

Additional File: Testing the Linguistic Niche Hypothesis in Large Language Models with a Multilingual Wug Test

Example question from the Wug Test

“One bird-like animal, then two. 'This is a wug. Now there is another one. There are two of them. There are two ___.”

Exemplary translated questions

German

“Ein Vogel, dann zwei. 'Dies ist eine Cra. Jetzt gibt es noch einen. Es gibt zwei von ihnen. Es gibt zwei ___.”

Vietnamese

“Một con chim, sau đó là hai. 'Đây là một con cra. Bây giờ lại có thêm một con nữa. Có hai con. Có hai con ___.”

Prompt prefix pattern

“Assuming that {word} is a {Language} word, read the following paragraph and replace the underscores with a suitable word form of {word}. {translated question}”

Exemplary language model completions

German

Success: “Ein vogelähnliches Tier, dann zwei. 'Das ist ein Wug. Jetzt gibt es noch einen. Es gibt zwei von ihnen. Es gibt zwei Wuge.”

Failure: “Ein Vogel, dann zwei. 'Dies ist eine Cra. Jetzt gibt es noch einen. Es gibt zwei von ihnen. Es gibt zwei Cras.”

Vietnamese

Success: 'Một con chim, sau đó là hai. 'Đây là một con cra. Bây giờ lại có thêm một con nữa. Có hai con. Có hai con cra.'

Failure: 'Đây là một người biết cách bing. Anh ta đang bing. Anh ta đã làm điều tương tự ngày hôm qua. Anh ta đã làm gì hôm qua? Hôm qua anh ta binged.'