

model.layers.5.mlp	Bulgarian	12333	"The highlighted tokens are predominantly suffixes, prefixes, or inflectional endings in Bulgarian (and occasionally other languages), marking grammatical features such as tense, number, gender, case, or aspect, as well as forming participles, adjectives, and nouns. These morphemes are essential for word formation and grammatical agreement in the language."	ja, *
model.layers.5.mlp	Bulgarian	30265	"The highlighted tokens are predominantly suffixes, inflections, or short morphemes in Bulgarian (and some other languages), such as grammatical endings for tense, case, gender, number, or diminutives, as well as common function words and particles. These elements are crucial for the grammatical structure and meaning of words and sentences."	ja, *
model.layers.5.mlp	German	31366	"The highlighted tokens are primarily German (and some Russian) morphemes, word stems, and inflectional endings, often marking grammatical features such as case, number, tense, or forming compound nouns and verbs. Many are parts of longer words or compounds, reflecting the agglutinative and inflectional nature of German, and are often found at the beginning, middle, or end of words."	*, ist
model.layers.5.mlp	Vietnamese	49501	"The highlighted tokens are characteristic morphemes, suffixes, or letter clusters commonly found in Romanian and related Eastern European place names, river names, and surnames, often marking grammatical or regional features."	Alexand
model.layers.6.mlp	Russian	121507	"The highlighted tokens are predominantly morphemes, roots, prefixes, suffixes, or short function words within Russian, Bulgarian, and related Slavic texts. These segments often mark grammatical features, word formation, or serve as connectors, and are frequently found at the beginning, middle, or end of words, reflecting the agglutinative and inflectional nature of these languages."	ль, га, Александ, Г
model.layers.6.mlp	Thai	68498	"The highlighted tokens are morphemes or syllables at word boundaries across multiple languages, often marking inflections, derivations, or meaningful subword units. These include suffixes, prefixes, or roots that contribute to the grammatical or semantic structure of words."	spi, Gro
model.layers.7.mlp	Thai	70835	"The highlighted tokens are primarily morphemes, syllables, or short word fragments from multiple languages, especially Thai, Russian, and Bulgarian, often marking the start or core of content words such as nouns, verbs, or adjectives. These fragments frequently appear at the beginning or within important words, indicating their role in word formation and semantic content across diverse scripts and languages."	О, ль, Г, ир
model.layers.8.mlp	English	123973	"The highlighted spans often represent multi-word phrases or clauses that function as cohesive semantic units, such as prepositional phrases, idiomatic expressions, or descriptive noun phrases. These segments frequently provide contextual detail, specify relationships, or add nuance to the main clause, and are typically composed of function words combined with content words to form meaningful, context-dependent chunks."	ga
model.layers.8.mlp	German	79729	"The highlighted tokens are primarily German function words, verb forms, noun and adjective endings, and common prefixes or suffixes, often marking grammatical relationships, tense, plurality, or case, as well as frequent collocations and set phrases. These tokens are essential for sentence structure and meaning in German text."	Juni, ist
model.layers.8.mlp	Russian	109317	"The highlighted tokens are primarily morphemes, word stems, suffixes, and grammatical endings in Russian (and some in English), often marking case, number, tense, or forming nouns, adjectives, and verbs. There is a focus on functional elements that contribute to word formation, inflection, and syntactic structure, as well as on common connective words and phrase boundaries."	ja
model.layers.9.mlp	German	9083	"The highlighted tokens are primarily function words, verb forms, noun and adjective endings, and common collocations in German sentences, often marking grammatical structure, tense, or case, as well as frequent phrase boundaries and connectors. These tokens are crucial for sentence construction and meaning in German text."	*, Juni, ), ist, eine, ).
model.layers.9.mlp	Portuguese	77250	"The highlighted tokens are predominantly prefixes, suffixes, or stems within Portuguese words, often marking verb conjugations, noun/adjective forms, or common morphemes. These segments frequently appear at the start or end of words, reflecting morphological structure and inflectional patterns in the language."	eine
model.layers.9.mlp	Russian	59767	"The highlighted tokens are predominantly Russian word stems, prefixes, and suffixes that form the core of verbs, nouns, and adjectives, often marking grammatical or semantic roles such as actions, objects, or qualities. These morphemes are crucial for word formation and meaning in Russian, frequently appearing at the beginning or within words to indicate tense, aspect, subject, or object, and are central to the structure and interpretation of Russian sentences."	г, О, ль, га, Александ, ров, на, Г, ир, я