Integrating SLA evidence into DDL effects-oriented research

Luciana Forti
University for Foreigners of Perugia

International Perspectives on Corpus Technology for Language Learning Seminar Series | The University of Queensland, 10th November 2022

Outline of the talk

- What we mean by “Data-driven learning (DDL) effects-oriented research”
- How DDL effects-oriented research has developed so far
- Connection between SLA evidence and DDL potential
  (“reaching the parts other teaching can’t reach”; Boulton, 2008)
- Exploring DDL effects on L2 Italian phraseological competence development

Defining Data-driven learning

"Using the tools and techniques of corpus linguistics for pedagogical purposes."

Gilquin & Granger, 2010, p. 359.

Broad definition

Defining Data-driven learning

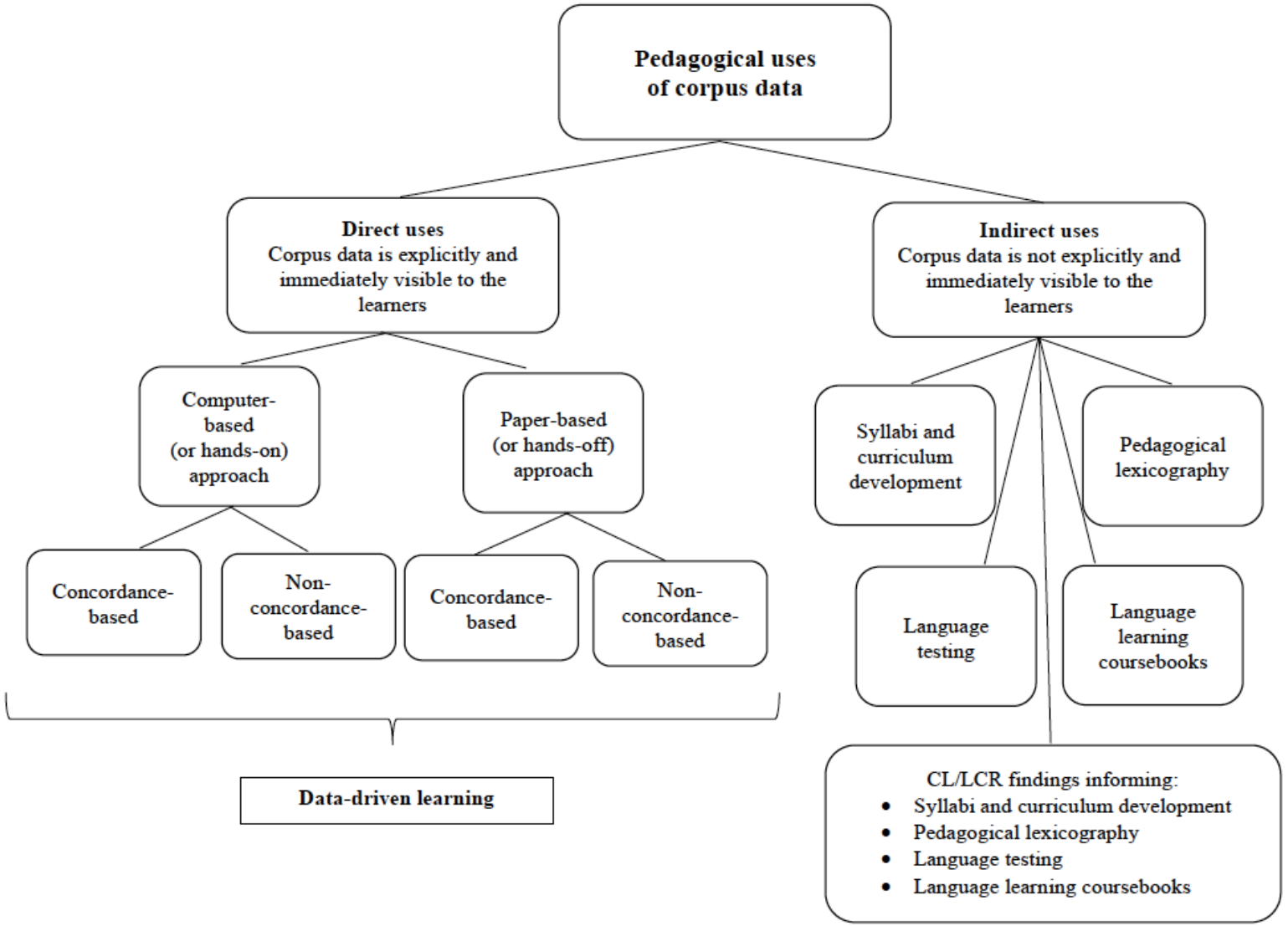
Direct vs. indirect uses
(Leech, 1997)

Immediate vs. delayed uses
(Granger, 2009)

Defining Data-driven learning

Direct vs. indirect uses
(Leech, 1997)

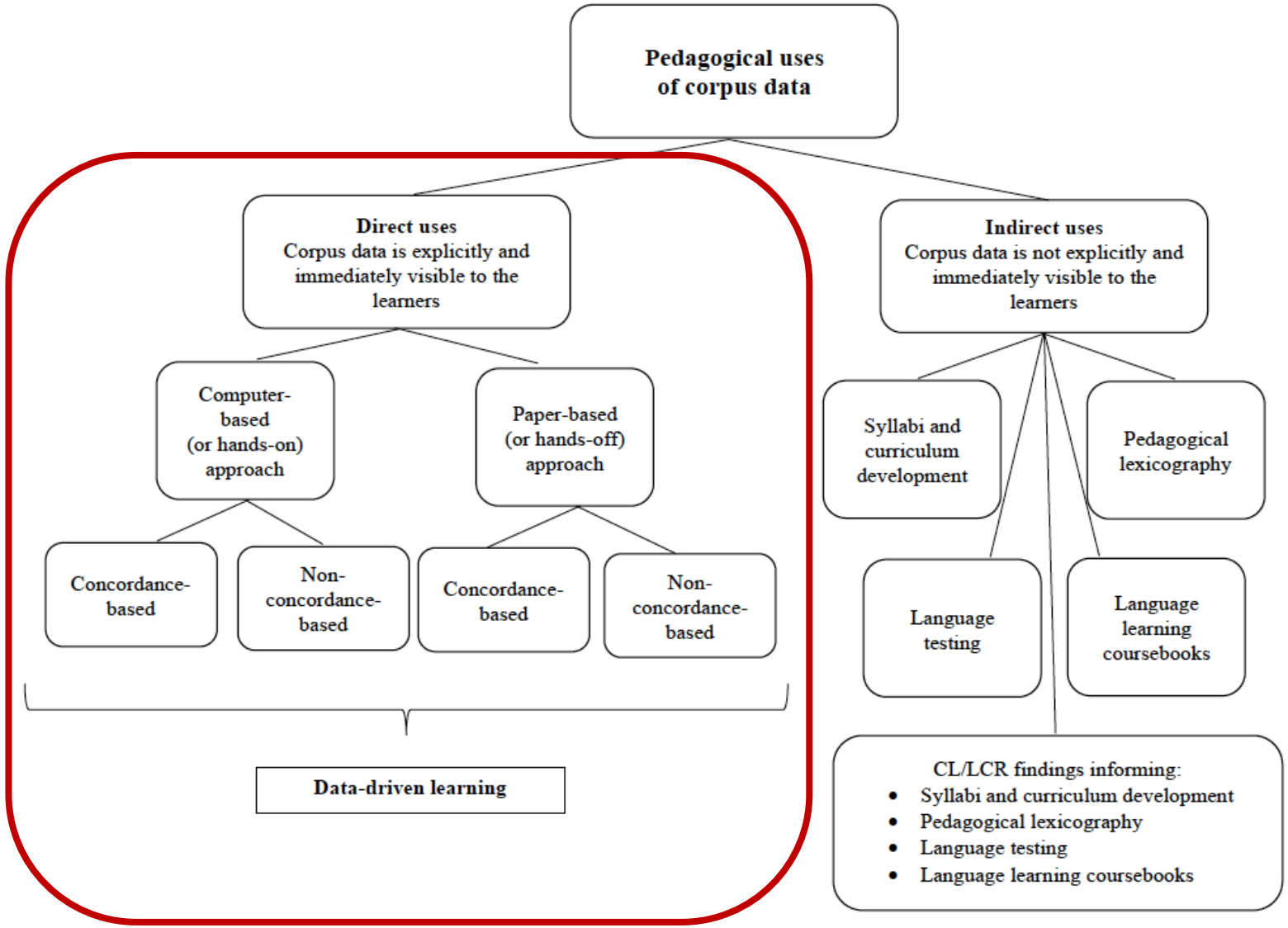
Immediate vs. delayed uses
(Granger, 2009)



Defining Data-driven learning

Direct vs. indirect uses
(Leech, 1997)

Immediate vs. delayed uses
(Granger, 2009)



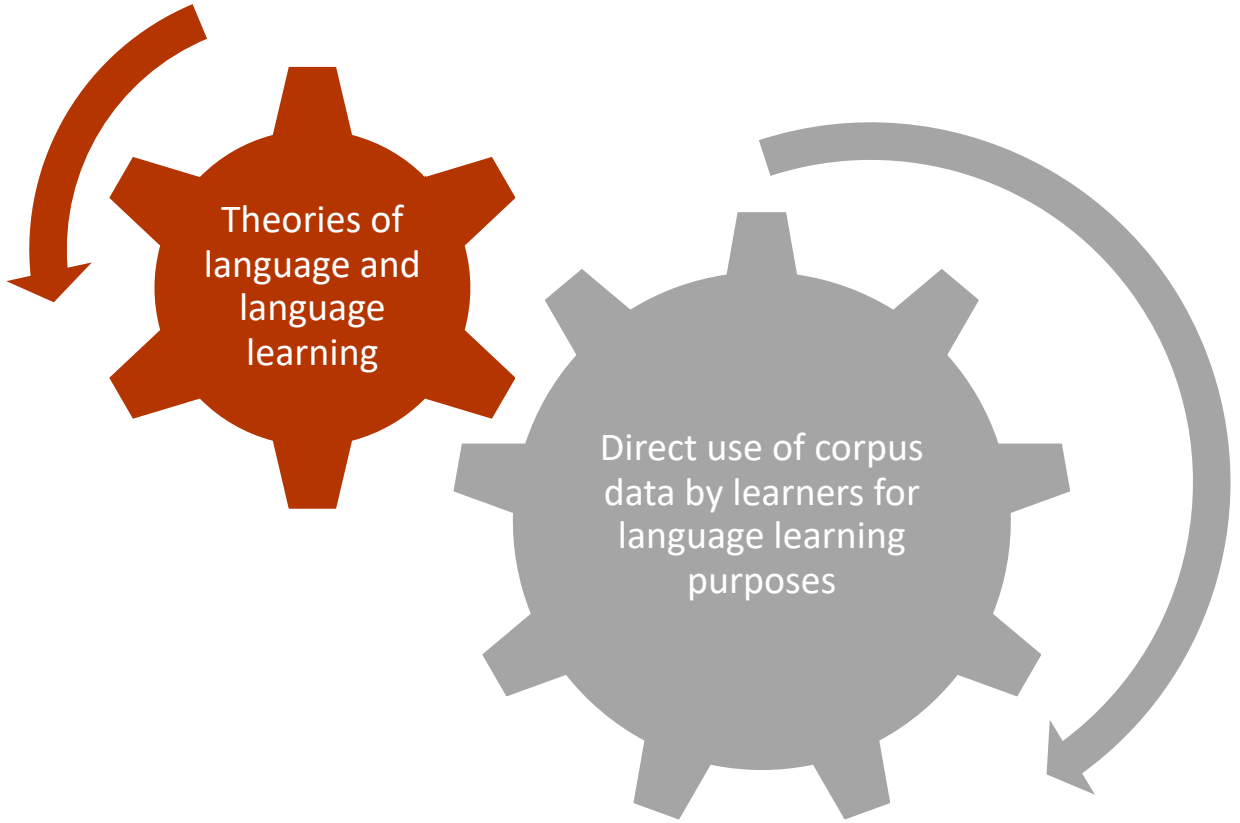
More specific definition and description

Defining Data-driven learning

Theoretical underpinnings of DDL:

- Constructivism
- Sociocultural theory
- Noticing hypothesis
- Usage-based approaches
- …

(Cobb, 1997; Flowerdew, 2005; O'Keeffe, 2021; Pérez-Paredes et al., 2020; Schmidt, 2001)



Defining “DDL effects-oriented research”

Research based on the collection, analysis, and interpretation of empirical data related to the effects of DDL on:

- **learning** with respect to a specific language focus;
- **attitudes** of the learners engaging in corpus-based activities;
- **behaviour** of the learners while engaging in corpus-based activities.

Defining “DDL effects-oriented research”

Research based on the collection, analysis, and interpretation of empirical data related to the effects of DDL on:

- **learning** with respect to a specific language focus;
- **attitudes** of the learners engaging in corpus-based activities;
- **behaviour** of the learners while engaging in corpus-based activities.

Learning aims in DDL effects-oriented research

Language skill

- listening
- speaking
- writing
- reading
- translating

Language aspect

- vocabulary
- lexicogrammar
- grammar
- discourse

Boulton & Cobb, 2017: Lee et al. 2018.

Learning aims in DDL effects-oriented research

- Which challenges do second language learners encounter in these areas?

Second language acquisition
empirical evidence

The case of phraseology

- **Semantic opacity**, e.g. *to make a stand*

(Altenberg & Granger, 2001; Gyllstad & Wolter, 2016)

- **Infrequency in input**, e.g. *to see red*

(Wolter & Gyllstad, 2013; Durrant, 2014)

- **Lack of congruency with the L1**, e.g. *to take a photo* vs. *fare una foto*

(no parallel use in the two languages, Jarvis, 2000, p. 255)

(Granger, 1998; Nesselhauf, 2003; Wang, 2016; Wolter & Gyllstad, 2011; Yamashita & Yang, 2010; Wolter & Gyllstad, 2013; Wolter & Yamashita, 2018)

The case of phraseology

- **Semantic opacity**, e.g. *to make a stand*

(Altenberg & Granger, 2001; Gyllstad & Wolter, 2016)

- **Infrequency in input**, e.g. *to see red*

(Wolter & Gyllstad, 2013; Durrant, 2014)

- **Lack of congruency with the L1**, e.g. *to take a photo* vs. *fare una foto*

(no parallel use in the two languages, Jarvis, 2000, p. 255)

(Granger, 1998; Nesselhauf, 2003; Wang, 2016; Wolter & Gyllstad, 2011; Yamashita & Yang, 2010; Wolter & Gyllstad, 2013; Wolter & Yamashita, 2018)

Can DDL help and make a difference?

The potential of concordance-based DDL

- **For opaque phraseological units**:
  - provides authentic examples, displaying a variety of phraseological uses;
  - learners can identify differences between literal and non-literal uses, by inferring meaning from context.

  (e.g. *raccontare una storia* vs. *raccontare storie*)

The potential of concordance-based DDL

- **For infrequent phraseological units:**
  - provides numerous examples for learners to explore;
  - this can trigger frequency effects, overcoming the potential limitations of infrequency.

The potential of concordance-based DDL

- **For incongruent phraseological units:**
  - provides numerous examples displaying the target language form-meaning mappings;
  - learners can build up awareness on differences with the L1, thus avoiding non-typical combinations (e.g. *dare un sorriso,* instead of *fare un sorriso;* incongruent for L1 Chinese learners).

The potential of concordance-based DDL

- **Overall, for opaque / infrequent / incongruent phraseological units**:

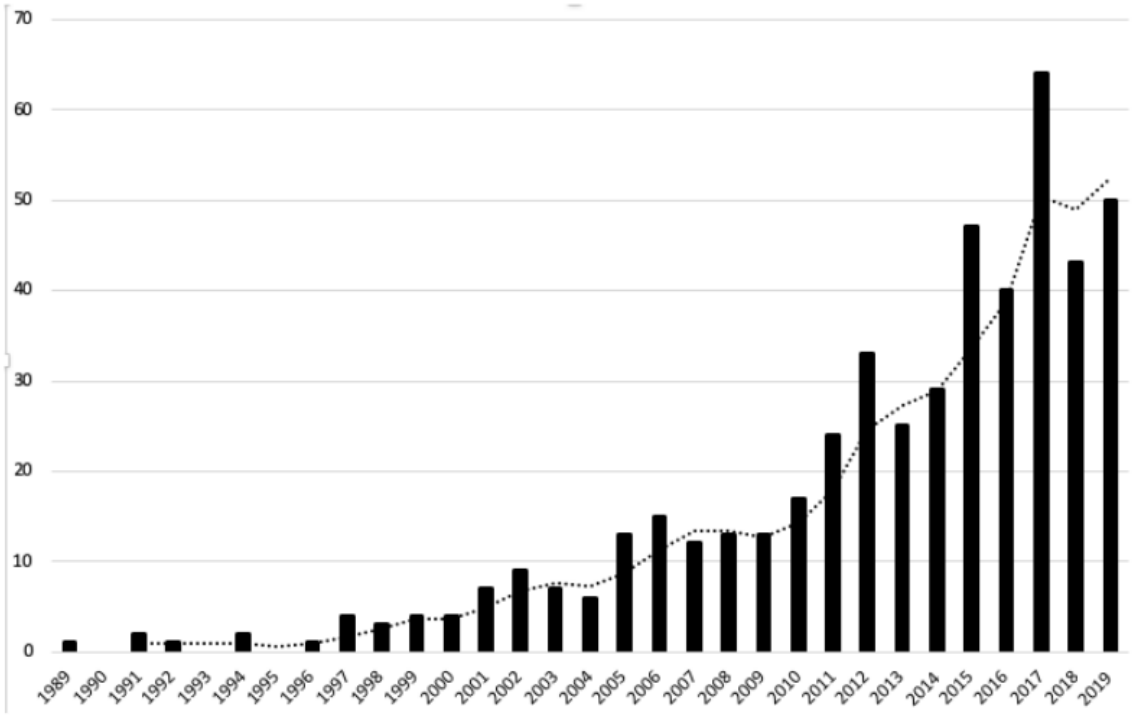
guided-discovery approach,

making learning experiences more memorable.

Setting out to explore DDL effects in an Italian L2 context…

The state of the art

489 papers (1989-2019)



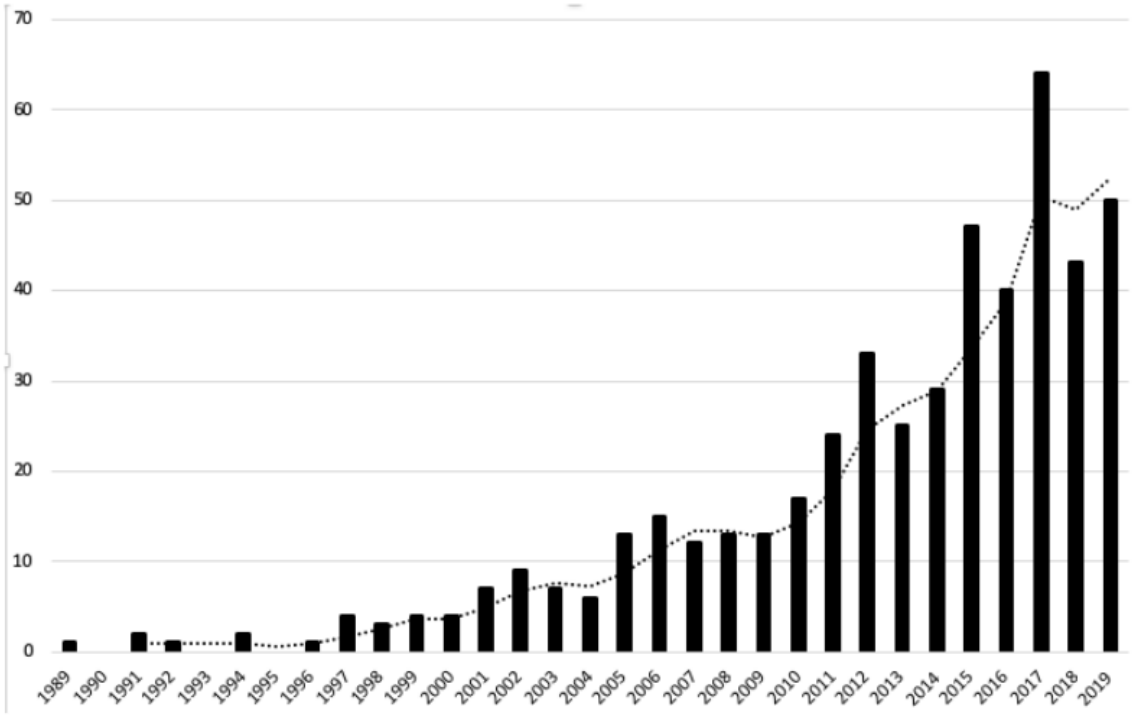
Boulton & Vyatkina, 2021, p. 72.

Setting out to explore DDL effects in an Italian L2 context…

The state of the art

489 papers (1989-2019)

How many focused on L2 Italian?

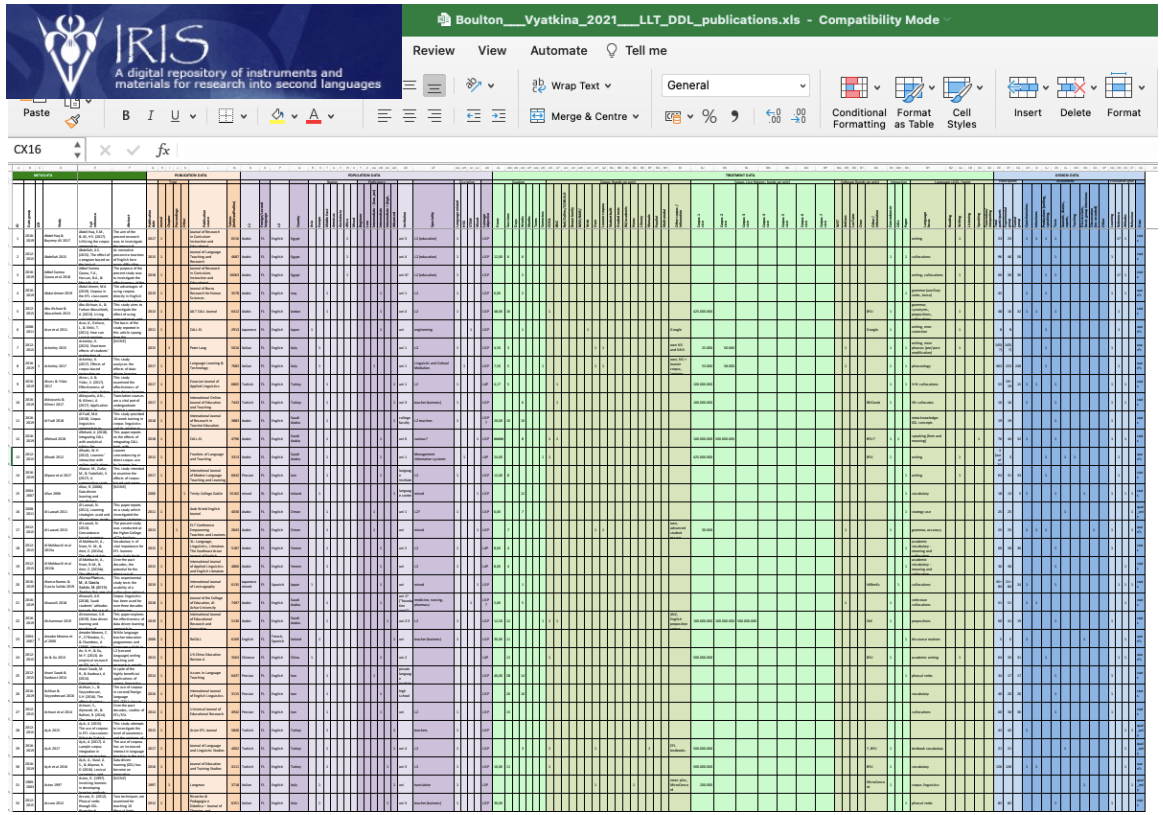


Boulton & Vyatkina, 2021, p. 72.

Setting out to explore DDL effects in an Italian L2 context…

The state of the art

489 papers (1989-2019)



Boulton & Vyatkina, 2021, p. 72.

How many focused on L2 Italian? **9 papers**

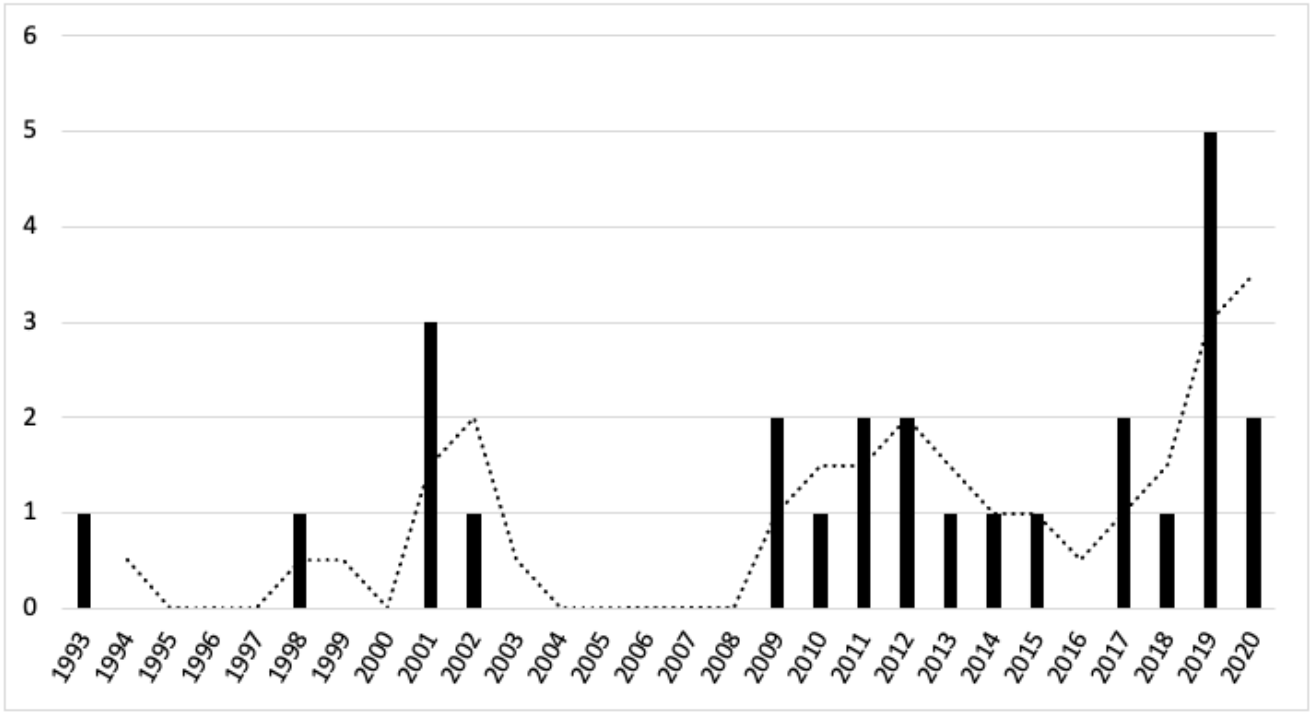
	A	B	C	D	E	F
1	META-DATA					
2	74	2004-2007		Caviglia 2004	Caviglia, F. (2004). Students' diverse	A medium-sized electronic corpus of
3	170	2016-2019		Forti 2017	Forti, L. (2017). Data-driven learning	This paper looks at how corpus data was
4	171	2016-2019		Forti 2019a	Forti, L. (2019a). Learner attitudes	Concordance-based Data-Driven
5	172	2016-2019		Forti 2019b	Forti, L. (2019). Evaluating data-driven	In this article we outline the results of
6	251	1989-2003	1	Kennedy & Miceli 2001 [see also Kennedy &	Kennedy, C., & Miceli, T. (2001). An	This paper reports on our experience
7	252	1989-2003		Kennedy & Miceli 2002 [see also Kennedy &	Kennedy, C., & Miceli, T. (2002). The	The paper discusses the compilation
8	253	2008-2011	1	Kennedy & Miceli 2010	Kennedy, C., & Miceli, T. (2010).	In much of the literature on the
9	254	2016-2019	1	Kennedy & Miceli 2017	Kennedy, C., & Miceli, T. (2017).	While there is widespread agreement
10	333	1989-2003		Miceli & Kennedy 2002 [see also	Miceli, T., & Kennedy, C. (2002). An	This paper discusses the composition

However, only papers published in English…

Setting out to explore DDL effects in an Italian L2 context…

The state of the art

DDL studies (not only effects-oriented) focusing on L2 Italian, published in either English or Italian



Forti, 2023

Total: **26 papers**
(mostly descriptive/introductory)

Setting out to explore DDL effects in an Italian L2 context...

Learner-friendly corpus resources

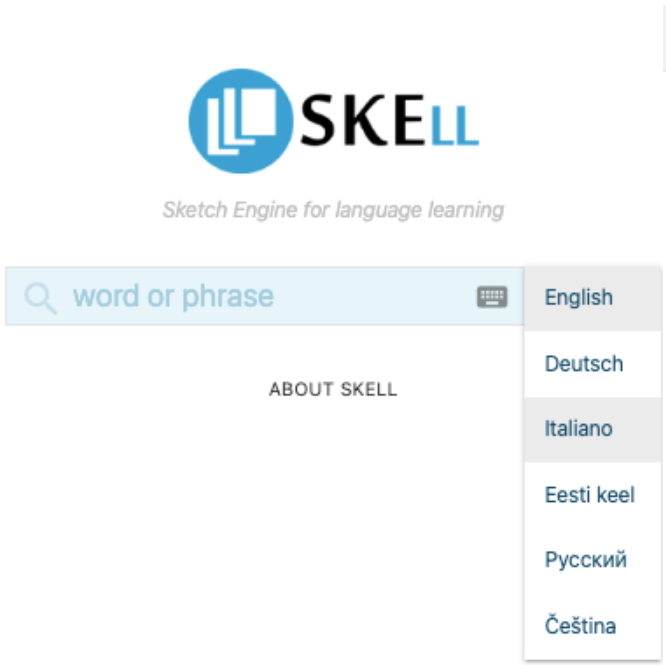
- SkELL (Sketch Engine for Language Learning).
- Selects 40 good examples for learners, from an Italian web corpus of ca. 320 million words.
- Provides “word sketches” showing co-occurrences and similar-word clouds.

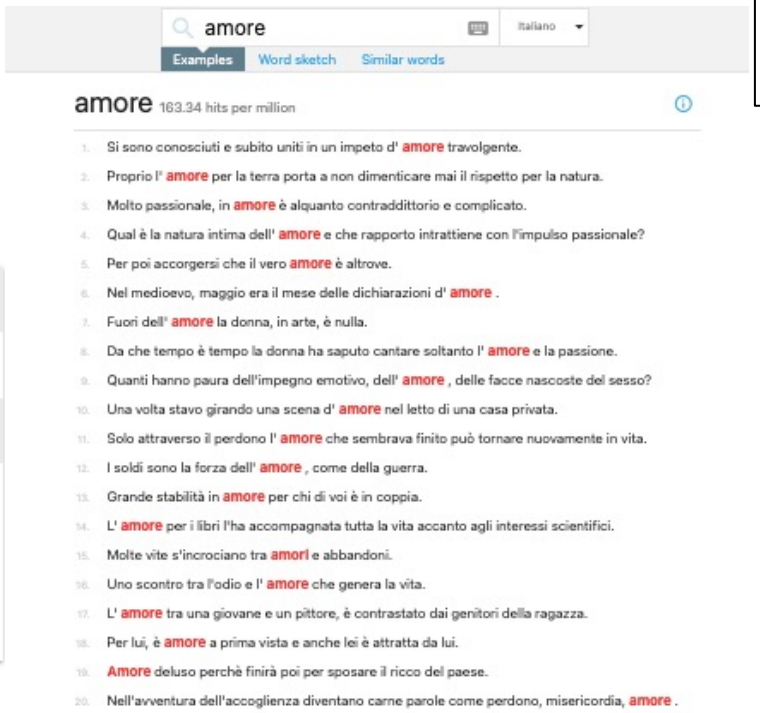
https://skell.sketchengine.eu/#home?lang=en

Setting out to explore DDL effects in an Italian L2 context...

Learner-friendly corpus resources

- SkELL (Sketch Engine for Language Learning).
- Selects 40 good examples for learners, from an Italian web corpus of ca. 320 million words.
- Provides “word sketches” showing co-occurrences and similar-word clouds.



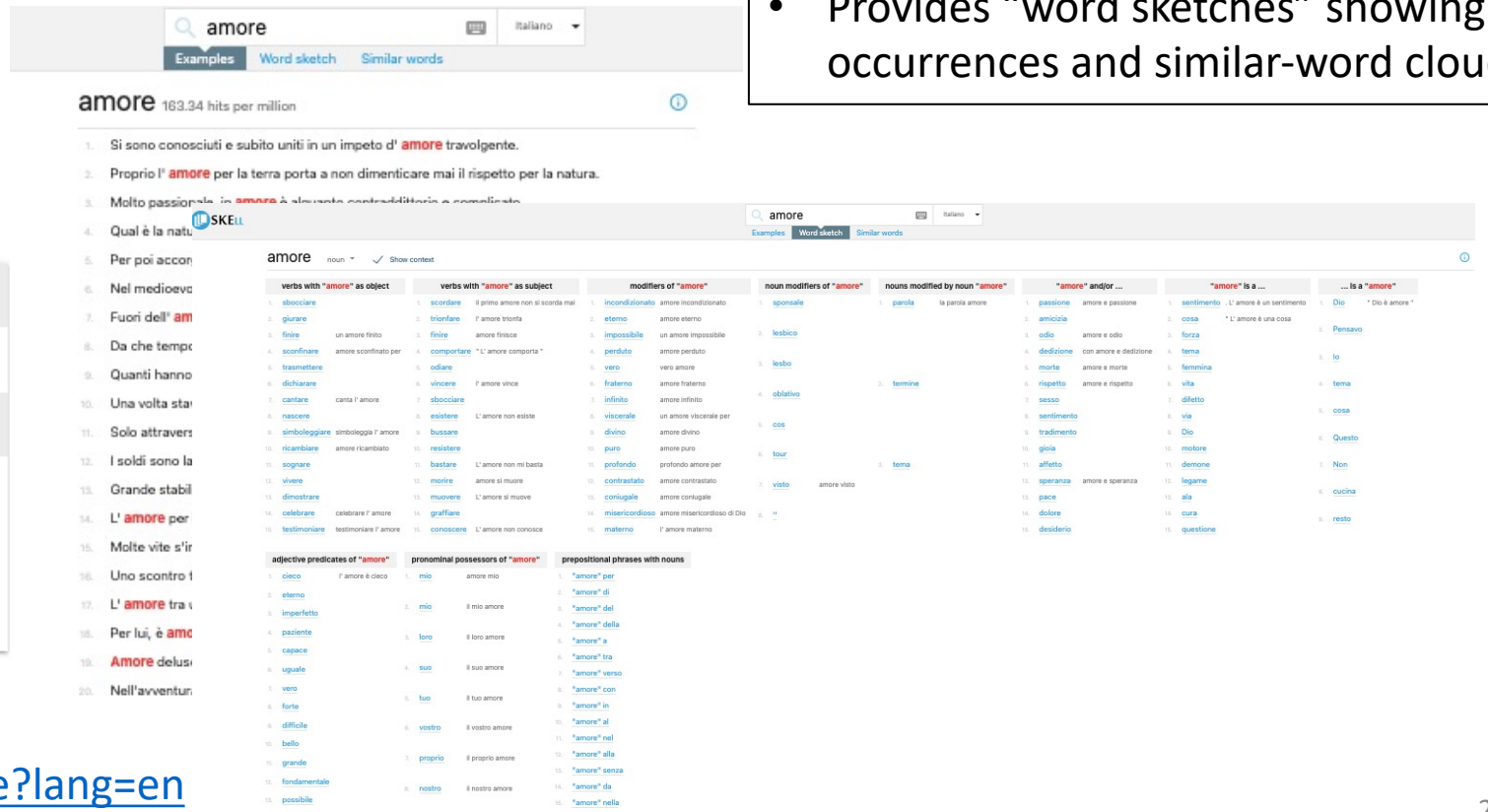


https://skell.sketchengine.eu/#home?lang=en

Setting out to explore DDL effects in an Italian L2 context…

- SkELL (Sketch Engine for Language Learning).
- Selects 40 good examples for learners, from an Italian web corpus of ca. 320 million words.
- Provides “word sketches” showing co-occurrences and similar-word clouds.

Learner-friendly corpus resources

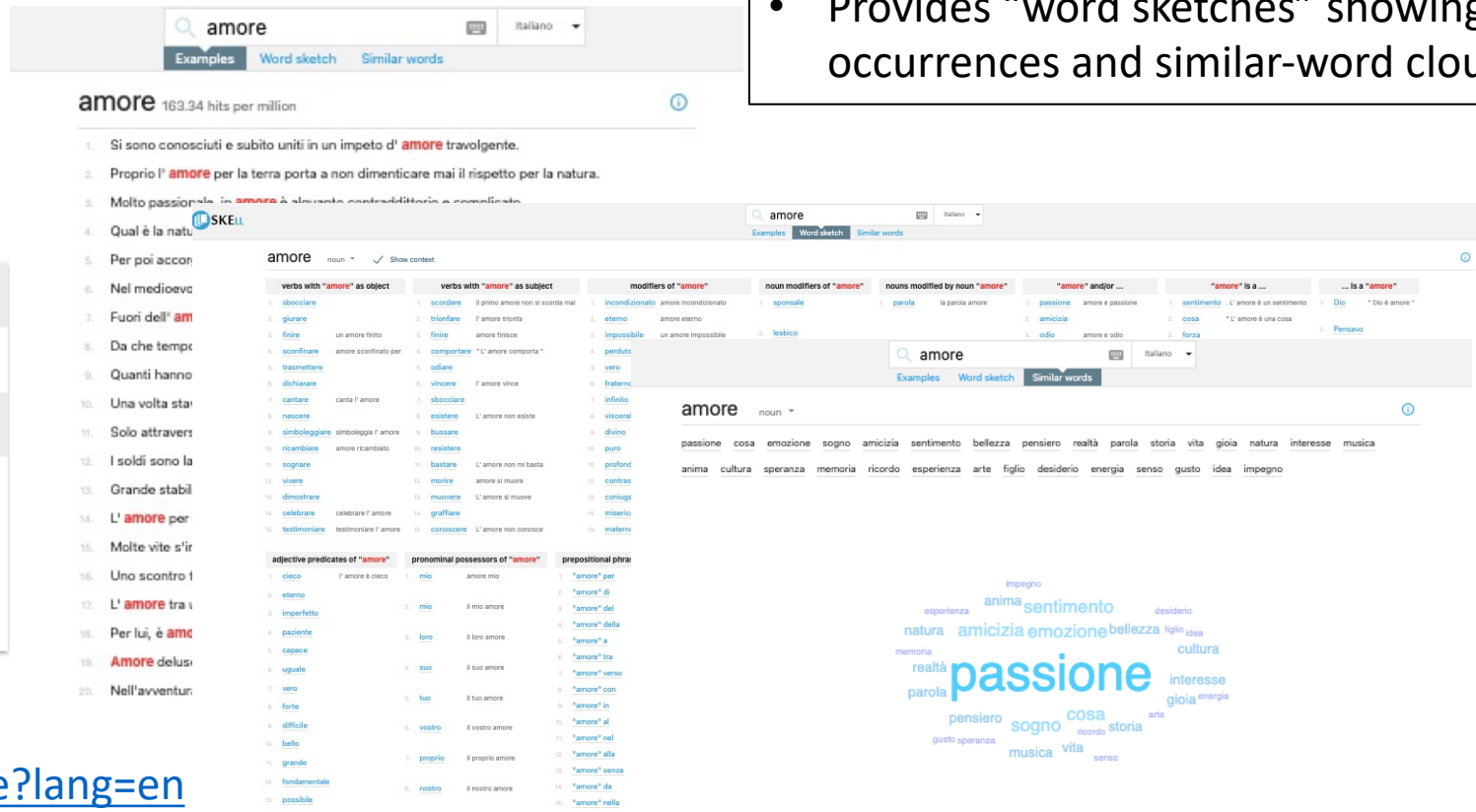


https://skell.sketchengine.eu/#home?lang=en

Setting out to explore DDL effects in an Italian L2 context…

- SkELL (Sketch Engine for Language Learning).
- Selects 40 good examples for learners, from an Italian web corpus of ca. 320 million words.
- Provides “word sketches” showing co-occurrences and similar-word clouds.

Learner-friendly corpus resources



https://skell.sketchengine.eu/#home?lang=en

Setting out to explore DDL effects in an Italian L2 context…

Other corpus resources

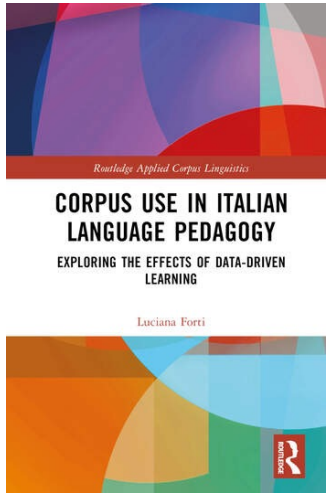
L1 corpora	L2 corpora
Perugia Corpus	Valico
Paisà	Merlin-IT
KiParla	CAIL2
La Repubblica	LIPS
CORDIC	LOCCLI
itTENTEN20	CELI

Require adaptation for most pedagogical contexts.

Exploring DDL effects on L2 Italian phraseological competence development

- Synthesis of research published in several articles and one forthcoming book.

- All are based on or are an extention of my PhD work.



https://www.routledge.com/Corpus-Use-in-Italian-Language-Pedagogy-Exploring-the-Effects-of-Data-driven/Forti/p/book/9780367683634

Research questions

Overall	1. How does DDL influence the learning of collocations overall?
In relation to semantic opacity and frequency	2 a. What effect does DDL have on learning semantically opaque and semantically transparent collocations? 2 b. How does DDL influence the learning of semantically opaque/transparent collocations in relation to their frequency?
In relation to congruency and frequency	3 a. What effect does DDL have on learning congruent and incongruent collocations? 3 b. How does DDL influence the learning of congruent/incongruent collocations in relation to their frequency?

Method: study design

- **Design:**

  Controlled, longitudinal, between-groups.

- **Data collection:**

  Phraseological competence test administered at 4-week intervals, over 12 weeks.

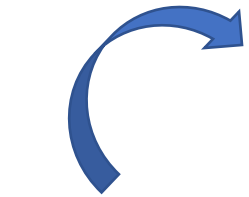
- **Treatment:**

  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 (tot. participants: 123 learners; 61 EXP, 62 CON)

- **Language focus:**

  Verb + Noun (Obj) collocations

Method: selection of learning aims



L2 corpus

LOCCLI
(Longitudinal Corpus of Chinese Learners of Italian; Spina, 2017)

Error analysis of all V+N(obj) combinations

Selection of 32 combinations.

Combinations grouped into 8 themes.

PEC
(Perugia corpus; Spina, 2014)

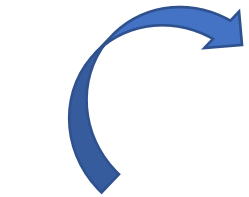
Extraction of all V+N(obj) combinations

Selection of 32 combinations thematically linked to previously identified topics.

L1 corpus

https://www.unistrapg.it/cqpwebnew/

Method: selection of learning aims



L2 corpus

LOCCLI
(Longitudinal Corpus of Chinese Learners of Italian; Spina, 2017)

Error analysis of all V+N(obj) combinations

Selection of 32 combinations.

Combinations grouped into 8 themes.

needs-driven selection

PEC
(Perugia corpus; Spina, 2014)

Extraction of all V+N(obj) combinations

Selection of 32 combinations thematically linked to previously identified topics.

target-driven selection

L1 corpus

https://www.unistrapg.it/cqpwebnew/

Method: weekly learning aims

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

Method: DDL treatment

EXPERIMENTAL VS. CONTROL LESSON STAGES		
5’	Gamified introduction to weekly collocations	
25’	EXPERIMENTAL GROUPS: paper-based DDL activities • Concordance based matching; • Concordance based 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: Sample DDL activity 1

La metà mancante (the missing half)
Qual è la parola che viene dopo *fare*, in ciascuna delle frasi tagliate a metà? (La parola è la stessa in ogni frase)

Era un ragazzo di vent'anni e aveva già **fatto**

Sai se Patrizia ha **fatto**

Credo che il metodo giusto per **fare**

Per me è facile **fare**

Chattare in internet con persone sconosciute è un ottimo modo per **fare**

Apriamo il nostro cuore e **facciamo**

Era tutto nuovo per me, ma pian piano ho cominciato a **fare**

Elvis era ormai adolescente, ma nella nuova città non riusciva a **fare**

L'università è un posto dove si studia, si **fa**

Vedo che state **facendo**

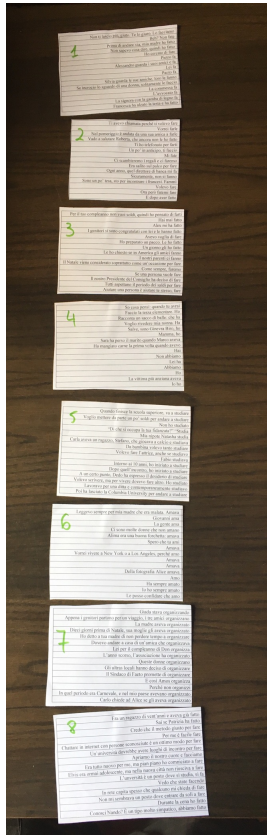
Method: Sample DDL activity 1

La metà mancante (the missing half)
Qual è la parola che viene dopo *fare*, in ciascuna delle frasi tagliate a metà? (La parola è la stessa in ogni frase)

Era un ragazzo di vent'anni e aveva già **fatto**	**amicizia** con quel prete dalla faccia buona.
Sai se Patrizia ha **fatto**	**amicizia** con qualcuno dell'azienda?
Credo che il metodo giusto per **fare**	**amicizia** sia quello di stare in compagnia.
Per me è facile **fare**	**amicizia** con persone straniere.
Chattare in internet con persone sconosciute è un ottimo modo per **fare**	**amicizia**, ma può essere pericoloso.
Apriamo il nostro cuore e **facciamo**	**amicizia** con persone di colore e sicuramente non ci pentiremo.
Era tutto nuovo per me, ma pian piano ho cominciato a **fare**	**amicizia** con tutti.
Elvis era ormai adolescente, ma nella nuova città non riusciva a **fare**	**amicizia** con nessuno.
L'università è un posto dove si studia, si **fa**	**amicizia**, si ride, e a volte si piange.
Vedo che state **facendo**	**amicizia**! Venite a pranzo con noi?

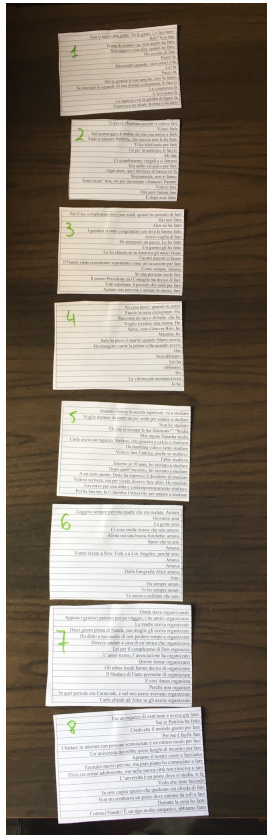
Method: Sample DDL activity 1

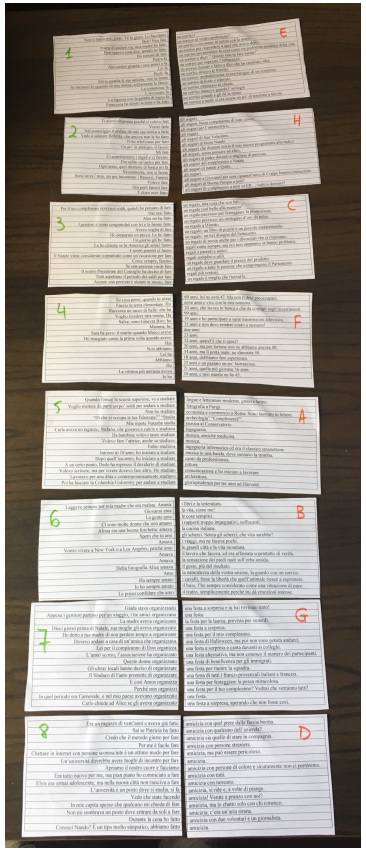
La metà mancante (the missing half)
Insieme agli altri membri del tuo gruppo, trova le metà mancanti per questi otto gruppi di frasi.



Method: Sample DDL activity 1

La metà mancante (the missing half)
Insieme agli altri membri del tuo gruppo, trova le metà mancanti per questi otto gruppi di frasi.





Method: Sample non-DDL activity 1

Quante sono le combinazioni possibili?

Lista di verbi:	Lista di nomi:
fare	economia
studiare	amicizia
avere	sorriso
amare	25 anni
organizzare	sport
	festa
	auguri
	regalo

Scrivile qui sotto:

Method: Sample non-DDL activity 1

Quante sono le combinazioni possibili?

Lista di verbi:

fare
studiare
avere
amare
organizzare

Lista di nomi:

economia
amicizia
sorriso
25 anni
sport
festa
auguri
regalo

Scrivile qui sotto:

Soluzione: 18!

Eccole:

1. Fare + economia
2. Fare + amicizia
3. Fare + sorriso
4. Fare + 25 anni
5. Fare + sport
6. Fare + festa
7. Fare + auguri
8. Fare + regalo
9. Studiare + economia
10. Avere + amicizia
11. Avere + sorriso
12. Avere + 25 anni
13. Avere + regalo
14. Amare + economia
15. Amare + sorriso
16. Amare + sport
17. Amare + regalo
18. Organizzare + festa

Method: Sample DDL activity 2

Multiple-sentence gap-fill

Qual è la parola che manca?

1. ……..

Io esco con Nadia, andiamo a	una passeggiata in collina.
Vorrei	una lunga passeggiata romantica con lui.
Non gli interessava mangiare, bere, dormire,	una passeggiata, voleva solo suonare: suonare e basta.
Ho lasciato la spesa a casa e sono andato a	una passeggiata.
Dopo cena	una passeggiata e siamo andati a bere un'altra birra.
Un vecchio giornalista che conosco da anni mi portò a	una passeggiata tra vari monumenti nel cuore di Washington.

2. ……..

Di giorno stavano sulla riva a	il sole e a nuotare.
Guarda che se io	il sole, sai come divento nero? Ma nero, nero, nero, nero!
Ci sono molte panchine dove le persone possono	il sole, leggere, parlare con gli amici…
Mandaci una cartolina e una foto di te che	il sole sulla spiaggia.
	il sole fa bene: può migliorare l' umore.
Fra un po' vado a	il sole in giardino.

3. ……

Quando uno	una gita all'estero, inizialmente vede tutto più bello.
Angela e Glauco stavano sempre insieme, andavano a	gite in motocicletta, su per le montagne, a nuotare, a sciare…
Marcolino e la sua fidanzata volevano	un viaggio da soli.
Sono ancora indecisa se	un viaggio nel Rajastan, o invece andare a Montecarlo.
Mi piacerebbe	un viaggio con Italia, dormire negli alberghi, fare l' amore, ripartire.

4. ……

Rocchi, vieni a	aria.
Quanto tempo è che non faccio	aria alla casa, tre mesi? Cinque?
Hai bisogno di	aria.
Padre Emanuele lo incitava a	aria fresca.
Volevano mandarla a	aria buona nelle fotezze alpine.
Esco fuori a	aria.

5. ……

Domani devo andare a lavorare e poi al ritorno dovrò	casa, lavare, stirare...
Mia nonna lavava i piatti o	la casa e poi si metteva sul divano e riposava.
Cantava le opere mentre	la casa.
Prima di andare via, ha	casa e sul tavolo in cucina un biglietto per me.
Vado a prepararmi,	casa, e poi andiamo fuori a festeggiare il compleanno di mio fratello.
Quella sera, mentre	la casa, Ernestina disse: "Brioschi sembra proprio una brava persona".

6. ……

Noi dobbiamo andare a casa.	fretta.
Salvatore, non è giornata! Fai guidare me! Levati!	fretta! Dai, fai guidare me! Guido io!
Potete rispondere quando volete, non	fretta. Grazie.
Ti richiamo domani. Adesso	fretta. Sto andando da un cliente
	fretta di abbandonare questo assurdo posto.
Io non	fretta. Ho solo la necessità di capire.

7. ……

La crisi economica ha messo in pericolo la tradizione natalizia di	soldi in regali costosi.
Quando i virus entrano nel computer, poi i nostri genitori	soldi per farlo riparare .
Quando mi regalano le figurine, le accetto perché così non devono	soldi in edicola .
A volte, si	soldi per cercare l'introvabile o l'inguardabile.
Purtroppo si continua a	soldi tra centrocampisti e attaccanti, e non per la difesa.
Si deve mostrare nei fatti come si	i soldi.

8. ……

Non ci possiamo sedere, devo	la spesa.
Io porto i bambini a scuola e poi vado a	la spesa con Eminè.
La domenica mattina lui va a	spese in un paesino vicino a dove abita.
Esco solo per	la spesa e così incontro qualche vecchio amico.
Chi deve lavare, pulire, stirare, cucinare,	la spesa? Io, sempre e solo io!
Senti cosa è successo.	la spesa per cinque persone e invece siamo in tre!

Method: Sample non-DDL activity 2

Scrivi la parola che manca.

1. Gli puoi inviare un fax, senza _____ soldi.
2. Porto i bambini a scuola e poi vado a _____ la spesa con Alessia.
3. Quand'è l'ultima volta che abbiamo fatto ____ passeggiata?
4. Se io _____ il sole, divento nero nero nero!
5. Angela e Glauco amavano fare _____ in motocicletta, su per le montagne.
6. Sono uscito a _____ aria.
7. Io non _____ fretta, voglio solo capire.
8. Questa sera ho ospiti cena, quindi adesso devo _____ casa.

Advantages of paper-based DDL

- Adaptation to proficiency level (curation of data) and context (lack of computers).
- Integration / normalisation with other paper-based materials.
- Flexible use.
- Corpus input tightly controlled (good for empirical studies and replications).

Method: data collection tool

Method: weekly learning aims

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

64-item phraseological competence test

32 target combinations based on LOCCLI

Multiple choice items

1. In estate vorrei…

a. fare viaggio
c. fare un viaggio
b. fare la viaggio
d. nessuna di queste

32 target combinations based on PEC

Gap-fill items

1. Tra poco, il film al cinema inizierà. _____ una doccia e andiamo.

Method: Data collection outline

Data collection point	Week	Collocation set	Experimental groups	Control groups
1	0	n/a	Getting to know each other activities Background questionnaire **Phraseological competence test 1**	
	1	1	DDL activities	Non-DDL activities
	2	2	DDL activities	Non-DDL activities
	3	3	DDL activities	Non-DDL activities
	4	4	DDL activities	Non-DDL activities
2	5	n/a	**Phraseological competence test 2**	
		5	DDL activities	Non-DDL activities
	6	6	DDL activities	Non-DDL activities
	7	7	DDL activities	Non-DDL activities
	8	8	DDL activities	Non-DDL activities
3		n/a	**Phraseological competence test 3**	
	9-12		No lessons	
4	12	n/a	**Phraseological competence test 4**	
			End-of-course questionnaire for experimental groups	End-of-course questionnaire for control groups

→ **Retention rates**

Method: data coding

- **Semantic transparency:** 13 raters asked to rate 64 collocations according to categories in Howarth's *Phraseological continuum model* (Howarth, 1998) – sample of 32 selected (Krippendorf's alpha = 0. 742)

  e.g.

  transparent: *ascoltare la musica,* 'to listen to music'

  opaque: *avere un'idea, '*to have an idea'

- **Congruency:** list of 64 target collocations categorised for L1 congruency by two different expert L1 Chinese speakers.

- **Phrasal frequency:** operationalised as overall number of co-occurring sequences present in a reference corpus of Italian (Spina, 2014)

  Example of CQP query:
  [lemma="fare"][pos="ADV.*"]?[pos="ART"]?[pos="ADJ|DET.*|NUM|PREDET"]?[lemma="passeggiata"]

Method: data analysis

Mixed-effects modeling

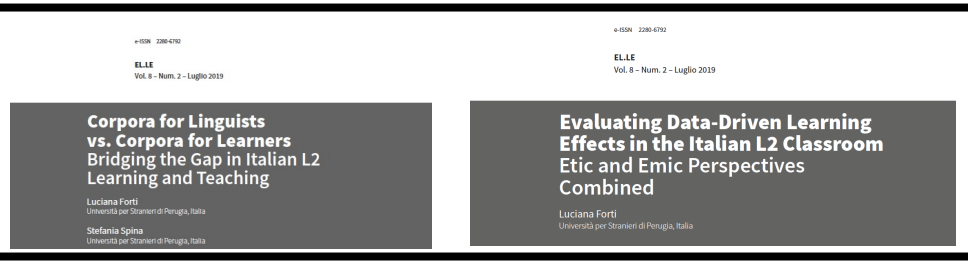
Outcome variable	Levels
Accuracy	Correct Incorrect

Interacting variables	Levels
Test Number	Test 1 Test 2 Test 3 Test 4
Item Type 1	Transparent Opaque
Item Type 2	Congruent Incongruent
Condition	Control Experimental
Frequency	(continuous)

Random effects
Student ID
Item ID

Main challenges in conducting the study

- Adapting the data from the Perugia Corpus to suit a pre-intermediate proficiency level (no it-SkELL at the time!).
- Expected level vs. actual level.
- Constructing DDL activities so that they varied enough within a lesson.
- Teachers teaching in language courses:
  - difficulties in getting accepted (1 hour of class a week devoted to my project…);
  - difficulties in getting teachers to collaborate in the (simultaneous!) administration of the end-of-course questionnaire.



«Data-driven Learning: la linguistica dei corpora al servizio della didattica delle lingue straniere e del CLIL» a cura di Elisa Corino
https://edizionicafoscari.unive.it/it/edizioni4/riviste/elle/2019/2/

Research questions

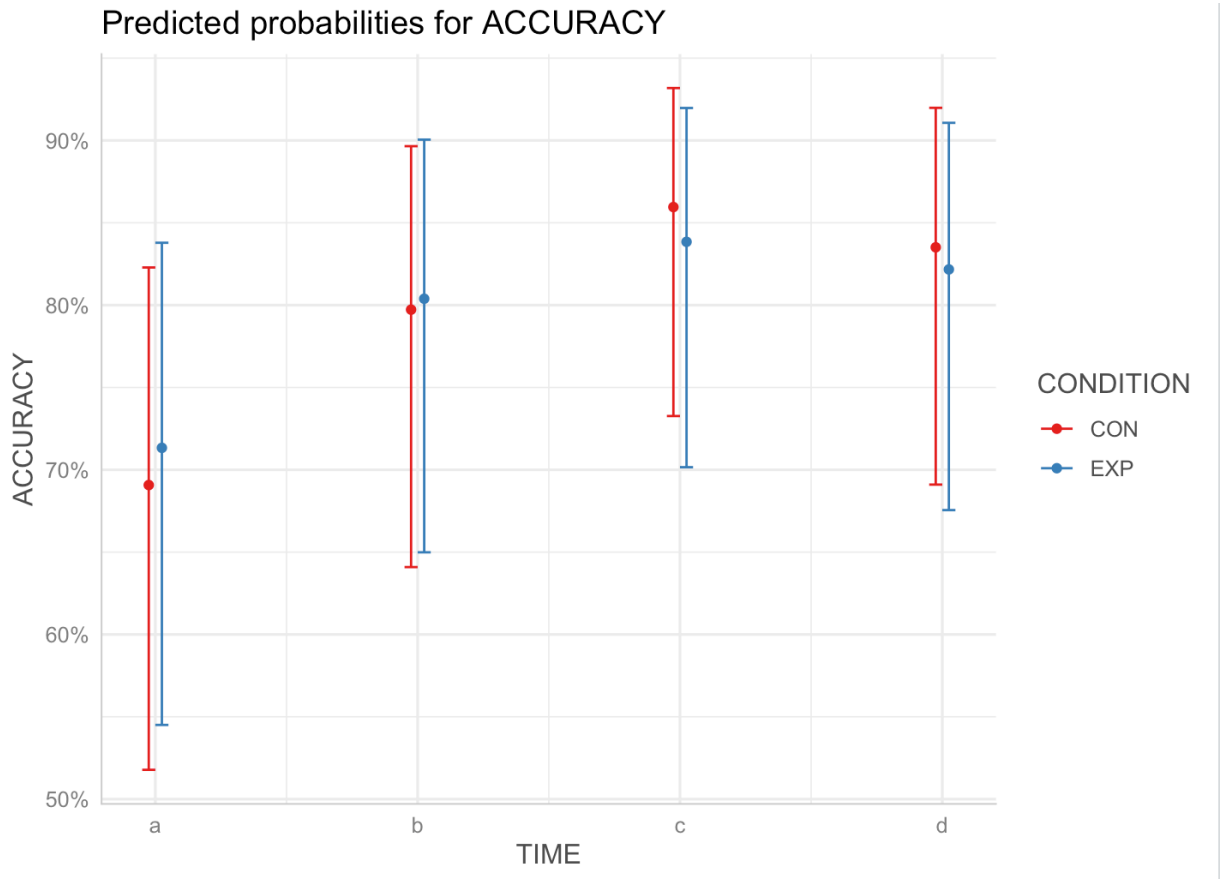
Overall	1. How does DDL influence the learning of collocations overall?
In relation to semantic opacity and frequency	2 a. What effect does DDL have on learning semantically opaque and semantically transparent collocations? 2 b. How does DDL influence the learning of semantically opaque/transparent collocations in relation to their frequency?
In relation to congruency and frequency	3 a. What effect does DDL have on learning congruent and incongruent collocations? 3 b. How does DDL influence the learning of congruent/incongruent collocations in relation to their frequency?

Results

Overall

RQ 1. How does DDL influence the learning of collocations overall?

- Very similar U-shaped learning curves in both conditions.



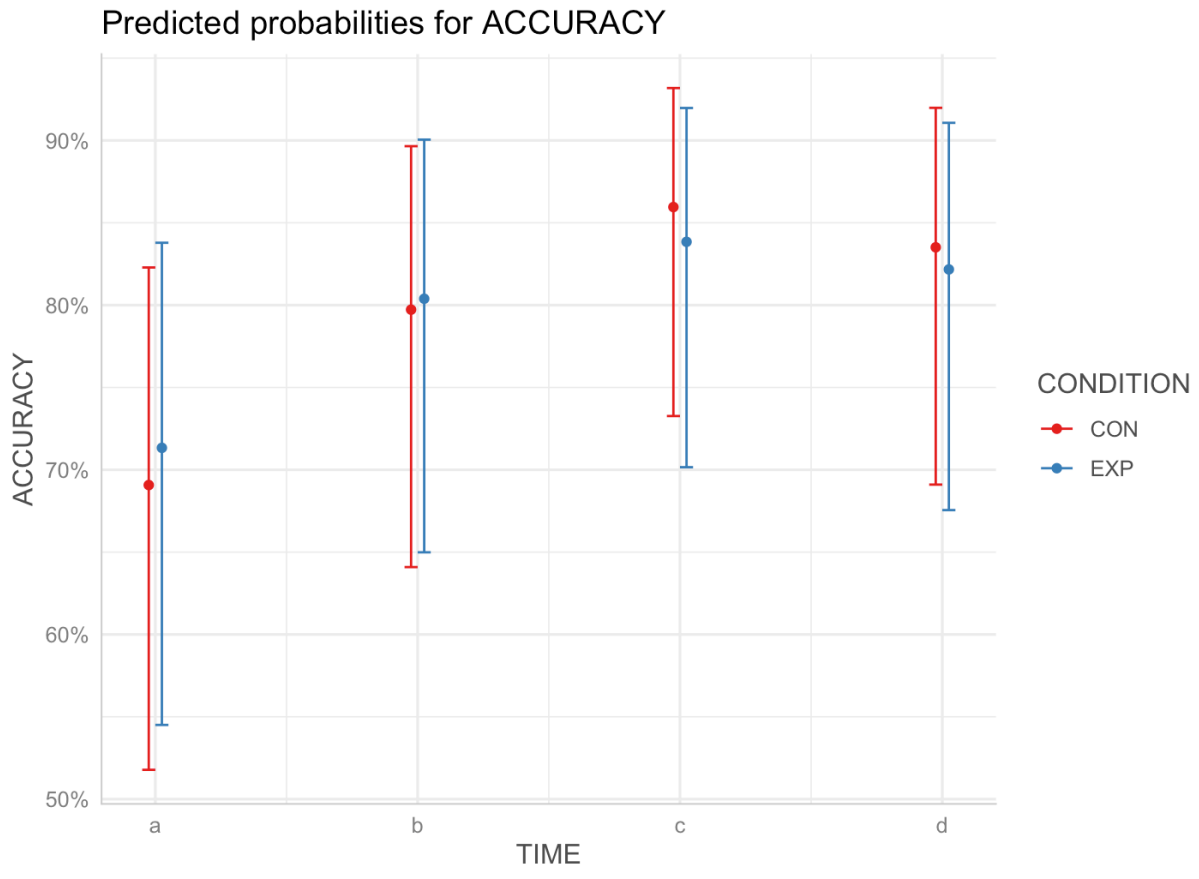
Results

Overall

RQ 1. How does DDL influence the learning of collocations overall?

- Very similar U-shaped learning curves in both conditions.

What if we include semantic transparency into the picture?

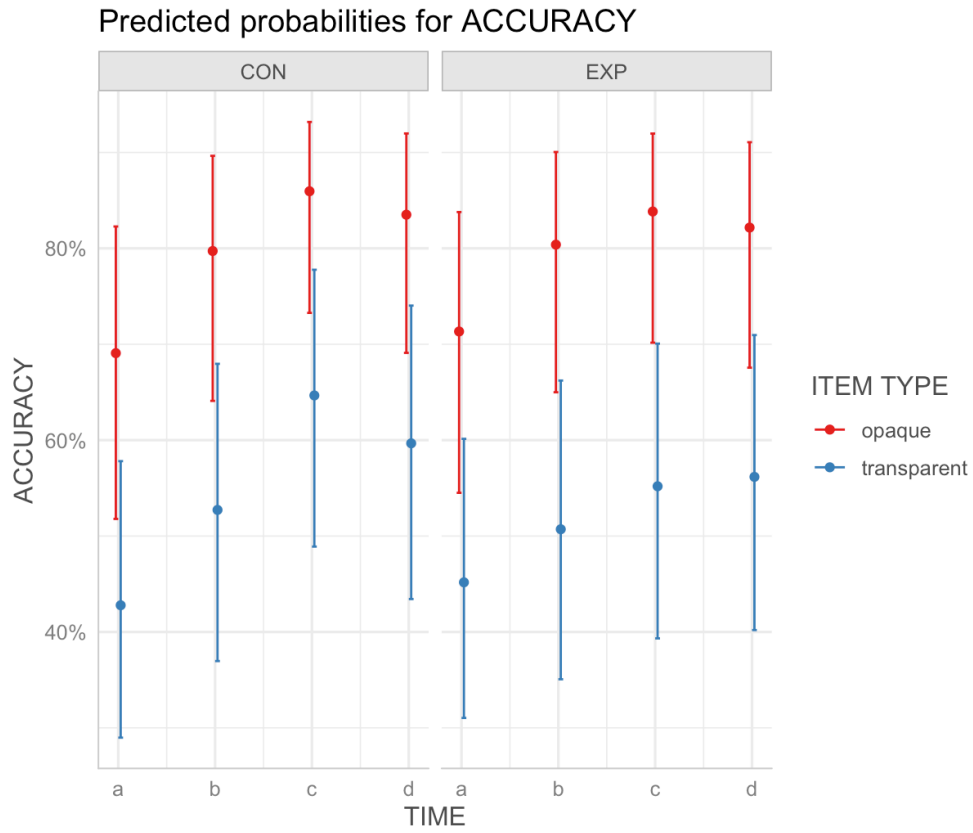


Results

In relation to
semantic opacity and frequency

RQ 2 a. What effect does DDL have on learning semantically opaque and semantically transparent collocations?

- U-shaped learning curves in both conditions, except for transparent items in the EXP condition.
- Opaque collocations with higher predicted probabilities for accuracy.



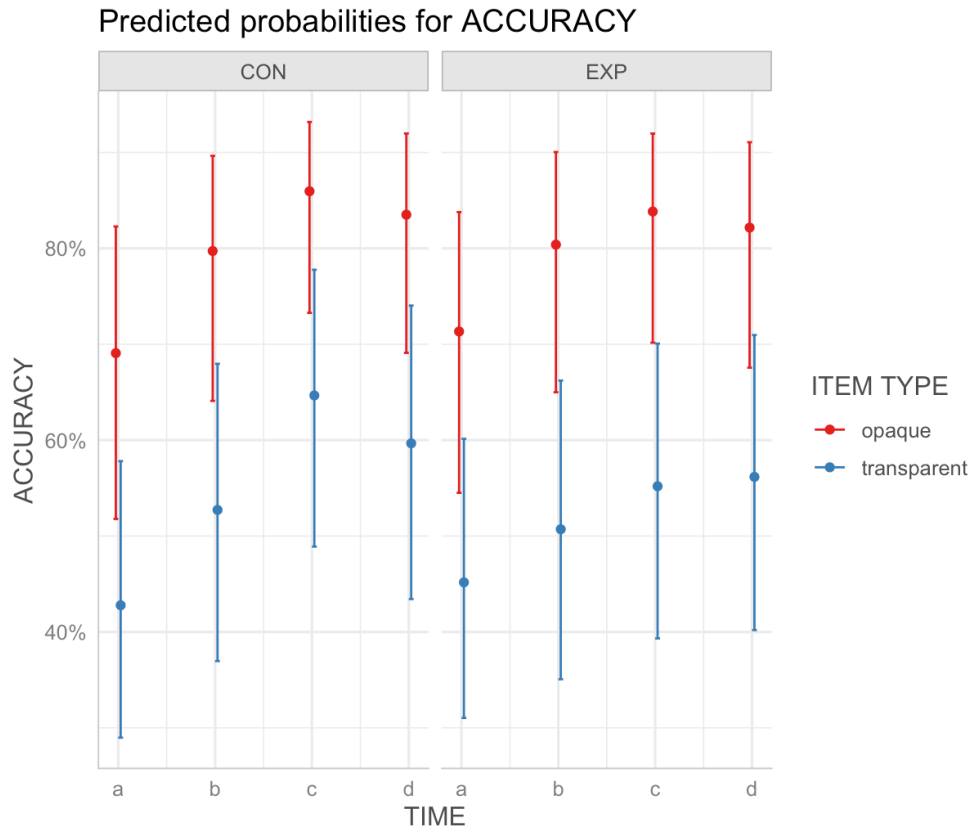
Results

In relation to
semantic opacity and frequency

RQ 2 a. What effect does DDL have on learning semantically opaque and semantically transparent collocations?

- U-shaped learning curves in both conditions, except for transparent items in the EXP condition.
- Opaque collocations with higher predicted probabilities for accuracy.

What if we also consider phrasal frequency?



Results

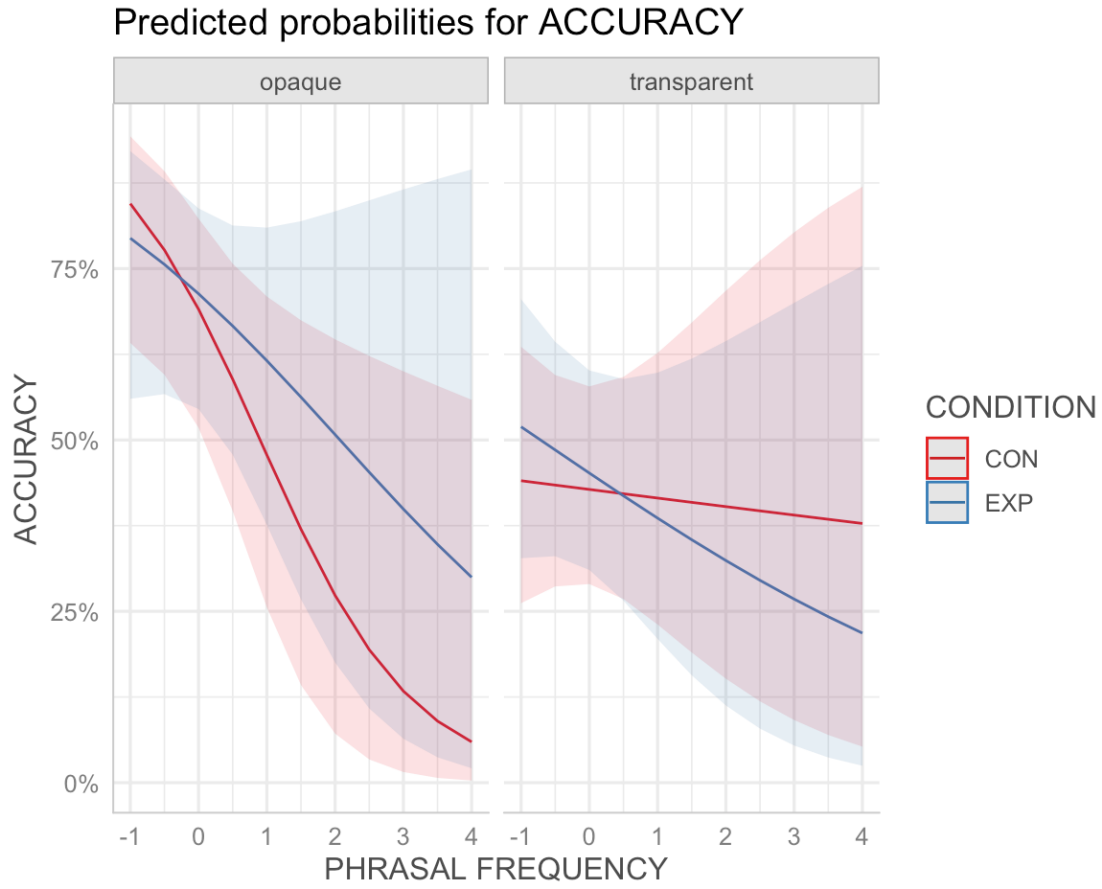
In relation to
semantic opacity and frequency

RQ 2 b. How does DDL influence the learning of semantically opaque/transparent collocations in relation to their frequency?

- As frequency increases, accuracy decreases in both opaque and trasparent collocations.

however:

- In opaque collocations, the decrease is not as steep in the EXP condition as it is in the CON condition;
- in the transparent collocations, the steeper decrease is seen in the EXP condition.

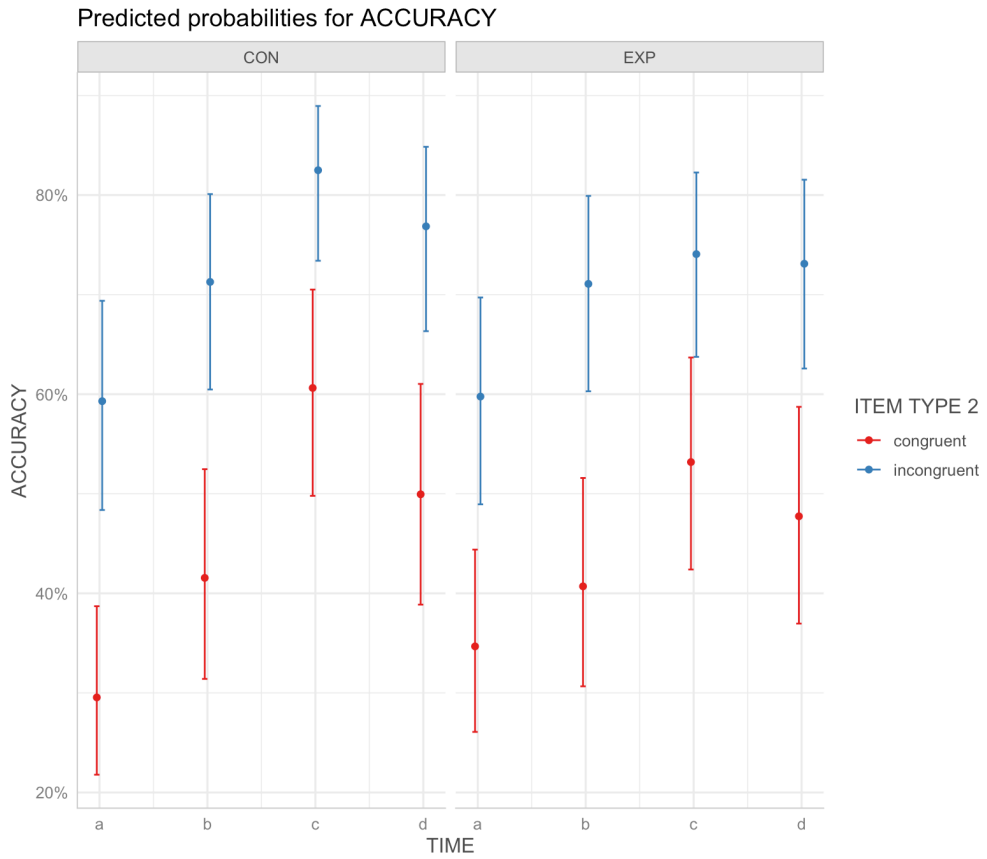


Results

In relation to
congruency and frequency

RQ 3 a. What effect does DDL have on learning congruent and incongruent collocations?

- Both conditions exhibit U-shaped learning patterns.
- Incongruent collocations are learned significantly better in both groups.
- The best retention rates (differences between time c and d) are observed for incongruent collocations in the EXP condition.

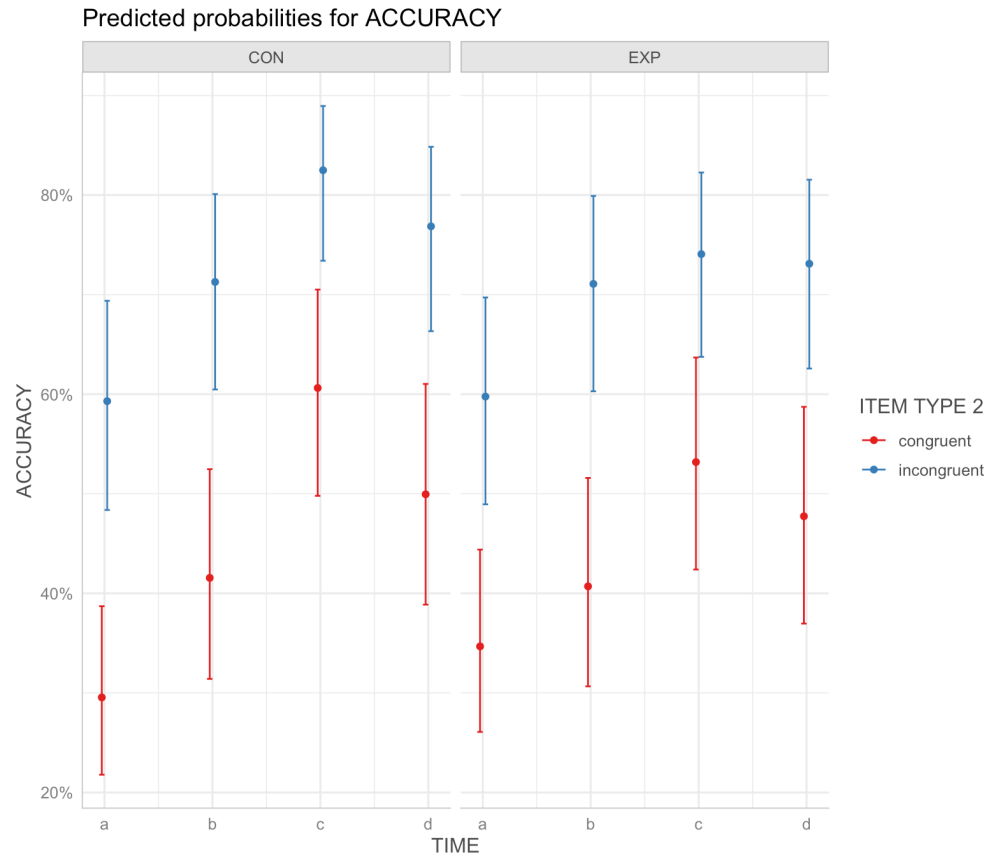


Results

In relation to
congruency and frequency

RQ 3 a. What effect does DDL have on learning congruent and incongruent collocations?

- Both conditions exhibit U-shaped learning patterns.
- Incongruent collocations are learned significantly better in both groups.
- The best retention rates (differences between time c and d) are observed for incongruent collocations in the EXP condition.



What if we also consider phrasal frequency?

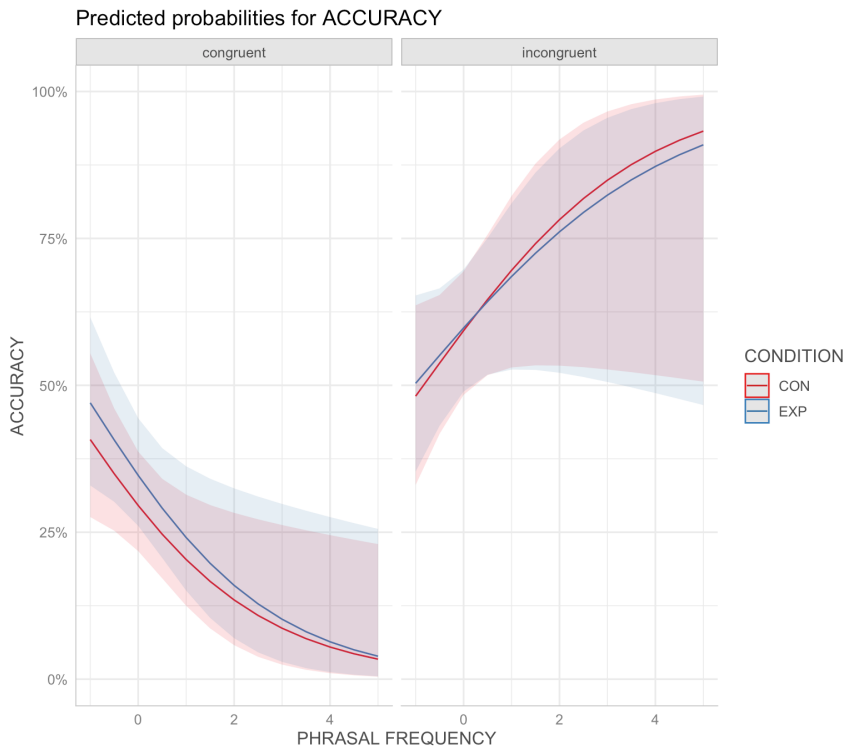
Results

In relation to
congruency and frequency

RQ 3 b. How does DDL influence the learning of congruent/incongruent collocations in relation to their frequency?

- Interaction between phrasal frequency and L2 congruency is significant.

- As phrasal frequency increases, predicted probabilities for accuracy:
  - decrease for congruent collocations;
  - increase for incongruent collocations.



Discussion

➢**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, p. 25).
- Absence of a preliminary training opportunity: no significant differences (Lee et al. 2018, p. 25).

Discussion

➢**Incongruent collocations are retained better in a DDL setting.**

- Engaging in DDL activities may lead to longer lasting effects in terms of the robustness of learning, thanks to repeated encounters with multiple examples containing the same combination and heavier cognitive load than traditional activities.

Discussion

➢**Opaque and infrequent collocations exhibit better accuracy rates in the DDL rather than the non-DDL setting.**

- Numerosity of examples -> frequency effects

- Deriving meaning from multiple contexts to overcome semantic opacity more effective than not having multiple contexts

Possible future steps

- Conduct the analysis with other quantitative measures of collocations (e.g. **Delta P, log-dice** etc.)
- More focus on **retention rates**, to see how the principles of concordance-based DDL affect memorability across the different properties of the learning aims.
- Investigate **differences among the 8 classes** of students, to see whether there are any changes in comparison to the overall patterns observed .
- Ultimately try to identify **the conditions that make concordance-based DDL more effective** than other approaches.

Conclusions

Informing effects-oriented DDL research with SLA evidence-driven research questions helps in:

- having a more nuanced view of DDL effects;
- making DDL research relevant to current L2 research findings related to second language learner needs;
- creating links that can more easily guide language teachers towards the observation of how DDL can help in language teaching practices, thus contributing to bridge the research/teaching gap (Chambers, 2019; Pérez-Paredes, 2019).

References 1/2

Altenberg, B., & Granger, S. (2001). The grammatical and lexical patterning of MAKE in native and non-native student writing. *Applied Linguistics*, *22*(2), 173–195.

Boulton, A. (2008). DDL: Reaching the parts other teaching can't reach? In A. Frankenberg-Garcia (Ed.), *Proceedings of the 8 Teaching and Language Corpora Conference* (pp. 38–44). Lisbon: Associação de Estudos e de Investigação Cientifíca do ISLA-Lisboa.

Boulton, A., & Vyatkina, N. (2021). Thirty years of data-driven learning: Taking stock and charting new directions. *Language Learning and Technology*, *25*(3), 66–89.

Chambers, A. (2019). Towards the corpus revolution? Bridging the research–practice gap. *Language Teaching*, 1–16.

Cobb, T. (1999). Applying constructivism: A test for the learner-as-scientist. *Educational Technology Research and Development*, *47*(3), 15–31.

Durrant, P. (2014). Corpus frequency and second language learners' knowledge of collocations: A meta-analysis. *International Journal of Corpus Linguistics*, *19*(4), 443–477.

Flowerdew, L. (2015). Data-driven learning and language learning theories: Whither the twain will meet. In A. Leńko-Szymańska & A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 15–36). Benjamins.

Forti, L. (forthcoming, 2023). *Corpus Use in Italian Language Pedagogy. Exploring the effects of Data-driven learning*. Routledge

Gilquin, G., & Granger, S. (2010). How can data-driven learning be used in language teaching? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (1st ed, pp. 359–370). Routledge.

Granger, S. (2009), The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation. Corpora and Language Teaching, Corpora and Language Teaching. Ed. Karin Aijmer. Studies in Corpus Linguistics 33. Amsterdam: John Benjamins, 13-33.

Granger, S. (Ed.). (1998). *Learner English on Computer*. Longman.

Gyllstad, H., & Wolter, B. (2016). Collocational Processing in Light of the Phraseological Continuum Model: Does Semantic Transparency Matter?: Collocational Processing and Semantic Transparency. *Language Learning*, *66*(2), 296–323.

Jarvis, S. (2000). Methodological rigor in the study of transfer: Identifying L1 influence in the interlanguage lexicon. *Language Learning*, *50*(2), 245–309.

References 2/2

Lee, H., Warschauer, M., & Lee, J. H. (2018). The Effects of Corpus Use on Second Language Vocabulary Learning: A Multilevel Meta-analysis. *Applied Linguistics*, *40*(5), 721–753.

Leech, G. (1997). Teaching and language corpora: A convergence. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and Language Corpora* (pp. 1–24). Addison Wesley Longman.

Nesselhauf, N. (2003). The Use of Collocations by Advanced Learners of English and Some Implications for Teaching. *Applied Linguistics*, *24*(2), 223–242.

O'Keeffe, A. (2021). Data-driven learning – a call for a broader research gaze. *Language Teaching*, *54*(2), 259–272.

Pérez-Paredes, P. (2019). A systematic review of the uses and spread of corpora and data-driven learning in CALL research during 2011–2015. *Computer Assisted Language Learning*, 36–61.

Pérez-Paredes, P., Mark, G., & O'Keeffe, A. (2020). *The impact of usage-based approaches on second language learning and teaching* [Cambridge Education Research Reports].

Schmidt, R. W. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction*. Cambridge University Press.

Spina, S. (2014). Il Perugia Corpus: Una risorsa di riferimento per l'italiano. Composizione, annotazione e valutazione. *Proceedings of the First Italian Conference on Computational Linguistics CLiC-It 2014 & the Fourth International Workshop EVALITA 2014*, *1*, 354–359.

Wang, Y. (2016). *The Idiom Principle and L1 Influence. A contrastive learner-corpus study of delexical verb+noun collocations*. John Benjamins Publishing Company.

Wolter, B., & Gyllstad, H. (2011). Collocational Links in the L2 Mental Lexicon and the Influence of L1 Intralexical Knowledge. *Applied Linguistics*, *32*(4), 430–449.

Wolter, B., & Gyllstad, H. (2013). Frequency of input and L2 collocational processing: A comparison of congruent and incongruent collocations. *Studies in Second Language Acquisition*, *35*(03), 451–482.

Wolter, B., & Yamashita, J. (2018). Word frequency, collocational frequency, L1 congruency, and proficiency in L2 collocational processing: What accounts for L2 performance? *Studies in Second Language Acquisition*, *40*(2), 395–416.

Thank you!

luciana.forti@unistrapg.it
@l_for_ti