WHAT’S IN A WORD?  

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?**  

Nick Ellis  
Vocab@Vic 2023  

ncellis@umich.edu
Shakespeare Romeo and Juliet (1597) What’s in a name?

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

**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.
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
WHAT’S IN A WORD?
Cognitive-linguistic, Neuroscientific, AI, Psycholinguistic, and Usage-based perspectives

Nick Ellis
Vocab@Vic 2023

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

- 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.

Routine Psycholinguistic determinants of Learning

The Construction to be learned

- cue: today
- at the present period of time

The Psychology of Learning

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


Where words make sense
A semantic atlas of the cerebral cortex

https://aeon.co/videos/see-how-our-brains-group-words-by-meaning-in-surprisingly-complex-semantic-maps

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

Sara F. Popham¹, Alexander G. Huth¹, Natalia Y. Bilenko¹, Fatma Deniz¹,², James S. Gao¹, Anwar O. Nunez-Elizalde¹ and Jack L. Gallant¹

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.
Signification - Current Neuroscience  Ev Fedoronko

The language system in the human brain

Ev Fedorenko (MIT)
December 1, 2021
Innovators in Cognitive Neuroscience Seminar Series

https://www.youtube.com/watch?v=sSr152-vOLc&ab_channel=InnovatorsCogNeuro

The language system in the human brain:
Parallels and differences with large language models

Ev Fedorenko (MIT)
December 13, 2022
MAIN2022

https://www.youtube.com/watch?v=uE9AiYuCwdE&t=2s&ab_channel=MAINConference

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
A word’s **company**....

Cognitive Linguistics
Usage-Based Grammar
We learn Constructions

- 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

---

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

ncellis@umich.edu
A word’s company - LLMs

DISSOCIATING LANGUAGE AND THOUGHT IN LARGE LANGUAGE MODELS

A PREPRINT

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

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

Man A. Blank 
The University of California Los Angeles 
blank@psych.ucla.edu

Susa Kaukola 
Massachusetts Institute of Technology 
kaukola@mit.edu

Joshua R. Tenenbaum 
Massachusetts Institute of Technology 
tenenbaum@mit.edu

Evelina Fedorenko 
Massachusetts Institute of Technology 
fedorenko@mit.edu

November 7, 2023

ABSTRACT

Large language models (LLMs) have come closest among all models 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 pattern—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. Although the success of the Turing test has since been questioned, it has undoubtedly shaped the way society today thinks of machine intelligence (Boehm et al., 2019; French, 1990, 2000; Marcus et al., 2016; Mee, 1975; Piir, 2018).

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

Can large language models produce expert-quality philosophical texts? To investigate this, we fine-tuned GPT-3 with the works of philosophers 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.”

KEYWORDS

artificial intelligence, Daniel C. Dennett, human-machine discrimination, language models, philosophical expertise

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 (Arora 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 public attention for its capacity to generate passable short student essays (Huang, 2023). But presumably expert-level professional philosophy
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

Between and through
The company they keep and the functions they serve
By Graeme Kennedy

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

Graeme Kennedy
First published 04 January 2013 | https://doi.org/10.5962/bhl.title.18808 | Citation: 6

Abstract
This study examines how adverts of degree tend to collocate with particular words in the 10th million-word British National Corpus and considers some possible implications for English language teaching. The mutual information measure used to show the strength of the bond between adverts and collocations such as anyone or greatly and other words typically adjectives or participles such as seem or appear to which result in collocations such as extremely peer greatly approved. Such a measure is shown to collocate most strongly with 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 learning dependent on the frequency with which they are experienced. It is the light of this that is based evidence on the nature of collocations presented in this study, the teaching of collocations might be expected to contribute a more explicit and permanent parent in the language reading 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.

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

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 traditional view 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 an extensive reading program.

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 English, my career 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.

Hannah McCulloch: You are 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 regard to vocabulary teaching?
Paul Nation: The major change that I have seen, is the 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 a 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. Now there is a considerable amount of research in this area and I hope that this research is translated into teachers setting up extensive reading programs.

Hannah McCulloch: A question that often plagues teacher’s 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?
Paul Nation: The four strands is a guideline for syllabus design. It says that a well-designed 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 is deliberate learning (represented by the language-focused input); reading, making up to more than 25% of a language course. The other 15% of the course should be the other three strands. In many language courses, there is too much teaching going on and we need to weight the balance of the time spent on learning through meaningful language use. Both deliberate learning and incidental learning should go on, 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 high-frequency vocabulary of English, around 3,000 words, owns 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.

Hannah McCulloch: Can vocabulary really be taught, or is it all incidental learning?
Paul Nation: 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 for the task, 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 misguided view. The teacher’s main job, in order of importance, should be to plan, to organize in tasks to assess, and then to teach. Applying the four strands principle and the frequency principle is a useful kind of planning for vocabulary learning. Organizing the learners to do extensive reading, extensive learning,

Averil Coxhead
The Academic Word List

The Academic Word List is a useful English measure for teachers and students.

The Academic Word List (AWL) is a list of 1,000 words that are considered to be essential for second language learners. The list is based on the frequency of occurrence in written and spoken language, which is (even) significantly more often in academic than in non-academic discourse, and (also) of relatively wide range of academic genres. It separates into 1000 formulas that are common in academic spoken and written lexicon, as well as those that are special in academic written language alone and academic written language alone. The list has further partitioned 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 mark the AWL, to be included in English for Academic Purposes examinations.

ncellis@umich.edu