

A Reproducibility Details

A.1 Data Collection

Tokens from the British National Corpus (Burnard, 2000) were collected using NLTK v3.4.5. The sentences in the corpus were shuffled to ensure a random sample of tokens for each word. Lists of token sentences for each word, along with their original BNC indices, are available in the `word_data` directory in the supplemental material. To generate token vector representations, we used the HuggingFace `pytorch_pretrained_bert` implementation of the pre-trained `bert-base-uncased`.

A.2 Evaluation Datasets

For a few words, the BNC did not contain enough tokens to generate multi-prototype embeddings for some choices of K . These words are not included in the analyses for that choice of K . For instance, if only 7 tokens of a word were collected, and $K=8$, predictions were not calculated for pairs containing that word. Fortunately, this was very rare, and most words in the evaluations datasets have at least 50 tokens in the BNC, if not more. Tables 1 and 2 give detailed information about which word-pairs from each dataset, if any, were not evaluated at each cluster. As a consequence of the pruning, the Spearman’s ρ correlation was sometimes calculated on minimally different data from one cluster to the next. However, the differences are so minimal as to make the issue negligible. Importantly, for the unioned models which we compare to other approaches, over 99% of all word-pairs were evaluated for each dataset.

A.3 Supplemental Material

The codebase for this project, including scripts for collecting data, generating BERT representations, calculating clusters, evaluating models, and visualizing results, is available at <https://github.com/gchronis/MProBERT>. Here one can also find lists of the tokens included in this analysis along with their BNC indices.

B Clusters

K-means clustering of token BERT representations captures polysemy as well as different usage types. Tables 3 and 4 shows a selection of sentences in clusters for *stream* and *river* at layer 8 for $k=3$ clusters. Tables 6 and 5 show a representative selection of tokens from each of the clusters at layer 8 for $k=7$. While the prototypes are often aligned

in part by grammatical constructions, there is an interesting concordance between form and semantic subsense or conceptual affordance. Cluster 2 for *stream* mostly contains examples of the phrase *stream of X*. This construction has the effect of focusing attention on the substance constituting the stream and emphasizing its movement. Interestingly, the examples in this cluster which do not contain this construction also emphasize the movement of a substance. Compare with *X stream*, where the stream is construed as a medium or channel through which other entities pass.

Dataset	K	Percentage Evaluated	Omitted Words	Omitted pairs
WordSim353_sim	8	99.51	{aluminum}	{aluminum-metal}
	9	99.01	{aluminum, kilometer}	{mile-kilometer, aluminum-metal}
	10	99.01	{aluminum, kilometer}	{mile-kilometer, aluminum-metal}
WordSim353_rel	9	99.6	{kilometer}	{territory-kilometer}
	10	99.6	{kilometer}	{territory-kilometer}
WordSim353	8	99.72	{aluminum}	{aluminum-metal}
	9	99.15	{kilometer, aluminum}	{mile-kilometer, territory-kilometer, aluminum-metal}
	10	99.15	{kilometer, aluminum}	{mile-kilometer, territory-kilometer, aluminum-metal}
SimLex999	4	99.8	{orthodontist}	{orthodontist-dentist, doctor-orthodontist}
	5	99.7	{disorganize, orthodontist}	{orthodontist-dentist, doctor-orthodontist, disorganize-organize}
	6	99.7	{disorganize, orthodontist}	{orthodontist-dentist, doctor-orthodontist, disorganize-organize}
	7	99.7	{disorganize, orthodontist}	{orthodontist-dentist, doctor-orthodontist, disorganize-organize}
	8	99.5	{aluminum, disorganize, orthodontist}	{metal-aluminum, tin-aluminum, orthodontist-dentist, doctor-orthodontist, disorganize-organize}
	9	99.5	{aluminum, disorganize, orthodontist}	{metal-aluminum, tin-aluminum, orthodontist-dentist, doctor-orthodontist, disorganize-organize}
	10	99.5	{aluminum, disorganize, orthodontist}	{metal-aluminum, tin-aluminum, orthodontist-dentist, doctor-orthodontist, disorganize-organize}

Table 1: Words from each dataset for which fewer than K tokens were collected, along with word pairs that were consequently omitted from evaluation. Where K is not listed, the number of tokens collected for each word was sufficient to construct K -prototype vectors. Each entry reports the percentage of word-pairs in the dataset evaluated for that K .

Dataset	K	Percentage Evaluated	Omitted Words	Omitted pairs
YP-130	3	99.23	{commercialize}	{distribute-commercialize}
	4	99.23	{commercialize}	{distribute-commercialize}
	5	99.23	{commercialize}	{distribute-commercialize}
	6	99.23	{commercialize}	{distribute-commercialize}
	7	99.23	{commercialize}	{distribute-commercialize}
	8	99.23	{commercialize}	{distribute-commercialize}
	9	99.23	{commercialize}	{distribute-commercialize}
MEN	10	99.23	{commercialize}	{distribute-commercialize}
	1	99.93	{ipod}	{chair-ipod, ipod-rope}
	2	99.93	{ipod}	{chair-ipod, ipod-rope}
	3	99.93	{ipod}	{chair-ipod, ipod-rope}
	4	99.87	{donut, ipod}	{cafe-donut, chair-ipod, ipod-rope, donut-panda}
	5	99.87	{donut, ipod}	{cafe-donut, chair-ipod, ipod-rope, donut-panda}
	6	99.87	{donut, ipod}	{cafe-donut, chair-ipod, ipod-rope, donut-panda}
	7	99.67	{donut, colorful, ipod}	{colorful-outfit, cafe-donut, colorful-toy, colorful-frame, colorful-duck, colorful-wood, colorful-lab, chair-ipod, ipod-rope, donut-panda}
	8	99.67	{donut, colorful, ipod}	{colorful-outfit, cafe-donut, colorful-toy, colorful-frame, colorful-duck, colorful-wood, colorful-lab, chair-ipod, ipod-rope, donut-panda}
	9	99.67	{donut, colorful, ipod}	{colorful-outfit, cafe-donut, colorful-toy, colorful-frame, colorful-duck, colorful-wood, colorful-lab, chair-ipod, ipod-rope, donut-panda}
SimVerb3500	10	99.67	{donut, colorful, ipod}	{colorful-outfit, cafe-donut, colorful-toy, colorful-frame, colorful-duck, colorful-wood, colorful-lab, chair-ipod, ipod-rope, donut-panda}
	1	99.94	{misspend}	{misspend-pass, pass-misspend}
	2	99.94	{misspend}	{misspend-pass, pass-misspend}
	3	99.8	{broil, misspend}	{bake-broil, broil-cook, broil-burn, broil-fry, broil-boil, misspend-pass, pass-misspend}
	4	99.8	{broil, misspend}	{bake-broil, broil-cook, broil-burn, broil-fry, broil-boil, misspend-pass, pass-misspend}
	5	99.69	{broil, plow, misspend}	{plow-dig, bake-broil, sow-plow, mow-plow, broil-cook, broil-burn, broil-fry, broil-boil, plow-hit, misspend-pass, pass-misspend}
	6	99.6	{intoxicate, broil, plow, misspend}	{plow-dig, bake-broil, sow-plow, drink-intoxicate, mow-plow, broil-cook, broil-burn, broil-fry, broil-boil, plow-hit, dislike-intoxicate, belong-intoxicate, misspend-pass, pass-misspend}
	7	99.6	{intoxicate, broil, plow, misspend}	{plow-dig, bake-broil, sow-plow, drink-intoxicate, mow-plow, broil-cook, broil-burn, broil-fry, broil-boil, plow-hit, dislike-intoxicate, belong-intoxicate, misspend-pass, pass-misspend}
	8	99.6	{intoxicate, broil, plow, misspend}	{plow-dig, bake-broil, sow-plow, drink-intoxicate, mow-plow, broil-cook, broil-burn, broil-fry, broil-boil, plow-hit, dislike-intoxicate, belong-intoxicate, misspend-pass, pass-misspend}
	9	99.6	{intoxicate, broil, plow, misspend}	{plow-dig, bake-broil, sow-plow, drink-intoxicate, mow-plow, broil-cook, broil-burn, broil-fry, broil-boil, plow-hit, dislike-intoxicate, belong-intoxicate, misspend-pass, pass-misspend}
	10	99.51	{intoxicate, hypnotize, broil, plow, misspend}	{plow-dig, bake-broil, sow-plow, drink-intoxicate, mow-plow, spell-hypnotize, hypnotize-control, broil-cook, broil-burn, broil-fry, broil-boil, plow-hit, hypnotize-remember, dislike-intoxicate, belong-intoxicate, misspend-pass, pass-misspend}

Table 2: Continued from last page. Words from each dataset for which fewer than K tokens were collected, along with word pairs that were consequently omitted from evaluation. Where K is not listed, the number of tokens collected for each word was sufficient to generate K clusters. Each entry reports the percentage of word-pairs in the dataset evaluated for that K .

Cluster ID	Sentences
0	<p>Care must be taken to select the correct neutralizing agent for the specific odorous gas to be treated and there are obvious difficulties when both acidic and alkaline compounds are present in the gas stream.</p> <p>Alcohol is absorbed into the blood stream via the stomach and takes effect within 5-10 minutes.</p> <p>Every infected stream has a 'parent' stream, and it may have more than one 'daughter' stream. That is to say, the infant must convert stimulation from light rays, sound waves, from the speech stream into the appropriate representational grist if it is to get the kind of information that it requires from the world; but this gleaning of information does not constitute thought.</p>
1	<p>Without waiting for the others he plunged down the bank into the stream, slipping and slithering heedlessly over the protruding roots and rocks.</p> <p>The horses were quietly cropping the rich grass by the stream.</p> <p>Using cement of their own manufacture, they skilfully build tubular houses for themselves out of materials that they pick up from the bed of the stream.</p> <p>A small lagoon is formed by the stream between a sandbank and the rock wall.</p>
2	<p>Thus a stream of pulses lasting 1 second each and given at 10 second intervals could be the 'background' (they could be sound pulses or pulses on a screen, for example); the 'signal' being sought could be the absence of a pulse, one that was shorter or longer than the standard value or one that appeared too soon or too late.</p> <p>Without stopping, the combine disgorged a stream of grain into the trailer.</p> <p>Burning straw was the best fun — it was poked through the grill at the front of the grate and, when it caught fire, smoke would stream out of the other end.</p> <p>As he ate and drank she found herself chattering away to him out of nervousness, a stream of things that went through her head, the small happenings of a day.</p>

Table 3: Example sentences from each *stream* cluster with $k=3$ and $l=8$. Cluster 0 is a bit of a catchall cluster, encompassing idioms like 'came on stream', but it contains all mentions of a stream as a medium through which other media pass. Cluster 1 represents the natural feature. Cluster 2 captures, but is not limited to, the usage 'stream of X'. Other usages in this cluster share with the construction a focus on the movement of the substance constituting the stream rather than the substance as a medium to move through. Consequently, cluster 2 contains more abstract streams, in that many of its arguments (often count nouns) are not typically thought of as fluids. Note the relationship between grammatical form and semantic subsense / conceptual affordance.

Cluster ID	Sentences
0	<p>In this tale, two weeds grew on a river bank; one of them conserved its energy, and grew low and small and brown, with its sights set on a long life, while the other put forth all its strength into growing tall and into colouring itself a beautiful green.</p> <p>The lights dazzled, but on the broad face of the water there were innumerable V shaped eddies, showing the exact position of whatever the river had not been able to hide.</p> <p>Across the river and through the streets of Cliffe men fought in close combat before the royalists scattered.</p> <p>How can he get all three safely over the river?</p>
1	<p>The River Doon flows north-west from Dalmellington, past Patna and Dalrymple, under the Auld Brig o' Doon at Alloway, where Tam o' Shanter escaped from pursuing witches in Burns's magnificent poem :</p> <p>The Malá Strana or Lesser Town spreads beneath the castle to the banks of the River Vltava.</p> <p>Turkey's Prime Minister Suleyman Demirel arrived in Nakhichevan on May 28 to attend the opening of a bridge between Azerbaijan and Turkey over the river Arax.</p> <p>The river Sol is the southernmost of the Empire's rivers.</p>
2	<p>A few low hovels that had once been homes to river people were now derelict, and an empty building which was once a sailmaker's and then a barge-builder's premises now stood empty after its last owner, a steam-traction engineer, foundered in the changing times.</p> <p>In 1972 the Government of Sind Province declared the river dolphin protected by law and prohibited its killing and trapping.</p> <p>As in India, a river, a hill, mountain or lake, in Celtic legend, is personified by a god-like person.</p> <p>The same accountants apparently proposed getting rid of river wardens and people in pollution control.</p>

Table 4: Four occurrences of *river* in the BNC belonging to each cluster with $k=3$ and $l=8$. Cluster 0 corresponds to a river as a natural feature, cluster 1 captures the construction *river X* where *X* is the name of the river, and cluster 2 contains river used as an adjective to describe things associated with rivers.

Cluster ID	Sentences
0	<p>Homer described it as a monster with the body of a goat, tail of a DRAGON and head of a lion, belching flames.</p> <p>With one sister slightly older and another two years younger there were real female ’’ spats’ at times with the sisters fighting like young lion cubs.</p> <p>A creature appeared, a lion, red and huge, bounding up the narrow winding streets of Edinburgh, splashing through rivers of blood which poured from the castle.</p> <p>The vertebrates found from that period include mammoth and other extinct elephants, extinct rhino, hippopotamus, giant deer, lion, spotted hyaena, tortoise and macaque (from the ‘monkey gravel’ of West Runton, Norfolk — where else?).</p>
1	<p>Finally, let us rekindle that vision in Isaiah 11 where the lion does not eat the lamb but lies down in a symbiotic relationship with it.</p> <p>Looking down from a height of ten or twelve feet, she saw an old friend, the MGM lion.</p> <p>He was like a man fearing his moment had come, he said, covering his eyes in silent prayer — yet astonished to find the lion in the same pose.</p> <p>The peace-keepers successfully tame the roaring lion.</p>
2	<p>The Blue Lagoon was the old Red Lion renamed, no one knew why, on the corner of Bankside and Trinity Street.</p> <p>Mrs Johnston, 35, was found dead at The Lion public house on Moorfields in Liverpool city centre last Thursday lunchtime.</p> <p>It was further established that Bacon had purchased some arsenic from a shop in Red Lion Square only days before, allegedly to kill rats.</p> <p>Glaxo sold its factory, as did Gresham Lion, and its successor, Dowty, which could not make the business succeed, sold to the Taiwanese.</p>
3	<p>The company, which claims the lion’s share of the object database market, has yet to record a profit.</p> <p>If this is done, care must be taken to ensure that each slice receives its proper priority in order of payment, otherwise one party may take the lion’s share of the income at the expense of the other.</p> <p>But he gave the lion’s share of the credit for the victory to Snodin, playing his first full 90 minutes for two and a half years after a series of hamstring and knee injuries put his career in doubt.</p> <p>Yet last week he had married Magda Tannenbaum, daughter of Sigmund Tannenbaum of Bradford and Hamburg, a wool merchant of legendary wealth, enormous possessions, and no son to inherit the lion’s share of them.</p>
4	<p>Three days after Fraser’s departure a large new flag bearing the arms of Dunbar and March, a white lion on red, flew from the castle’s topmost tower, indicative that the Earl had arrived.</p> <p>The arms granted to his chosen foundation were the fleur-de-lis of France and the royal lion of England, above the three lilies of the Virgin Mary.</p> <p>The sun shone through an elaborate crest of arms in coloured glass, with the lion of Venice rampant above a flurry of plumes and a Latin motto, the glass throwing dark Harlequin patterns on to his expressionless face.</p> <p>The green left sleeve brassard carries a red-on-yellow rampant Lion of Scotland patch, which we are told is special to the CO and his crew.</p>
5	<p>Henry the Lion</p> <p>At Acre, the ramparts of Richard Coeur de Lion’s massive fortress stretch down to a tideless Mediterranean while tiny Arab figures promenade in the dusk past the serail.</p> <p>The play was the true story of bachelor Mr Lewis, author of <i>The Lion</i>, <i>The Witch And The Wardrobe</i>, and his meeting at the age of 50 with writer Joy Davidman.</p> <p>They sent ambassadors to England to encourage marriage arrangements between two of Henry’s daughters and Henry the Lion and one of Barbarossa’s sons.</p>
6	<p>He was one of the first eminent European scientists to make a career in the USA, and rapidly became a lion : his lectures and books were popular, and he built up a school and museum at Harvard.</p> <p>She was one fine lion and I do n’t blame Raja, only it wasn’t me.</p> <p>He was only too well aware of the Talmudic dictum that a handful does not satisfy a lion, but he was neither apologetic nor guilty over it.</p> <p>He looked frightening and she had a momentary sensation of having caged herself in with an angry lion.</p>

Table 5: Four tokens from the BNC from each cluster for *lion* with $k=7$ and $l=8$. Cluster 0 corresponds to a lion as a wild animal, with respect to other animals and the features which distinguish lions from them. Cluster 1 corresponds to lions interacting with humans and especially acting in ways that are unstereotypically docile. Cluster 2 corresponds to place names containing *Lion*. Cluster 3 corresponds to the idiom *lion’s share*. Cluster 4 corresponds to lions on heraldic coats of arms. Cluster 5 corresponds to human or character names containing *Lion*. Cluster 6 corresponds to metaphorical senses of *lion* to describe a human.

Cluster ID	Sentences
0	Did you believe that, Thomas is part of an eight cat routine at Circus International at, it's been at Sutton Coldfield, it's at Northfield over the weekend, but these are just domestic cats, ordinary domestic cats.
1	<p>'Cat got yer' tongue?'</p> <p>Paul says he saw a little cat swallowing a big dog.</p> <p>The list consisted of an assortment of well known mammals, birds, reptiles, fish and invertebrates, and also included three domestic species : dog, horse and cat.</p> <p>I, I know that one hey diddle diddle the cat and the fiddle the cow jumped over the moon, the little dog laughed to see such fun and the dish ran away with the spoon</p>
2	<p>He sat looking at the fire with lowered eyelids, a contented expression on his face, looking like a big overfed cat.</p> <p>Ecstatic, the boy tried on one mask after another, roaring like a lion or mewling like a cat.</p> <p>He took off his clothes, and Isobel curled up on the bed watching him, like a little cat.</p> <p>Lacuna was looking like a cat that had seen its prey.</p>
3	Obviously, if you are eating more fibre-rich food you are likely to cat more grams of fibre.
4	<p>A FAMILY gave up their holiday to pay £700 for a life-saving operation on their cat Tilly.</p> <p>She lives in a terraced house in Lancashire with her mum and dad and her cat, Arthur.</p> <p>He sat up and stroked the cat gently, scratching between the backs of its ears, making Bonaventure purr with pleasure.</p> <p>She took out her pen and paper and wrote a very angry letter to the doctor about the death of her valuable cat.</p>
5	<p>A cat does not want to die with the smell of humanity in his nostrils and the noise of humanity in his delicate peaked ears.</p> <p>'All right,' said the Cat.</p> <p>After this, he became a nicer cat, + was n't so proud of himself all the time.</p> <p>'There you are !' she said, and the cat lifted its tail up with pleasure and rubbed its head against the branch.</p>
6	<p>Cat.</p> <p>Cat.</p> <p>Cat.</p>

Table 6: Four tokens from the BNC from each cluster for *cat* with $k=7$ and $l=8$. Cluster 0 corresponds to a unique circus context. Cluster 1 corresponds (loosely) to a cat in relation to other animals. Cluster 2 corresponds to similes likening people to cats by way of stereotypical catlike behavior. Cluster 3 contains a colloquial use of the word to mean 'obtain'. Cluster 4 corresponds to domestic pets. Cluster 5 is not very cohesive, but includes many examples of cats as grammatical subjects or in more agentive roles. Cluster 6 is the word *Cat* in isolation.