Linguistic diversity as a barrier to disease spread

Anonymous Author 1[*1] and Anonymous Author 2[2]

[*]Corresponding Author: name@domain.com
[1]This Department, University X, City, Country
[2]That Department, University Y, City, Country

We test the hypothesis in the literature that linguistic diversity hinders disease spread: people speaking different languages contact less. We first analyze Covid-19 transmission data in countries differing in linguistic diversity, controlling for factors such as development level, geography, and climate. We then simulate the disease transmission in societies varying in size, degree of contact, and network density - three sociopolitical factors impacting language diversity. Both studies suggest that language diversity correlates with reduced transmission rates.

Moro (2016) hypothesized that language diversity might have helped mitigate the problems of population growth when food supply and healthcare were limited. Combined with the more general idea that language diversity is shaped by physical and social factors external to language (Nettle, 1999; Lupyan and Dale, 2016), here we investigate potential correlations between language diversity and disease spread. We follow two complementary approaches. First, we compared the spread of COVID-19 in regions that show opposite linguistic landscapes. Following Kirkeby et al. (2017), we calculated the average daily COVID-19 transmission coefficient β for 101 countries using data from Johns Hopkins University. We then compared disease transmission in countries differing in linguistic diversity index (LDI) (from Eberhard et al. (2024)), while controlling for factors that could impact disease transmission: temperature, precipitation, elevation, latitude, human development index, road and population density, drawn from Global Data Lab, Wikipedia, and the World Bank. A linear regression was done between the difference in LDI (ΔLDI) and the difference in COVID-19 transmission rate $\Delta\beta$, with other factors as predictors. Second, to clarify the mechanism of language diversity on disease spread, we simulate the behavior of a pathogenic vector in the 6 society types discussed in (Trudgill, 2011), hypothesized to impact differentially on language features and language diversity. These 6 types vary in community size, network topology, and degree of contact. Specifically, we implemented an SEIR disease transmission model (Bjørnstad et al., 2020) for each of the 6 societal types. In our model, each type consisted of 4 societies, each with 4 communities. Size was modeled by changing the population number. Network topology and the degree of contact were modeled by modifying the average number of people traveling from one community to another, or from one society to another, respectively (See Table 1). We implemented the model outlined in Lee and Jung (2015) with MATLAB (R2024b). To quantify disease spread, we assumed a single exposure case in community 1 from society 1 and measured the number of days for the number of susceptible people in each community to drop below 50%.

Table 1.: The six basic types of society simulated in the study

Type	Size	Network	Contact	Population	Within-society contact	Between-society contact
1	small	tight	low	1000	$[1, 1.5]$	$[0.1, 0.15]$
2	small	tight	high	1000	$[1, 1.5]$	$[1, 1.5]$
3	small	loose	low	1000	$[0.1, 0.15]$	$[0.1, 0.15]$
4	small	loose	high	1000	$[0.1, 0.15]$	$[1, 1.5]$
5	large	loose	low	4000	$[0.1, 0.15]$	$[0.1, 0.15]$
6	large	loose	high	4000	$[0.1, 0.15]$	$[1, 1.5]$

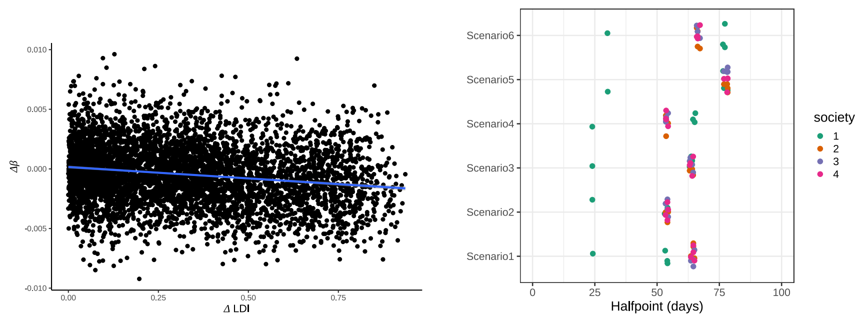

Figure 1.: **a**: Difference in transmission coefficient ($\Delta\beta$) vs. difference in Language Diversity Index (ΔLDI). **b**: Time (unit: days, x-axis) for the susceptible population to drop below 50% in each community across scenarios (y-axis). Communities from the same society are represented by the same color.

Our first analysis (Figure 1a) found a significant, negative effect of ΔLDI on $\Delta\beta$ (r =-2.665*10^{-4}, p = 0.021), suggesting language diversity reduces disease transmission independently of physical environment or social development. Our simulation (Figure 1b) showed in general, larger population sizes, looser social networks, and reduced contacts slow disease spread. Interestingly, scenarios 1 and 4/6 correspond to esoteric and exoteric languages, respectively (Wray and Grace, 2007; Chen et al., 2024). Esoteric languages are typically spoken in areas with high linguistic diversity. As shown, disease spread proceeds differently in both cases, with slower ingroup transmission in the esoteric scenario but slower intergroup transmission in the exoteric scenario. Overall, our findings give support to the view that a barrier effect to disease spread could be one of the factors promoting language diversity.

References

Bjørnstad, O. N., Shea, K., Krzywinski, M., and Altman, N. (2020). The seirs model for infectious disease dynamics. *Nature Methods*, 17(6):557–558.

Chen, S., Gil, D., Gaponov, S., Reifegerste, J., Yuditha, T., Tatarinova, T., Progovac, L., and Benítez-Burraco, A. (2024). Linguistic correlates of societal variation: A quantitative analysis. *PloS one*, 19(4):e0300838.

Eberhard, D. M., Simons, G. F., and Fennig, C. D., editors (2024). *Ethnologue: Languages of the World*. SIL International, Dallas, Texas, twenty-seventh edition edition. Online version: `http://www.ethnologue.com/`.

Kirkeby, C., Halasa, T., Gussmann, M., Toft, N., and Græsbøll, K. (2017). Methods for estimating disease transmission rates: Evaluating the precision of poisson regression and two novel methods. *Scientific reports*, 7(1):9496.

Lee, J. and Jung, E. (2015). A spatial–temporal transmission model and early intervention policies of 2009 a/h1n1 influenza in south korea. *Journal of theoretical biology*, 380:60–73.

Lupyan, G. and Dale, R. (2016). Why are there different languages? the role of adaptation in linguistic diversity. *Trends in cognitive sciences*, 20(9):649–660.

Moro, A. (2016). *Impossible languages*. MIT Press.

Nettle, D. (1999). *Linguistic diversity*. Oxford University Press.

Trudgill, P. (2011). *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*. Oxford University Press, Oxford.

Wray, A. and Grace, G. W. (2007). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*, 117(3):543–578.