Input determinants of L2 construction learning

Yevgen Matusevych
Afra Alishahi
Ad Backus
Tilburg University, the Netherlands

ATiLA-2014, Ghent
Input determinants of construction learning

Yevgen Matusevych
Afra Alishahi
Ad Backus
Tilburg University, the Netherlands

ATiLA-2014, Ghent
Overview

- Computational model of bilingual construction learning
- Distributional properties of input → construction learning
- Which properties are really important?
Constructions

- Constructions are pairings of form and meaning [Langacker, 2008]:
  - a *jar*, a *jar lid*, sat
  - a *N*, a *N N*, V-ed
Constructions

- Constructions are pairings of form and meaning [Langacker, 2008]:
  - a jar, a jar lid, sat
  - a N, a N N, V-ed

- Argument structure constructions is a subclass of these, which represents clauses [Goldberg, 1995]:
  - ditransitive transfer: Agent causes Patient to receive Theme
    A1 V A2 A3
    John sent me a package.
  - intransitive motion: Theme moves to Location
    A1 V A2
    The bottle floated into the cave.
Learning argument structure constructions
Learning argument structure constructions

Yevgen Matusevych
Learning argument structure constructions
Learning argument structure constructions

The bear gives you the ball!
Learning argument structure constructions

The bear gives you the ball
The bear gives you the ball

Daddy's coming home!
Learning argument structure constructions

The bear gives you the ball

Daddy's coming home
Learning argument structure constructions

Grandma sent you some cookies.

John passed you the ball!

Mr. Rich donated us a thousand dollars.

The bear gives you the ball

Daddy's coming home
Learning argument structure constructions

Grandma sent you some cookies

The bear gives you the ball

Mr. Rich donated us a thousand dollars

John passed you the ball

Daddy's coming home
Learning argument structure constructions

<table>
<thead>
<tr>
<th>Predicate meaning</th>
<th>cause to receive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of arguments</td>
<td>3</td>
</tr>
<tr>
<td>Word order</td>
<td>X verb Y Z</td>
</tr>
<tr>
<td>Argument meanings</td>
<td>{human}; {human}; {object}</td>
</tr>
<tr>
<td>Argument roles</td>
<td>Giver; Recipient; Theme</td>
</tr>
</tbody>
</table>

Grandma sent you some cookies

The bear gives you the ball

Mr. Rich donated us a thousand dollars

John passed you the ball

Daddy's coming home
Learning argument structure constructions

Dittransitive transfer construction

Daddy's coming home
Learning argument structure constructions: L2

- Ditransitive transfer construction
- Resultative construction
- ...

Yevgen Matusevych
Learning argument structure constructions: L2

- Ditransitive transfer construction
- Resultative construction
- ...

Meine Schwester lieh mir Geld.
(My sister lent me some money.)
Learning argument structure constructions: L2

Ditransitive transfer construction

Resultative construction

Meine Schwester lieh mir Geld. (My sister lent me some money.)
Das Geld gab ich meiner Mutter.
(I gave the money to my mother.)
Learning argument structure constructions: L2

Ditransitive transfer construction

Resultative construction

... Das Geld gab ich meiner Mutter

Das Geld gab ich meiner Mutter. (I gave the money to my mother.)
Learning argument structure constructions: L2
Learning argument structure constructions

- Based on a probabilistic model of early argument structure learning
  [Alishahi & Stevenson, 2008]

- Simulates the process of learning constructions in two languages
Evaluating language knowledge

<table>
<thead>
<tr>
<th>Predicate</th>
<th>give</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicate meaning</td>
<td>cause to receive</td>
</tr>
<tr>
<td>Number of arguments</td>
<td>3</td>
</tr>
<tr>
<td>Word order</td>
<td>X verb Y Z</td>
</tr>
<tr>
<td>Argument meanings</td>
<td>{human}; {human}; {object}</td>
</tr>
<tr>
<td>Argument roles</td>
<td>Giver; Recipient; Theme</td>
</tr>
</tbody>
</table>
Evaluating language knowledge

<table>
<thead>
<tr>
<th>Predicate</th>
<th>give</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicate meaning</td>
<td>cause to receive</td>
</tr>
<tr>
<td>Number of arguments</td>
<td>3</td>
</tr>
<tr>
<td>Word order</td>
<td>X verb Y Z</td>
</tr>
<tr>
<td>Argument meanings</td>
<td>{human}; {human}; {object}</td>
</tr>
<tr>
<td>Argument roles</td>
<td>Giver; Recipient; Theme</td>
</tr>
</tbody>
</table>
Evaluating language knowledge

Giver | verb | Recipient | Theme
---|---|---|---
The bear | gives | you | the ball!

Yevgen Matusevych
Evaluating language knowledge

- Elicited production task

The bear ____ you the ball!

Giver  verb  Recipient  Theme

The bear  you  the ball!

Yevgen Matusevych
Input-related determinants of construction learning
Input-related determinants of construction learning

- [Ellis, O'Donnell, & Römer, 2014]:

  Determinants of learning argument structure constructions:

  (1) verb frequency

  (2) strength of association between verb and construction (ΔP)

  (3) semantic centrality
1. Frequency

Frequency of verbs in a certain argument structure construction

- E.g., prepositional dative (transfer) construction:

  He ____ it to someone.

  
<table>
<thead>
<tr>
<th>Verb</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>give</td>
<td>1000</td>
</tr>
<tr>
<td>show</td>
<td>150</td>
</tr>
<tr>
<td>send</td>
<td>50</td>
</tr>
<tr>
<td>lend</td>
<td>10</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
How strong is the association between a verb and a construction?

<table>
<thead>
<tr>
<th>Verb</th>
<th>Association strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>give</td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Verb</th>
<th>Association strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>give</td>
<td>100</td>
</tr>
</tbody>
</table>
2. Association strength

How strong is the association between a verb and a construction?

<table>
<thead>
<tr>
<th></th>
<th>Prepositional dative (transfer) construction</th>
<th>Other constructions</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>give</em></td>
<td>100</td>
<td>500</td>
</tr>
<tr>
<td>other verbs</td>
<td>120</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Ditransitive (transfer) construction</th>
<th>Other constructions</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>give</em></td>
<td>100</td>
<td>500</td>
</tr>
<tr>
<td>other verbs</td>
<td>900</td>
<td></td>
</tr>
</tbody>
</table>
2. Association strength

How strong is the association between a verb and a construction?

\[ \Delta P (\text{construction} \rightarrow \text{verb}) = \frac{a}{a+b} - \frac{c}{c+d} \]

<table>
<thead>
<tr>
<th></th>
<th>Prepositional dative (transfer) construction</th>
<th>Other constructions</th>
</tr>
</thead>
<tbody>
<tr>
<td>give</td>
<td>a</td>
<td>c</td>
</tr>
<tr>
<td>other verbs</td>
<td>b</td>
<td>d</td>
</tr>
</tbody>
</table>
3. Meaning centrality

How central, or prototypical, is the verb meaning for a construction?

[Ellis et al., 2014]
Input-related determinants of construction learning

- [Ellis et al., 2014]:

  “constructions”:

  \( N \ V \ about \ N \)
  \( N \ V \ across \ N \)
  \( N \ V \ among \ N \)
  ...
  \( N \ V \ with \ N \)

  grammar patterns
  [Francis, Hunston, & Mannin, 1996]
Input-related determinants of construction learning

- [Ellis et al., 2014]: “constructions”:
  - $N V$ about $N$
  - $N V$ across $N$
  - $N V$ among $N$
  - ...
  - $N V$ with $N$

  grammar patterns
  [Francis, Hunston, & Mannin, 1996]

- this study “constructions”:
  - $A1 V A2 \leftrightarrow \text{killing}$
  - $A1 V \leftrightarrow \text{event_start}$

  syntactic pattern + frame semantics
  [Goldberg, 1995] + FrameNet
Input-related determinants of construction learning

- [Ellis et al., 2014]:
  grammar patterns
  
  $N \text{ V about } N$
  $N \text{ V across } N$
  $N \text{ V among } N$
  ...
  $N \text{ V with } N$

- this study
  syntactic pattern + frame semantics
  
  $A1 \text{ V A2} \iff \text{killing}$
  $A1 \text{ V} \iff \text{event_start}$

- **RQ.** Do the same input properties affect the learning of constructions:
  
  1) if constructions are represented differently?
  
  2) in the second language?
  
  3) in terms of other measures of language knowledge?
Data

Syntactic structure
- TIGER
  - Penn Treebank, WSJ part

Argument structure
- SALSA
  - PropBank

Filtering data
- FrameNet
  - FrameNet, SemLink

Semantics
- WordNet, VerbNet, FrameNet–WordNet mapping

Final data
- 3370 instances
- 3803 instances

Legend
- English
- German
Learning scenario

Language exposure → Test
Evaluating language knowledge

- Elicited production task

The bear ____ you the ball!

Giver: The bear
Recipient: you
Theme: the ball!

Yevgen Matusevych
<table>
<thead>
<tr>
<th>Items</th>
<th>Variables</th>
<th>Probability of use</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Construction</strong></td>
<td><strong>Verb</strong></td>
<td><strong>Frequency</strong></td>
</tr>
<tr>
<td>Constr1</td>
<td>Verb1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Verb2</td>
<td></td>
</tr>
<tr>
<td>Constr2</td>
<td>Verb1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Verb2</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Design

<table>
<thead>
<tr>
<th>Items</th>
<th>Variables</th>
<th>Probability of use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construction</td>
<td>Verb</td>
<td></td>
</tr>
<tr>
<td>'Process start'</td>
<td>begin</td>
<td></td>
</tr>
<tr>
<td>ARG1 VERB</td>
<td>start</td>
<td></td>
</tr>
<tr>
<td></td>
<td>erupt</td>
<td></td>
</tr>
</tbody>
</table>

- mixed effect models
- fixed effects of frequency, association and centrality
- random effect of construction
Results: statistics

<table>
<thead>
<tr>
<th>Language</th>
<th>Predictor</th>
<th>Coefficient $\beta$</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>frequency</td>
<td>2.55</td>
<td>$&lt;.001$ ***</td>
</tr>
<tr>
<td></td>
<td>association</td>
<td>0.01</td>
<td>$&gt;.05$</td>
</tr>
<tr>
<td></td>
<td>centrality</td>
<td>0.71</td>
<td>$&lt;.001$ ***</td>
</tr>
<tr>
<td>German</td>
<td>frequency</td>
<td>0.36</td>
<td>$&lt;.001$ ***</td>
</tr>
<tr>
<td></td>
<td>association</td>
<td>0.10</td>
<td>$&lt;.05$ *</td>
</tr>
<tr>
<td></td>
<td>centrality</td>
<td>0.68</td>
<td>$&lt;.001$ ***</td>
</tr>
</tbody>
</table>
Summary

Partly in line with the existing results for L1 learning:

[Ellis et al., 2014]: frequency, association strength, centrality
this study: frequency, association strength (?), centrality
Summary

Partly in line with the existing results for L1 learning:

Ellis et al., 2014]: frequency, association strength, centrality
this study: frequency, association strength (?), centrality

How about L2?
Learning scenario

L1 exposure → Mixed L1 + L2 exposure → Test
# Results: L2

<table>
<thead>
<tr>
<th>L2</th>
<th>Predictor</th>
<th>Coefficient $\beta$</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>frequency</td>
<td>1.00</td>
<td>$&lt; .001$ ***</td>
</tr>
<tr>
<td></td>
<td>association</td>
<td>-0.08</td>
<td>$&gt; .05$</td>
</tr>
<tr>
<td></td>
<td>centrality</td>
<td>1.09</td>
<td>$&lt; .001$ ***</td>
</tr>
<tr>
<td>German</td>
<td>frequency</td>
<td>0.39</td>
<td>$&lt; .001$ ***</td>
</tr>
<tr>
<td></td>
<td>association</td>
<td>0.01</td>
<td>$&gt; .05$</td>
</tr>
<tr>
<td></td>
<td>centrality</td>
<td>0.73</td>
<td>$&lt; .001$ ***</td>
</tr>
</tbody>
</table>
Summary

Partly in line with the existing results for L1 learning:

[Ellis et al., 2014]: frequency, association strength, centrality
this study: frequency, association strength (?), centrality

Our findings are consistent in L1 and L2:

L1: frequency, association strength (?), centrality
L2: frequency, association strength, centrality
Summary

Partly in line with the existing results for L1 learning:

[Ellis et al., 2014]: frequency, association strength, centrality
this study: frequency, association strength (?), centrality

Our findings are consistent in L1 and L2:

L1: frequency, association strength (?), centrality
L2: frequency, association strength, centrality

How about other measures of language knowledge?
Representing language knowledge

1. Distribution

Input properties

\[ \text{distribution of verbs within a certain construction} \]

Open task

[Ellis et al., 2014]
Representing language knowledge

1. Distribution
   - Input properties: distribution of verbs within a certain construction
   - Open task
   - [Ellis et al., 2014]

2. Proficiency score
   - Input properties: proficiency score for verbs within a certain construction
   - Closed task
   - [Goldschneider & DeKeyser, 2001]
Evaluating language knowledge

1. Distribution
(elicited production)

The bear ____ you the ball!

Giver verb Recipient Theme

1. Verb production probability
Evaluating language knowledge

1. Distribution
   (elicited production)
   The bear ___ you the ball!
   Giver    verb    Recipient    Theme
   1. Verb production probability

2. Proficiency score
   (comprehension)
   The bear gives you the ball!
   Giver    verb    Recipient    Theme
   2. Verb comprehension score
# Results: L2

<table>
<thead>
<tr>
<th>L2</th>
<th>Predictor</th>
<th>Coefficient $\beta$</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>frequency</td>
<td>0.13</td>
<td>$&lt; .001$ ***</td>
</tr>
<tr>
<td></td>
<td>association</td>
<td>$&lt; 0.01$</td>
<td>$&gt; .05$</td>
</tr>
<tr>
<td></td>
<td>centrality</td>
<td>0.07</td>
<td>$&lt; .001$ ***</td>
</tr>
<tr>
<td>German</td>
<td>frequency</td>
<td>0.11</td>
<td>$&lt; .001$ ***</td>
</tr>
<tr>
<td></td>
<td>association</td>
<td>0.01</td>
<td>$&gt; .05$</td>
</tr>
<tr>
<td></td>
<td>centrality</td>
<td>0.06</td>
<td>$&lt; .001$ ***</td>
</tr>
</tbody>
</table>
Summary

Partly in line with the existing results for L1 learning:

[Ellis et al., 2014]: frequency, association strength, centrality
this study: frequency, association strength (?), centrality

Our findings are consistent in L1 and L2:

L1: frequency, association strength (?), centrality
L2: frequency, association strength, centrality

Our findings are consistent for different language tasks in L2:

elicited production: frequency, association strength, centrality
verb comprehension: frequency, association strength, centrality
Conclusions

- For the construction representations that we used, verb frequency and centrality of verb meaning are important, but not association strength.

- The findings are consistent for German and English, for L1 and L2, for elicited production and verb comprehension tasks.
References


Formal model

1. Find most likely construction for a given frame:

   $\text{BestConstruction}(F) = \arg\max_k P(k|F)$

2. For this, use prior and conditional probability:

   $P(k|F) = \frac{P(k)P(F|k)}{P(F)} \propto P(k)P(F|k)$

3. Prior probability = entrenchment:

   $P(k) = \frac{N_k}{N+1}, \quad P(0) = \frac{1}{N+1}$

4. Conditional probability = similarity in terms of each feature:

   $P(F|k) = \prod_{i \in \text{Features}(F)} P(F_i|k)$
An example frame

I ate a tuna sandwich.

<table>
<thead>
<tr>
<th>predicate</th>
<th>eat</th>
</tr>
</thead>
<tbody>
<tr>
<td>event properties</td>
<td>consume, take in, prep</td>
</tr>
<tr>
<td>arg. count</td>
<td>2</td>
</tr>
<tr>
<td>arg1</td>
<td>I</td>
</tr>
<tr>
<td>arg2</td>
<td>sandwich</td>
</tr>
<tr>
<td>arg1 lexical props</td>
<td>self, person, ..., entity</td>
</tr>
<tr>
<td>arg2 lexical props</td>
<td>snack food, dish, ..., entity</td>
</tr>
<tr>
<td>arg1 role props</td>
<td>living thing, entity, ..., organism</td>
</tr>
<tr>
<td>arg2 role props</td>
<td>solid, substance, ..., entity</td>
</tr>
<tr>
<td>arg1 case</td>
<td>N/A</td>
</tr>
<tr>
<td>arg2 case</td>
<td>N/A</td>
</tr>
<tr>
<td>syntactic pattern</td>
<td>ARG1 VERB ARG2</td>
</tr>
<tr>
<td>prepositions</td>
<td>N/A</td>
</tr>
</tbody>
</table>
Existing models

- DevLex family of connectionist models [Zhao & Li, 2010]: semantics + phonology
- Model of entrenchment and memory development [Monner et al., 2013]: phonology + morphology
- Model of bilingual semantic memory [Cuppini et al., 2013]: lexis + semantics
- Other models [Li, 2013]