L1 CONGRUENCY IN THE EVALUATION OF DATA-DRIVEN LEARNING EFFECTIVENESS: A STUDY ON ITALIAN VERB-NOUN COLLOCATIONS

LUCIANA FORTI

13th Teaching and Language Corpora Conference
18-21 July 2018 – Faculty of Education, University of Cambridge
OUTLINE

- Background
- Aims
- Method
- Preliminary findings
- Discussion
- Conclusions and next steps
BACKGROUND

DATA-DRIVEN LEARNING (DDL)

Direct use of corpus data in the classroom

Constructivist learning theory
Background

Data-Driven Learning (DDL)

Corpus data is visible to the learners (Leech, 1997)

Direct use of corpus data in the classroom

Constructivist learning theory

13th TaLC Conference – Faculty of Education, University of Cambridge
BACKGROUND

DATA-DRIVEN LEARNING (DDL)

“Knowledge encoded from data by learners themselves will be more flexible, transferable, and useful than knowledge encoded and transmitted to them by an instructor.” (Cobb, 1999)

Corpus data is visible to the learners (Leech, 1997)

Constructivist learning theory

Direct use of corpus data in the classroom

13th TaLC Conference – Faculty of Education, University of Cambridge
BACKGROUND

THE EFFECTIVENESS OF DDL

IN RELATION TO ALL TYPES OF LEARNING AIMS

Overall, *large effect sizes found*, but need for better research design and reporting.
BACKGROUND

THE EFFECTIVENESS OF DDL

IN RELATION TO ALL TYPES OF LEARNING AIMS
Overall, *large effect sizes found*, but need for better research design and reporting.

IN RELATION TO VOCABULARY LEARNING ONLY
Overall, *medium effect sizes found*; DDL more effective with concordance lines purposely selected.
BACKGROUND

THE EFFECTIVENESS OF DDL

IN RELATION TO ALL TYPES OF LEARNING AIMS
Overall, *large effect sizes found*, but need for better research design and reporting.

IN RELATION TO VOCABULARY LEARNING ONLY
Overall, *medium effect sizes found*; DDL more effective with concordance lines purposely selected.

IN RELATION TO THE LINGUISTIC PROPERTIES OF THE LEARNING AIMS
BACKGROUND

THE EFFECTIVENESS OF DDL

IN RELATION TO ALL TYPES OF LEARNING AIMS
Overall, large effect sizes found, but need for better research design and reporting.

IN RELATION TO VOCABULARY LEARNING ONLY
Overall, medium effect sizes found; DDL more effective with concordance lines purposely selected.

IN RELATION TO THE LINGUISTIC PROPERTIES OF THE LEARNING AIMS

13th TaLC Conference – Faculty of Education, University of Cambridge
AIMS

THE PRESENT STUDY

• To evaluate the effectiveness of DDL in learning Italian verb-noun collocations (Bestgen & Granger, 2014; Wang, 2016)
• To take into account two properties of the learning aims: semantic transparency and L1 congruency.
• To consider different dimensions of collocational knowledge.
• To integrate the findings with emic data deriving from an end-of-course student questionnaire.
AIMS

THE PRESENT STUDY

• To evaluate the effectiveness of DDL in learning Italian verb-noun collocations (Bestgen & Granger, 2014; Wang, 2016)

• To take into account two properties of the learning aims: semantic transparency and L1 congruency.

• To consider different dimensions of collocational knowledge.

• To integrate the findings with emic data deriving from an end-of-course student questionnaire.
AIMS

RESEARCH QUESTIONS

1. How does collocational proficiency develop over time when comparing a DDL versus a non-DDL learning approach?

2. How does L1 congruency influence the development of collocational proficiency in the two conditions?
Based on CLI and DDL literature

➢ The richer input that is typical in DDL will lead to:
  • overall higher language gains;
  • facilitating effect of congruent collocations; incongruent collocations being learned better than in control condition.
AIMS

PREDICTIONS

Based on CLI and DDL literature

➢ The richer input that is typical in DDL will lead to:
  • overall higher language gains;
  • facilitating effect of congruent collocations; incongruent collocations being learned better than in control condition.

➢ The guided-discovery approach that is typical of constructivism will lead to:
  • better retention rates.
Based on CLI and DDL literature

- The richer input that is typical in DDL will lead to:
  - overall higher language gains;
  - facilitating effect of congruent collocations; incongruent collocations being learned better than in control condition.

- The guided-discovery approach that is typical of constructivism will lead to:
  - better retention rates.
AIMS

PREDICTIONS

Based on CLI and DDL literature

➢ The richer input that is typical in DDL will lead to:
  • overall higher language gains;
  • facilitating effect of congruent collocations; incongruent collocations being learned better than in control condition.

➢ The guided-discovery approach that is typical of constructivism will lead to:
  • better retention rates.
OVERVIEW OF THE STUDY

➤ **Design:**

Controlled, longitudinal, between-groups.

➤ **Treatment:**

One 1-hour lesson a week for 8 weeks, in 8 classes of pre-intermediate Chinese learners of Italian, aged 18-27; random assignment to EXP or CON conditions.

➤ **Data collection:**

Collocational proficiency test administered at 4 week intervals over 12 weeks.
METHOD

SELECTION OF LEARNING AIMS

**LOCCLI**
(Italian learner corpus)
- Error analysis of all verb-noun collocations.
- Selection of 32 collocations.
- Collocations groups into 8 themes.

**PEC**
(Italian native corpus)
- Extraction of all verb-noun collocations.
- Selection of 32 collocations thematically linked to previously identified topics.
# Method

## Weekly Themes and Collocations

<table>
<thead>
<tr>
<th>Week</th>
<th>Theme</th>
<th>Collocations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>At a party</td>
<td>LOCCLI: fare amicizia; fare un sorriso; avere [numero] anni; studiare [materia]; amare [attività]. PEC: organizzare una festa; fare gli auguri; fare un regalo.</td>
</tr>
<tr>
<td>2</td>
<td>The weekend</td>
<td>LOCCLI: fare una passeggiata; prendere il sole; fare una gita; prendere aria. PEC: avere fretta; pulire casa; spendere soldi; fare la spesa.</td>
</tr>
<tr>
<td>3</td>
<td>My typical day</td>
<td>LOCCLI: prendere l'autobus; fare colazione; mettersi la giacca; avere lezione. PEC: rifare il letto; mettere la musica; fare la doccia; mandare un messaggio.</td>
</tr>
<tr>
<td>4</td>
<td>My house</td>
<td>LOCCLI: avere fame; preparare la cena. PEC: sbagliare la strada; trovare la strada; trovare casa; affittare una casa; dividere un appartamento; dividere una spesa.</td>
</tr>
<tr>
<td>5</td>
<td>My hobbies</td>
<td>LOCCLI: suonare la chitarra; fare sport; fare shopping; ascoltare musica; dipingere quadri; fare una foto; leggere un romanzo; vedere un film.</td>
</tr>
<tr>
<td>6</td>
<td>My last holidays</td>
<td>LOCCLI: gustare i cibi; visitare la città; ampliare le conoscenze; ricordare un’esperienza. PEC: organizzare un viaggio; prendere un treno; fare la fila; fare la valigia.</td>
</tr>
<tr>
<td>7</td>
<td>A friendship</td>
<td>LOCCLI: raccontare una storia. PEC: diventare amico; avere un dubbio; chiedere un consiglio; dare un consiglio; ascoltare un consiglio; trovare una soluzione; cambiare opinione.</td>
</tr>
<tr>
<td>8</td>
<td>Plans for the future</td>
<td>LOCCLI: fare l’artista; fare un viaggio; risparmiare soldi; fare esperienze. PEC: fare un esame; avere un’idea; cambiare casa; avere successo.</td>
</tr>
</tbody>
</table>
## METHOD

### EXPERIMENTAL VS. CONTROL

#### LESSON STAGES

<table>
<thead>
<tr>
<th>Time</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’</td>
<td>Gamified introduction to weekly collocations</td>
</tr>
</tbody>
</table>
| 25’  | **EXPERIMENTAL GROUPS:** paper-based DDL activities  
    - Multiple sentence matching;  
    - Multiple sentence gap-fill;  
    - Concordance based pattern-hunting;  
    - Concordance-based matching. |
|      | **CONTROL GROUPS:** non-DDL activities  
    - Matching single split sentences;  
    - Single sentence gap-fill;  
    - Single sentence error correction;  
    - Single sentence transformation exercise. |
| 15’  | Practice and production activities |
| 1’   | Homework assignment |
| 4’   | End-of-lesson game |
METHOD

DATA COLLECTION TOOL

64-item collocational proficiency test

32 target collocations based on LOCCLI
Multiple choice items

32 target collocations based on PEC
Gap-fill items
# METHOD

## DATA COLLECTION OUTLINE

<table>
<thead>
<tr>
<th>Data collection point</th>
<th>Week</th>
<th>Collocation set</th>
<th>Experimental groups</th>
<th>Control groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>n/a</td>
<td>Getting to know each other activities</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Background questionnaire</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Collocational proficiency test 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>DDL activities</td>
<td>Non-DDL activities</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>DDL activities</td>
<td>Non-DDL activities</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>DDL activities</td>
<td>Non-DDL activities</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4</td>
<td>DDL activities</td>
<td>Non-DDL activities</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>n/a</td>
<td>Collocational proficiency test 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td></td>
<td>DDL activities</td>
<td>Non-DDL activities</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>6</td>
<td>DDL activities</td>
<td>Non-DDL activities</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>7</td>
<td>DDL activities</td>
<td>Non-DDL activities</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>8</td>
<td>DDL activities</td>
<td>Non-DDL activities</td>
</tr>
<tr>
<td>3</td>
<td>9-12</td>
<td>n/a</td>
<td>Collocational proficiency test 3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>No lessons</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>n/a</td>
<td>Collocational proficiency test 4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>End-of-course questionnaire for experimental groups</td>
<td>End-of-course questionnaire for control groups</td>
</tr>
</tbody>
</table>
List of 64 target collocations categorised for L1 congruency by two different expert Chinese native speakers:

- 35 congruent collocations;
- 29 incongruent collocations.
METHOD

DATA ANALYSIS

The data:

63 students for the experimental condition & 61 students for the control condition.

The analysis:

<table>
<thead>
<tr>
<th>mixed-effects modeling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variable</td>
</tr>
<tr>
<td>Predictor variables</td>
</tr>
<tr>
<td>Random effects</td>
</tr>
</tbody>
</table>
## METHOD

### DATA ANALYSIS

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td>Correct Incorrect</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Predictor variables</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test Number</strong></td>
<td>Test 1 Test 2 Test 3 Test 4</td>
</tr>
<tr>
<td><strong>Item Type</strong></td>
<td>Congruent Incongruent</td>
</tr>
<tr>
<td><strong>Condition</strong></td>
<td>Control Experimental</td>
</tr>
</tbody>
</table>
DATA ANALYSIS

Generalised mixed-effects model

Best model fit:

ACCURACY predicted by
an interaction between TEST NUMBER and CONGRUENCY
+ CONDITION
How does collocational proficiency develop over time when comparing a DDL versus a non-DDL learning approach?
PRELIMINARY FINDINGS

RESEARCH QUESTION 1

Predicted probabilities for ACCURACY

<table>
<thead>
<tr>
<th>TEST NUMBER</th>
<th>ACCURACY</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>40%</td>
</tr>
<tr>
<td>2</td>
<td>50%</td>
</tr>
<tr>
<td>3</td>
<td>60%</td>
</tr>
<tr>
<td>4</td>
<td>70%</td>
</tr>
</tbody>
</table>

CONDITION
- CON
- EXP
PRELIMINARY FINDINGS

RESEARCH QUESTION 1

Predicted probabilities for ACCURACY

<table>
<thead>
<tr>
<th>TEST NUMBER</th>
<th>CONDITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CON</td>
</tr>
<tr>
<td>2</td>
<td>EXP</td>
</tr>
<tr>
<td>3</td>
<td>CON</td>
</tr>
<tr>
<td>4</td>
<td>EXP</td>
</tr>
</tbody>
</table>
RESEARCH QUESTION 1

Predicted probabilities for ACCURACY

CONDITION
- CON
- EXP

no lessons
PRELIMINARY FINDINGS

RESEARCH QUESTION 1

Predicted probabilities for ACCURACY

CONDITION
- CON
- EXP

TEST NUMBER

ACCURACY

1 2 3 4

no lessons
PRELIMINARY FINDINGS

RESEARCH QUESTION 1

Predicted probabilities for ACCURACY

CONDITION
- CON
- EXP

no lessons
PRELIMINARY FINDINGS

RESEARCH QUESTION 1

Predicted probabilities for ACCURACY

ACCURACY

TEST NUMBER

CONDITION

CON
EXP

no lessons
PRELIMINARY FINDINGS

RESEARCH QUESTION 1

Predicted probabilities for ACCURACY

<table>
<thead>
<tr>
<th>TEST NUMBER</th>
<th>CONDITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CON</td>
</tr>
<tr>
<td>2</td>
<td>EXP</td>
</tr>
<tr>
<td>3</td>
<td>CON</td>
</tr>
<tr>
<td>4</td>
<td>EXP</td>
</tr>
</tbody>
</table>

no lessons
How does L1 congruency influence the development of collocational proficiency in the two conditions?
PRELIMINARY FINDINGS

RESEARCH QUESTION 2

Predicted probabilities for ACCURACY

ITEM TYPE 2
- congruent
- non-congruent
PRELIMINARY FINDINGS

RESEARCH QUESTION 2

Predicted probabilities for ACCURACY

ITEM TYPE 2
- congruent
- non-congruent

I3th TaLC Conference – Faculty of Education, University of Cambridge
PRELIMINARY FINDINGS

RESEARCH QUESTION 2

Predicted probabilities for ACCURACY

ITEM TYPE 2
- congruent
- non-congruent
PRELIMINARY FINDINGS

RESEARCH QUESTION 2

Predicted probabilities for ACCURACY

ITEM TYPE 2
- congruent
- non-congruent
PRELIMINARY FINDINGS

RESEARCH QUESTION 2

Predicted probabilities for ACCURACY

CON

EXP

ITEM TYPE 2

congruent

non-congruent

13th TaLC Conference – Faculty of Education, University of Cambridge
DISCUSSION

PREDICTIONS

Based on CLI and DDL literature

- The richer input that is typical in DDL will lead to:
  - overall higher language gains;
  - facilitating effect of congruent collocations; incongruent collocations being learned better than in control condition.

- The guided-discovery approach that is typical of costructivism will lead to:
  - better retention rates.
DISCUSSION

PREDICTIONS

Based on CLI and DDL literature

- The richer input that is typical in DDL will lead to:
  - overall higher language gains; ❌
  - facilitating effect of congruent collocations; incongruent collocations being learned better than in control condition.

- The guided-discovery approach that is typical of costructivism will lead to:
  - better retention rates.
DISCUSSION

PREDICTIONS

Based on CLI and DDL literature

➢ The richer input that is typical in DDL will lead to:
  • overall higher language gains; X
  • facilitating effect of congruent collocations; incongruent collocations being learned better than in control condition. X

➢ The guided-discovery approach that is typical of constructivism will lead to:
  • better retention rates.
DISCUSSION

PREDICTIONS

Based on CLI and DDL literature

➢ The richer input that is typical in DDL will lead to:
  • overall higher language gains; ❌
  • facilitating effect of congruent collocations; incongruent collocations being learned better than in control condition. ❌

➢ The guided-discovery approach that is typical of constructivism will lead to:
  • better retention rates. ✓
The richer input typical of DDL does not necessarily lead to overall better language gains in comparison to a non-DDL approach.

- Duration: interventions of 10 sessions or more lead to better results (Lee et al. 2018: 25).
- Absence of a preliminary training opportunity: no significant differences (Lee et al. 2018: 25).
Incongruent collocations are not necessarily being learned better in a DDL setting.

- Are incongruent collocations significantly more frequent than congruent ones?
- Are they semantically transparent?
- In general, could there be a Kellerman effect (Kellerman, 1979)?
DDL consistently determines better retention rates when compared to non-DDL approach.

- Higher cognitive load.
- Autonomous discovery of patterns in language use.
- Interaction with peers.
- Novelty of the approach.
• Continue the statistical analysis (Growth curve analysis – Mirman, 2014)

• Investigate possible relationship between the different properties of the learning aims in determining the effectiveness of DDL.

• Consider different dimensions of collocational knowledge (Lee et al., 2018; Supatranont, 2005).

• Integrate etic data analysis with emic data from the end-of-course student questionnaire (pedagogical effectiveness evaluated through a mixed-methods approach – Riazi, 2017).
REFERENCES


THANK YOU!

Email: luciana.forti@unistrapg.it
Twitter: @l_for_ti