How many protolanguages were there?

When we talk about protolanguage, it is typically in the singular. The name of this conference series is Exhibit A here, but also e.g. the authoritative Oxford Handbook of Language Evolution talks about “the protolanguage” (Tallerman, 2012, my emphasis). This leads us to imagine a single protolanguage being spoken (or signed or mimed or whatever) at some point in prehistory. At most we imagine multiple successive stages of more and more advanced protolanguages (e.g. Jackendoff & Wittenberg, 2014, Author, submitted a).

But rarely has the likely diversity of protolanguages been considered. A literature search for “protolanguages” in the plural mainly returns results referring to the proto-forms ancestral to modern language families (proto-Indoeuropean, etc.); see Robbeets et al. (2020) for an example. There are rare exceptions referring to plural protolanguages in the sense used at this conference (e.g. Donald, 1999, Dowman, 2008) but they still do not have diversity in focus.

Is there any reason to believe that linguistic diversity in the world was any less in prehistory than it is today? The answer to that question depends on whether the forces driving linguistic diversification were operative already at the protolinguistic stage. If two groups speaking the same language get separated, their dialects will gradually drift apart over time until mutual intelligibility is lost, the linguistic equivalent of genetic drift in biology; this must have been operative for as long as we had conventionalized language at all. Differential language contact can also cause language change, with different subgroups accumulating different loanwords etc. In this way, diversity begets more diversity, and this would have been operative as soon as some diversity was present. Finally, if language is used as a marker of group identity, a group may intentionally steer their way of speaking away from the neighbors’. This requires a level of meta-linguistic awareness that need not have been present from the beginning.

But even without the latter force, linguistic drift and language contact would have been quite sufficient to drive diversity among protolanguages. And even in the absence of synchronic diversity, linguistic drift would create diachronic diversity: over a millennium or so, a modern language changes beyond intelligibility, and across 10,000 years it changes beyond recognition even at the family level. Even if we assume for the sake of the argument that protolanguage was a single lineage, every 1,000 years what is spoken deserves to be called a new language.

A single lineage is, however, extremely unlikely. From the dawn of language, people have spread out over multiple continents, far out of touch with each other. Protolanguages must have diverged. But populations were sparse, estimates for world population at relevant times are typically below one million, sometimes well below (e.g. Huff et al. 2010), living in small tribes. From such numbers, a rough order-of-magnitude estimate of the total number of protolanguages ever spoken can be calculated (Author, submitted b). With any reasonable assumptions, the number is huge, much larger than current linguistic diversity.

References


