

RelIE	FeatID	Interpreted Function	Languages
6B specific			
[0.94, 0.02, 0.04]	3672	Detects ellipsis and question/exclamation marks	arb,eng,fra,hin,por,spa
[1.00, 0.00, 0.00]	5273	Detects commas and full stop	eng,fra,hin,por,spa
[0.45, 0.28, 0.26]	849	Plural nouns / noun compounds that typically are the subject at BOS, promotes plural verb completion	eng
[0.50, 0.24, 0.26]	10974	3rd person plural pronoun <i>they</i> detector	eng
[0.59, 0.24, 0.17]	9163	Noun-phrase head detector, activates on the key content word (noun or adjective) that carries the meaning of multi-word noun chunk	eng,fra,por,spa
[0.70, 0.23, 0.07]	15758	Detects head of multi token or compound nouns	eng
6B-55B shared			
[0.44, 0.31, 0.25]	10235	Adverbial connectives (, <i>then</i> , <i>tambien</i> , <i>also</i> , <i>tambem</i>)	eng,por,spa
[0.40, 0.35, 0.26]	4332	Detects regular plural nouns	eng
[0.49, 0.37, 0.15]	5819	Activates most on new beginning of clauses right after a punctuation and wanes until a new clause	arb,eng,fra,spa
6B-341B shared			
[0.40, 0.21, 0.39]	5704	-	-
[0.45, 0.23, 0.32]	7007	People-related regular plural nouns (, <i>workers</i> , <i>investors</i> , <i>experts</i>)	eng
55B specific			
[0.19, 0.52, 0.29]	14073	Detects noun beginnings related to academic write-ups (, <i>dissertation</i> , <i>thesis</i> , EnbD - Hindi, <i>essai</i> - French)	eng,fra,hin
[0.00, 1.00, 0.00]	8086	Detects repeated interpuncts, most likely a training data artifact	-
341B specific			
[0.00, 0.00, 1.00]	7694	-	-
[0.00, 0.00, 1.00]	9110	-	-
[0.00, 0.00, 1.00]	15193	Detects punctuation and parenthesis	arb,eng,fra,hin,spa,zh
[0.10, 0.30, 0.60]	11280	-	-
[0.00, 0.00, 1.00]	6461	-	-
[0.01, 0.05, 0.93]	14020	-	-
[0.00, 0.04, 0.96]	2598	-	-
[0.00, 0.00, 1.00]	1638	Detects brackets/parenthesis/punctuation	eng,fra,por,spa
[0.00, 0.00, 1.00]	7066	-	-
[0.00, 0.00, 1.00]	12151	-	-

Table 1: **3-way L1-Sparsity Crosscoder MultiBLiMP English Annotation for BLOOM-1B — 6B 55B 341B.** While this annotation focuses on finding features with just English examples, many features in BLOOM still activate across multiple languages due to cross-lingual representations, showing that the model often leverages multilingual patterns such as shared connectives, pronouns, and beginning of clause detectors.