

Data-driven learning and second language acquisition – it's time to connect

Dr Anne O'Keeffe,
Mary Immaculate College
University of Limerick



DDL works well for:



Undergraduates
and graduates



Local and large
corpora



Hands-on and
paper-based



Intermediate to
advanced levels

Boulton and Cobb (2017: 386) based on meta-study analysed 64 empirical studies

DDL works



Why doesn't DDL work?

How can DDL work across levels?



Language experience + world experience

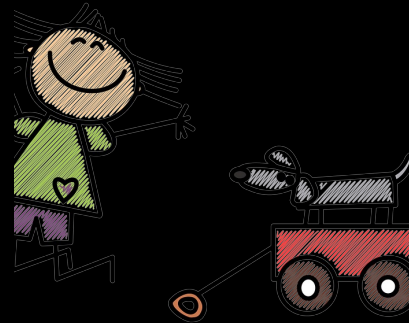
Mammy's gone.

Mammy's gone in the car.

All gone.

Where's daddy gone?

It's gone!



The car's gone,

Where's the car gone?

The juice is gone,

Where's the Lego gone?

**Usage-Based
Model**

Usage-based model

Our knowledge of language comes from experiencing and using it as part of a communicatively-rich human social environment (Ellis & Larsen-Freeman 2006)



Language knowledge *is* language use. (Pérez-Paredes, Mark & O’Keeffe 2020)

Constructions

a construction is entrenched when it has become *automatized* as a routine chunk of language (De Smet and Cuyckens 2007: 188).

The more a speaker encounters a particular construction, the more *entrenched* it becomes (Wulff and Ellis 2018)

As language users, we have ‘a huge warehouse of constructions’ of varying degrees of complexity and abstraction (Wulff and Ellis 2018: 39)



Resonance with DDL

Frequency is a key determinant of acquisition because “rules” of language, at all levels of analysis from phonological, through syntax to discourse, are **structural regularities** which emerge from learners’ lifetime analysis of the distributional characteristics of language input

Ellis (2012: 196)

DDL offers learners a type of “condensed exposure” (Gabrielatos 2005: 10) that can aid lexical and pattern awareness





DDL works well for:

- Undergraduates and graduates
- Local and large corpora
- Hands-on and paper-based
- **Intermediate to advanced levels**

Boulton and Cobb (2017: 386) based on meta-study analysed 64 empirical studies

ional levels are closely associated (Thrift , 1986). 2.5.1 Changing **patterns** Superficially, there is little disagreement between theoretical approaches of researchers, policy-makers and service-providers alike. Changing **patterns** of life expectancy and advances in medical technology, treatment and rehabilitation, drainage basin characteristics, hydraulic geometry, river channel **patterns** , theoretical approaches, dynamic contributing areas, and the palaeohydrology of a river type one sees the picture alternating randomly between the chaotic **pattern** and the ordered one. The typical duration that it spends in either mode is shown in Fig. 2 b shows, the asymptotic fraction of C is as for the chaotic **pattern** typified by Figs 1 b and 2 a. Many of the dynamic features of the chaotic **pattern** why the approximation also works for the irregular, spatially chaotic **patterns** (Fig. 2 a) we do not know. Our 'spatial dilemmas' game obviously is sufficient to generate astonishingly complex and spatially chaotic **patterns** in which cooperation and defection persist indefinitely. The parameters (the value of b), but a wide range of values leads to chaotic **patterns** , whose nature is almost always independent of the value of b . The result is a characteristic **pattern** of light and dark fringes. The result is a characteristic **pattern** of light and dark fringes. The result is a characteristic **pattern** of light and dark fringes.

...we can hypothesise that grappling with patterns in DDL, whether consciously or sub-consciously, is not daunting to more advanced learners because they have already abstracted many patterns and have a critical level of understanding of these patterns from their earlier language experiences with both their L1 and L2. (O’Keeffe 2021: 46)

Rather than write off learners below Intermediate, there is a need to **curate better DDL language experiences** for them so that they can build up patterns and meanings.

to hearing how **much** I

it **much better** later.

l **much better** having

much better I feel.

ch better it is today t

How can usage-based theory help us better curate language for lower levels?

Formula -> slot and frame -> formulaic

A2 Level pattern

past simple + preposition + determiner + noun

went	for	a	walk
------	-----	---	------

looked	in	the	shop
--------	----	-----	------

Formula -> slot and frame -> formulaic

A2 Level pattern

past simple + preposition + determiner + noun

went

from

the

restaurant

looked

in

the

shop

There are many things near my house.

I'm happy

I go to work at 9 o'clock

My name is Anna

They live in Cologne.

A1 GERMAN LEARNERS

Some of the most
frequent Verb
Argument
Constructions

Römer (2019: 276)

	A1 CLC	BNC Written
1	there is a concert in	there are a number of
2	there are a lot of	there is no doubt that
3	there is a concert .	there is no need to
4	there were a lot of	there is no reason why
5	there were all my friends	there is no evidence that
6	there is a concert on	there is a need for
7	there is a concert of	there is no reason to
8	there is my house .	there are a lot of
9	there 's a concert in	there is no need for
10	there is a concert ,	there is no such thing
11	there were a lot of	there is more than one
12	there was my family and	there was a lot of
13	there was a lot of	there is a danger that
14	there is a concert next	there is a great deal
15	there is a concert of	there is a lot of
16	there were all our friends	there is little doubt that
17	there is a concert at	there were a number of
18	there is a rock concert	there 's a lot of
19	there is a concert near	there is no point in
20	there was all my family	there is likely to be

	A1 CLC	BNC Written
1	there is a concert in	there are a number of
2	there are a lot of	there is no doubt that
3	there is a concert .	there is no need to
4	there were a lot of	there is no reason why
5	there were all my friends	there is no evidence that
6	there is a concert on	there is a need for
7	there is a concert of	there is no reason to
8	there is my house .	there are a lot of
9	there 's a concert in	there is no need for
10	there is a concert ,	there is no such thing
11	there were a lot of	there is more than one
12	there was my family and	there was a lot of
13	there was a lot of	there is a danger that
14	there is a concert next	there is a great deal
15	there is a concert of	there is a lot of
16	there were all our friends	there is little doubt that
17	there is a concert at	there were a number of
18	there is a rock concert	there 's a lot of
19	there is a concert near	there is no point in
20	there was all my family	there is likely to be

	A1	B2	C2	BNC Written
1.	there is a concert in	there are a lot of	there are a lot of	there are a number of
2.	there are a lot of	there is a lot of	there is no doubt that	there is no doubt that
3.	there is a concert .	there were a lot of	there are many people who	there is no need to
4.	there were a lot of	there is a new direct	there was no money left	there is no reason why
5.	there were all my friends	there were over 5,000 people	there is a lot of	there is no evidence that
6.	there is a concert on	there is a new collection	there are some people who	there is a need for
7.	there is a concert of	there is a new shop	there is no need to	there is no reason to
8.	there is my house .	there is no doubt that	there is no point in	there are a lot of
9.	there 's a concert in	there is no need to	there is a lack of	there is no need for
10.	there is a concert ,	there are many things to	there are more and more	there is no such thing
11.	there were a lot of	there were no discounts available	there is no need for	there is more than one
12.	there was my family and	there are advantages and disadvantages	there were a lot of	there was a lot of
13.	there was a lot of	there are some things that	there are people who are	there is a danger that
14.	there is a concert next	there is a lack of	there are people who do	there is a great deal
15.	there is a concert of	there are a few things	there are a number of	there is a lot of
16.	there were all our friends	there are too many cars	There are , of course	there is little doubt that
17.	there is a concert at	there was a lot of	There are , however ,	there were a number of
18.	there is a rock concert	there is a need for	there are people who believe	there 's a lot of
19.	there is a concert near	there was a different actor	there is a number of	there is no point in
20.	there was all my family	there 's a lot of	there is a wide range	there is likely to be

How to bridge the gap?

	A1 CLC
1	there is a concert in
2	there are a lot of
3	there is a concert .
4	there were a lot of
5	there were all my friends

There is a lot of

There is a lack of

There is a need for

There is no need for

There is no shortage of

BNC

there is no doubt that

there is no evidence that

there is a need for

there is no need for

there is a danger that

there is a lot of

there is no point in

there is no way of

there is no guarantee that

there is no evidence of

there is no question of

there is the question of

there is an element of

there is a risk of

there is no room for

there is some evidence that

there is no shortage of

there is a possibility that

there is a limit to



There's

There is/are

There was/were

a lot +

a lot +

a lot +

a lot +

many +

over +

some +

more than +

a great deal +

a number of

+

Three central factors in the usage-based model of acquisition

frequency

- the amount of times a construction is experienced and used

recency

- the more recently we experience a construction, the stronger our memory of it

context

- context triggers an association and mental categorisation of a frequently experienced construction

Through frequent and meaningful encounters with patterns of **phonology, syntax and discourse, regularities** emerge for a child acquiring a first language and for a learner acquiring a second one

As Ellis (2002, p. 144) notes: structural regularities “emerge from learners’ lifetime analysis of the distributional characteristics of **language input**”.



Meunier (2020: 19)



- There is room for more creativity in the DDL tasks ... especially keeping in mind the affordances of current digital tools (multimodality, gaming options, easy access, intuitive use, etc).
- Concordances are not the only possible “triggers of frequency effects and form-meaning mappings in focus on form activities”.
- In natural first and second language acquisition, form-frequency mapping is a multi-modal experience (usually with audio, visual, verbal, non-verbal, prosodic information).

“To put it starkly, we need to engage with theories of learning and models of second language acquisition if we are to move beyond driving text-based data at learners in the hope that some of it will stick in their subconscious store.”
O’Keeffe (2021: 53)

Thank you for listening

thank you for listening

thank you for listening

Thank you for listening

Thank you for listening

Thank you for listening

thank you for listening

thank you for listening

thank you for listening

The end!

References

- De Smet, H. and Cuyckens, H. 2007. Diachronic aspects of complementation: Constructions, entrenchment and the matching-problem. In SHELL, Date: 2004/05/06-2004/05/07, Location: Ann Arbor, MI. Mouton: Berlin, 1-37.
- Ellis, N. C. 2012. Frequency-based accounts of second language acquisition. In S. M. Gass and A. Mackey (Eds). *The Routledge Handbook of Second Language Acquisition*. Abingdon: Routledge, 193-216.
- Gabrielatos, C. 2005. Corpora and language teaching: Just a fling or wedding bells? *Teaching English as a Second or Foreign Language* 8(4): 1–34.
- O’Keeffe, A. 2021. “Data-Driven Learning, Theories of Learning and Second Language Acquisition: in search of intersections”. In P. Pérez-Paredes and Mark, G. (Eds.) *Beyond the Concordance: Multiple Applications of Language Corpora for Language Education*. Amsterdam: John Benjamins, pp. 35-55.
- O’Keeffe, A. 2020. “[Data-Driven Learning – a call for a broader research gaze](#)”. *Language Teaching*, 1–14.
- O’Keeffe, A., Pérez-Paredes, P. and Mark, G. 2020. ‘[Aligning usage-based theory and Data-Driven Learning](#)’. Paper presented at TaLC 2020, University of Perpignan, 15-18 July, 2020.
- Pérez-Paredes, P., Mark, G. and O’Keeffe, A. (2020) “[The impact of usage-based approaches on second language learning and teaching](#)”. *Cambridge Education Research Reports*. Cambridge: Cambridge University Press.
- Römer, U. 2019. A corpus perspective on the development of verb constructions in second language learners. *International Journal of Corpus Linguistics*, 24(3): 268-290.
- Wulff, S. and Ellis, N. C. (2018). Usage-based approaches to SLA. In D. Miller, F. Bayram, J. Rothman, L. Serratrice (Eds.), *Bilingual Cognition and Language: The state of the science across its subfields*. (pp. 37-56). Amsterdam: John Benjamins.