Individual variation and the roles of L1 and proficiency in the longitudinal L2 development of English grammatical morphemes

Akira Murakami
University of Cambridge / University of Birmingham
a.murakami@bham.ac.uk
Background

• Describing and explaining individual variation in longitudinal development are critical for a theory of L2 development.

• In order to tackle the issue, we need large empirical data.

• With the advent of large-scale learner corpora, we now have enough data to address the issue.

• Few studies have investigated longitudinal development at a large scale.
Longitudinal Development of Articles

![Graph showing longitudinal development of articles with TLU on the y-axis and Window on the x-axis. The graph includes a loess trend line with different colors for ABSENT, PRESENT, and L1 Chinese.]
Clustered Development

A

B

C

1.0

0.8

0.6

0.4

0.2

TLU

n = 367

(35.2%)

n = 357

(34.2%)

n = 320

(30.7%)

Window

Iloess (trend line)

ABSENT

PRESENT

L1 Chinese

Learner Corpus Research 2013 (29 September, 2013)
Aims of the Study

• Investigate whether learners’ L1 backgrounds and proficiency affect the longitudinal accuracy transition of L2 English grammatical morphemes.

• Reveal inter-learner variation.
Predictors

• L1 influence is pervasive in L2 acquisition (Jarvis & Pavlenko, 2007). The same is true for the L2 acquisition of English grammatical morphemes (Luk & Shirai, 2009; Murakami, 2013).

• The effect of proficiency on accuracy is prevalent (Thewissen, 2013).

• Given these, the two variables may affect the within-learner developmental patterns as well.
Target Morphemes

• articles
• past tense -ed
• plural -s
Target L1 Groups

- Typologically diverse L1s
  - Japanese
  - Korean
  - Spanish
  - Russian
  - Turkish
  - German
  - French
  - Brazilian-Portuguese
  - Mandarine-Chinese
  - Italian
EF-Cambridge Open Language Database (EFCamDat)

- Essays written at Englishtown, the online school of Education First
- 16 Lessons × 8 Units (A1-C2 in CEFR)
- Each student writes one essay per unit
- Teachers’ feedback available on some essays (≈ error tags)
- 140,000 essays by 52,000 learners, totaling 10 million words
- Available at http://corpus.mml.cam.ac.uk/efcamdat/
Accuracy Measure

TLU (Target-Like Use) score was used

\[
\text{TLU score} = \frac{\text{number of correct suppliance}}{\text{number of obligatory contexts} + \text{number of incorrect suppliance}}
\]
Two Types of Regression Models

• Mixed-effects model
  - Take into account the dependency of data within individual learners

• Generalised additive model
  - Take into account the nonlinearity of the effects of proficiency

• See if the two analyses converge
Mixed-Effects Model

- A mixed-effects logistic regression model
- Dependent variable: TLU
- Independent variables
  - L1
  - Morpheme
  - Proficiency (standardised)
  - Essay number (standardised)
  - Their two-way interactions
- Random-effects
  - Learner (random-intercept)
  - Morpheme (random-slope)
  - Essay number (random-slope)
Multi-Model Inference

• Compare the following models in order to test the effects of L1 and proficiency.
  • Model 1: Maximal model with the full structure just described
  • Model 2: Model 1 - EssayNum-L1 interaction
  • Model 3: Model 1 - EssayNum-Proficiency interaction
  • Model 4: Model 1 - EssayNum-Morpheme interaction
  • Null Model: Random-effects only
# Model Selection

<table>
<thead>
<tr>
<th>Model</th>
<th>AIC</th>
<th>Likelihood ratio test against</th>
<th>Null Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>72,607</td>
<td>$\chi^2(54) = 2462.675$ $p &lt; 0.001$</td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td>72,603</td>
<td>$\chi^2(9) = 14.210$ $p = 0.115$</td>
<td>$\chi^2(45) = 2448.465$ $p &lt; 0.001$</td>
</tr>
<tr>
<td>Model 3</td>
<td>72,614</td>
<td>$\chi^2(1) = 9.225$ $p = 0.002$</td>
<td>$\chi^2(53) = 2453.450$ $p &lt; 0.001$</td>
</tr>
<tr>
<td>Model 4</td>
<td>72,614</td>
<td>$\chi^2(2) = 11.256$ $p = 0.004$</td>
<td>$\chi^2(52) = 2451.419$ $p &lt; 0.001$</td>
</tr>
<tr>
<td>Null Model</td>
<td>74,961</td>
<td>$\chi^2(54) = 2462.675$ $p &lt; 0.001$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>Model description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>Maximal model</td>
</tr>
<tr>
<td>Model 2</td>
<td>EssayNum–L1 interaction excluded</td>
</tr>
<tr>
<td>Model 3</td>
<td>EssayNum–Proficiency interaction excluded</td>
</tr>
<tr>
<td>Model 4</td>
<td>EssayNum–Morpheme interaction excluded</td>
</tr>
<tr>
<td>Null Model</td>
<td>No fixed–effects</td>
</tr>
</tbody>
</table>
# Model Selection

<table>
<thead>
<tr>
<th>Model</th>
<th>AIC</th>
<th>Model 1 likelihood ratio test against</th>
<th>Null Model likelihood ratio test against</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>72,607</td>
<td>$\chi^2(54) = 2462.675$ $p &lt; 0.001$</td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td>72,603</td>
<td>$\chi^2(9) = 14.210$ $p = 0.115$</td>
<td>$\chi^2(45) = 2448.465$ $p &lt; 0.001$</td>
</tr>
<tr>
<td>Model 3</td>
<td>72,614</td>
<td>$\chi^2(1) = 9.225$ $p = 0.002$</td>
<td>$\chi^2(53) = 2453.450$ $p &lt; 0.001$</td>
</tr>
<tr>
<td>Model 4</td>
<td>72,614</td>
<td>$\chi^2(2) = 11.256$ $p = 0.004$</td>
<td>$\chi^2(52) = 2451.419$ $p &lt; 0.001$</td>
</tr>
<tr>
<td>Null Model</td>
<td>74,961</td>
<td>$\chi^2(54) = 2462.675$ $p &lt; 0.001$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>Model description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>Maximal model</td>
</tr>
<tr>
<td>Model 2</td>
<td>EssayNum–L1 interaction excluded</td>
</tr>
<tr>
<td>Model 3</td>
<td>EssayNum–Proficiency interaction excluded</td>
</tr>
<tr>
<td>Model 4</td>
<td>EssayNum–Morpheme interaction excluded</td>
</tr>
<tr>
<td>Null Model</td>
<td>No fixed-effects</td>
</tr>
</tbody>
</table>
Random-Effects

<table>
<thead>
<tr>
<th>Factor</th>
<th>Random Effects</th>
<th>Model 2 SD</th>
<th>Null Model SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learner Intercept</td>
<td></td>
<td>0.501</td>
<td>0.758</td>
</tr>
<tr>
<td>Morpheme</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past tense -ed</td>
<td></td>
<td>0.919</td>
<td>1.140</td>
</tr>
<tr>
<td>Plural -s</td>
<td></td>
<td>0.510</td>
<td>1.071</td>
</tr>
<tr>
<td>EssayNum.Standardized</td>
<td></td>
<td>0.297</td>
<td>0.327</td>
</tr>
</tbody>
</table>
## Random-Effects

<table>
<thead>
<tr>
<th>Factor</th>
<th>Random Effects</th>
<th>Model 2</th>
<th>Null Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learner Intercept</td>
<td>0.501</td>
<td></td>
<td>0.758</td>
</tr>
<tr>
<td>Morpheme</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past tense $-ed$</td>
<td>0.919</td>
<td></td>
<td>1.140</td>
</tr>
<tr>
<td>Plural $-s$</td>
<td>0.510</td>
<td></td>
<td>1.071</td>
</tr>
<tr>
<td>EssayNum.Standardized</td>
<td>0.297</td>
<td></td>
<td>0.327</td>
</tr>
</tbody>
</table>
Demonstration of Random-Intercept
Demonstration of Random-Intercept

TLU

Essay Number
Demonstration of Random-Intercept
Demonstration of Random-Intercept
Demonstration of Random-Intercept

Essay Number

TLU

0.34

0.72 0.70 0.68 ... 0.36
Demonstration of Random-Intercept

```
0.72 0.70 0.68 . . . 0.36 0.34
```
Demonstration of Random-Intercept

![Graph showing trend lines for essay numbers with TLU on the y-axis and Essay Number on the x-axis.]

Learner Corpus Research 2013 (29 September, 2013)
Demonstration of Random-Intercept
Random-Effects

<table>
<thead>
<tr>
<th>Factor</th>
<th>Random Effects</th>
<th>Model 2</th>
<th>Null Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learner</td>
<td>Intercept</td>
<td>0.501</td>
<td>0.758</td>
</tr>
<tr>
<td>Morpheme</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past tense -ed</td>
<td>0.919</td>
<td>1.140</td>
<td></td>
</tr>
<tr>
<td>Plural -s</td>
<td>0.510</td>
<td>1.071</td>
<td></td>
</tr>
<tr>
<td>EssayNum.Standardized</td>
<td>0.297</td>
<td>0.327</td>
<td></td>
</tr>
</tbody>
</table>

Individual differences (IDs) in absolute accuracy
## Random-Effects

<table>
<thead>
<tr>
<th>Factor</th>
<th>Random Effects</th>
<th>Model 2</th>
<th>Null Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learner Intercept</td>
<td>0.501</td>
<td>0.758</td>
<td></td>
</tr>
<tr>
<td>Morpheme</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past tense $-ed$</td>
<td>0.919</td>
<td>1.140</td>
<td></td>
</tr>
<tr>
<td>Plural $-s$</td>
<td>0.510</td>
<td>1.071</td>
<td></td>
</tr>
<tr>
<td>EssayNum.Standardized</td>
<td>0.297</td>
<td>0.327</td>
<td></td>
</tr>
</tbody>
</table>

Individual differences (IDs) in absolute accuracy

*cf. EssayNum.Standardized in fixed-effects = 0.140 (1 SD of EssayNum $\approx$ 2/3 CEFR level)*
Demonstration of Random-Slopes
Demonstration of Random-Slopes
# Random-Effects

<table>
<thead>
<tr>
<th>Factor</th>
<th>Random Effects</th>
<th>Model 2</th>
<th>Null Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learner</td>
<td>Intercept</td>
<td>0.501</td>
<td>0.758</td>
</tr>
<tr>
<td>Morpheme</td>
<td>Past tense $-ed$</td>
<td>0.919</td>
<td>1.140</td>
</tr>
<tr>
<td></td>
<td>Plural $-s$</td>
<td>0.510</td>
<td>1.071</td>
</tr>
<tr>
<td></td>
<td>EssayNum.Standardized</td>
<td>0.297</td>
<td>0.327</td>
</tr>
</tbody>
</table>

IDs in the rate of accuracy increase
Findings Based on Mixed-Effects Modeling

• Large individual variation in
  - overall accuracy
  - accuracy difference between articles and the other morphemes, and
  - the rate of development

• Despite large individual differences, proficiency and morpheme explain the variation to a certain degree
  - Developmental shape is different depending on learners’ overall proficiency and on the morpheme concerned
  - But no evidence for differing rate of development by L1
Fitted Values of Model 2 for Low- vs High-Proficiency Learners

Low Proficiency Learners

High Proficiency Learners
Cross-Sectional View of the Development

Articles

Past tense -ed

Plural -s

TLU

Proficiency

ABSENT
PRESENT
L1 Chinese (in Articles)
Generalised Additive Model (GAM)

- The relationship between independent and dependent variables is estimated by a nonlinear function.
- GAMs are semi-parametric, allowing both parametric and nonparametric terms in one model.
- In the present study,
  - L1, morpheme, and their interaction were entered as parametric terms.
  - A tensor product spline for the interaction between the overall proficiency and the within-learner development was obtained for each L1-Morpheme pair.
Models Considered

- Similar to the mixed-effects models
- Model 1: Maximal model
- Model 2: Tests the effect of L1
- Model 3: Tests the effect of Proficiency
- Model 4: Tests the effect of Morpheme
Models Considered

• Similar to the mixed-effects models

• **Model 1: Maximal model**

• Model 2: Tests the effect of L1

• Model 3: Tests the effect of Proficiency

• Model 4: Tests the effect of Morpheme
Article Development

L1 French

L1 Italian

L1 Spanish

L1 Korean

L1 Japanese

L1 Turkish

L1 Brazilian

L1 Chinese

L1 German

L1 Russian

Proficiency (Standardised)
Findings Based on the GAM

- Striking nonlinearity in morpheme accuracy development
- The nonlinear effects further interact with L1, proficiency, and morpheme, that is, the developmental patterns vary across learners' L1s, proficiency levels, and morphemes.
- When nonlinearity is taken into account, both L1 and proficiency affect the developmental patterns of accuracy.
Regression Summary

- We observe significant individual variation. However, the developmental patterns are not random.

- Proficiency systematically affects the developmental pattern
  - e.g., Ceiling effect

- L1 influence not clear
  - The mixed-effects analysis failed to show the significance of L1 influence
Interpretation

- Dynamic systems theory (DST; Verspoor, de Bot, & Lowie, 2011)

- variability
  - one’s linguistic system = dynamic system
  - Dynamic systems are in complex interactions
    ➝ constant change with chaotic variation
    - complete interconnectedness

- stability
  - attractor state
Conclusion

• Significant individual differences are present in L2 morpheme development.
• Systematicity is present at the same time.
• They are in line with DST.
References


