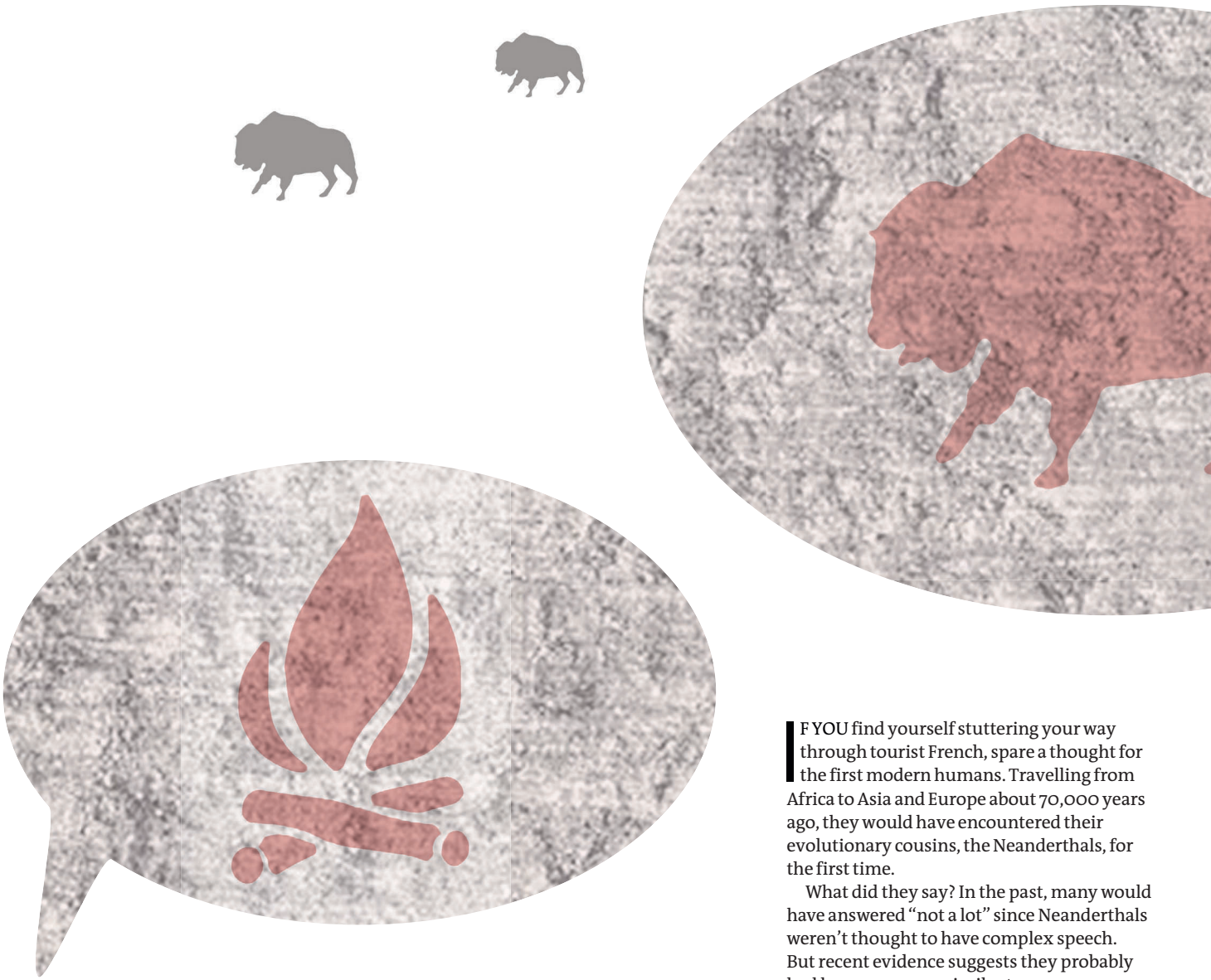


How to speak Neanderthal

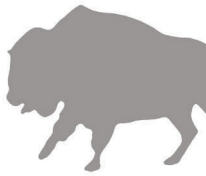
Traces of our ancient cousins' words are harder to find than a needle in a haystack – but that's not going to stop some linguists from trying

IF YOU find yourself stuttering your way through tourist French, spare a thought for the first modern humans. Travelling from Africa to Asia and Europe about 70,000 years ago, they would have encountered their evolutionary cousins, the Neanderthals, for the first time.

What did they say? In the past, many would have answered "not a lot" since Neanderthals weren't thought to have complex speech. But recent evidence suggests they probably had languages very similar to our own. Surprisingly, we may now have the means to glimpse those utterances in the words we speak today, with huge consequences for our understanding of language evolution.

The argument that Neanderthals spoke like us comes from many discoveries. Archaeological remains show that they had a sophisticated lifestyle, with human traits like caring for the infirm and the sick, and an advanced toolkit, including bone tools and body paint – complex behaviour that should only be possible if they had language. We also have some more direct anatomical evidence: traces of nerve pathways through bones in the skull suggest Neanderthals could control their vocalisations, for instance – an adaptation necessary for language that other apes lack. It also looks as if Neanderthals had many gene variants associated with processing language.

So it seems reasonable to assume that their speech would have been similar to our own, with the differences either being down to their vocal anatomy, the way their brains were wired, or simply cultural evolution around the



"We must act quickly, given the saddening rate at which languages are dying"

time they diverged from modern humans. The question is, can we guess what it sounded like?

Unlikely as it may seem, there is a way. Here's the rationale: when two groups that speak different languages come into contact, they exchange bits and pieces of language, like words or grammatical rules. Linguists can detect traces of such interactions even after thousands of years have passed. We know that once modern humans left Africa, they lived alongside Neanderthals and sometimes bred with them. They may have shared cultures, and there is evidence that Neanderthals gave our ancestors the idea for certain tools – so it seems likely they conversed too. The task, then, is to find out whether languages differ between the populations, mostly in Africa, that never came in contact with Neanderthals, and those that would have met them.

Eroded influences

The traces will be very faint and are probably reflected in a combination of features, just as differences between human populations are usually caused by variations in hundreds of genes, rather than just one or two. To complicate matters further, these exchanges happened thousands of years earlier than most historical linguists would even dream of investigating – meaning that time could have eroded away the influence. Never mind looking for a needle in a haystack, it is like searching for a small patch of straw in a barn full of hay.

We certainly aren't ready to build a Neanderthal dictionary, but we have begun to investigate whether modern linguistics could, in principle, find any remains of our relatives' speech in today's languages. Then we can focus our search on more specific features.

Our starting point was the World Atlas of Language Structures, a database that documents hundreds of languages. We used a statistical method to split these into two groups, so that languages within one group were more similar to each other than to languages in the other group. We then tested whether there was a geographical divide between them, perhaps with one group mostly containing the African languages – as you might expect from our theory. Results were mixed, but comparing the overall structures – including things like word order and gender – showed a greater difference between African and non-African languages than simply comparing the vocabulary. This suggests that some kind of Neanderthal influence might linger in the grammar of non-African languages.

Along similar lines, we applied a separate technique that uses linguistic data to predict how populations must have migrated and mixed in order to arrive at today's language diversity. The best-fitting model supported the idea of two main founding populations, one in Africa, and a second that had outside influence from the Neanderthals.

Finally, we turned to methods originally used to study the divergence of species, to map out the family trees of different languages based on their related features. The trees predict when those features first emerged, so we can then look for aspects that change slowly and could still reflect interactions thousands of years ago. We could then find out if there are different patterns in the African and non-African language families. If so, they might be evidence of Neanderthal contact.

It is very tempting to jump on initial results. For instance, the way different languages mark possession proved to be one possible candidate. In African languages, possession is marked by an inflection that depends on the class of word – words about humans would have a different rule from words about inanimate objects, for instance. Eurasian languages don't make that distinction – "my dog" follows the same rule as "my son" or "my computer" – perhaps because the Neanderthals didn't either. But this could easily be a fluke result.

However, rather than a single feature, we expected there to be a more general "fingerprint" left on the languages touched by Neanderthal interactions. So we trained a machine learning algorithm to rank how well different combinations of features could predict whether a language came from Africa, or elsewhere. African and non-African languages could be distinguished with over 90 per cent accuracy, but only by using a large number of features. This makes it difficult to say what caused this difference, but it's possible that something, such as conversations with Neanderthals, pushed the evolution of European and Asian languages in a different direction to those in Africa.

Race against time

Before celebrating these results, we must make sure the statistics don't pick up on other confounding factors. For instance, we are missing information on many of the world's languages, especially those with few speakers. Since the choice of data isn't random, any patterns that seem to emerge could be influenced by biases in the selection.

But the crucial point is that the methods seem to offer a way to test these ideas, and we won't even need a time machine to get the extra data we need; the secrets may be hidden in undocumented languages. Several large-scale language databases are already being put together, although we must act quickly given the saddening rate at which languages are dying. If that helps amplify the faint echoes of our cousin's voices, we will then be able to pick apart more specific features of their speech.

That could have important implications. The traditional view, championed by Noam Chomsky among others, is that the variation we see in world languages is constrained by our innate biases. But if these variations are, at least partly, the result of two different trajectories, one of which reflects Neanderthal biases as well as our own, we may be able to find new insights into the way genes and cultures interact to shape the words we speak.

The prospect may seem audacious, but 10 years ago, probing the Neanderthal genome was also a distant dream. Stranger things have certainly happened in science. ■

Sean Roberts, Dan Dediu and Scott Moisik research language evolution at the Max Planck Institute for Psycholinguistics in Nijmegen, the Netherlands