VP idioms in Norwegian: A subconstructional approach

Petter Haugereid

Department of Linguistic, Literary and Aesthetic Studies,
University of Bergen

HPSG 2014
Buffalo
August 27–29, 2014
Outline

1. Introduction
3. Analysis
   - Lexical representation
   - Phrasal subconstructions
   - Analysis of VP idioms
4. Discussion and future work
NorGram (Dyvik, 2000)

- Norwegian LFG grammar
- 56 VP idioms in the lexicon
- distributed over 20 templates
- Four main kinds of idioms
Four kinds of idioms

- Two intransitive with only one argument, the subject.
  - In the first an object is selected:
    
    (1) Han **gikk** konkurs.
    
    he went bankrupt
    
    *He went bankrupt.*

  - In the second a PP is selected:
    
    (2) De **løftet i** flokk.
    
    They lifted in flock
    
    *They worked together.*
The last two kinds of idioms are transitive, hence they take two arguments.

- In one an **object** and the **preposition** of a PP are selected:

  (3) **Han la ikke skjul på sin glede.**
  he laid not hiding on his joy
  *He didn’t hide his joy.*

- In the other a **PP** is selected:

  (4) **Han brakte temaet på bane.**
  he brought topic.the on track
  *He brought up the topic.*
LFG representations

- VP idiom frames are listed among the other frames of the verbs.
- A lexical entry is allowed to have more than one argument frame.
  → Disjunctions of frames in the lexical entries:
    - @(V-SUBJ-POBJrefl-OBJ bringe med)
    - @(V-SUBJ-PRT-OBJ bringe inn)
    - @(V-SUBJ-OBJ-OBJ bringe)
    - @(V-SUBJ-OBJ-OBLBEN bringe)
    - @(V-SUBJ-OBJ bringe)
    - @(VPIDIOM-PSELOBJ-OBJ bringe på bane)

- Disjunctions are expanded into full lexical entries during parsing.
  → 6 disjunctive argument frames computationally equals six lexical entries.
Outline

1. Introduction


3. Analysis
   - Lexical representation
   - Phrasal subconstructions
   - Analysis of VP idioms

4. Discussion and future work
Treatment of idioms in Sag et al. (2003)

- Special lexical entries for words that constitute an idiom.
- The idiomatic verb *keep* in *keep tabs on* has three items on the SUBCAT list:
  - the NP subject
  - an idiomatic noun *tabs*
  - a constituent marked by the preposition *on*
In HPSG, the lexicon is very constrained
→ Separate lexical entries for idiomatic verbs
Motivated from a semantic point of view
- Idiomatic meaning deviates from the compositional meaning
But no morphological motivation
- Stem identical to compositional version
- Inflectional pattern identical to compositional version
Outline

1. Introduction

3. Analysis
   - Lexical representation
   - Phrasal subconstructions
   - Analysis of VP idioms

4. Discussion and future work
Among the frames of the verb *bringe* are a transitive frame and a ditransitive frame:

(5)  

a. Han brakte maten.  
   he brought food.the
   *He brought the food.*

b. Han brakte henne maten.  
   he brought her food.the
   *He brought her the food.*
## Lexical entry

### Lexical entry for *bringe* ‘bring’

Only one lexical entry is assumed:

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>STEM</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bringe-v</td>
<td>&quot;bringe”</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HEAD</strong></td>
<td>verb</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>VAL</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-ARG1</td>
<td>HEAD noun</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-ARG2</td>
<td>HEAD noun</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-ARG3</td>
<td>HEAD noun</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-ARG4</td>
<td>HEAD compl-noun</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>KEYREL</strong></td>
<td>1</td>
<td>PRED bringe</td>
<td></td>
</tr>
<tr>
<td><strong>RELS</strong></td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Four valence features

- **C-ARG1**
  - → external subject
- **C-ARG2**
  - → (deep) direct object
- **C-ARG3**
  - → (deep) indirect obj.
- **C-ARG4**
  - → oblique object
The position of *bringe* in the *link* type hierarchy
Phrasal subconstructions

- Rule for linking (external) subjects:

  \[\arg1-struct\]
  \[\text{VAL} \begin{bmatrix} \text{C-ARG1} | \text{LINK} \ arg1- \end{bmatrix}\]
  \[\text{KEYREL} \begin{bmatrix} \text{ARG1} \end{bmatrix}\]

  \[\text{ARGS} \begin{bmatrix} \text{VAL} \begin{bmatrix} \text{C-ARG1} \begin{bmatrix} \text{LINK} \ arg1+ \end{bmatrix}\end{bmatrix}, \text{INDEX} \begin{bmatrix} \text{ARG1} \end{bmatrix}\end{bmatrix}\]

- The grammar also has subconstructions that in the same fashion link:
  - (deep) direct objects \((\arg2-struct)\)
  - (deep) indirect objects \((\arg3-struct)\)
  - oblique objects \((\arg4-struct)\)
The **LINK** values have two functions:

1. Keeping track of arguments that are realized
   - Like empty/non-empty valence lists in HPSG

2. Narrowing down what kind of construction the clause has
   - Each subconstruction that applies in a clause leaves a trace
     - either a **LINK** value is switched from negative to positive
     - or a selected item leaves a mark by unifying its **FORM** value with the respective **LINK** value (later ...)
   - At the bottom of the tree, all this information is present
The grammar has a rule that adds the verb:

\[
(6) \begin{align*}
\text{vbl-struct} & \\
\text{VBL} & \quad 1 \\
\text{KEYREL} & \quad 2 \\
\text{ARGS} & \langle \left[ \begin{align*}
\text{VBL} & \quad 3 \\
\text{KEYREL} & \quad 2 \\
\text{verb-word} & \end{align*} \right],
\left[ \begin{align*}
\text{VBL} & \quad 1 \\
\text{KEYREL} & \quad 2 
\end{align*} \right] \rangle 
\end{align*}
\]
Analysis of a transitive sentence
Analysis of VP idioms

Linking types in the *START* sign

```
START
C-ARG1|LINK  arg1+
C-ARG2|LINK  arg2+
C-ARG3|LINK  arg3–
C-ARG4|LINK  arg4–
KEYREL|PRED  bringe
```

Unification of linking types

```
uni-link
C-ARG1|LINK   1
C-ARG2|LINK   1
C-ARG3|LINK   1
C-ARG4|LINK   1
KEYREL|PRED   1
```
Subconstruction types for transitive *bringe*
Analysis of VP idioms

The analysis of idioms includes special idiom words:

- one type for selected prepositions:
  \[
  \begin{array}{c}
  \text{idiom-prep-word} \\
  \text{HEAD} \quad \text{prep} \\
  \text{FORM} \quad \text{link} \\
  \text{RELS} \quad \langle \text{!!} \rangle
  \end{array}
  \]

- one type for idiomatic nouns:
  \[
  \begin{array}{c}
  \text{idiom-noun-word} \\
  \text{HEAD} \quad \text{noun} \\
  \text{FORM} \quad \text{link} \\
  \text{RELS} \quad \langle \text{!!} \rangle
  \end{array}
  \]
The analysis also includes three subconstructions:

- one for selected prepositions (*prepsel-struct*)
  \[
  \text{prepsel-struct} \\
  \text{ARGS} \left( \left[ V \mid C-\text{ARG4} \mid \text{LINK} \right], \left[ \text{idiom-prep-word} \right] \right)
  \]

- two subconstructions for idiomatic nouns:
  - *arg2-idiom-struct*
  - *arg4-idiom-struct*

  \[
  \text{arg2-idiom-struct} \\
  V \mid C-\text{ARG2} \mid \text{LINK} \ arg2- \\
  \text{ARGS} \left( \left[ V \mid C-\text{ARG2} \mid \text{LINK} \right], \left[ \text{idiom-noun-word} \right] \right)
  \]
Analysis of VP idioms

- **arg4-idiom-str**
  - C-ARG1
  - LINK arg1–
  - C-ARG2
  - LINK arg2–
  - C-ARG3
  - LINK arg3–
  - C-ARG4
  - LINK arg4–

- **prepsel-str**
  - C-ARG1
  - LINK arg1–
  - C-ARG2
  - LINK arg2–
  - C-ARG3
  - LINK arg3–
  - C-ARG4
  - LINK arg4–

- **arg2-str**
  - C-ARG1
  - LINK arg1–
  - C-ARG2
  - LINK arg2–
  - C-ARG3
  - LINK arg3–
  - C-ARG4
  - LINK arg4–

- **arg1-str**
  - C-ARG1
  - LINK arg1–
  - C-ARG2
  - LINK arg2–
  - C-ARG3
  - LINK arg3–
  - C-ARG4
  - LINK arg4–

- **vbl-str**
  - C-ARG1
  - LINK arg1–
  - C-ARG2
  - LINK arg2–
  - C-ARG3
  - LINK arg3–
  - C-ARG4
  - LINK arg4–

- **verb-word**
  - KEYREL
  - ARG1

- **PREP**
  - KEYREL
  - ARG1

- **noun-word**
  - FORM
  - bane

- **start**
  - C-ARG1
    - LINK arg1–
    - C-ARG2
      - LINK arg2–
      - C-ARG3
        - LINK arg3–
        - C-ARG4
          - LINK arg4–

- **brakte**
  - PRED
  - rels
  - brake

- **han**
  - NP
  - temaet

- **bringe**
  - NP
  - han

- **på**
  - prep-word
  - FORM
  - bane

- **bane**
  - noun-word
  - FORM
  - bane
Analysis of VP idioms

```
[START
C-ARG1|LINK arg1+
C-ARG2|LINK arg2+
C-ARG3|LINK arg3–
C-ARG4|LINK 4på*bane
[ PRED 0bringe
KEYREL 7 ARG1 8
ARG2 9 ]
```
Idiom subconstruction types

Idiom subconstruction types:
- \( \text{arg}_3^+ \text{arg}_4^- \)
- \( \text{arg}_1^+ \text{arg}_2^+ \)
- \( \text{arg}_3^- \text{arg}_4^+ \)

- \( \text{bringe} \)
- \( \text{på*bane} \)

- \( \text{bringe}_123_{\text{rel}} \)
- \( \text{bringe}_12_{\text{rel}} \)
- \( \text{bringe*på*bane}_{\text{rel}} \)
VP idiom MRS (Copestake et al. 2005)

\[
\begin{align*}
\text{LTOP} & \quad h_1 \\
\text{INDEX} & \quad e_2 \{ \text{pres yes-no-question} \} \\
\text{RELS} & \quad \begin{cases} 
  h_3: \text{pron}._\text{rel}(x_4) \\
  h_5: \text{pronoun}._\text{q}._\text{rel}(x_4,h_6,h_7) \\
  h_8: \text{bringe}^\ast \text{på}^\ast \text{bane}._\text{rel}(e_2,x_4,x_9) \\
  h_{10}: \text{tema}._\text{n}._\text{rel}(x_9) \\
  h_{11}: \text{def}._\text{q}._\text{rel}(x_9,h_{12},h_{13}) 
\end{cases} \\
\text{HCONS} & \quad \begin{cases} 
  h_6 =q h_3 \\
  h_{12} =q h_{10} 
\end{cases}
\end{align*}
\]
### VP idioms – combinations of subconstructions

<table>
<thead>
<tr>
<th>Intransitive with idiomatic noun</th>
<th>Transitive with idiomatic noun</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>vbl-struct</strong></td>
<td><strong>vbl-struct</strong></td>
</tr>
<tr>
<td><strong>arg1-struct</strong></td>
<td><strong>arg1-struct</strong></td>
</tr>
<tr>
<td><strong>arg2-idiom-struct</strong></td>
<td><strong>arg2-idiom-struct</strong></td>
</tr>
<tr>
<td><strong>prepsel-struct</strong></td>
<td><strong>prepsel-struct</strong></td>
</tr>
<tr>
<td><strong>arg4-struct</strong></td>
<td><strong>arg4-struct</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Intransitive with idiomatic PP</th>
<th>Transitive with idiomatic PP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>vbl-struct</strong></td>
<td><strong>vbl-struct</strong></td>
</tr>
<tr>
<td><strong>arg1-struct</strong></td>
<td><strong>arg1-struct</strong></td>
</tr>
<tr>
<td><strong>prepsel-struct</strong></td>
<td><strong>prepsel-struct</strong></td>
</tr>
<tr>
<td><strong>arg2-idiom-struct</strong></td>
<td><strong>arg2-struct</strong></td>
</tr>
<tr>
<td><strong>arg4-idiom-struct</strong></td>
<td><strong>arg4-idiom-struct</strong></td>
</tr>
</tbody>
</table>
Outline

1 Introduction

2 Treatment of idioms in Sag et al. (2003)

3 Analysis
   - Lexical representation
   - Phrasal subconstructions
   - Analysis of VP idioms

4 Discussion and future work
The hierarchy of linking types that results from this account is huge, but finite.

This kind of hierarchy is interesting in that it reflects what kinds of phrasal subconstructions are needed in order to express all grammaticalized concepts in a given grammar.

The grammar presented in this paper, so far only has a small hand-built type hierarchy.

The aim is to generate a full hierarchy from the lexicon of the LFG grammar NorGram.
