		Genetic		Syntactic		Geographic		Phonological	
		D.	C.	D.	C.	D.	C.	D.	C.
LLaMA2 -chat	7B	-0.04	**0.77**	-0.12	**0.63**	-0.25	**0.21**	-0.03	-0.06
	13B	-0.17	**0.53**	-0.12	**0.65**	-0.17	**0.35**	0.09	**0.24**
	70B	-0.07	**0.78**	-0.12	**0.66**	-0.26	**0.30**	0.00	0.01
Qwen -chat	1B8	0.06	**0.42**	0.07	**0.32**	-0.03	0.00	-0.02	0.05
	7B	0.03	**0.39**	0.07	**0.33**	-0.04	0.04	-0.01	0.17
	14B	0.01	**0.42**	0.01	**0.50**	-0.03	0.14	0.01	0.14
BLOOMZ	560M	**0.20**	**0.43**	0.13	**0.55**	-0.03	**0.38**	-0.12	-0.29
	1B7	**0.23**	**0.45**	**0.21**	**0.67**	-0.01	**0.43**	-0.13	-0.28
	7B1	0.16	**0.36**	0.09	**0.52**	-0.06	**0.31**	-0.11	-0.26

Table 1: Pearson correlation between cross-lingual concept consistency and linguistic similarity for all language pairs. Scores greater than or equal to 0.2 are highlighted in bold. "D." refers to results obtained through direct computation; "C." pertains to the average results derived by first categorizing languages based on language resources and then computing correlations within different language categories.

a high average cosine similarity might raise concerns when dealing with unrelated representations. However, the results in Appendix G.1 indicate that, in our specific context, cosine similarity between concept vectors could reflect their genuine correlation. For comprehensive results on each value concept and further discussions, please refer to Appendix G.2 and G.3.

4.3.2 Trait 2: Linguistic Relationships Distortion due to the Imbalance of Language Data

Figure 3 also suggests that LLMs may learn linguistic correlations between languages and reflect them in cross-lingual concept consistency. Regarding BLOOMZ-7B1, although the cross-lingual consistency between the low-resource languages ta, te and other languages is low, the consistency between these two languages is very high because they both belong to the Dravidian language family. A similar pattern is observed for sw and ny, both of which are from the Niger-Congo family.[3] From this observation, we hypothesize that cross-lingual concept consistency may be influenced by both the amount of language resources and linguistic relationships between languages. In this section, we further explore this phenomenon, specifically investigating to what extent cross-lingual concept consistency reflects natural linguistic relationships between languages and how language resources affect their correlation.

To explore the correlation between cross-lingual concept consistency and linguistic similarity, following Qi et al. (2023), we used lang2vec[4] to compute four types of linguistic similarity (genetic, syntactic, geographic, and phonological) between languages. We then calculated the Pearson correlation between cross-lingual concept consistency and linguistic similarity for all language pairs. We employed two calculation methods to estimate the correlation. The first method directly computes the Pearson correlation on all language pairs (Direct), while the second starts by categorizing language pairs based on language resources. Subsequently, correlations are computed within different categories and averaged (Category). Such categorization aims to mitigate the influence of language resources. Please refer to Appendix F for details of the latter method.

Table 1 presents the correlation results. First, we observe that neglecting differences in language resources (Direct), there is no significant correlation between cross-lingual concept consistency with all types of linguistic similarity. However, upon considering disparities in language resources (Category), the correlation becomes apparent. These findings highlight that the multilingual concept representations embedded by LLMs can distinctly reflect linguistic relationships between languages. Nevertheless, these relationships are influenced by language discrepancies in the pre-training data of LLMs, deviating from the natural patterns.

In terms of linguistic variations, cross-lingual concept consistency exhibits the strongest correlation with genetic and syntactic similarity. In contrast, there is a weak positive correlation between cross-lingual concept consistency with geographic similarity, while no correlation is observed with phonological similarity. The results suggest that LLMs embed more consistent value concepts for language pairs with similar syntactic structures, genetic relations, and geographic proximity, align-

[3]This trend also applies to LLaMA2-chat-7B, where the cross-lingual consistency between en and fr, es, pt, ca is higher because they all belong to the Indo-European language family.

[4]https://github.com/antonisa/lang2vec