WHAT'S IN A WORD?

Nick Ellis

Vocab@Vic 2023

ncellis@umich.edu

Shakespeare (1597) Romeo and Juliet. What's in a **name**?

de Saussure (1916) Thought-Sound. **Signification**

Firth (1957) "You shall know a word by the **company** it keeps"

Wittgenstein (1953) "In most cases meaning is **use**"

Where are these ideas now? Educational implications?   VUW

Shakespeare Romeo and Juliet (1597) What's in a name?

JULIET

O Romeo, Romeo! wherefore art thou Romeo?
Deny thy father and refuse thy name;
Or, if thou wilt not, be but sworn my love,
And I'll no longer be a Capulet.

ROMEO

[Aside] Shall I hear more, or shall I speak at this?

JULIET

'Tis but thy name that is my enemy;
Thou art thyself, though not a Montague.
What's Montague? it is nor hand, nor foot,
Nor arm, nor face, nor any other part
Belonging to a man. O, be some other name!
What's in a name? that which we call a rose
By any other name would smell as sweet;
So Romeo would, were he not Romeo call'd,
Retain that dear perfection which he owes
Without that title. Romeo, doff thy name,
And for that name which is no part of thee
Take all myself.

Bally Gill as Romeo and Karen Fishwick as Juliet in Erica Whyman's 2018 production of Romeo and Juliet with the Royal Shakespeare Company.

ncellis@umich.edu

Shakespeare (1597) What's in a name?

Rose by any other name. Symbols arbitrary

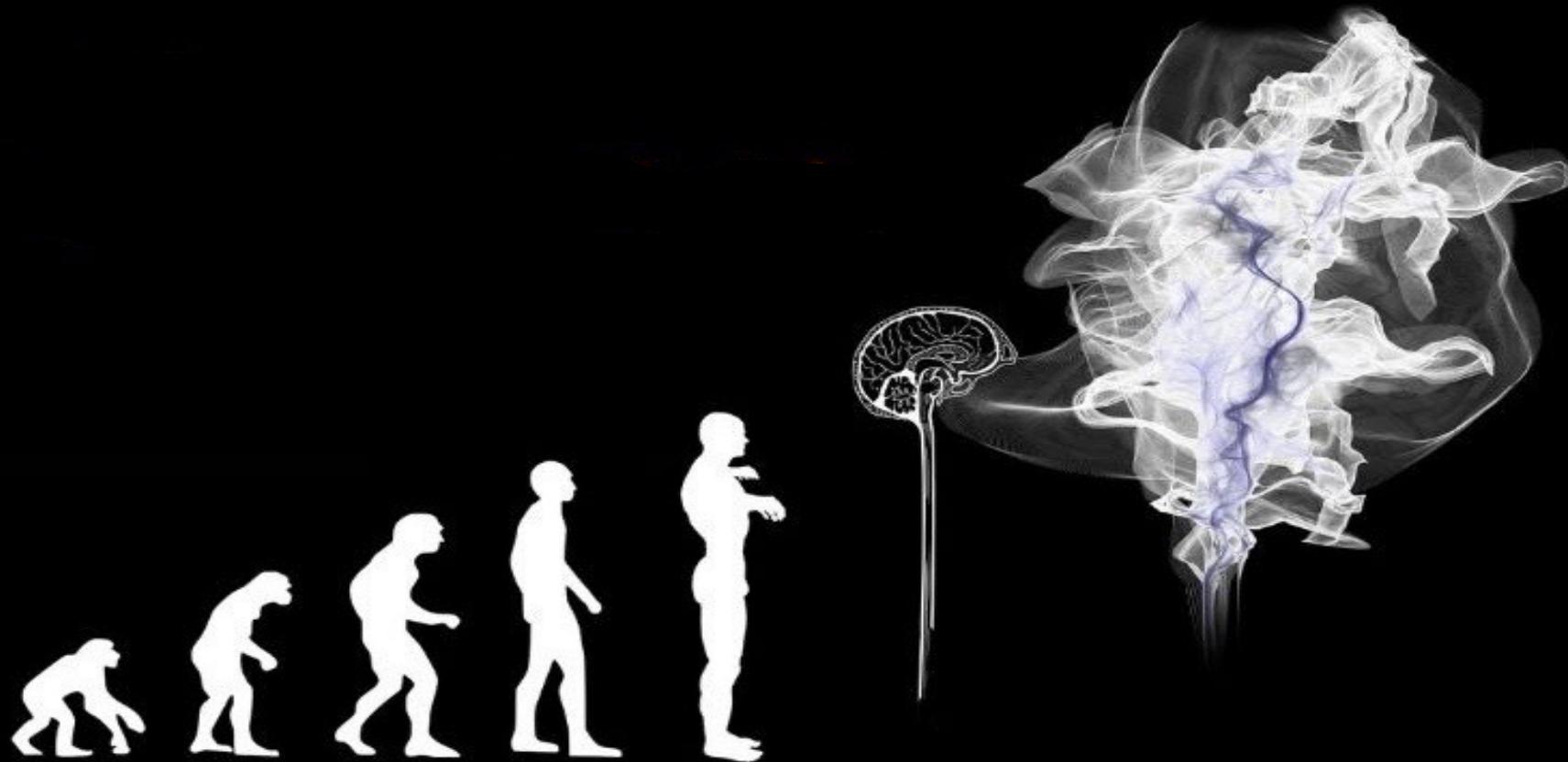Associative learning

Not all words are equal – (hand / Montague)

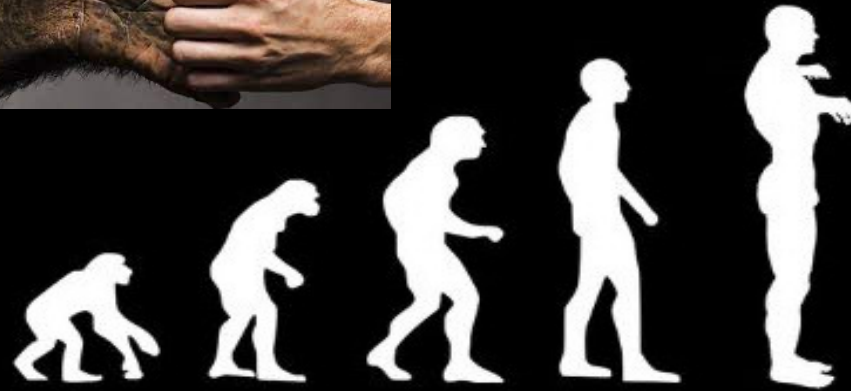The 'learning burdens' can be quite different

ncellis@umich.edu

WHAT'S IN A WORD?

Cognitive-linguistic, Neuroscientific, AI, Psycholinguistic, and Usage-based perspectives

Nick Ellis

Vocab@Vic 2023

ncellis@umich.edu

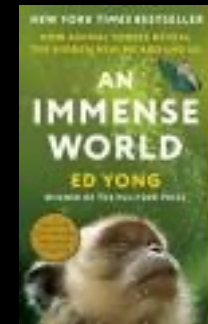https://www.insightsmix.com/artificial-intelligence/gpt-alternative-top-large-language-models/
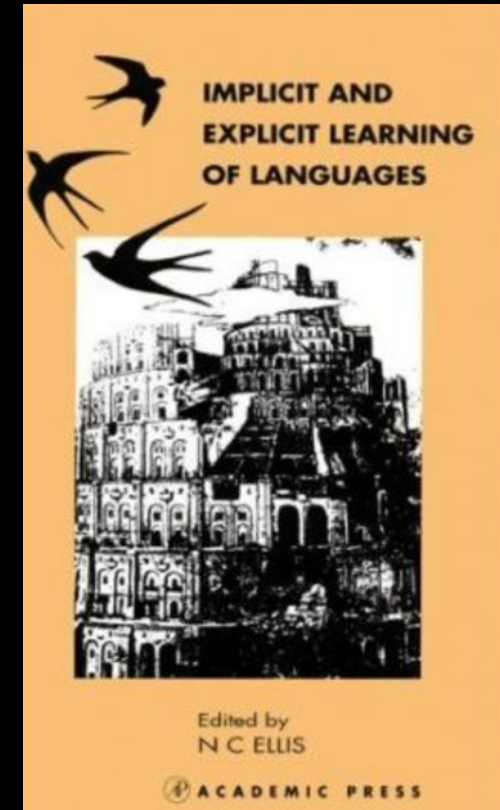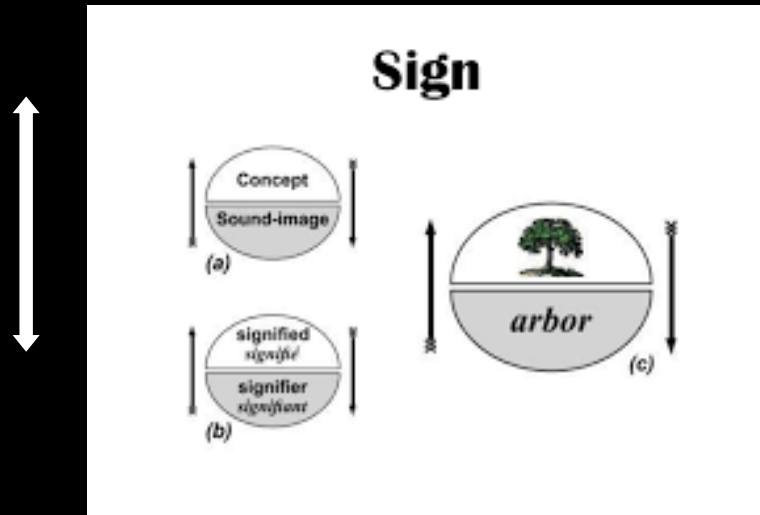
Embodiment
Sight
Perceptual symbol systems
Hand
Umwelt
Poke push put... VOL
Social Cognition

Primates 55 million years. Human language 200,000 years

Ellis, N. C. (2019). Essentials of a theory of language cognition. *Modern Language Journal, 103* (Supplement 2019), 39-60.

ncellis@umich.edu

de Saussure (1916) Thought-Sound. **Signification**
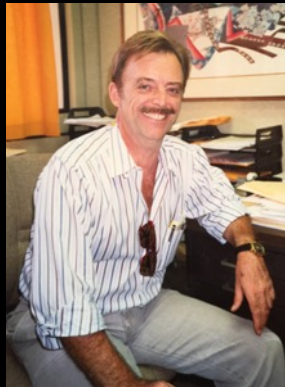
Arbitrariness of Signs - Associative learning

Factors affecting Explicit learning & Memory

Attention

Depth of Processing

Desirable difficulties – self-testing, spaced practice, interleaving, ..

Implicit and Explicit Vocabulary Learning

Noticing
Attention

- The hippocampal system subserves rapid **EXPLICIT** memory: one-off learning,
- the establishment of new conjunctions of arbitrarily different elements into a unitized representation
- the learning of separate discrete episodic memories where we do not want an average, an abstraction, or a gist:
- There is benefit in being able to keep some records straight, complete, and distinct.
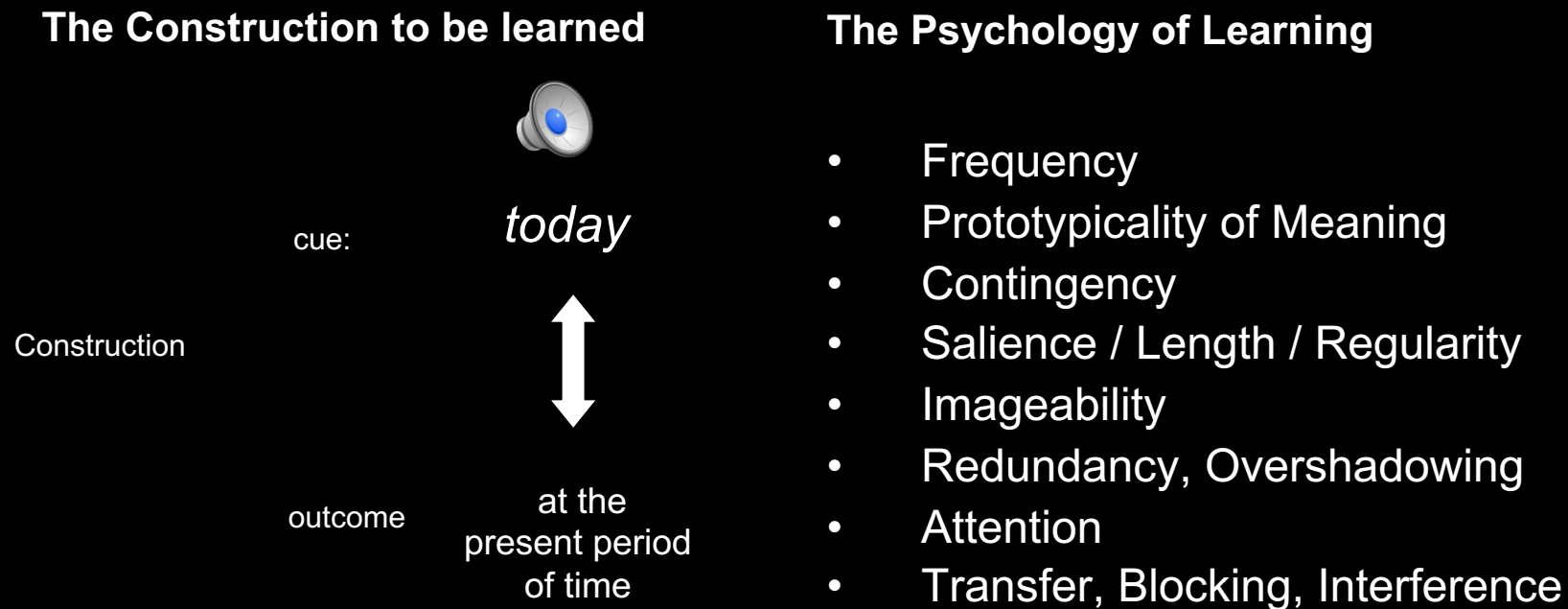
- The neocortical systems subserve **IMPLICIT** memory
- the tuning of associative systems to reflect repeated patterns of local activity and to generalize from them,
- generalizations rather than episodic memory.
- To operate efficiently in the world we need to be able to identify general patterns by abstracting from instances - to classify and categorize.

Ellis, N. C. (1994). Vocabulary acquisition: The implicit ins and outs of explicit cognitive mediation In N. Ellis (Ed.) *Implicit and explicit learning of languages* (pp. 211-282). London: Academic Press.
Ellis, N. C. (2006). Language acquisition as rational contingency learning. *Applied Linguistics. 27 (1)*, 1-24.
Ellis, N. C. (2015). Implicit AND explicit learning: Their dynamic interface and complexity. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 3-23). Amsterdam: John Benjamins

ncellis@umich.edu

Routine Psycholinguistic determinants of Learning

The Construction to be learned

The Psychology of Learning

cue:

today

Construction

⬍

outcome

at the
present period
of time

- Frequency
- Prototypicality of Meaning
- Contingency
- Salience / Length / Regularity
- Imageability
- Redundancy, Overshadowing
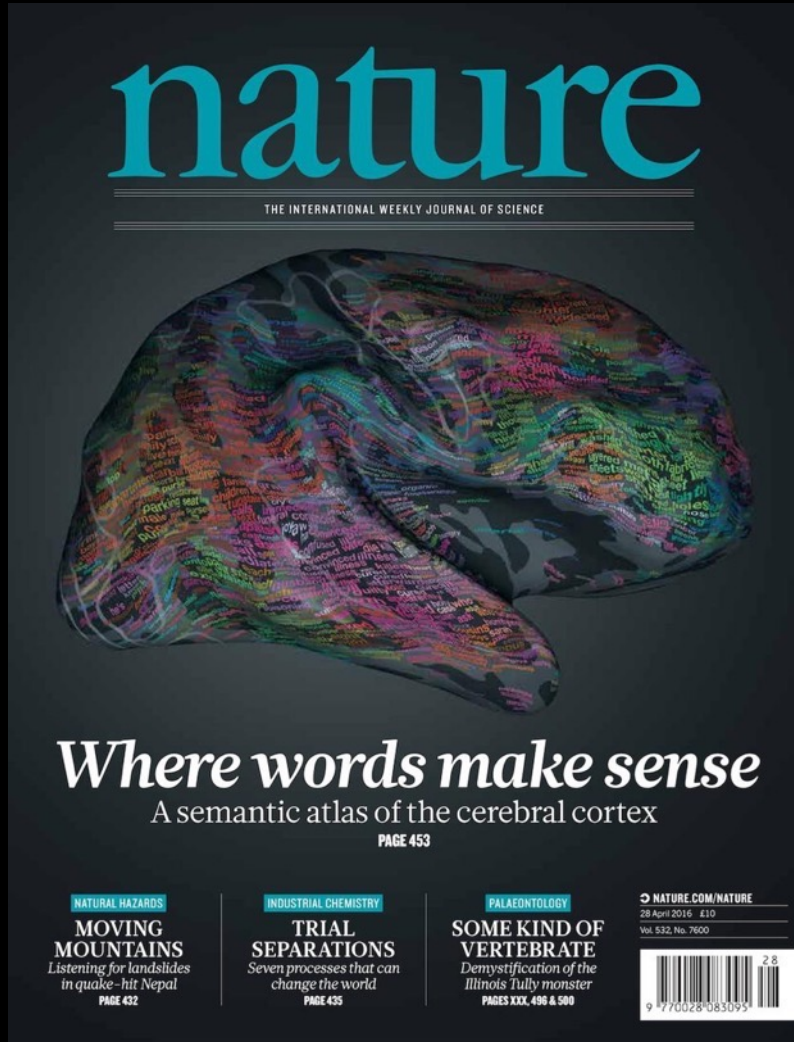- Attention
- Transfer, Blocking, Interference

Ellis, N. C. (1995). The psychology of foreign language acquisition: Implications for CALL. *International Journal of Computer Assisted Language Learning (CALL), 8*, 103-128.
Ellis, N. C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning, 43*, 559-617.
Ellis, N. C. (2006). Language acquisition as rational contingency learning. *Applied Linguistics. 2*7 (1), 1-24.
Berger, C., Crossley, S., Skalicky ,S.. (2019). Using lexical features to investigate second language lexical decision performance. *Studies in Second Language Acquisition, 41*, 911-935.

ncellis@umich.edu

Signification Neuroscience Jack Gallant

nature

THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

Where words make sense

A semantic atlas of the cerebral cortex
PAGE 453

NATURAL HAZARDS
MOVING MOUNTAINS
Listening for landslides in quake-hit Nepal
PAGE 432

INDUSTRIAL CHEMISTRY
TRIAL SEPARATIONS
Seven processes that can change the world
PAGE 435

PALAEONTOLOGY
SOME KIND OF VERTEBRATE
Demystification of the Illinois Tully monster
PAGES XXX, 496 & 500

NATURE.COM/NATURE
28 April 2016 £10
Vol. 532, No. 7600

ARTICLES
https://doi.org/10.1038/s41593-021-00921-6

nature neuroscience

Check for updates

Visual and linguistic semantic representations are aligned at the border of human visual cortex

Sara F. Popham[1], Alexander G. Huth[1,3], Natalia Y. Bilenko[1], Fatma Deniz[1,4], James S. Gao[1], Anwar O. Nunez-Elizalde[1] and Jack L. Gallant[1,2]
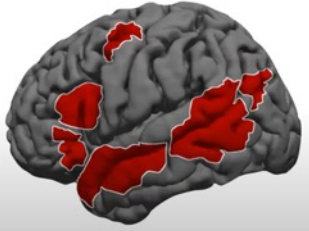
Semantic information in the human brain is organized into multiple networks, but the fine-grain relationships between them are poorly understood. In this study, we compared semantic maps obtained from two functional magnetic resonance imaging experiments in the same participants: one that used silent movies as stimuli and another that used narrative stories. Movies evoked activity from a network of modality-specific, semantically selective areas in visual cortex. Stories evoked activity from another network of semantically selective areas immediately anterior to visual cortex. Remarkably, the pattern of semantic selectivity in these two distinct networks corresponded along the boundary of visual cortex: for visual categories represented posterior to the boundary, the same categories were represented linguistically on the anterior side. These results suggest that these two networks are smoothly joined to form one contiguous map.

ncellis@umich.edu

Signification - Current Neuroscience   Ev Fedoronko

ncellis@umich.edu

Firth (1957) "You shall know a word by the company it keeps"

Sinclair (1991, p. 110) Principle of Idiom:
"a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analyzable into segments.

Pattern Grammar. Cobuild project. Collocation, Colligation, Semantic Prosody, Lexical Priming

Sinclair – 'The phrase, the whole phrase, and nothing but the phrase'

Corpus Linguistics

Cognitive Linguistics

ncellis@umich.edu

A word's company….

Cognitive Linguistics
Usage-Based Grammar
We learn Constructions

la "pensée-son"
de Saussure (1916)

- ❑ Constructions as basic symbolic units of language representation:
  - ❑ Form meaning mappings
  - ❑ Conventionalized in the speech community
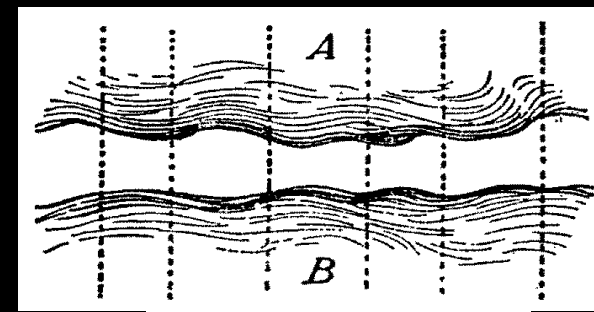 - ❑ Entrenched as language knowledge in the learner's mind
- ❑ Usage-based acquisition
  - ❑ We learn constructions through using language, engaging in communication.
- ❑ Emergence
  - ❑ Creative linguistic competence emerges from the collaboration of the memories of all of the utterances in a learner's entire history of language use and the frequency-biased abstraction of regularities within them

morphological, syntactic, lexical, phrasal form

↕

semantic, pragmatic, discourse functions

- ❑ Cognitive Linguistics
- ❑ Functional Linguistics
- ❑ Psycholinguistics
- ❑ Corpus Linguistics
- ❑ Can't separate:
  - ❑ Grammar from lexis
  - ❑ Form from meaning
  - ❑ Meaning from context
  - ❑ Structure from Usage
- ❑ Applied Linguistics

prediction →

Ellis, N. C. (2006). Cognitive perspectives on SLA: The Associative-Cognitive CREED. *AILA Review, 19*, 100-12
Ellis, N. C., Römer, U. & O'Donnell, M. B. (2016). *Usage-based Approaches to Language Acquisition and Processing: Cognitive and Corpus Investigations of Construction Grammar*. Language Learning Monograph Series. Wiley-Blackwell.

ncellis@umich.edu

A word's company…
            LLMs, Deep Learning, GenAI, Thinking Machines, GPT-4
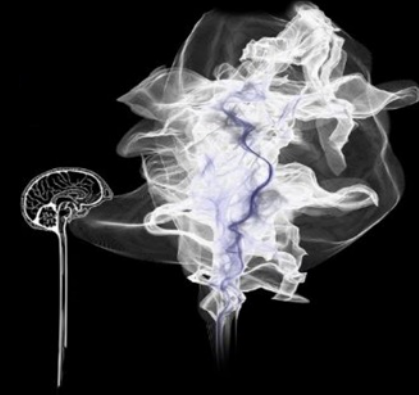
No embodiment – No skin in the game

LLMs as castles in the sky   ……………………………………………………..

Hallucinations

LLMs produce essays that are Turing-test-indistinguishable from human authors'

The language produced by LLMs is meaningful to the humans that read it, but not to the LLMs themselves.

The better they are at bullshitting, the better they will replicate

Batesian mimicry in the memeosphere

BUT – They have cracked Language form

They can certainly talk the talk   prediction

Parrot, stochastic — generated via Midjourney (free for all use)

Stochastic Parrots in the Chinese Room

Coming to Terms with Thinking Machines

David Rostcheck · Follow
10 min read · May 23

ncellis@umich.edu

A word's company - LLMs

DigiDan

WILEY

SUBMITTED ARTICLE

Creating a large language model of a philosopher

Eric Schwitzgebel[1] | David Schwitzgebel[2] | Anna Strasser[3,4]

[1]Department of Philosophy, University of California, Riverside, California, USA

[2]Institut Jean Nicod, École Normale Supérieure, Université Paris Sciences et Lettres, Paris, France

[3]Faculty of Philosophy, Ludwig-Maximilians-Universität München, München, Germany

[4]DenkWerkstatt Berlin, Berlin, Germany

Correspondence
Eric Schwitzgebel, Department of Philosophy, University of California, Riverside, CA 92521, USA.
Email: eschwitz@ucr.edu

Can large language models produce expert-quality philosophical texts? To investigate this, we fine-tuned GPT-3 with the works of philosopher Daniel Dennett. To evaluate the model, we asked the real Dennett 10 philosophical questions and then posed the same questions to the language model, collecting four responses for each question without cherry-picking. Experts on Dennett's work succeeded at distinguishing the Dennett-generated and machine-generated answers above chance but substantially short of our expectations. Philosophy blog readers performed similarly to the experts, while ordinary research participants were near chance distinguishing GPT-3's responses from those of an "actual human philosopher".

1 | INTRODUCTION

Artificial Intelligence can now outperform even expert humans in games such as chess, go, and poker and in practical domains such as lung cancer screening, predicting protein structure, and discovering novel matrix multiplication algorithms (Ardila et al., 2019; Brown & Sandholm, 2019; Campbell et al., 2002; Fawzi et al., 2022; Jumper et al., 2021; Silver et al., 2016, 2018). ChatGPT has received considerable popular attention for its capacity to generate passable short student essays (Huang, 2023). But presumably expert-level professional philosophy

---

The DISSOCIATING LANGUAGE AND THOUGHT IN LARGE LANGUAGE MODELS

A PREPRINT

Kyle Mahowald*
The University of Texas at Austin
mahowald@utexas.edu

Anna A. Ivanova*
Massachusetts Institute of Technology
annaiv@mit.edu

Idan A. Blank
University of California Los Angeles
iblank@psych.ucla.edu

Nancy Kanwisher
Massachusetts Institute of Technology
ngk@mit.edu

Joshua B. Tenenbaum
Massachusetts Institute of Technology
jbt@mit.edu

Evelina Fedorenko
Massachusetts Institute of Technology
evelina9@mit.edu

November 7, 2023

ABSTRACT

Large language models (LLMs) have come closest among all models to date to mastering human language, yet opinions about their linguistic and cognitive capabilities remain split. Here, we evaluate LLMs using a distinction between formal linguistic competence—knowledge of linguistic rules and patterns—and functional linguistic competence—understanding and using language in the world. We ground this distinction in human neuroscience, showing that formal and functional competence rely on different neural mechanisms. Although LLMs are surprisingly good at formal competence, their performance on functional competence tasks remains spotty and often requires specialized fine-tuning and/or coupling with external modules. In short, LLMs are good models of language but incomplete models of human thought.

* The two lead authors contributed equally to this work.

1 Introduction

When we hear a sentence, we typically assume that it was produced by a rational, thinking agent (another person). The sentences that people generate in day-to-day conversations are based on their world knowledge ("Not all birds can fly."), their reasoning abilities ("You're 15, you can't go to a bar."), and their goals ("Would you give me a ride, please?"). Thus, we often use other people's statements as a window into their minds.

In 1950, Alan Turing leveraged this tight relationship between language and thought to propose his famous test (Turing, 1950). The Turing test uses language as an interface between two agents, allowing human participants to probe the knowledge and reasoning capacities of two other agents to determine which of them is a human and which is a machine.[1] Although the utility of the Turing test has since been questioned, it has undoubtedly shaped the way society today thinks of machine intelligence (Boneh et al., 2019; French, 1990, 2000; Marcus et al., 2016; Moor, 1976; Pinar Saygin et al., 2000).

---

[1]In later versions of the test, the number of conversation partners has been reduced to one.

https://aeon.co/videos/see-how-our-brains-group-words-by-

Prediction

Wittgenstein (1953) "In most cases meaning is **use**"

We've moved on to ideas which are much more complex and abstract than 'hand':

'Montague', 'LLMs', 'memeosphere' / 'symbolosphere', 'desirable difficulties'….
We can only sort their meaning through usage

Educational implications?:
Read
Communities of practice
Choose the right school
Choose the right conversation partner
Choose the right summer school & conference

Douglas Fir Group (2016). A transdisciplinary framework for SLA in a multilingual world. *Modern Language Journal, 100*, 19-47
Ellis, N. C. (2019). Essentials of a theory of language cognition. *Modern Language Journal, 103* (Supplement 2019), 39-60.

ncellis@umich.edu

VUW

Graeme Kennedy

Averil Coxhead

Professor Paul Nation
On the four strands, extensive reading, and more

Hannah McCulloch's conversation with Professor Paul Nation covers a wide range of issues, from his early exposure to the tradition of vocabulary control and wordlists to the use of extensive reading programs as a valuable part of language learning, to his current work with a colleague on a book on extensive reading.

Hannah McCulloch: Tell us a little bit about how your interest in vocabulary began?

Paul Nation: I became interested in vocabulary because when I began teaching at university, my senior colleagues H.V. George and Helen Barnard, had worked in India and were strongly in the tradition of vocabulary control and wordlists. They were very well acquainted with the work of Michael West. So, very early on in my career, I became familiar with vocabulary counts, corpus linguistics, graded readers and other simplified material, and speed-reading.

HM: You have taught in many countries around the world — Indonesia, Finland, Japan, the United States to name just a few. What are some of the major changes you have seen over the years with regards to vocabulary teaching?

PN: The major change that I have seen is the very substantial growth in published studies of research and thinking on vocabulary. When I wrote the second edition of my book *Learning Vocabulary in Another Language*, I worked out that of all the research on vocabulary that had appeared in the last hundred years, 30% of it had appeared in the last ten years. This trend continues.

Perhaps an aspect of growth rather than change is the growing interest in graded readers and extensive reading. The idea of using simplified material has been around for a long time, largely because of the efforts of people like Michael West and Harold Palmer in promoting graded readers and extensive reading. However, recently with the setting up of the Extensive Reading Foundation there has been a new impetus to extensive reading and the use of vocabulary-controlled material. There is now a considerable amount of research in this area and I hope that this research is translated into teachers setting up extensive reading programs.

HM: A question that often plays on teachers' minds is, "What vocabulary do I teach?", and in your work you have suggested good planning using the "four strands" technique. Could you tell us a bit more about this?

PN: The four strands is a guideline for syllabus design. It says that a well-balanced language course should consist of four equally sized strands – meaning-focused input, meaning-focused output, language-focused learning, and fluency development. The value of this principle is that it sees deliberate learning (represented by the language-focused learning strand) as making up no more than 25% of a language course. The other 75% of the course should be the other three communicative strands. In many language courses, there is too much teaching going on and we need to weight the balance back in favour of incidental learning through meaningful language use. Both deliberate learning and incidental learning should be part of a language course, but they have to be present in the right proportions. There is more about the four strands in the book that I most enjoyed writing, *What Should Every EFL Teacher Know?*

The four strands [technique] does not talk about which vocabulary to teach but is directed at how vocabulary can be learnt. The major principle guiding what vocabulary to learn is the frequency principle. This principle says that in general the most frequent vocabulary should be learnt before less frequent vocabulary. The justification for this principle is that the high frequency vocabulary of English, around 3 000 words, covers 80% to 90% of the running words in most spoken and written texts. It makes good sense to learn this useful, very high frequency vocabulary first.

If learners are learning English for special purposes, then they need to consider the vocabulary which is frequent within their special purposes area. An example of this is the survival vocabulary for foreign travel. This consists of 120 words and phrases which are really useful for people who are going to visit another country for a short time. We have had that vocabulary translated into several different languages, and these survival word lists are available from my website.

HM: Can vocabulary really be taught or is it all incidental learning?

PN: Vocabulary can be taught, but vocabulary teaching should make up a rather small proportion of a vocabulary-focused course. The most important job of the language teacher is to plan. One aspect of planning is making sure that there is a balance of opportunities to learn through the four strands. Another aspect is choosing the right vocabulary to focus on, and to do this a teacher needs to be aware of how many words the learners know and how well they know those words. Thus, another important job of the teacher is to test, in order to see how many and what words the learners know.

Most teachers feel that their number one job should be teaching. I think this is a misdirected view. The teacher's main jobs, in order of importance, should be to plan, to organise, to train, to assess, and then to teach. Applying the four strands principle and the frequency principle is a useful kind of planning for vocabulary learning. Organising the learners to do extensive reading, extensive listening,

---

Graeme Kennedy materials

Chapter
Between and through
The company they keep and the functions they serve
By *Graeme Kennedy*
Book English Corpus Linguistics

Edition	1st Edition
First Published	1991
Imprint	Routledge
Pages	16
eBook ISBN	9781315845890

tesol QUARTERLY

🔓 Full Access

Amplifier Collocations in the British National Corpus: Implications for English Language Teaching

GRAEME KENNEDY

First published: 04 January 2012 | https://doi.org/10.2307/3588400 | Citations: 6

Abstract
This study examines how adverbs of degree tend to collocate with particular words in the 100-million-word British National Corpus and considers some possible implications for English language teaching. The mutual information measure is used to show the strength of the bond between 24 selected amplifiers such as *extremely* and *greatly* and other words (typically adjectives or participles such as *rare* or *appreciated*, which result in collocations such as *extremely rare* or *greatly appreciated*). Each amplifier is shown to collocate most strongly with particular words having particular grammatical and semantic characteristics. Research in cognitive science has shown the extent to which words and collocations become established as units of language depending on the frequency with which they are experienced. In the light of the corpus-based evidence on the nature of collocations presented in this study, the teaching of collocations might be expected to have a more explicit and prominent place in the language teaching curriculum. In class, teachers can draw attention to collocations not only through direct teaching but also by maximizing opportunities to acquire them through an emphasis on autonomous implicit learning activities such as reading.

Averil Coxhead materials

The Academic Word List

The Academic Word List is a useful English resource for lecturers and students.

Averil Coxhead from the School of Linguistics and Applied Language Studies developed and evaluated The Academic Word List (AWL) for her MA thesis. This list is a very useful resource for English for Academic Purposes teachers and learners.

— AWL Information
— AWL Sublist Families
— AWL Headwords
— AWL Most Frequent Words in Sublists
— AWL Tests
— Useful links

JOURNAL ARTICLE

An Academic Formulas List: New Methods in Phraseology Research Get access ›

Rita Simpson-Vlach, Nick C. Ellis

Applied Linguistics, Volume 31, Issue 4, September 2010, Pages 487–512,
https://doi.org/10.1093/applin/amp058
Published: 12 January 2010

❝ Cite   🔑 Permissions   ◂ Share ▾

Abstract
This research creates an empirically derived, pedagogically useful list of formulaic sequences for academic speech and writing, comparable with the Academic Word List (Coxhead 2000), called the Academic Formulas List (AFL). The AFL includes formulaic sequences identified as (i) frequent recurrent patterns in corpora of written and spoken language, which (ii) occur significantly more often in academic than in non-academic discourse, and (iii) inhabit a wide range of academic genres. It separately lists formulas that are common in academic spoken *and* academic written language, as well as those that are special to academic written language alone and academic spoken language alone. The AFL further prioritizes these formulas using an empirically derived measure of utility that is educationally and psychologically valid and operationalizable with corpus linguistic metrics. The formulas are classified according to their predominant pragmatic function for descriptive analysis and in order to marshal the AFL for inclusion in English for Academic Purposes instruction.

---

four equal strands of meaning focused input, meaning focused output, language focused learning, and fluency development. 95% coverage

ncellis@umich.edu