged as a representation format of lexical and conceptual knowledge. [6,12,22]

Frames, FrameNet frames [26]

Frame semantics with quantification: see Kallmeyer, Osswald, Pogodalla (this conference)
Frames emerged as a representation format of lexical and conceptual knowledge.\[6,12,22]\]

Frames can be formalized as (extended) typed feature structures.\[18,27]\]
Frames emerged as a representation format of lexical and conceptual knowledge.\[6,12,22\]

Frames, FrameNet frames \[26\]

Frames can be formalized as (extended) typed feature structures.\[18,27\]

Frames ≠ FrameNet frames \[26\]
Frames emerged as a representation format of lexical and conceptual knowledge.\cite{6,12,22}

Frames can be formalized as (extended) typed feature structures.\cite{18,27}

Frames ≠ FrameNet frames\cite{26}

Frame semantics with quantification: see Kallmeyer, Osswald, Pogodalla (this conference)
Kallmeyer & Osswald [18]:

- **lexicon**: pairs of elementary trees and frames

Elementary trees are enriched with **interface features**, which contain base labels from the frame representation.

- unification of interface features $\sim$ unification of frames
- parallel composition of derived trees and larger frames
TAG + frame semantics: Example
Outline

1. Tree-Adjoining Grammar + frame semantics

2. Former work
   - Syntactic ambiguity approaches with TAG
   - Semantic ambiguity approaches

3. New: Semantic ambiguity approach with TAG

4. Summary
Syntactic ambiguity approaches with TAG
(idea from Abeillé & Schabes)$^{[1,3,4]}$

**Idiomaticity through multiple anchoring:** Components of an MWE jointly anchor an elementary tree.

$$S_{[E = 0]}$$

$$NP_{[I = 1]}$$
$$VP_{[E = 0]}$$

$$V$$
$$N_{[E = 0]}$$

$kicked$
$$D$$

$\begin{bmatrix}
\text{dying} \\
\text{PATIENT}
\end{bmatrix}$

$\begin{bmatrix}
\text{the} \\
\text{bucket}
\end{bmatrix}$
Syntactic ambiguity approaches with TAG

(idea from Abeillé & Schabes)

The literal meaning is evoked by regular single-anchored elementary trees:

\[
\begin{align*}
S_{[E = 0]} & \quad \text{NP}_{[I = 1]} \quad \text{VP}_{[E = 0]} \\
\text{V} & \quad \text{NP}_{[I = 2]} \\
& \quad \text{NP}_{[I = 3]} \\
& \quad \text{bucket}
\end{align*}
\]

\[
\begin{bmatrix}
\text{kicking} \\
0 & \text{ACTOR} & 1 \\
0 & \text{PATIENT} & 2 \\
3 & \text{container} \\
\end{bmatrix}
\]
Syntactic ambiguity approaches with TAG

Example with “decomposable” *spill the beans*:

```
S[\(E = 0\)]
    / \  
   /   \  
NP[\(I = 1\)] VP[\(E = 0\)]
     / \  
V    NP[\(I = 2\)]
       / \  
spilled N[\(I = 2\)]
       |   
beans
```

```
[divulging
  \[0\] \[ACTOR \ 1\]
  \[1\] \[THEME \ 2\]\[information\]]
```

idiomatic
Syntactic ambiguity approaches with TAG

Example with “decomposable” *spill the beans*:

```
S[E = 0]
   NP[I = 1]
   VP[E = 0]
      V
      NP[I = 2]
         spilled
      NP[I = 3]
         beans

[spilling 0]
[spilled 1]
[ACCTOR 2]
[PATIENT 3]
```

Literal
Syntactic ambiguity approaches elsewhere

Syntactic ambiguity approach

There are different syntactic derivations/representations for literal and idiomatic meanings.

Also found in:[29]

- Transformational Grammar (Chomsky 1980)
- Lexical-functional Grammar (Bresnan 1982)
- Head-driven Phrase Structure Grammar (Sailer 2000)[30,33]
- Sign-based Construction Grammar (Kay & Sag To appear)
Syntactic ambiguity approaches elsewhere

Syntactic ambiguity approach

There are different syntactic derivations/representations for literal and idiomatic meanings.

Also found in:[29]

- Transformational Grammar (Chomsky 1980)
- Lexical-functional Grammar (Bresnan 1982)
- Head-driven Phrase Structure Grammar (Sailer 2000)[30,33]
- Sign-based Construction Grammar (Kay & Sag To appear)

But there are (general?) problems …
Syntactic ambiguity approaches: Problems

- bad for parsing: non-delayable ambiguity resolution
bad for parsing: non-delayable ambiguity resolution

missing compatibility with psycholinguistic results (Müller & Wechsler): MWEs cause an increased semantic rather than syntactic processing load. [28,34,35]
Syntactic ambiguity approaches: Problems

- bad for parsing: non-delayable ambiguity resolution
- missing compatibility with psycholinguistic results (Müller & Wechsler): MWEs cause an increased semantic rather than syntactic processing load.\[^{28,34,35}\]
- missing connection between literal and idiomatic meaning
Syntactic ambiguity approaches: Problems

- bad for parsing: non-delayable ambiguity resolution
- missing compatibility with psycholinguistic results (Müller & Wechsler): MWEs cause an increased semantic rather than syntactic processing load.\[^{[28,34,35]}\]
- missing connection between literal and idiomatic meaning
- missing account of the “extendability” of literal senses (Egan):

\[^{[28,34,35]}\]
Syntactic ambiguity approaches: Problems

- bad for parsing: non-delayable ambiguity resolution
- missing compatibility with psycholinguistic results (Müller & Wechsler): MWEs cause an increased semantic rather than syntactic processing load. [28,34,35]
- missing connection between literal and idiomatic meaning
- missing account of the “extendability” of literal senses (Egan):
  
  (3) *If you let this cat out of the bag, a lot of people are going to get scratched.*
Syntactic ambiguity approaches: Problems

- bad for parsing: non-delayable ambiguity resolution
- missing compatibility with psycholinguistic results (Müller & Weckssler): MWEs cause an increased semantic rather than syntactic processing load.\(^{28,34,35}\)
- missing connection between literal and idiomatic meaning
- missing account of the “extendability” of literal senses (Egan):
  (3) *If you let this cat out of the bag, a lot of people are going to get scratched.*
- missing generalizations on lexical variability (Pulman):
  \{put/lay/spread\} the cards on the table
  \{let the cat / the cat is\} out of the bag
Syntactic ambiguity approaches: Problems

- bad for parsing: non-delayable ambiguity resolution
- missing compatibility with psycholinguistic results (Müller & Wechsler): MWEs cause an increased semantic rather than syntactic processing load.\,[28,34,35]
- missing connection between literal and idiomatic meaning
- missing account of the “extendability” of literal senses (Egan):
  (3) *If you let this cat out of the bag, a lot of people are going to get scratched.*
- missing generalizations on lexical variability (Pulman):
  \{*put/lay/spread*\} *the cards on the table*
  \{*let the cat / the cat is*\} *out of the bag*
- difficult to deal with partial uses:
Syntactic ambiguity approaches: Problems

- bad for parsing: non-delayable ambiguity resolution
- missing compatibility with psycholinguistic results (Müller & Wechsler): MWEs cause an increased semantic rather than syntactic processing load.\[28,34,35\]
- missing connection between literal and idiomatic meaning
- missing account of the “extendability” of literal senses (Egan):
  (3) *If you let this cat out of the bag, a lot of people are going to get scratched.*
- missing generalizations on lexical variability (Pulman):
  \{*put/lay/spread*\} *the cards on the table*
  \{*let the cat / the cat is*\} out of the bag
- difficult to deal with partial uses:
  (4) *Eventually she spilled all the beans. But it took her a few days to spill them all.* (Riehemann)
Syntactic ambiguity approaches: Problems

- bad for parsing: non-delayable ambiguity resolution
- missing compatibility with psycholinguistic results (Müller & Wechsler): MWEs cause an increased semantic rather than syntactic processing load.\textsuperscript{[28,34,35]}
- missing connection between literal and idiomatic meaning
- missing account of the “extendability” of literal senses (Egan):
  (3) \textit{If you let this cat out of the bag, a lot of people are going to get scratched.}
- missing generalizations on lexical variability (Pulman):
  \{\textit{put/lay/spread} the cards on the table \}
  \{\textit{let the cat / the cat is} out of the bag \}
- difficult to deal with partial uses:
  (4) \textit{Eventually she spilled all the beans. But it took her a few days to spill them all.} (Riehemann)
  (5) \textit{Pat pulled some strings for Chris. But Alex didn’t have access to any strings.} (Manfred Sailer, pc)
Semantic ambiguity approaches

There is one syntactic derivation/representation for literal and idiomatic meanings.

⇒ There is no special lexical entry for MWEs; *kick* and *spill* each have only one lexical entry.

![Diagram]

- semantic ambiguity
  - lexicon-/disjunction-based
  - inference-based
    - compositional
    - non-compositional
Components of decomposable MWEs are assigned disjunctions over meaning constants (of intensional logic):

(6) a. \( \text{spill} \rightarrow \text{spill’} \lor \text{spill-idiom’} \)
    \( \text{beans} \rightarrow \text{beans’} \lor \text{beans-idiom’} \)

b. spill-idiom’ (beans-idiom’): defined
    spill-idiom’ (beans’): undefined
    spill’ (beans-idiom’): undefined
Components of decomposable MWEs are assigned disjunctions over meaning constants (of intensional logic):

(6) a. $spill \sim spill' \lor spill$-idiom'  
    $beans \sim beans' \lor beans$-idiom'  

b. spill-idiom' (beans-idiom'): defined  
    spill-idiom' (beans'): undefined  
    spill' (beans-idiom'): undefined

Also applicable to non-decomposable idioms (not in Gazdar et al. 1985):

(7) a. $kick \sim kick' \lor kick$-idiom'  
    $bucket \sim bucket' \lor bucket$-idiom'  

b. kick-idiom' (bucket-idiom'): defined  
    kick-idiom' (bucket'): undefined  
    kick' (bucket-idiom'): undefined
Lexicon-/disjunction-based: Gazdar et al. (1985)

Advantages of Gazdar et al.’s partial function approach:
- unified syntax of literal and idiomatic readings

Drawbacks:
Lexicon-/disjunction-based: Gazdar et al. (1985)

Advantages of Gazdar et al.’s partial function approach:

- unified syntax of literal and idiomatic readings
- delayable ambiguity resolution

Drawbacks:

Lichte & Kallmeyer (Düsseldorf)
Advantages of Gazdar et al.’s partial function approach:

- unified syntax of literal and idiomatic readings
- delayable ambiguity resolution
- adequate in terms of human processing

Drawbacks:
Advantages of Gazdar et al.’s partial function approach:

- unified syntax of literal and idiomatic readings
- delayable ambiguity resolution
- adequate in terms of human processing

(Prediction: increased semantic processing load; no categorical difference between lexical and idiomatic meanings)

Drawbacks:
Advantages of Gazdar et al.’s partial function approach:

- unified syntax of literal and idiomatic readings
- delayable ambiguity resolution
- adequate in terms of human processing
  (Prediction: increased semantic processing load; no categorical difference between lexical and idiomatic meanings)
- closer connection between literal and idiomatic meanings

Drawbacks:
Advantages of Gazdar et al.’s partial function approach:

- unified syntax of literal and idiomatic readings
- delayable ambiguity resolution
- adequate in terms of human processing
  (Prediction: increased semantic processing load; no categorical difference between lexical and idiomatic meanings)
- closer connection between literal and idiomatic meanings

Drawbacks:

- invention of masses of meaning constants that essentially reflect morphological properties
Advantages of Gazdar et al.’s partial function approach:

- unified syntax of literal and idiomatic readings
- delayable ambiguity resolution
- adequate in terms of human processing
  (Prediction: increased semantic processing load; no categorical difference between lexical and idiomatic meanings)
- closer connection between literal and idiomatic meanings

Drawbacks:

- invention of masses of meaning constants that essentially reflect morphological properties
- partial functions have to be defined explicitly
The idiomatic meaning is deduced from the literal one by means of “quasi-inference”. Hence MWE-components are equipped with their literal meaning only!

(8) kick’(x,y) ∧ bucket’(y) ≈ die’(x)
The idiomatic meaning is deduced from the literal one by means of “quasi-inference”. Hence MWE-components are equipped with their literal meaning only!

\[(8) \text{kick}'(x,y) \land \text{bucket}'(y) \approx \text{die}'(x)\]

**Drawbacks of Pulman’s quasi-inference approach:**

- poorly constrained surface: *The bucket was kicked.*
The idiomatic meaning is deduced from the literal one by means of “quasi-inference”. Hence MWE-components are equipped with their literal meaning only!

\[(8) \text{kick}'(x,y) \land \text{bucket}'(y) \approx \text{die}'(x)\]

**Drawbacks of Pulman’s quasi-inference approach:**

- poorly constrained surface: *The bucket was kicked.*
  - Pulman: due to information structure!
The idiomatic meaning is deduced from the literal one by means of “quasi-inference”. Hence MWE-components are equipped with their literal meaning only!

\[(8) \text{kick}'(x,y) \land \text{bucket}'(y) \approx \text{die}'(x)\]

**Drawbacks of Pulman’s quasi-inference approach:**

- poorly constrained surface: *The bucket was kicked.*
  
  ⇒ Pulman: due to information structure!
  
  *(The bucket will be kicked. (Manfred Sailer))*
The idiomatic meaning is deduced from the literal one by means of “quasi-inference”. Hence MWE-components are equipped with their literal meaning only!

\[(8) \text{kick}'(x,y) \land \text{bucket}'(y) \approx \text{die}'(x)\]

**Drawbacks of Pulman’s quasi-inference approach:**

- poorly constrained surface: *The bucket was kicked.*
  
  \[\Rightarrow\] Pulman: due to information structure!
  
  *(The bucket will be kicked. (Manfred Sailer))*

- MWEs with bounded/cranberry words: leave sb. in the lurch
The idiomatic meaning is deduced from the literal one by means of “quasi-inference”. Hence MWE-components are equipped with their literal meaning only!

\[(8) \quad \text{kick}'(x, y) \land \text{bucket}'(y) \approx \text{die}'(x)\]

**Drawbacks of Pulman’s quasi-inference approach:**

- poorly constrained surface: *The bucket was kicked.*
  
  ⇒ Pulman: due to information structure!
  
  *(The bucket will be kicked. (Manfred Sailer))*

- MWEs with bounded/cranberry words: *leave sb. in the lurch*

- MWEs with ill-formed syntax: *trip the light fantastic*
The idiomatic meaning is deduced from the literal one by means of “quasi-inference”. Hence MWE-components are equipped with their literal meaning only!

(8) kick′(x,y) \land bucket′(y) \approx die′(x)

**Drawbacks of Pulman’s quasi-inference approach:**

- poorly constrained surface: *The bucket was kicked.*
  
  ⇒ Pulman: due to information structure!
  
  *(The bucket will be kicked.* (Manfred Sailer))

- MWEs with bounded/cranberry words: *leave sb. in the lurch*

- MWEs with ill-formed syntax: *trip the light fantastic*

- computationally very powerful: non-monotonic inference rules.
Outline

1. Tree-Adjoining Grammar + frame semantics

2. Former work
   - Syntactic ambiguity approaches with TAG
   - Semantic ambiguity approaches

3. New: Semantic ambiguity approach with TAG

4. Summary
Main problem of Gazdar et al. (1985): tons of extra meaning constants; partial functions have to be defined explicitly.
Main problem of Gazdar et al. (1985): tons of extra meaning constants; partial functions have to be defined explicitly.

Our proposal: decompose meaning constants + constraint-based composition!

\[ \text{kick-idiom'} \sim \begin{bmatrix} \text{FRAME} \\ \text{MORPH} \end{bmatrix} = \begin{bmatrix} \text{dying} \\ \text{PATIENT} \end{bmatrix} \begin{bmatrix} 1 \\ \text{LEMMA} \end{bmatrix} \text{kick} \]
A lexicon-/disjunction-based approach with TAG

Main problem of Gazdar et al. (1985): tons of extra meaning constants; partial functions have to be defined explicitly.

Our proposal: decompose meaning constants + constraint-based composition!

\[
\text{kick-idiom'} \sim \left[ \begin{array}{c}
\text{FRAME} \\
\text{MORPH}
\end{array} \right]
\left[ \begin{array}{c}
dying \\
patient \ [1] \\
lemma \ kick
\end{array} \right]
\]

\[
\text{bucket-idiom'} \sim \left[ \begin{array}{c}
\text{FRAME} \\
\text{MORPH}
\end{array} \right]
\left[ \begin{array}{c}
dying \\
lemma \ bucket \\
def \ + \\
num \ sing
\end{array} \right]
\]
Main problem of Gazdar et al. (1985): tons of extra meaning constants; partial functions have to be defined explicitly.

Our proposal: decompose meaning constants + constraint-based composition!

\[
\begin{align*}
\text{kick-idiom'} & \leadsto \\
& \begin{bmatrix}
\text{FRAME} & \begin{bmatrix}
\text{dying} \\
\text{PATIENT} & 1
\end{bmatrix} \\
\text{MORPH} & \begin{bmatrix}
\text{LEMMA} & \text{kick}
\end{bmatrix}
\end{bmatrix} \\
\text{bucket-idiom'} & \leadsto \\
& \begin{bmatrix}
\text{FRAME} & \begin{bmatrix}
\text{dying}
\end{bmatrix} \\
\text{MORPH} & \begin{bmatrix}
\text{DEF} & + \\
\text{NUM} & \text{sing}
\end{bmatrix}
\end{bmatrix}
\end{align*}
\]

⇒ How to combine those two?
A lexicon-/disjunction-based approach with TAG

Lichte & Kallmeyer (Düsseldorf)
A lexicon-/disjunction-based approach with TAG
A lexicon-/disjunction-based approach with TAG

S_{[E = 0]}

NP_{[I = 1]}

VP_{[E = 0]}

V

kicked

NP_{[I = 2], [E = 0]}

FRAMES:

0

FRAME
ACTOR
kicking

MORPH
LEMMAs:
kick

NP_{[I = 3], [E = 4]}

NP_{[I = 3]}

FRAME
[container]
MORPH
[LEMMAs: bucket]

FRAME
[dying]
MORPH
[LEMMAs: bucket]

FRAME
[dying]
MORPH
[LEMMAs: kick]

NP_{[I = 3]}

MORPH
[LEMMAs: bucket]

N_{[I = 3]}

FRAME
[container]
MORPH
[LEMMAs: bucket]

FRAME
[dying]
MORPH
[LEMMAs: bucket]

FRAME
[dying]
MORPH
[LEMMAs: kick]
A lexicon-/disjunction-based approach with TAG

Lichte & Kallmeyer (Düsseldorf)
A lexicon-/disjunction-based approach with TAG

kicked-idiom' (bucket-idiom')
A lexicon-/disjunction-based approach with TAG

\[ S_{E = 0} \]

\[ \text{NP}[I = 1] \quad \text{VP}[E = 0] \]

\[ V \quad \text{NP}[I = 2, E = 0] \]

\[ \text{NP}[I = 3, E = 4] \quad \text{N}[I = 3] \quad \text{bucket} \]

\[ \text{kicked} \]

\[ \text{kicked-idiom'(bucket')} \]

\[ \text{VP}_{E = 0} \]

\[ \text{FRAME} \quad \text{ACTOR} \quad \text{LEMMAN kick} \]

\[ \text{FRAME} \quad \text{PATIENT} \quad \text{NUM sing} \]

\[ \text{FRAME} \quad \text{PATIENT} \quad \text{LEMMAN kick} \]

\[ \text{FRAME} \quad \text{DEF} + \text{NUM sing} \]

\[ \text{FRAME} \quad \text{dying} \quad \text{LEMMAN bucket} \]

\[ \text{FRAME} \quad \text{dying} \quad \text{LEMMAN kick} \]
A lexicon-/disjunction-based approach with TAG

NP[I = 1]

VP[E = 0]

V

NP[I = 2, E = 0]

kicked

NP[I = 3, E = 4]

bucket

VP

FRAME

FRAME

FRAME

MORPH

MORPH

MORPH

MORPH

kicked’(bucket-idiom’)

NP

V

VP

MORPH

Lichte & Kallmeyer (Düsseldorf)
A lexicon-/disjunction-based approach with TAG

Result of combining *kicked* and *bucket*:

```
S_{E = 0}
   NP_{I = 1}   VP_{E = 0}
      V          NP_{I = 3, E = 0}
         kicked   NP_{I = 3, E = 0}
            bucket

V_{E = 0}
   FRAME
      ACTOR
          0
         PATIENT
          3
        LEMMA
            kicking
            MORPH
              LEMMA
                kick

VP_{E = 0}
   FRAME
      ACTOR
          0
         PATIENT
          1
        LEMMA
            dying
            MORPH
              LEMMA
                kick

V_{E = 0}
   FRAME
      ACTOR
          0
         PATIENT
          3
        LEMMA
            bucket
            MORPH
              DEF
                +
       NUM
           sing
```

Lichte & Kallmeyer (Düsseldorf)
Bargmann’s challenge

Here is a challenge from Bargmann (2015):

(9) *The whole idea of the really talented/successful person in their 20s isn’t a real thing. Or at the very least, it isn’t an actual attainable thing. All those people have people behind them pulling string after string for them.*
Bargmann’s challenge

Here is a challenge from Bargmann (2015):

(9) The whole idea of the really talented/successful person in their 20s isn’t a real thing. Or at the very least, it isn’t an actual attainable thing. All those people have people behind them pulling string after string for them.

- pull combines with a plurality of strings (pull a string).
- string after string is syntactically singular, but semantically plural (Matsuyama, Jackendoff).
Here is a challenge from Bargmann (2015):

(9) *The whole idea of the really talented/successful person in their 20s isn’t a real thing. Or at the very least, it isn’t an actual attainable thing. All those people have people behind them pulling string after string for them.*

- *pull* combines with a plurality of *strings* (??*pull a string*).
- *string after string* is syntactically singular, but semantically plural (Matsuyama, Jackendoff).

⇒ Analyses with purely morpho-syntactic constraints fail.
Bargmann’s challenge

Here is a challenge from Bargmann (2015):

(9) *The whole idea of the really talented/successful person in their 20s isn’t a real thing. Or at the very least, it isn’t an actual attainable thing. All those people have people behind them **pulling string after string** for them.*

- *pull* combines with a plurality of *strings* (?*pull a string*).
- *string after string* is syntactically singular, but semantically plural (Matsuyama, Jackendoff).

⇒ Analyses with purely morpho-syntactic constraints fail.
⇒ We need some intermediate level between surface and pure semantics to capture the constraints on *pull strings*!
Here is a challenge from Bargmann (2015):

(9) The whole idea of the really talented/successful person in their 20s isn’t a real thing. Or at the very least, it isn’t an actual attainable thing. All those people have people behind them pulling string after string for them.

- pull combines with a plurality of strings (pull a string).
- string after string is syntactically singular, but semantically plural (Matsuyama, Jackendoff).

⇒ Analyses with purely morpho-syntactic constraints fail.
⇒ We need some intermediate level between surface and pure semantics to capture the constraints on pull strings!
Here is a challenge from Bargmann (2015):

(9) The whole idea of the really talented/successful person in their 20s isn’t a real thing. Or at the very least, it isn’t an actual attainable thing. All those people have people behind them **pulling string after string** for them.

- **pull** combines with a plurality of **strings** (? pull a string).
- **string after string** is syntactically singular, but semantically plural (Matsuyama, Jackendoff).

⇒ Analyses with purely morpho-syntactic constraints fail.
⇒ We need some intermediate level between surface and pure semantics to capture the constraints on **pull strings**!

Working with HPSG, Bargmann proposes a “Semantic Representation approach”:

- idiom constants pull′_id and string′_id have to co-occur
- string′_id is in the scope of a “non-specific plural quantifier” (Mel’čuk)
Bargmann’s challenge: Analysis with TAG

\[
S_{[E = 0]}
\]

\[
NP_{[I = 1]}
\]

\[
VP_{[E = 0]}
\]

\[
V
pull
\]

\[
NP_{[I = 2, E = 0]}
\]

\[
\begin{array}{l}
\text{FRAME} \\
\text{ACTOR} \\
\text{INSTR} \\
\text{LEMMAS} \\
\text{NUM} \\
\text{LEMMA} \\
\text{string} \\
\text{pull} \\
\text{pl}
\end{array}
\]

\[
\lor \ldots
\]
Bargmann’s challenge: Analysis with TAG

Lichte & Kallmeyer (Düsseldorf)
Bargmann’s challenge: Analysis with TAG

S[\text{E} = 0]

NP[I = 1]

VP[\text{E} = 0]

V

NP[I = 1, \text{E} = 0, \text{sg}=+]

COORD

NP[I = 1, \text{E} = 0, \text{det}=-, \text{sg}=+]

after

[\text{assistance-activity}]

FRAME 0

MORPH 2

[\text{pull}]

[\text{actor} 1]

[\text{instr} 2]

FRAME 4

MORPH 5

[\text{string}]

FRAME 4

MORPH 5

[\text{lemma}]

[\text{num} pl]

Lichte & Kallmeyer (Düsseldorf)
Bargmann’s challenge: Analysis with TAG

```
S[E = 0]
  NP[I = 0]
    V
    pull
    NP[I = 0, E = 0, det=-, sg=+]
      COORD
        after
        NP[I = 3, E = 0, det=-, sg=+]
          NP[I = 3, det=-, sg=+]
            COORD
              a/f_ter
                NP[I = 3, det=-, sg=+]
                  string
                    string
                      string
                        string
                          string
                            string
                                string
                                    string
                                        string
                                            string
                                                string
                                                    string
                                                        string
                                                            string
                                                                string
                                                                    string
                                                                        string
                                                                            string
                                                                                string
                                                                                    string
                                                                                        string
                                                                                            string
                                                                                                string
                                                                                                      string
                                                                                                            string
                                                                                                                string
                                                                                                                    string
                                                                                                                        string
                                                                                                                            string
                                                                                                                                string
                                                                                                                                    string
                                                                                                                                        string
                                                                                                                                            string
                                                                                                                                                string
                                                                                                                                                    string
                                                                                                                                                           string
                                                                                                                                                    \ldots
```

```
NP[I = 0, E = 0]
  N[I = 0]
    string
      string
        string
          string
            string
              string
                  string
                      string
                          string
                              string
                                  string
                                      string
                                          string
                                              string
                                                  string
                                                      string
                                                          string
                                                              string
                                                                  string
                                                                      string
                                                                              string
                                                                                      string
                                                                                                      string
                                                                                                            string
                                                                                                                string
                                                                                                                    string
                                                                                                                        string
                                                                                                                            string
                                                                                                                                    string
                                                                                                                                        string
                                                                                                                                                string
                                                                                                                                                    string
                                                                                                                                                            string
                                                                                                                                                           string
                                                                                                                                                    \ldots
```

```
VP[E = 0]
  V
    pull
    NP[I = 2, E = 0, sg=+]
```

```
assistance-activity
  FRAME [actor 1]
  INSTR [frame 8]
  LEMMA [pull]
```

```
string
  FRAME [num sg 1]
  LEMMA [string 5]
  LEMMA [string 5]
  LEMMA [string 5]
  LEMMA [string 5]
  LEMMA [string 5]
  LEMMA [string 5]
```

Lichte & Kallmeyer (Düsseldorf)
Bargmann’s challenge: Analysis with TAG

S[E = 0]

NP[I = 1]

VP[E = 0]

V

pull

NP[I = 2] E = 0, det=-, sg=+]

NP[I = 3]

after

NP[I = 3, det=-, sg=+]

N[I = 3]

string

COORD

assistance-activity

FRAME 0

ACTOR 1

FRAME 2

INSTR 2

FRAME 4

LEMMA pull

FRAME 4

LEMMA string

FRAME 4

LEMMA string

V ...

Lichte & Kallmeyer (Düsseldorf)
Advantages:

- unified syntax of literal and idiomatic readings
- delayable ambiguity resolution
- adequate in terms of human processing
  (Prediction: increased semantic processing load; no categorical difference between lexical and idiomatic meanings)
- closer connection between literal and idiomatic meanings

+ constraint-based composition
Outline

1. Tree-Adjoining Grammar + frame semantics

2. Former work
   - Syntactic ambiguity approaches with TAG
   - Semantic ambiguity approaches

3. New: Semantic ambiguity approach with TAG

4. Summary
The landscape of approaches to idiomatic MWEs from a TAG perspective:

- literal vs. idiomatic readings
- syntactic ambiguity: non-compositional
- semantic ambiguity: compositional vs. non-compositional
- lexicon-/disjunction-based vs. inference-based

⇒ One approach for all types of MWEs?
⇒ Connection between literal and idiomatic meaning?
⇒ Multi-dimensional approach following Ernst (1981)?


