Isolating second language learning factors

Yevgen Matusevych
Afra Alishahi
Ad Backus

Tilburg University, the Netherlands

CogSci – 2014, Quebec City
Methodological problems in SLA

- Research question: What is the exact relation between learners' L2 proficiency and their length of residence in L2 environment?
Methodological problems in SLA

- Research question: **What is the exact relation between learners' L2 proficiency and their length of residence in L2 environment?**
- Recommendations by DeKeyser, 2013:
Methodological problems in SLA

- Research question: **What is the exact relation between learners' L2 proficiency and their length of residence in L2 environment?**
- Recommendations by DeKeyser, 2013:
  1. Participants should be native speakers of the same language.
Methodological problems in SLA

- Research question: **What is the exact relation between learners' L2 proficiency and their length of residence in L2 environment?**

- Recommendations by DeKeyser, 2013:
  1. Participants should be native speakers of the same language.
  2. L1 should be distant from L2.
Methodological problems in SLA

- Research question: What is the exact relation between learners' L2 proficiency and their length of residence in L2 environment?

- Recommendations by DeKeyser, 2013:
  1. Participants should be native speakers of the same language.
  2. L1 should be distant from L2.
  3. Participants must have spent most time in immigration using L2.
Methodological problems in SLA

- Research question: **What is the exact relation between learners' L2 proficiency and their length of residence in L2 environment?**

- Recommendations by DeKeyser, 2013:
  1. Participants should be native speakers of the same language.
  2. L1 should be distant from L2.
  3. Participants must have spent most time in immigration using L2.
  4. Participants should be spread over all socioeconomic levels.
Methodological problems in SLA

- Research question: **What is the exact relation between learners' L2 proficiency and their length of residence in L2 environment?**

- Recommendations by DeKeyser, 2013:
  1. Participants should be native speakers of the same language.
  2. L1 should be distant from L2.
  3. Participants must have spent most time in immigration using L2.
  4. Participants should be spread over all socioeconomic levels.
  5. Length of residence should be 10+ years.
Methodological problems in SLA

- Research question: **What is the exact relation between learners' L2 proficiency and their length of residence in L2 environment?**

- Recommendations by DeKeyser, 2013:
  1. Participants should be native speakers of the same language.
  2. L1 should be distant from L2.
  3. Participants must have spent most time in immigration using L2.
  4. Participants should be spread over all socioeconomic levels.
  5. Length of residence should be 10+ years.
  6. Age at testing should not go beyond middle age.
Methodological problems in SLA

- Research question: **What is the exact relation between learners' L2 proficiency and their length of residence in L2 environment?**

- Recommendations by DeKeyser, 2013:
  1. Participants should be native speakers of the same language.
  2. L1 should be distant from L2.
  3. Participants must have spent most time in immigration using L2.
  4. Participants should be spread over all socioeconomic levels.
  5. Length of residence should be 10+ years.
  6. Age at testing should not go beyond middle age.
  7. Sample size should be approx. 20 per age.
Methodological problems in SLA

- Research question: **What is the exact relation between learners' L2 proficiency and their length of residence in L2 environment?**

- Recommendations by DeKeyser, 2013:
  1. Participants should be native speakers of the same language.
  2. L1 should be distant from L2.
  3. Participants must have spent most time in immigration using L2.
  4. Participants should be spread over all socioeconomic levels.
  5. Length of residence should be 10+ years.
  6. Age at testing should not go beyond middle age.
  7. Sample size should be approx. 20 per age.
  8. Participants should be spread homogeneously over the age ranges.
Methodological problems in SLA

- Variation in L2 learners
- Many confounds
- Approximate measures
Computational modeling

- Explicit assumptions
- Controlled input
- Observable behavior
- Testable predictions
Existing models

- DevLex family of connectionist models (e.g., Zhao & Li, 2010): semantics + phonology

- Model of entrenchment and memory development (Monner et al., 2012): phonology + morphology

- Model of bilingual semantic memory (Cuppini et al., 2012): lexis + semantics

- etc. (see Li, 2013)
What is missing?

- Models of SLA or bilingualism that would go beyond the word level.
What is missing?

- Models of SLA or bilingualism that would go beyond the word level.
- Language structure, syntax, constructions
The model

- Usage-based linguistics, construction grammar

- Learning argument structure constructions from language input

- Adapted from the original model of child argument structure acquisition (Alishahi & Stevenson, 2008)
Learning argument structure constructions
Learning argument structure constructions

Yevgen Matusevych
Learning argument structure constructions
Learning argument structure constructions

Look! The bear gives you the ball!
Learning argument structure constructions

Look! The bear gives you the ball!
Learning argument structure constructions

The bear gives you the ball
Learning argument structure constructions

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

The bear gives you the ball
Daddy's coming home

Grandma sent you some cookies.
John passed you the ball!
Mr. Rich donated us a thousand dollars.
Learning argument structure constructions

The bear gives you the ball
Daddy's coming home
Grandma sent you some cookies
John passed you the ball
Mr. Rich donated us a thousand dollars
Learning argument structure constructions

- The bear gives you the ball
- Daddy’s coming home
- Grandma sent you some cookies
- John passed you the ball
- Mr. Rich donated us a thousand dollars

<table>
<thead>
<tr>
<th>Predicate meaning</th>
<th>changing object possessor</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>
Learning argument structure constructions

The bear gives you the ball
Grandma sent you some cookies
John passed you the ball
Mr. Rich donated us a thousand dollars
Daddy's coming home

Predicate meaning: changing object possessor
Number of arguments: 3
Word order: X {verb} Y Z
Argument meanings: {human}; {human}; {object}
Argument roles: Giver; Recipient; Theme
Learning argument structure constructions

The bear gives you the ball  ditransitive
Grandma sent you some cookies  X {verb} Y Z
John passed you the ball
Mr. Rich donated us a thousand dollars
Daddy’s coming home
Learning argument structure constructions

ditransitive
X {verb} Y Z

transitive
X {verb} Y
Learning argument structure constructions

 ditransitive
 X {verb} Y Z

 transitive
 X {verb} Y

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

<table>
<thead>
<tr>
<th>Predicate meaning</th>
<th>giving</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>

Meine Schwester lieh mir Geld.
(My sister lent me some money.)
Input type

- A number of instances (frames):

Table 1: An example frame extracted from a verb usage *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>1</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>
Formal model

1. Find most likely construction for a given frame:

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

\[
P(k|F) = \frac{P(k)P(F|k)}{P(F)} \propto P(k)P(F|k)
\]
**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:

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

   \[ P(F|k) = \prod_{i \in \text{Features}(F)} P(F_i|k) \]
Model: evaluation

Evaluation on language use: various comprehension/production tasks (predicting missing features in a frame).
Model: evaluation

Evaluation on language use: various comprehension/production tasks (predicting missing features in a frame).

\[ P(F_i|F) = \sum_k P(F_i|k)P(k|F) \propto \sum_k P(F_i|k)P(k)P(F|k) \]
Model: evaluation

Evaluation on language use: various comprehension/production tasks (predicting missing features in a frame).

\[ P(F_i|F) = \sum_k P(F_i|k)P(k|F) \propto \sum_k P(F_i|k)P(k)P(F|k) \]

<table>
<thead>
<tr>
<th>Test frame number</th>
<th>Original value F₀</th>
<th>Predicted value Fᵢ</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X</td>
<td>X</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Y</td>
<td>Z</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>abc</td>
<td>abd</td>
<td>0.66</td>
</tr>
</tbody>
</table>
Model: evaluation

Evaluation on language use: various comprehension/production tasks (predicting missing features in a frame).

\[ P(F_i|F) = \sum_k P(F_i|k)P(k|F) \propto \sum_k P(F_i|k)P(k)P(F|k) \]

<table>
<thead>
<tr>
<th>Test frame number</th>
<th>Original value $F_o$</th>
<th>Predicted value $F_i$</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X</td>
<td>X</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Y</td>
<td>Z</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>abc</td>
<td>abd</td>
<td>0.66</td>
</tr>
</tbody>
</table>

*Prediction accuracy (PA) for feature $F_i$*
Model: evaluation

Evaluation on language use: various comprehension/production tasks (predicting missing features in a frame).

Giver    verb    Recipient    Theme
The bear   gives   you    the ball!
Model: evaluation

Evaluation on language use: various comprehension/production tasks (predicting missing features in a frame).

The bear ____ you the ball!

1. PA (predicate)
Model: evaluation

Evaluation on language use: various comprehension/production tasks (predicting missing features in a frame).

1. PA (predicate)

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

2. PA (predicate meaning)

Giver: The bear
Verb: gives
Recipient: you
Theme: the ball!
Model: evaluation

Evaluation on language use: various comprehension/production tasks (predicting missing features in a frame).
Model: evaluation

Evaluation on language use: various comprehension/production tasks (predicting missing features in a frame).

1. PA (predicate)
   - Giver: The bear
   - Verb: ____
   - Recipient: you
   - Theme: the ball!

2. PA (predicate meaning)
   - Giver: The bear
   - Verb: gives
   - Recipient: you
   - Theme: the ball!

3. PA (argument roles)
   - Verb: gives
   - Recipient: you
   - Theme: the ball!

\[ LPI = \frac{\sum_{i=1}^{n} PA_i}{n} \]
## Data

<table>
<thead>
<tr>
<th>Language</th>
<th>Data Sources</th>
<th>Frame Instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>PropBank + Penn Treebank + SemLink + FrameNet + WordNet</td>
<td>~ 3,800</td>
</tr>
<tr>
<td>German</td>
<td>SALSA + TIGER corpus + dict.cc + WordNet</td>
<td>~ 3,400</td>
</tr>
</tbody>
</table>

Yevgen Matusevych
## Data

<table>
<thead>
<tr>
<th>Language</th>
<th>Dataset</th>
<th>Frame Instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>PropBank + Penn Treebank + SemLink + FrameNet + WordNet</td>
<td>~3,800</td>
</tr>
<tr>
<td>German</td>
<td>SALSA + TIGER corpus + dict.cc + WordNet</td>
<td>~3,400</td>
</tr>
</tbody>
</table>

Both datasets can be used as either L1 or L2.

The actual input in each simulation is sampled from the data.
Learning scenarios

L1 training

L1 training

L2 training
Learning scenarios

- L1 training
- L2 training
Learning scenarios

L1 training

L1 training

L2 training

Late L2 learner, immersion
Learning scenarios

L1 training

Late L2 learner, immersion

Late L2 learner, classroom
Learning scenarios

Late L2 learner, immersion

Late L2 learner, classroom

Early bilingual
L1/L2 ratio

L2 English

Input frames

LPI
L1/L2 ratio

L2 English

R=1  R=2  R=5  R=10  R=20

Input frames

LPI

0.6

0.4

0.2
L1/L2 ratio

L2 English

L1 training
L2 training

L1 training
L2 training

R=1  R=2  R=5  R=10  R=20

LPI

Input frames

0  100  200  300  400
**L1/L2 ratio**

**L2 English**

**L2 German**

Yevgen Matusevych
Replicating the effect of L2 amount on L2 proficiency
Age of L2 onset

L2 English

LPI

Input frames
Age of L2 onset

L2 English

Input frames

LPI

AO=0  AoO=100  AoO=200  AoO=300  AoO=400
Age of L2 onset

L2 English

LPI vs. Input frames for L2 English

- AO=0
- AO=100
- AO=200
- AO=300
- AO=400
Age of L2 onset

L2 English

L2 German

Yevgen Matusevych
Age of L2 onset

L2 English

L2 German

Did not replicate the age effect
Frequency distribution in the input

- Language learners are sensitive to input frequencies

- Certain verbs occur in a certain construction much more often than other verbs (Ellis & Ferreira-Junior, 2009)
Frequency distribution in the input

- Language learners are sensitive to input frequencies

- Certain verbs occur in a certain construction much more often than other verbs (Ellis & Ferreira-Junior, 2009)

I _____ it to someone.
Frequency distribution in the input

• Language learners are sensitive to input frequencies

• Certain verbs occur in a certain construction much more often than other verbs (Ellis & Ferreira-Junior, 2009)

I ____ it to someone.

give
show
send
lend
...

Yevgen Matusevych
Frequency distribution in the input

• Language learners are sensitive to input frequencies

• Certain verbs occur in a certain construction much more often than other verbs (Ellis & Ferreira-Junior, 2009)

I _____ it to someone.

give
show
send
lend
...

Yevgen Matusevych
Frequency distribution in the input

- Language learners are sensitive to input frequencies

- Certain verbs occur in a certain construction much more often than other verbs (Ellis & Ferreira-Junior, 2009)

- What is better? Balanced? Skewed? Or none?
Frequency distribution in the input

- German ditransitive with reversed order of arguments:

  THEME  PREDICATE  AGENT  PATIENT
  das    gab       ich    dem Herren
  it     gave      I      the gentlemen
  “I gave it to the gentlemen.”
**Frequency distribution in the input**

- German ditransitive with reversed order of arguments:
  
<table>
<thead>
<tr>
<th>THEME</th>
<th>PREDICATE</th>
<th>AGENT</th>
<th>PATIENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>das</td>
<td>gab</td>
<td>ich</td>
<td>dem Herren</td>
</tr>
<tr>
<td>it</td>
<td>gave</td>
<td>I</td>
<td>the gentlemen</td>
</tr>
</tbody>
</table>

  “I gave it to the gentlemen.”

- 15 different predicates appearing in this construction (10 training + 5 testing):
  - balanced: 1:1:1: … :1
  - skewed: 20:20:1: … :1
Frequency distribution in the input

Yevgen Matusevych
Frequency distribution in the input

Balanced input facilitates learning
Conclusions

- Statistical learner, bilingual construction learning

- Simulating different populations, fast hypothesis testing, easy replication

- A framework for studying concurrent acquisition of 2+ languages
Future work

- Larger-scale and more diverse simulations
- Data for other language pairs
References


Thank you for attending our talk!
SLA modeling: opinions

● “We need models of acquisition that relate such … measures to longitudinal patterns of child language and second language acquisition” (Nick Ellis)

● “Future research on implicit learning must implement computer simulations of language learning” (Jan Hulstijn)

● “Modeling SLA is hardest problem in linguistics. It is only hoped that computational models may contribute” (Rens Bod)