

Text Generation using LLMs

Job Schepens

job.schepens@uni-koeln.de

Example text (GPT 3.5)

Generated text: Das rote Phantom

Es war einmal ein kleiner **Junge namens Max**. Max war sieben Jahre alt und lebte in einer kleinen Stadt am Rande des Waldes. Er liebte es, draußen zu spielen und **Abenteuer** zu erleben.

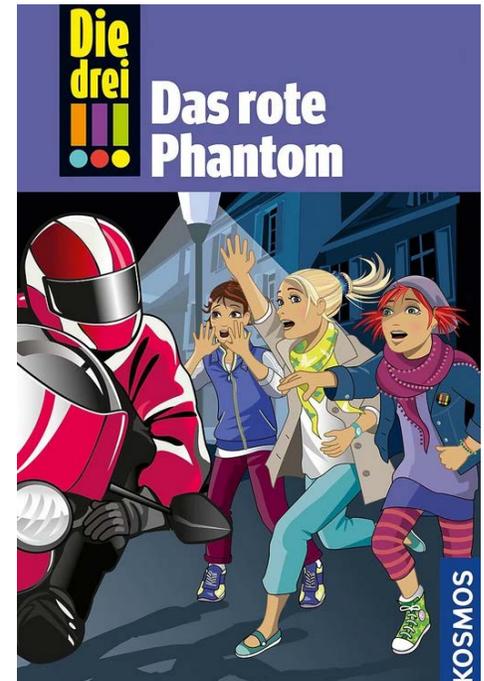
Eines Tages, als Max im Wald spielte, hörte er ein seltsames Geräusch. Es war ein leises Knurren, das immer lauter wurde. Max drehte sich um und sah ein rotes Phantom auf ihn zukommen. Es hatte riesige Augen und scharfe Krallen.

Max war zuerst erschrocken, aber dann erkannte er, dass das rote Phantom nur ein kleiner Hund war. Der Hund war verängstigt und hungrig. Max beschloss, ihm zu helfen.

Klappentext

Kim, Franzi und Marie sind Die drei !!! Mutig und clever ermitteln die drei Freundinnen und sind jedem Fall gewachsen.

Ein Phantom geht um! Nachts stellt ein **vermummter Motorradfahrer** auf seiner roten Maschine Verbrecher und vereitelt ihre Straftaten. Als ein dreister Dieb Franzis Rucksack klaut, ist das rote Phantom sofort zu Stelle. Der Ehrgeiz der drei Detektivinnen ist geweckt – wer ist das rote Phantom? Bei einer wilden Verfolgungsjagd durch die Nacht kommen Kim, Franzi und Marie hinter das Geheimnis und erleben eine erstaunliche Überraschung.



Example text (Llama 3.3 70B)

Generated text: Mein erster Schultag und der Eisbär-Schreck

Es war ein sonniger Morgen im September. Ich, Tim, war so aufgeregt, dass ich heute meinen ersten Schultag hatte. Meine Mutter half mir, mein neues Schulranzen zu packen, und mein Vater machte mir ein leckeres Frühstück.

Als wir zur Schule kamen, sah ich viele neue Gesichter. Ich fühlte mich ein bisschen schüchtern, aber meine Mutter gab mir einen großen Kuss und sagte: "Alles wird gut, Tim. Du wirst viele neue Freunde finden."

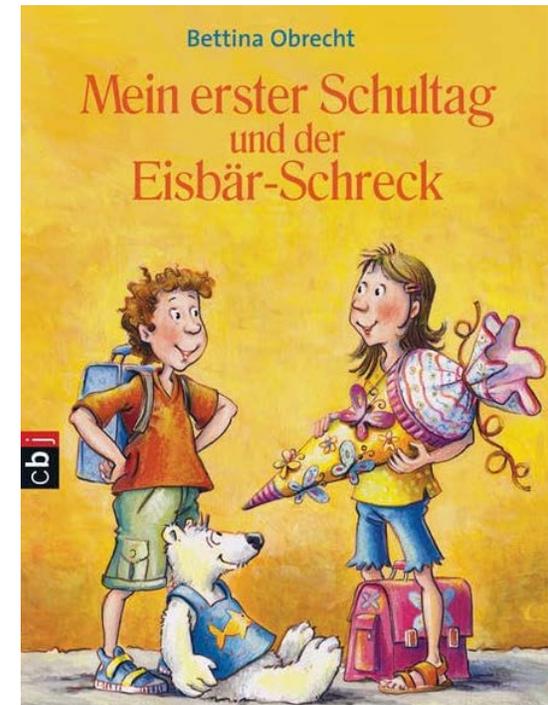
Ich ging in meine Klasse und sah, dass alle anderen Kinder auch ein bisschen nervös waren. Unser Lehrer, Herr Müller, stellte sich vor und erzählte uns, dass wir heute viele spannende Dinge lernen würden.

Klappentext

Ein fröhliches Mut-mach-Buch für den ersten Schultag

Ben freut sich riesig: Endlich kommt er in die Schule! Aber er hat auch ganz schön Lampenfieber, schließlich wird nach den Ferien alles anders sein: Er muss Plüschisbärin Sardine zum ersten Mal alleine lassen! Und was wird aus seiner Schultüte – schließlich kann Mama doch gar nicht basteln! Wie gut, dass er Florence kennenlernt, die Bens beste Freundin wird. Denn wenn man zusammen in die Schule kommt, kann gar nicht mehr viel schiefgehen ...

- Mit farbenfrohen Illustrationen von Heike Wiechmann
- Einfühlsame Schilderung der aufregenden Zeit vor dem ersten Schultag
- Ein Vorlesebuch für alle, die auf den ersten Schultag warten



Example text (DeepSeek V3 – long-form text prompt)

Generated text: King-Kong, das Schulschwein, Kapitel 1: Ein ungewöhnlicher Neuzugang

Es war ein sonniger Montagmorgen in der kleinen Stadt Grünenfeld, als in der Grundschule am Waldrand etwas Ungewöhnliches geschah. Die Kinder der Klasse 4a waren gerade dabei, ihre Hefte und Bücher auszupacken, als ihre Lehrerin, Frau Müller, mit einem breiten Lächeln ins Klassenzimmer trat. Hinter ihr folgte ein kleines, rosafarbenes Ferkel, das neugierig mit seinem Ringelschwänzchen wedelte.

„Guten Morgen, Kinder!“, begrüßte Frau Müller die Klasse. „Ich möchte euch heute jemanden ganz Besonderen vorstellen. Das hier ist King-Kong, unser neues Schulschwein.“

Die Kinder starrten das **Ferkel** mit großen Augen an. Ein Schwein in der Schule? Das hatte es noch nie gegeben!

„King-Kong wird ab jetzt bei uns leben und uns helfen, mehr über Tiere und die Natur zu lernen“, erklärte Frau Müller. „Er wird auch ein Teil unserer Klasse sein, und wir werden uns alle gemeinsam um ihn kümmern.“

Klapptext: Hunde sind das Wunderbarste auf der Welt. Aber Jan-Arne hat nur ein **Meerschweinchen**. Das heißt King-Kong und ist total lieb. Ein Hund ist King-Kong natürlich nicht. Mit einem Hund könnte man so viel machen! Und mit einem Meerschweinchen? Überhaupt nichts, sagt Michi aus seiner Klasse. Weil Meerschweinchen echt bescheuert sind. Das kann Jan-Arne natürlich nicht auf King-Kong sitzen lassen...

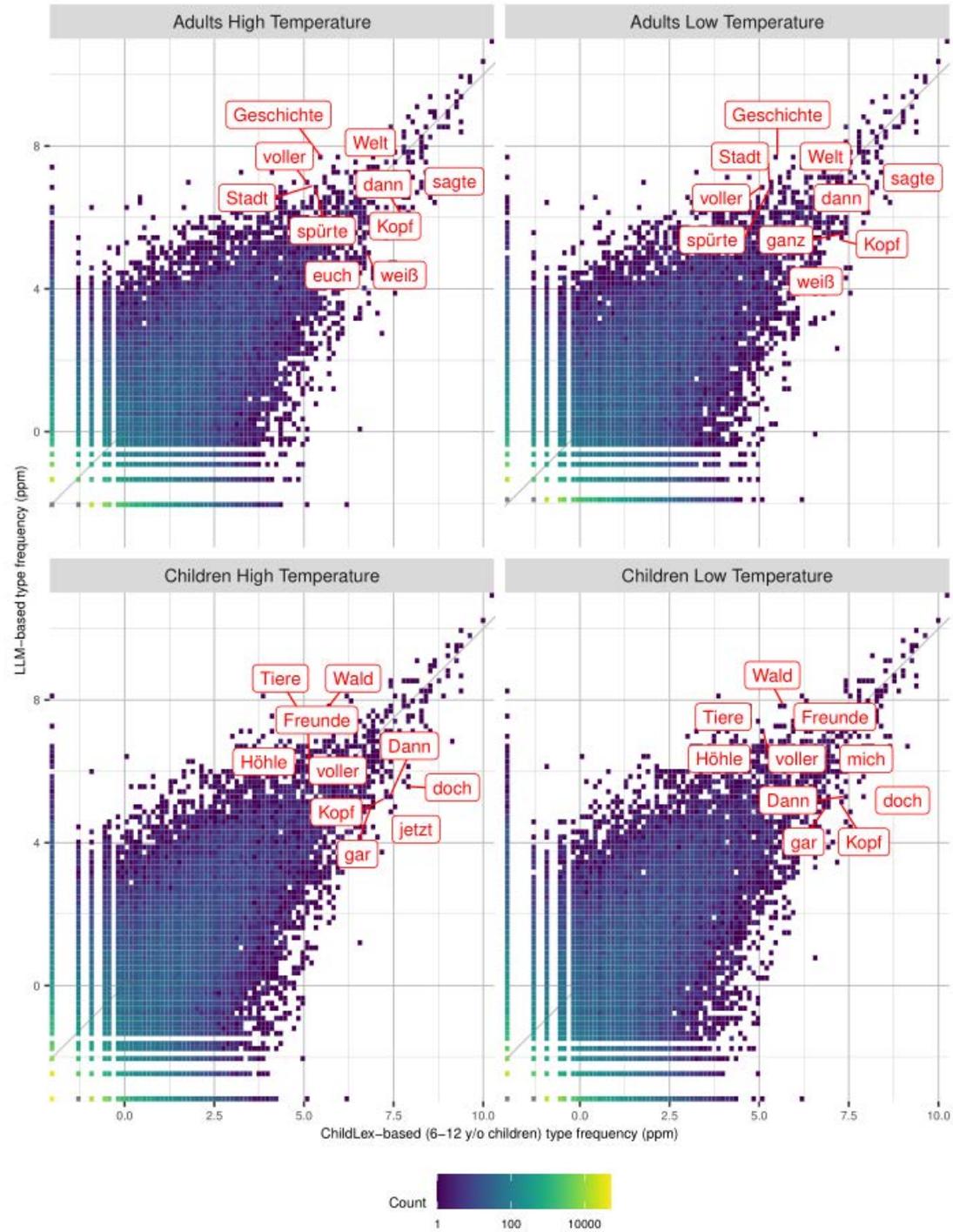


Example text (DeepSeek V3 – long-form text prompt – “ending”)

Mögen wir niemals vergessen, was er uns gelehrt hat: dass die Welt voller Geheimnisse ist, die es zu entdecken gilt, und dass die wahre Stärke oft in den unerwartetsten Gestalten zu finden ist. King-Kong lebt weiter, nicht nur auf der Leinwand, sondern in unseren Herzen und in den Geschichten, die wir von Generation zu Generation weitergeben. Er bleibt ein **Symbol für die ungezähmte Kraft der Natur**, für die Schönheit des Unbekannten und für die unerschütterliche Entschlossenheit, die in jedem von uns schlummert. **King-Kong ist mehr als nur ein Monster oder eine Kreatur** – er ist ein Spiegel unserer eigenen Ängste, Träume und Sehnsüchte.

In einer Welt, die oft von Rationalität und Technologie dominiert wird, erinnert er uns daran, dass es noch Raum für das Mystische, das Unerklärliche und das Wunderbare gibt. Er zeigt uns, dass selbst in der Konfrontation mit dem Unbekannten Respekt und Mitgefühl die mächtigsten Werkzeuge sind.

Mögen wir King-Kongs Erbe ehren, indem wir mutig in die unbekanntesten Welten unserer eigenen Leben vordringen, die Geheimnisse der Natur schützen und niemals aufhören, an die Magie zu glauben, die in jedem Winkel dieser Erde verborgen liegt. Denn solange wir uns an ihn erinnern, wird King-Kong immer bei uns sein – ein stummer Wächter, ein Freund in der Dunkelheit und ein Zeichen dafür, dass die größten Abenteuer oft dort beginnen, wo wir es am wenigsten erwarten.



All corpora: Corpus comparison

Measure	childLex	LLM-corpus
n Books	500	500
Tokens	9,850,786	6,252,808
Types	182,454	46,409
Lemmas	117,952	34,519

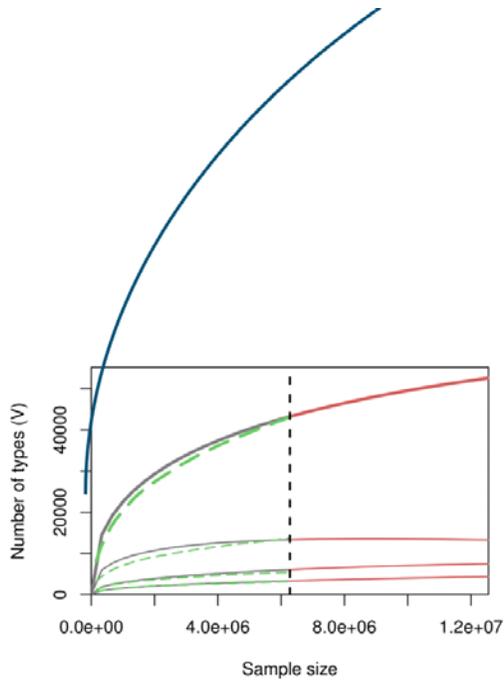
Measure	childLex	Adult Corpus		Child Corpus	
		Low Temp.	High Temp.	Low Temp	High Temp
n Books	500	500	500	500	500
Tokens	9,850,786	7,191,531	7,368,921	23,320,466	23,887,118
Types	182,454	71,423	83,921	84,978	110,603
Lemmas	117,952	52,528	61,318	63,552	82,126

Measure	childLex	Llama long	Llama short	DS-V3 long	DS-V3 short
n Books	500	500	500	500	500
Tokens	9,850,786	10,332,850	7,215,565	9,763,062	7,162,685
Types	182,454	51,320	40,660	239,830	95,321
Lemmas	117,952	39,272	39,272	191,309	74,695

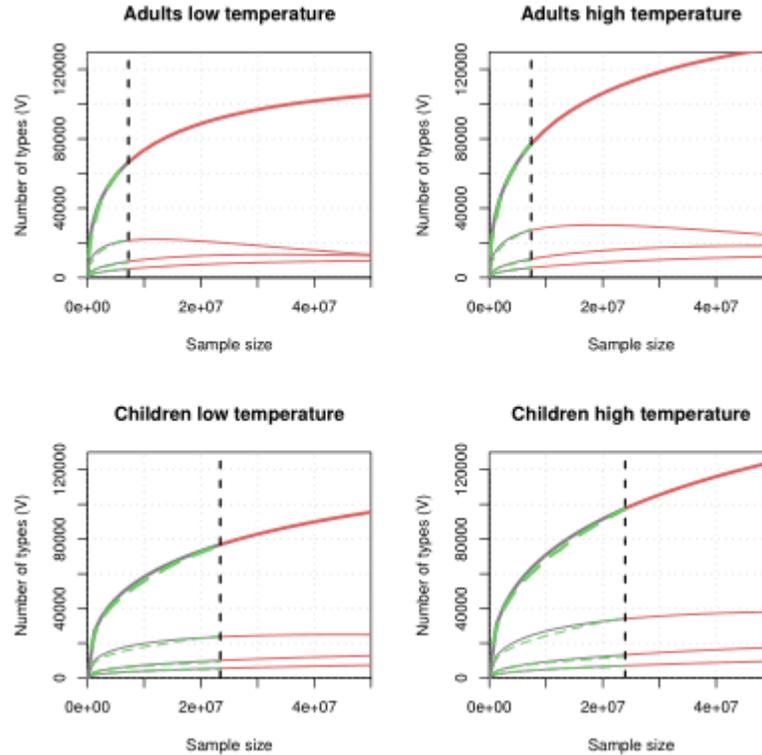
Increase in lexical richness with higher temperature and DeepSeek

All corpora: Corpus comparison

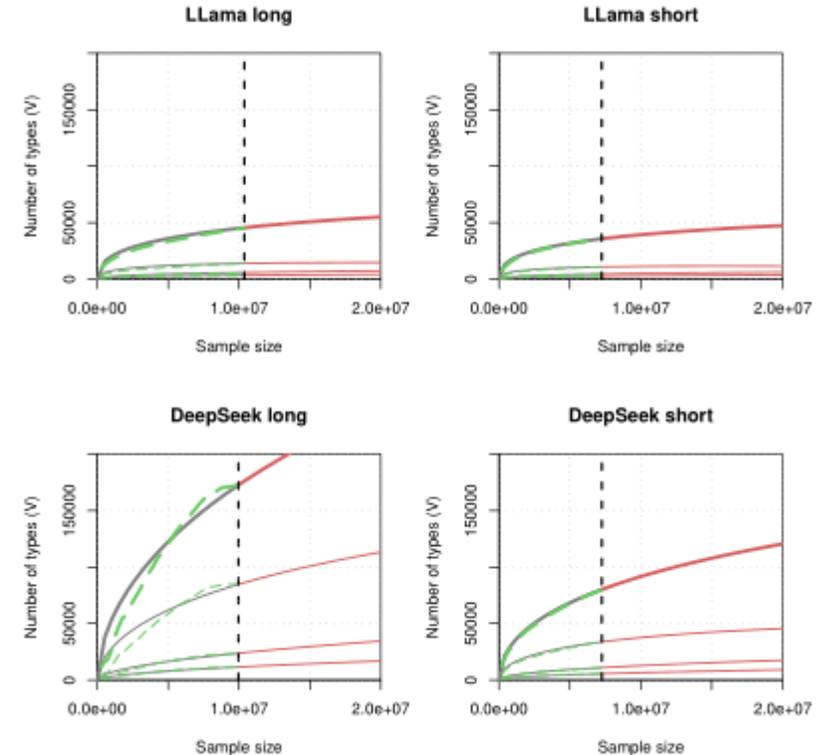
Exp 1



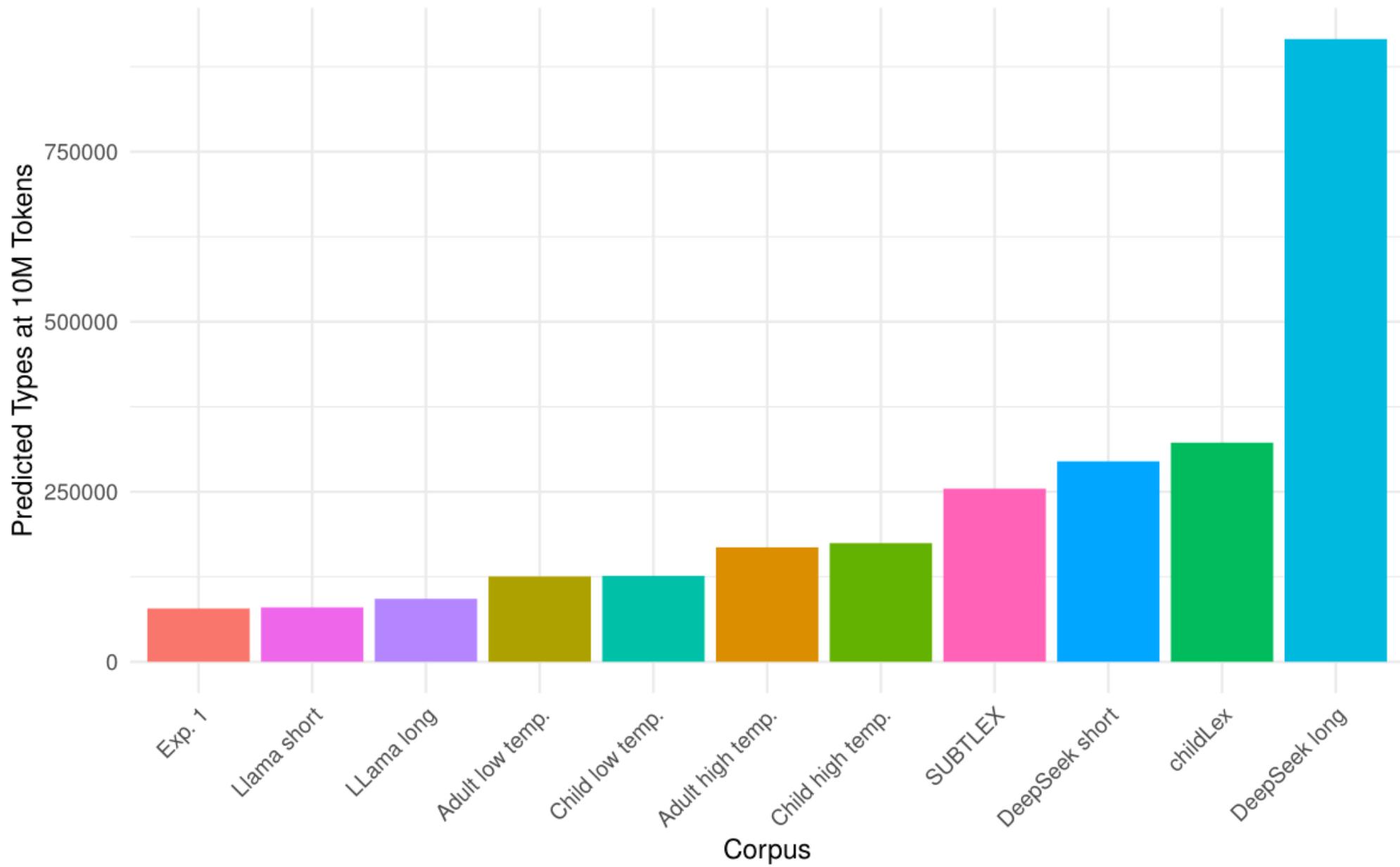
Exp 2

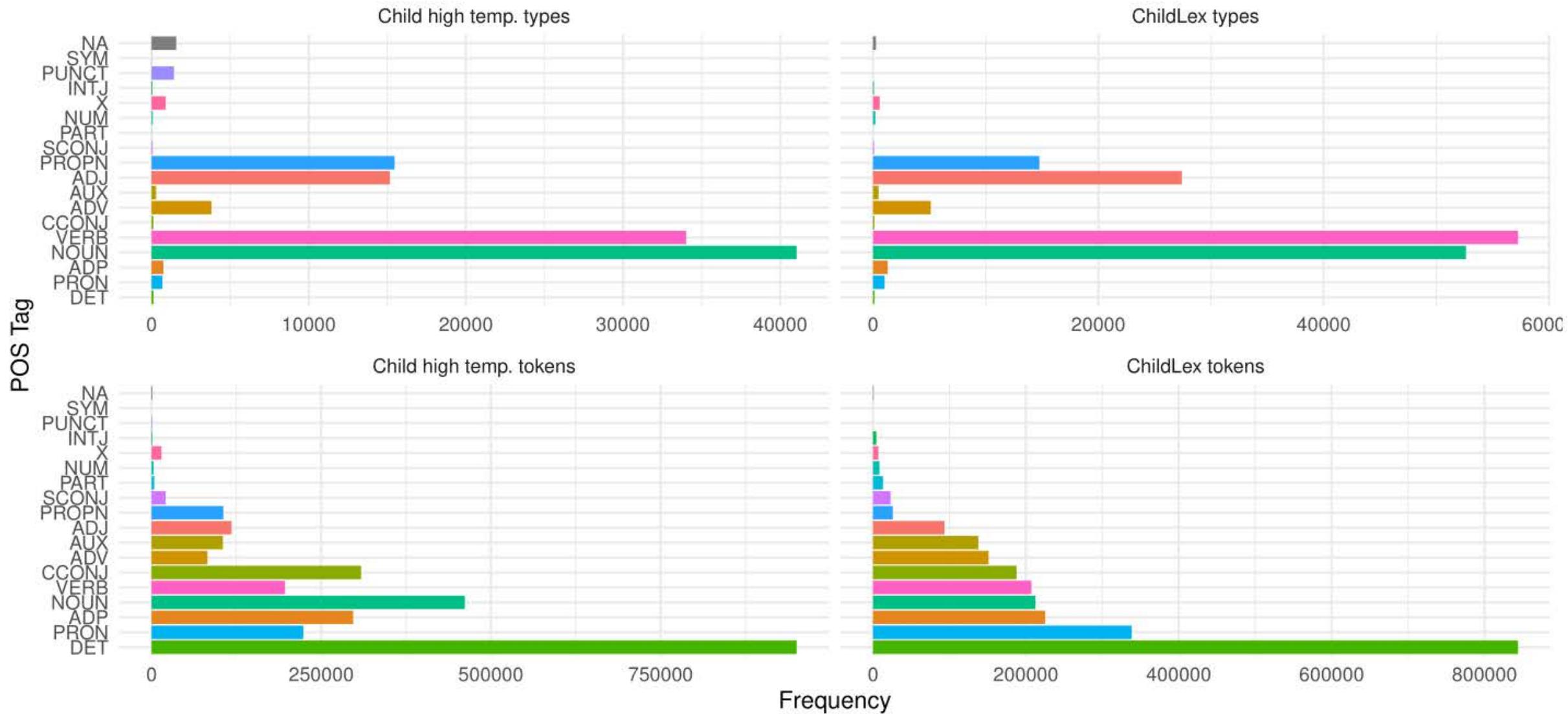


Exp 3



Increase in lexical richness with higher temperature and DeepSeek





LLM > 50, ChildLex < 10

ChildLex > 50, LLM < 10



- POS
- a ADJ
 - a ADP
 - a ADV
 - a AUX
 - a CCONJ
 - a INTJ
 - a NOUN
 - a NUM
 - a PRON
 - a PROPN
 - a SCONJ
 - a VERB
 - a X

Distributions of word embeddings using dimensionality reduction with UMAP to word vectors from FastText-German. Embeddings were reduced from 300-dimensional vectors. Color coding corresponds to different parts-of-speech (POS) tags