

RelIE	FeatID	Interpreted Function	Languages
6B specific			
[0.79, 0.00, 0.21]	14483	Locative/in-marker m	hin
[0.67, 0.18, 0.16]	4192	Present-habitual marker using jAnA / honA for singular-masculine subjects	hin
[0.67, 0.22, 0.11]	1174	The -n participle ending marking habitual aspect, here with a masculine-plural (respectful) participle	hin
[0.66, 0.22, 0.13]	8471	Subordinating conjunction Ek (“that”) introducing subordinate clauses	hin
[0.51, 0.24, 0.25]	2539	Perfective participle plural (and oblique-singular) ending used adjectively	hin
6B-55B shared			
[0.45, 0.36, 0.20]	6215	Detects abstract singular nouns (chance, permission, responsibility, order, signal, shelter, danger, etc.)	hin
[0.40, 0.54, 0.07]	4338	Light-verb root kr used as “do” auxiliary in compound verbs (root + another light-verb + conjugation)	hin
6B-341B shared			
[0.64, 0.09, 0.27]	3969	Feminine possessive marker kF (“of”/’s)	hin
[0.68, 0.05, 0.27]	4361	Masculine possessive marker kA (“of”/’s)	hin
55B specific			
[0.19, 0.61, 0.20]	11884	Marker detecting the second noun or second element in a compound	hin
55B-341B shared			
[0.26, 0.31, 0.42]	4579	Nominalizer of “to be,” functioning as a gerund (“its being X,” “because of X being”)	hin
[0.05, 0.43, 0.52]	643	Detects first token of verbs / verb roots that appear before subject number conjugation	hin
[0.04, 0.37, 0.58]	2526	Perfective aspect marker in compounds like EkyA gyA (“did/gave” in the perfective)	hin
[0.02, 0.54, 0.44]	1082	Inflection of the verb “to be” (ho) in the subjunctive/continuous mood (, ho sktA, ho ckA) indicating possibility or completed action	hin
341B specific			
[0.20, 0.21, 0.59]	11856	Negation marker nhF placed before verbs	hin
[0.00, 0.00, 1.00]	2598	-	-
[0.00, 0.00, 1.00]	14020	-	-
6B-55B-341B shared			
[0.48, 0.26, 0.26]	2563	Plural pronoun marker for “you/you all” or “us/we”	hin

Table 1: 3-way L1-Sparsity Crosscoder MultiBLiMP Hindi Annotation for BLOOM-1B — 6B 55B 341B. When using just Hindi subject–verb agreement examples to find highest IE features, most features remain strongly language-specific, focusing on tense, aspect, case, and possessive markers, with fewer cross-lingual activations compared to English and French. This may be due to Hindi’s lower representation in the BLOOM training data (2%) relative to French (15%), which limits the emergence of more shared, language-agnostic detectors.