RelIE	FeatID	Interpreted Function	Languages
6B specific			
[1.00, 0.00, 0.00]	3672	Detects ellipsis and question/exclamation marks	arb,eng,fra,hin,por,spa
[0.78, 0.14, 0.07]	7122	Main-verb head detector	eng
[0.83, 0.14, 0.03]	10388	Plural noun detector for several languages (, *players*, *usários* - users, *al naas* - the people, mEhlAe - ladies)	arb,eng,fra,hin,por,spa
[0.72, 0.23, 0.06]	9163	Noun-phrase head detector of multi-word noun chunk, activates on the key content (noun or adjective) that carries the meaning	eng,fra,por,spa
[0.62, 0.17, 0.20]	15288	Detects *Ev, ev, év* subtokens in different languages	eng,fra,por,spa
[0.59, 0.17, 0.24]	5704	-	-
6B-55B shared			
[0.35, 0.41, 0.24]	15248	Detects the token *that* only in English	eng
[0.36, 0.36, 0.28]	12525	Boss concept detector (, *chief*, *jefe*, *chefs*, *chefe*, mHy - boss)	eng,fra,hin,por,spa
[0.56, 0.27, 0.17]	15758	Detects head of multi-token or compound nouns	eng
55B specific			
[0.15, 0.63, 0.22]	10862	Predicts tokens related to the concept or form of *consult* (, , *fasr* - explain)	eng,fra,por,spa
55B-341B shared			
[0.05, 0.31, 0.64]	6997	Proper-noun/ID detector that activates on named-entity heads	arb,eng,fra,hin,por,spa
341B specific			
[0.00, 0.00, 1.00]	15193	Detects punctuation and parenthesis	arb,eng,fra,hin,spa,zh
[0.00, 0.00, 1.00]	2598	-	-
[0.00, 0.00, 1.00]	12151	-	-
[0.00, 0.00, 1.00]	9110	-	-
[0.00, 0.00, 1.00]	7066	-	-
[0.00, 0.00, 1.00]	6461	-	-
6B-55B-341B shared			
[0.35, 0.32, 0.32]	12140	Multilingual relative pronoun detector (, *que*, *that*, *who*, *aladhi*)	arb,eng,fra,por,spa
[0.32, 0.31, 0.37]	4610	Phrasal-verb/PP-complement detector that fires on the first token of a verb-plus-particle or adjective-plus-preposition pattern	eng,fra,por,spa
[0.39, 0.26, 0.35]	5819	Activates most on new beginning of clauses right after a punctuation and wanes until a new clause	arb,eng,fra,spa

Table 1: **3-way L1-Sparsity Crosscoder CLAMS French/English Annotation for BLOOM-1B — 6B 55B 341B.** The languages column depicts which languages appeared to also use this feature when observing the feature's top-activating sentences. Early checkpoints often rely on language-specific patterns (, English token detectors, suffix patterns), but later checkpoints increasingly learn cross-lingual and language-shared features, such as multilingual pronoun, noun-phrase, and clause detectors.