Executive Summary: LINKS: Generate linguistically grounded mnemonics by fine-tuning and distilling from reasoning LLMs

Problem: Learning advanced vocabulary (abstract or academic words) is a challenge for language learners. When preparing for standardized tests like GRE, SAT, or TOEFL, students need effective strategies to learn and retain hundreds of challenging words and phrases.

Why mnemonics? Mnemonic devices are memory techniques that help learners connect new vocabulary to something already familiar, making words easier to remember. They could be laborious or difficult to create helpful mnemonics by yourself without being instructed.

What are *helpful* mnemonics? Based on literature, mnemonics should be effective, which must consist of vocabulary + association + mnemonic, with characteristics listed in Figure 1.

Good Mnemonic

preposterous: comes from pre- ("before") + post ("after") + -ous, meaning reversed or absurd. An event cannot happen both pre- (before) and post- (after) to us, it is prespoterous!

- v is used correctly in m
- Clear a linking v and m
- Strong a
- m uses similar or lower vocabulary than v
- m is memorable

Incorrect Definition
Preposterous means very important or significant.

Circular Association
Preposterous sounds like preposterous.

Weak Association
Preposterous sounds like prosperous. Preposterous people usually make prosperous business decisions.

Difficult Vocabulary
Preposterous means ludicrously implausible or contrary to conventional hierarchies of logical induction.

Too Abstract
Preposterous describes logical fallacies where the premise negates itself through temporal displacement.

Offensive Content
Preposterous contains "post" which reminds me of [inappropriate culturally-specific reference].
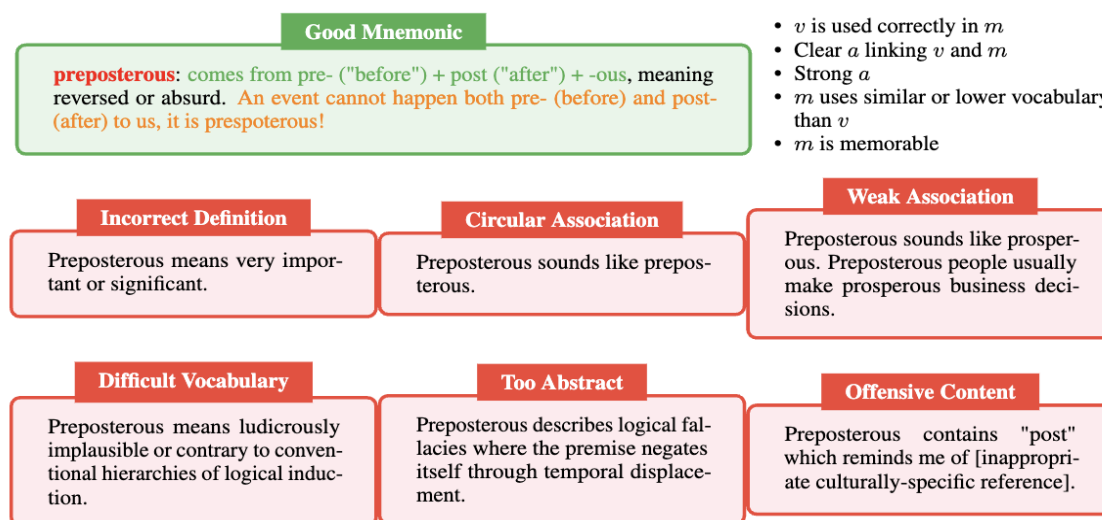
Figure 1: Characteristics of good mnemonics, and examples of bad mnemonics. We propose VAM/VEM model, where a good mnemonic must have three components: vocabulary (v), association (a) (or explanation (e)), and mnemonic (m), with characteristics listed above. These characteristics are also available in list (Appendix A.1)

What are *linguistically grounded* mnemonics (LINKS)? For abstract, advanced vocabulary, simple mnemonic techniques are not enough. LINKS are memory aids that leverage a word's linguistic features such as etymology, sound patterns, or word structure (Table 2). They create meaningful connections that make words easier to remember by helping learners understand *both* what the words mean *and* why they mean what they do, while also tapping into students' prior vocabulary knowledge.

feature	description	example
phonetics	sound patterns	*apparent* sounds like "a bare Asian."
orthography	written/spelling patterns	*abet* looks like "a + bet."
morphology	modern English forms, including free and bound morphemes	*aggrandize* = a + grand + –ize, to mean to make grander.
etymology	origin and history	*adumbrate* comes from Latin ad- (to, on) + umbra (shade) + ate, to mean foreshadow or outline.
semantics	meaning and semantic relationships	*confound* has similar meaning and history with 'confuse'.

How can AI help? Large Language Models (LLMs) acquired linguistic knowledge and demonstrated creative writing capability, so with the right tuning, it could generate linguistically grounded mnemonics for language learners with more ease

My approach: I developed LINKSYS, an AI system that creates LINKS where users can interact with the familiar interface. I distilled linguistic reasoning from a teacher model (DeepSeekR1) into a student model (Google's Gemma 3) that is accessible on a consumer laptop. At this point, there were two models: the baseline student model and LINKSYS.

Results: I scored the mnemonics generated by both models on metrics of helpful mnemonics above and a set of 50 unseen vocabulary. With the same user instruction, **LINKSYS-generated mnemonics that are more linguistically grounded, clear, memorable, with stronger associations between vocabulary and mnemonics**. The results are statistically and practically significant, but the study sample size is small (n=3).

Table 1: Metric comaparison between the base (Gemma-3-1b-it) and LINKSYS models

Evaluation Metric	Base (μ)	LINKSYS (μ)	μ diff.	p-value	Effect size
Correct vocabulary usage (bool)	0.755	0.795	+0.040	0.077	–
Linguistic grounding (bool)	0.670	0.735	+0.065	**0.009**	–
Association strength (1-5)	2.845	3.070	+0.225	**<0.001**	0.464
Clarity (1-5)	3.155	3.410	+0.255	**<0.001**	0.481
Memorability (1-5)	2.600	2.900	+0.300	**<0.001**	0.566

Note: Bold p-values indicate statistically significant differences ($p < 0.05$). Effect sizes are Cohen's d, where 0.2 is small, 0.5 is medium, and 0.8 is large. Boolean metrics (first two rows) do not have effect sizes.