Multi-word Expressions in HPSG

Dan Flickinger
Center for the Study of Language and Information
Stanford University
danf@stanford.edu

PARSEME Training School
Charles University, Prague
January 2015
Overview for the week

• Day One
  Brief introduction to Head-driven Phrase Structure Grammar
  Implementation in the English Resource Grammar (ERG)
  Meaning representation in Minimal Recursion Semantics

• Day Two
  Classification of Multi-word Expressions (MWEs)
  Implementation of MWEs in the ERG
  Strengths and weaknesses of the approach

• Day Three
  Case study of one class of MWEs: idioms with possessives
  Interactions with other linguistic phenomena and processing
  Disambiguation challenges

• Day Four
  Lab session using the ERG to identify and analyse MWEs
Possessive Idioms

Joint work with Francis Bond, Jia Qian, and Christiane Fellbaum

- One constituent contains a possessive pronoun co-indexed with a different constituent (typically the subject)
- *wrack one’s brains*: “think hard”
- *He wracked his brains. She wracked her brains.*
- *She wracked his brains. “She made him think hard”*
Motivation

- Machine translation, to parse or generate near-equivalent idioms Japanese-English, where often no possessive in Japanese idiom
- Error correction in language learning
  Immediate use in existing commercial online ELA course
  10,000 primary school students composing paragraphs
  Automatic error analysis with the ERG
- Corpus research to find instances and measure frequency of use
Collection of possessive idioms

- Consulted WordNet and online dictionaries, and corpus observation
- 324 expressions classified so far
- 290 involve locally controlled possession
  - 20 types cover all but 23 of these
- 34 have externally controlled possession
  - 6 types cover all but 9 of these
Internally controlled possessive idiom patterns

<table>
<thead>
<tr>
<th>Structure</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>XNP V1 X’s N1</td>
<td>118</td>
</tr>
<tr>
<td>XNP V1+P1 X’s N1</td>
<td>12</td>
</tr>
<tr>
<td>XNP V1 [PP P1 X’s N1]</td>
<td>29</td>
</tr>
<tr>
<td>XNP V1 X’s N1 P1 X</td>
<td>2</td>
</tr>
<tr>
<td>XNP V1 X’s N1 [PP P1 YNP]</td>
<td>30</td>
</tr>
<tr>
<td>XNP V1 X’s N1 [PP P1 D1 N2]</td>
<td>13</td>
</tr>
<tr>
<td>XNP V1 X’s N1 [PP P1 X’s N2]</td>
<td>6</td>
</tr>
<tr>
<td>XNP V1 X’s N1 A1</td>
<td>23</td>
</tr>
<tr>
<td>XNP V1 X’s N1 Adj1</td>
<td>3</td>
</tr>
<tr>
<td>XNP V1 X’s N1 and V2 N2</td>
<td>2</td>
</tr>
<tr>
<td>XNP V1 X’s own N1</td>
<td>3</td>
</tr>
<tr>
<td>XNP V1 N1 [PP P1 X’s N2]</td>
<td>3</td>
</tr>
<tr>
<td>XNP V1 D1 N1 [PP P1 X’s N2]</td>
<td>2</td>
</tr>
<tr>
<td>XNP V1 YNP [PP P1 X’s N1]</td>
<td>4</td>
</tr>
<tr>
<td>XNP V1 YNP D1 N1 [PP P1 X’s N2]</td>
<td>2</td>
</tr>
<tr>
<td>XNP aux+neg V1 X’s N1</td>
<td>3</td>
</tr>
<tr>
<td>XNP aux+neg V1 X’s N1 [PP P1 YNP]</td>
<td>3</td>
</tr>
<tr>
<td>XNP be [PP P1 X’s N1]</td>
<td>4</td>
</tr>
<tr>
<td>XNP be Adj1 [PP P1 X’s N1]</td>
<td>3</td>
</tr>
<tr>
<td>XNP be A1 P1 X’s N1 P2 YNP</td>
<td>2</td>
</tr>
</tbody>
</table>
Externally controlled possessive idiom patterns

XNP V1 YNP’s N1 9
XNP V1 [PP P1 YNP’s N1] 6
XNP V1 YNP’s N1 P1 3
XNP V1 ZNP [PP P1 YNP’s N2] 2
XNP V1 YNP D1 N1 [PP P1 Y’s N2] 2
XNP V1 D1 N1 [PP P1 YNP’s N2] 3
Some examples of possessed idioms

- *She bit her tongue* and avoided insulting her guest.
  
  XNP V1 X’s N1

- *He has already made up his mind* about that topic.
  
  XNP V1+P1 X’s N1

- *They finally came to their senses.*
  
  XNP V1 [PP P1 X’s N1]

- *She is going to pull her hair out* in frustration.
  
  XNP V1 X’s N1 A1

- *You always seem to have your head in the clouds.*
  
  XNP V1 X’s N1 [PP P1 D1 N2]

- *They just can not manage to get their head around that idea.*
  
  XNP aux+neg V1 X’s N1 [PP P1 YNP]
An example: Dependencies from MRS

She wracked her brains.

{e3:
  x5:pron<0:2>[]
  e3:_wrack_v_i<3:10>[ARG1 x5, ARG2 x6]
  i8:id<3:10>[ARG1 x5, ARG2 x7]
  _1:def_explicit_q<11:14>[BV x6]
  e13:poss<11:14>[ARG1 x6, ARG2 x7]
  x7:pron<11:14>[]
  x6:_brain_n_1<15:22>[]
}
An example: Dependency MRS (DMRS)

She wracked her brains.

don pron _wrack_v_i id def_explicit_q poss pron _brain_n_1

ARG1/NEQ /EQ

ARG1/NEQ

ARG2/NEQ

ARG1/EQ

RSTR/H

ARG2/NEQ

ARG2/NEQ

ARG2/NEQ

ARG1/NEQ

ARG2/NEQ

ARG2/NEQ

ARG2/NEQ
Implementation: The idiom lexicon in the ERG

She wracked her brains.

\[
\text{wrack+brains} := \text{v\_reflnp\_idiom\_mtr} \& \\
\begin{array} {l}
\text{[ INPUT.RELS.LIST } < \text{ [ PRED "\_wrack\_v\_i\_rel" ],} \\
\text{ [ PRED "\_brain\_n\_1\_rel" ], }\ldots > ].
\end{array}
\]
Possessive idiom type definition

\[
\text{v\_reflnp\_idiom\_mtr} := \text{monotonic\_mtr} \ & \ [ \ \text{INPUT\_RELS} < [ \ \text{ARG1 \ #verb-arg1}, \\
\text{ARG2 \ #verb-arg2 } ], \\
[ \ \text{ARG0 \ #verb-arg2 } ], \\
[ \ \text{PRED id\_rel}, \\
\text{ARG1 \ #verb-arg1}, \\
\text{ARG2 \ #poss-arg2 } ], \\
[ \ \text{PRED poss\_rel}, \\
\text{ARG1 \ #verb-arg2}, \\
\text{ARG2 \ #poss-arg2 } ] > ].
\]
Possessive idiom rule, instantiated

\[
v_{\text{reflnp idioms}} := \text{monotonic mtr} \land \\
[\text{INPUT.RELS} < [ \text{PRED } \text{"_w rack_v_i_re l"}], \\
\text{ARG1 } \#\text{verb-arg1}, \\
\text{ARG2 } \#\text{verb-arg2}], \\
[\text{PRED } \text{"_brain_n_1_re l"}], \\
\text{ARG0 } \#\text{verb-arg2}], \\
[\text{PRED } \text{id_re l}], \\
\text{ARG1 } \#\text{verb-arg1}, \\
\text{ARG2 } \#\text{poss-arg2}], \\
[\text{PRED } \text{poss_re l}], \\
\text{ARG1 } \#\text{verb-arg2}, \\
\text{ARG2 } \#\text{poss-arg2}] > ].
\]
Lexical type for idiomatic verb

```
wrack_v1_i := v_np_refl-idm_le &
[ ORTH < "wrack" >,
  SEMPREDE "_wrack_v_i_rel" ].

v_np_refl-idm_le := np_nontrans_verb &
[ CAT.VAL.COMPS < [ LOCAL.CONT.HOOK.XARG #arg2 ] >,
  CONT [ HOOK.XARG #arg1,
    RELS < [ PRED #pred ],
      [ PRED id_rel,
        ARG1 #arg1,
        ARG2 #arg2 ] > ],
  SEMPREDE #pred,
  IDIOM + ].
```
Lexical type for possessive pronouns

her_pos := d_--poss_le &
[ ORTH < "her" >,
    AGR [ PERNUM 3sing, GENDER fem ] ].

d_--poss_le := det_word &
[ CONT [ HOOK [ INDEX #arg1,
            XARG #arg2 ],
    RELS < [ PRED def_explicit_q_rel,
             ARG0 #arg1 ],
             [ PRED poss_rel,
             ARG1 #arg1,
             ARG2 #arg2 ],
             [ PRED pron_rel
             ARG0 #arg2 ] > ] ]
Some variants of the same idiom

She wracked her brains.
She wracked her brain.
She wracked that brain of hers.
She wracked those brains of hers.
Some variants of the same idiom

She wracked her brains.
She wracked her *brain*.
She wracked that brain *of hers*.
She wracked those brains *of hers*.
Some variants of the same idiom

She wracked her brains.
She wracked her brain.
She wracked that brain of hers.
She wracked those brains of hers.
She **racked** her brains.
She racked her brains.
She racked her brain.
She racked that brain of hers.
She racked those brains of hers.
Some variants of the same idiom

She wracked her brains.
She wracked her brain.
She wracked that brain of hers.
She wracked those brains of hers.
She racked her brains.
She racked her brain.
She racked that brain of hers.
She racked those brains of hers.
She wracked her big brain.
She wracked that big brain of hers.
Another example of the same type

We'll wait our turn. ("We will wait until it is our turn.")

{e3:
  x5:pron<0:2>[]
  e3:_wait_v_i<8:12>[ARG1 x5, ARG2 x6]
  i8:id<8:12>[ARG1 x5, ARG2 x7]
  _1:def_explicit_q<13:16>[BV x6]
  e13:poss<13:16>[ARG1 x6, ARG2 x7]
  x7:pron<13:16>[]
  x6:_turn_n_of<17:22>[
}

wait+turn := v_reflnp_idiom_mtr &
  [ INPUT.RELS.LIST < [ PRED "_wait_v_i_rel" ],
    [ PRED "_turn_n_of_rel" ], ... > ].
A syntactically more complex example

You always seem to have your head in the clouds.
“You are not paying attention to the immediate situation.”
A second type of idiom rule

You always seem to have your head in the clouds.

\[
\text{have+head+in+clouds := v_reflnp-pp_idiom_mtr} \& \\
[ \text{INPUT.RELS < [ PRED "_have_v_prd_rel" ]}, \\
[ \text{PRED "_head_n_of_rel" ]}, \\
[ \text{PRED _in_p_rel }], \\
[ \text{PRED "_clouds_n_i_rel" ], ... > ]].
\]
The second idiom rule type, instantiated

You always seem to have your head in the clouds.

\[
\text{v_reflnp-pp_idiom_mtr} := \text{monotonic_mtr} \land \\
[ \text{INPUT.RELS} < ] \quad \begin{array}{l}
\text{PRED "have_v_prd_rel" }, \\
\text{ARG1 \#verb-arg1,} \\
\text{ARG2 \#verb-arg2, ARG3 \#verb-arg3 }, \\
\text{PRED "head_n_of_rel"}, \\
\text{ARG0 \#verb-arg2 }, \\
\text{PRED in_p_rel,} \\
\text{LBL \#verb-arg3,} \\
\text{ARG1 \#verb-arg2, ARG2 \#prep-arg2 }, \\
\text{PRED "clouds_n_i_rel"}, \\
\text{ARG0 \#prep-arg2 }, \\
\text{PRED id_rela,} \\
\text{ARG1 \#verb-arg1,} \\
\text{ARG2 \#poss-arg2 }, \\
\text{PRED poss_rela,} \\
\text{ARG1 \#verb-arg2,} \\
\text{ARG2 \#poss-arg2 } > ] .
\end{array}
\]
Interactions with other phenomena

• Unbounded dependencies

   Those large brains of yours, you definitely ought to wrack immediately.
Interactions with other phenomena

- Unbounded dependencies
  Those large brains of yours, you definitely ought to wrack immediately.
- Coordination
  The students and the teachers should all rack their brains.
Interactions with other phenomena

- Unbounded dependencies
  Those large brains of yours, you definitely ought to wrack immediately.
- Coordination
  The students and the teachers should all rack their brains.
- Imperatives
  Wrack your brains!
Interactions with other phenomena

- Unbounded dependencies
  
  *Those large brains of yours, you definitely ought to wrack immediately.*

- Coordination
  
  *The students and the teachers should all rack their brains.*

- Imperatives
  
  *Wrack your brains!*

- Modification
  
  *You should wrack as quickly as possible those excellent brains of yours.*
Interactions with other phenomena

- Unbounded dependencies
  Those large brains of yours, you definitely ought to wrack immediately.

- Coordination
  The students and the teachers should all rack their brains.

- Imperatives
  Wrack your brains!

- Modification
  You should wrack as quickly as possible those excellent brains of yours.

- Subordinate clauses
  I think you should at least try to wrack your brains.
  Those very large brains of yours, I definitely think you ought to try to wrack as quickly as possible.