Teaching with corpora in the human-machine era: from the core to the peripheries

Prof. Fanny Meunier, UCLouvain, Institute for Language and Communication,

International Perspectives on Corpus Technology for Language Learning

Seminar Series

25 Oct. 2022

TeAMM

UCLouvain iLC

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

STOP : importance of 'taking stock' but also of reconsidering some key aspects of 'teaching with corpora'

THINK : NLP and AI-based technologies prompt us to further question some aspects of using corpora in instructed settings

ACT : change our perspective

shift from a mainly applied linguistics mindset to an applied (S)LA one

leave some well-trodden paths and actively explore new peripheries

STOP : importance of 'taking stock' but also need to reconsider some key aspects of 'teaching with corpora'
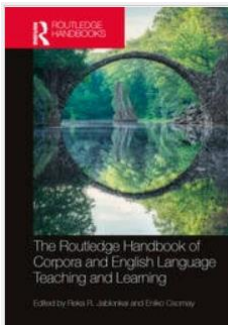
Taking stock

30 years of incredible achievements and progress!

nearly 30 years after the first TaLC conference (1994, in Lancaster)

one recent illustration of range of topics covered : The Routledge Handbook of Corpora and English Language Teaching and Learning *Edited By Reka R. Jablonkai, Eniko Csomay,* 2022 - types of topics covered (selection)

- and more...

  ✓Numerous corpora and corpus types for **LOTE** (Languages Other Than English)

  ✓With similar types of topics but also with **specific agendas** (e.g. acquisition of agglutinative languages, revitilization of ('minority'/'small'/'less-spoken'/'endangered') languages

  ✓Corpora and **literary criticism**

  ✓Corpora and **sociological research**

  ✓Corpora and **language change/evolution**

  ✓...

Culture/language biases in society
Politics
Social and mass media
Gender stereotypes
Technologies
Social movements
...

Taking stock and reconsidering what teaching with corpora means
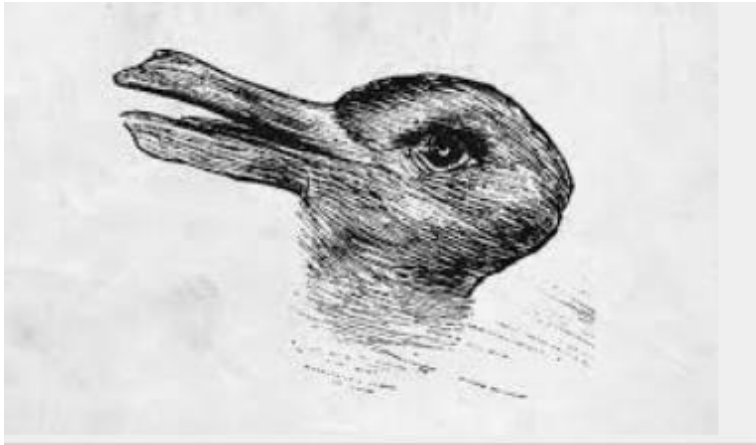
'Teaching with corpora' = ….

- To whom?
- What for?
- Based on which identified needs?
- For which learning outcomes?
- Are those learning outcomes assessed?
- Is it constructively aligned with/in the curriculum?

'Teaching with corpora' = ...

- To whom? kids / teenagers / university graduates / post-grads…

- What for? language awareness raising activities / vocabulary acquisition / autonomous learning…

- Based on which identified needs? the corpus linguist who sees potential value of the tools/findings made for 'others' (students? learners? teachers? readers of journals in linguistics/education/sociology? / actual learners' needs that have been voiced by learners themselves? actual teachers' needs that have been collected and taken into account?

- For which learning outcomes? … not always clearly identified…

- Are those learning outcomes assessed (pre-test? Immediate/delayed post-test?) … not often enough…

- Constructive aligment in the curriculum?

… learning outcome presented not in line with tasks proposed… (e.g. language awareness as LO – voc retention test as assessment…

… often completely disregarded… (hands up challenge at TaLC2022 not very conclusive… explain)

… interim conclusion: *sometimes*

- there is a lack of alignement between learners' or teachers' needs, claims in terms of LO, assessment, tasks suggested;

- generic statements are made in papers (e.g. 'corpora should be used more often by teachers', 'teachers should be trained to use corpora') but they are not supported by clearly situated/ecological pedagogical alignment or discussions.

- at times, this gives a *'first you jump, then you get wings attitude'* (Meunier, this presentation ;-)), which is not to be recommended… unless you're a Simple Minds fan

also…

- some taken-for-granted options/habits in CL are critically challenged (e.g. Tribble (2017: 30) 'EAP programmes are described as unfairly imposing national or native models on non-native speakers of English, and the genre-based paradigm which informs many EAP writing instruction programmes is characterised as conforming to rather than challenging the status quo')

There is ROOM FOR IMPROVEMENT…

ELFA vs. Genre: A new paradigm war in EAP writing instruction?

Christopher Tribble ✉

THINK : NLP and AI-based technologies prompt us to further question some aspects of using corpora in instructed settings
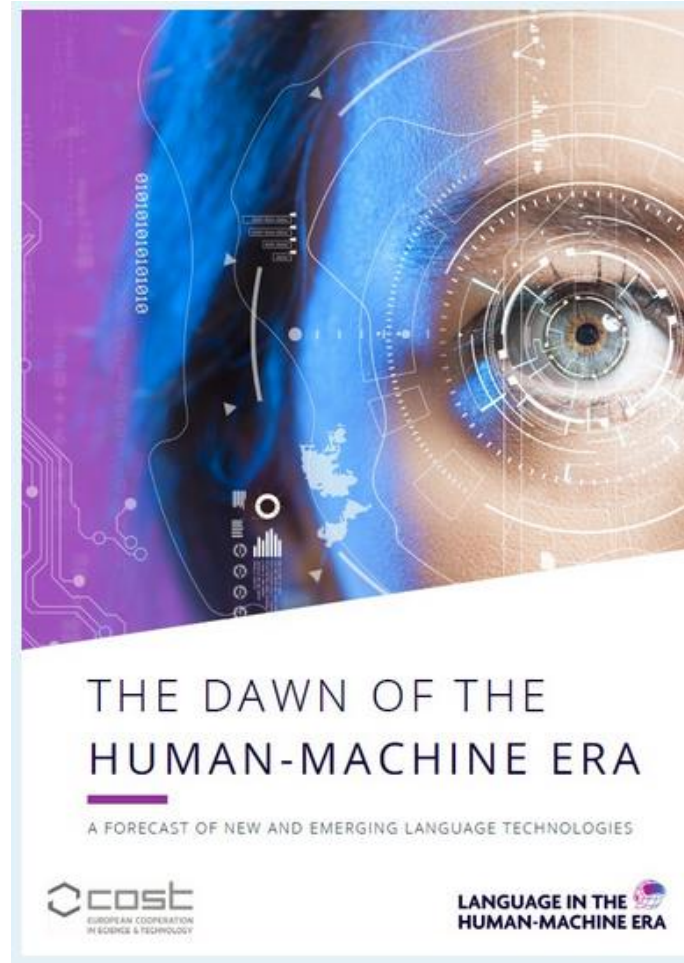
The man-machine era... poses additional challenges

How will pervasive, augmented, virtual, and/or real time technology affect areas such additional language teaching and learning?

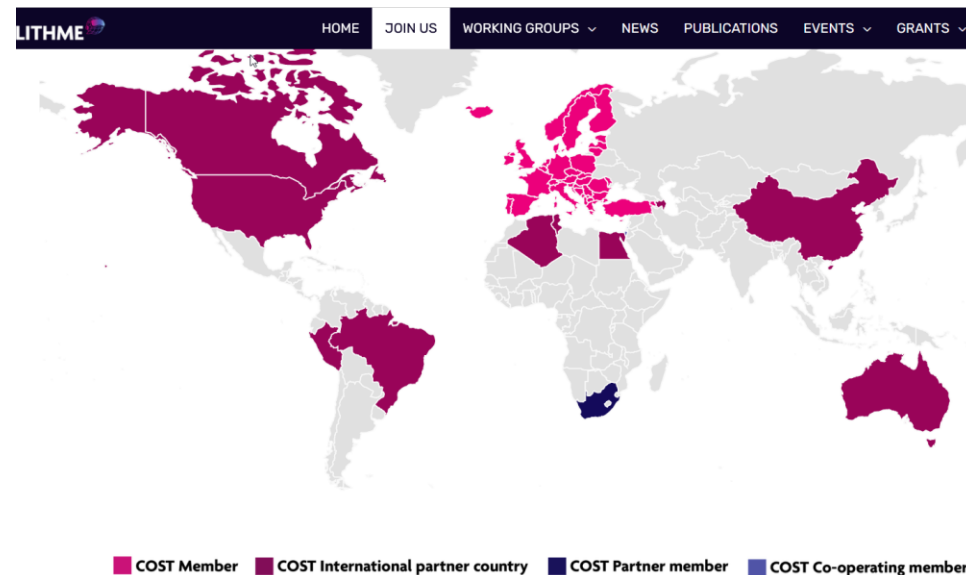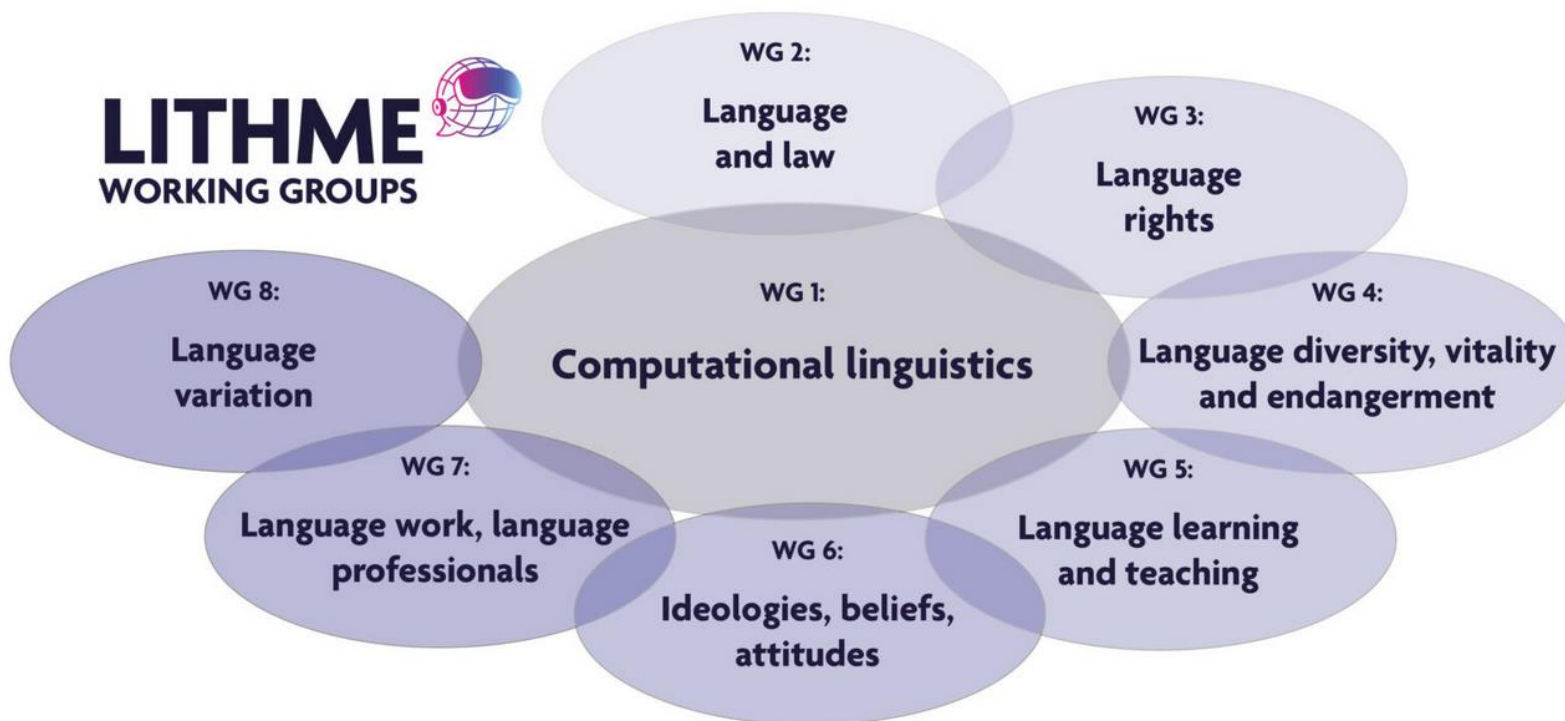*and **teaching with corpora** more specifically ? (our focus today)*

Teaching with corpora... in the human-machine era

Sayers, D., R. Sousa-Silva, S. Höhn et al. (2021). **The Dawn of the Human-Machine Era: A forecast of new and emerging language technologies**. Report for EU COST Action CA19102 'Language In The Human-Machine Era'. www.lithme.eu.

COST Member   COST International partner country   COST Partner member   COST Co-operating member

LITHME WORKING GROUPS

WG 2: Language and law

WG 3: Language rights

WG 1: Computational linguistics

WG 4: Language diversity, vitality and endangerment

WG 8: Language variation

WG 5: Language learning and teaching

WG 7: Language work, language professionals

WG 6: Ideologies, beliefs, attitudes

Advances in the human-machine era perceived as potentiel threats to LT... by the media

franca, qu'on met à toutes les sauces, puisque des technologies performantes seront disponibles pour d'autres langues. Il sera important d'être attentif à ce qu'on finance ce type de technologie pour plusieurs langues, afin qu'on n'appauvrisse pas la richesse langagière".

■ Dans laquelle oreillettes, lunettes et logiciels nous traduiront de façon automatique et en temps réel les langues étrangères.

24 La Libre Belgique - samedi 26 et dimanche 27 juin 2021 -

Quand la technologie cherche à abattre la tour de Babel

Avec ces technologies présentées dans le rapport, sera-t-il encore nécessaire d'apprendre les langues étrangères?
Oui, il faudra continuer à apprendre les langues. Mais il faudra les apprendre différemment. Comme toutes les technologies, ces technologies auront beaucoup de points forts mais aussi beaucoup de points faibles. Par exemple, il est

Ces technologies amèneront-elles à une meilleure ou moins bonne communication entre humains? Ici, on resterait isolé dans sa propre langue, un peu comme les réseaux sociaux nous amènent à être avec des gens qui pensent comme nous...
Comme toute technologie, elle n'est aussi bonne ou mauvaise que par l'utilisation qu'on en fait, ou ce que les hu-

...by students

- (EAP course UCLouvain, Master's level students)

«je ne vois pas pourquoi je devrais passer 1h à chercher des équivalences dans des corpus alors que deepl me donne une super traduction en 30 secondes »

...by educational stakeholders

Can technology replace teachers?

Aanchal goel @My musings
APR 25, 2019, 20:29 IST

It is a much known fact that in this modern age, technology is replacing almost everything. We do not need papers and pens, as software can be used to write on, we do not need physical books because of the advent of e-books and PDFs. So, the question arises – Can this modern technology replace teachers too? Is this technology sophisticated enough to teach children the way virtual teachers do ~~~d of human teachers being eliminated?

But remember, time is a precious commodity for teachers. Fortunately, HP Reveal's easy drag-and-drop interface allows users to create a trigger image in less than a minute. Companies looking to create something similar should look to minimize teachers' time investment with elements like user-generated content.

Of course, teachers often have their hands tied when it comes to bringing new technologies into the classroom. It's usually up to districts and administrative leaders to decide which new tools to invest in, and AR can be a tough sell without supporting research and low-cost options.

But once augmented reality developers and marketers realize the opportunity of the education sector, expect classroom content to skyrocket. AR is too exciting and educational for educators and students to stay away for long.

AI in Education: Will Tech Destroy Education or Save It?

Some teachers like to joke that teaching is a walk in the park. If the park is Jurassic Park, that is.

But soon it can become more similar to another fictional amusement park—Westworld. Whether you're more afraid of killer dinosaurs or intelligent robots is another story.

# [Teaching]{.underline} with corpora, says the title…

- what about our current usages of corpora for language teaching?

- do we need to reconsider some aspects of 'teaching with corpora'?

- how can we convince 'people' (beyond the CL community) that corpora are useful resources for teaching? because there's some convincing needed, and increasingly so.

Let's act !

2 suggestions I'd like to share with you
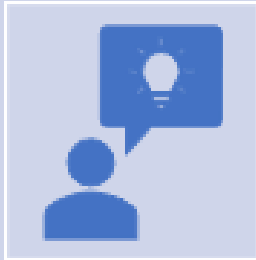
ACT : change our perspective

shift from a mainly applied linguistics mindset to an applied (S)LA one

leave some well-trodden paths and actively explore new peripheries

ACTION #1: shifting from an applied linguistics mindset to an applied (S)LA one

(Illustration with DDL)

- Boulton (2020) : DDL work in applied linguistics remains relatively unknown in educational circles
  - Meunier (2019, 2020, 2022) : argues for adopting a reversed perspective. Applied linguists promoting DDL often tend to ignore, or at least fail to comment on, key educational principles (constructive alignment, curricular alignment, learning outcomes, trained versus assessed competences, etc.). Extreme example: 4 pages in a paper discussing the various options regarding the stats… but nothing on whether or not it made sense to spend two hours of classroom time doing DDL activities to learn 4 new words and 2 collocates for each…

(Illustration with typical comments when lack of pedagogical alignement or ecological classroom approach are pointed)

- (a colleague, personal conversation) 'you cannot say that I have no interest in pedagogy, I'm an applied linguist!'
  - (Meunier, personal conversation) 'Well, applied linguistics and applied SLA are two different constructs... let me try to explain'

Applied linguistics includes aspects related to real world relevance, yes

APPLIED LINGUISTICS

ISSN 0142-6001 (PRINT)
ISSN 1477-450X (ONLINE)

Volume 43 Number 4 August 2022

Published in cooperation with
AILA International Association of Applied Linguistics
B.A.A.L. British Association for Applied Linguistics

OXFORD
UNIVERSITY PRESS

Latest Issue

Volume 43, Issue 4

August 2022

Impact Factor
4.155

5 year Impact Factor
6.178

Linguistics
15 out of 194

Editor
Christina Higgins

About the journal

Applied Linguistics publishes research into language with relevance to real-world issues. The journal is keen to help make connections between scholarly discourses, theories, and research methods…

Find out more

Articles

Reviews

Notes on Contributors

'Who Is My Research for?': Researcher Perceptions of the Research–Practice Relationship Get access >

Masatoshi Sato, Shawn Loewen, Dmitrii Pastushenkov

Abstract ▲

This study investigated second language (L2) researchers' perceptions of the research–practice relationship. While research use (or lack thereof) in the classroom has been debated by researchers for decades, researchers have rarely investigated themselves as stakeholders in the relationship. However, it is vital to know researchers' approaches to the issue so that collaborative researcher-practitioner relationships may be fostered. Hence, the current survey study explored: (a) L2 researchers' identities; (b) their perceptions of the research–practice relationship; and (c) predictors of their perceptions. Participants were 217 researchers from 25 countries with 31 different first-language backgrounds. The results showed that L2 researchers' identities often crossed multiple boundaries, as L2 learners, teachers, as well as researchers. A factor analysis showed that researchers held distinct beliefs for their own research (individual self) and for researchers in general (collective self). The regression models revealed that L2 teaching experience, pre-service teacher training experience, and institutional rewards positively predicted researchers' perceptions of L2 practice. We argue that it is largely researchers' responsibility to facilitate the research–practice dialogue.

But... in

- « Applied linguistics », linguistics in the NP head... (+ previous slide, see how researchers in applied linguistics often perceive their agency in terms of real world relevance...)

- « Applied second language acquisition », acquisition is the NP head...

Han (2016) on ASLA

- Need to address gaps in current SLA research resulting from :
  (a) persistent narrowness (as seen, e.g., in the design scope of individual studies)
  (b) a tendency towards isolation (e.g., fragmentation, the yawning distance from the real world)

- Proposes a 3-tiered division
  - three subfields differentially accountable to context and real world relevance
  - and yet equally valid sources of knowledge on SLA
  - relationship among the subfields = cooperative, not mutually exclusive, and not hermetic from one another

FSLA : focusing on issues and phenomena that are unique to SLA, with little regard for how the insights and findings would translate into practical ideas or solutions.

ISLA: carrying out experimental work formulating and testing hypotheses related to pedagogical intervention without being overly concerned with ecological validity.

ASLA: real life educational concerns front and center, with L2 learning and instruction situated in circumstances with their own set of contingencies and investigated accordingly.

# Three subfields differentially accountable to context and real-world relevance but working cooperatively

If FSLA, then FSLA issues and claims are central

If ISLA, then ISLA issues and claims are central

If ASLA, then ASLA issues and claims are central – hopefully and ideally informed by FSLA and ISLA // and active and on-site ecological teacher-researcher collaboration essential !
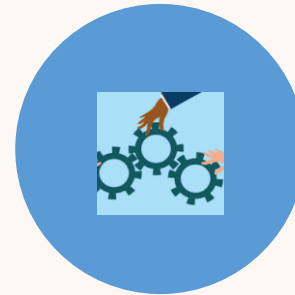
relationship among the subfields = cooperative, not mutually exclusive, and not hermetic from one another

- ASLA is now an existing term with a clear definition putting real life educational concerns front and center, with L2 learning and instruction situated in circumstances with their own set of contingencies and investigated accordingly

- So let's adopt an ASLA mindset!

On 'the periphery and the heart of language.
# ACTION #2: leave some well-trodden paths and actively explore new peripheries

Ellis, Nick C. (2008). Phraseology: The periphery and the heart of language. Preface to F. Meunier; S. Granger (Eds.), Phraseology in language learning and teaching, pp. 1-13. Amsterdam: John Benjamins.

- Corpora and NLP have been instrumental in helping machines access the heart of the language in terms of patterns, schemas, constructions, prototypicality, frequencies, text types, etc.

- But some 'treasures' present in corpora have perhaps been neglected: ideosyncrasies, content, speech acts, language awareness issues beyond the typical lexical and grammatical foci addressed in mainstream DDL activities, the 'complete texts' asset (vs short often decontextualized concordance lines)

- Let us, humans/linguists/teachers, occupy and take care of those peripher<u>ies </u>that are largely inaccessible to technology even in our tech-saturated world and « let's think HOTS » (in reference to higher order thinking skills, see Krathwohl (2002). A Revision of Bloom's Taxonomy: An Overview, Theory Into Practice, 41:4, 212-218, DOI: 10.1207/s15430421tip4104_2)
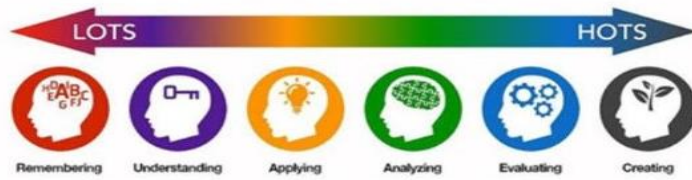
A few suggestions only…

Think HOTS…

- Doing corpus-driven lexical analysis of MT and HT to unveil fascinating discourse features that set them apart (Frankenberg-Garcia, 2022) ttps://benjamins.com/online/target/articles/target.20065.fra

- Revamping DDL to exploit the affordances of digital technology (using high tech, low tech, and wild tech tools) (Meunier, in Jablonkai & Csomay 2022, see slide 6)

- Much needed life-long '(Re)awakening to languages' (Sayers et al., 2021, see slide 15; Meunier & Van de Vyver, 2022 https://www.linguanum.eu/)

- 'Moving toward an additive approach' (Staples, 2022, https://ulsites.ul.ie/cals/talc-conference-plenaries-2022) :
  - promote learner texts as models for other language learners (mentor texts),
  - use texts as sites for discussion of functional language use in and beyond the concordance line
  - address language choices in contexts relevant to learners.

- Using corpora as resources for narrative inquiry (Tyne, 2022, https://ulsites.ul.ie/cals/talc-conference-plenaries-2022) :
  - the experiences of an individual or small group, revealing their lived experience or particular perspective through the analysis of their language productions