

Linguists need preservation of languages to study human language

February 25 2019, by Jeroen Willemsen, Kristoffer Friis Bøegh, Aarhus University



Linguist and researcher Jeroen Willemsen is recording a story told by a Reta speaking man as part of his field work in Indonesia. Credit: Jeroen Willemsen

In a recent article here at ScienceNordic & Forskerzonen, we argued that this year's U.N. initiative International Year of Indigenous Languages is urgently needed in light of the global decline of minority and indigenous languages.

In this article, we will delve into how and why we study these languages.

Besides the human rights aspect of [language](#) extinction explained in the previous article, [linguists](#) are generally also concerned with the very loss of languages themselves. This is because we study languages in a scientific fashion, just like animals, fungi, neutrons and tectonic plates are studied by other scientists.

Language is an everyday miracle

Language is in fact a remarkable phenomenon: by producing sequences of sounds with your mouth (or by using gestures in the case of sign languages), you can declare your love for someone, explain quantum theory or order a drone strike.

Language is an everyday miracle, as the Danish linguist Hans Arndt called it, and this miracle is exclusive to humans: while some animals like whales, chimps and bees exhibit a fairly complex form of communication, no bee will ever be able to signal to another bee that a movie they saw yesterday didn't quite live up to their expectations.

Only humans do this, and -barring physiological disorders- all humans do this. Linguists study why we have language, how language works, what the limits of linguistic variation are and how language develops.

If linguists are to advance our collective knowledge of language as a uniquely human phenomenon, we need to learn about as many languages as possible. You can never get a good understanding of what language actually is if you only compare a handful of them.

If we based everything we know about language on, say, English, Danish, and German, we would miss so many insights about the structural flexibility and diversity of language as a global phenomenon.

Specific languages reveal the working of human language

Languages differ immensely across countries and cultures, but they also show remarkable similarities and display traits that correlate with each other.

If, for instance, in a particular language the direct object comes after the verb (such in English 'I ate the fish'), this language will most likely also have prepositions (after, in, to, on, with etc.) which are placed before a noun.

But if in a language the direct object comes before the verb (such in Reta nang 'aab 'anga kede, literally 'I the fish ate'), the language will probably have "postpositions".

These have the same function as prepositions, but are placed after the noun rather than before it. This is what is called a correlation in [word order](#).

In those same languages, where the direct object comes before the verb (like Latin, Turkish or Reta) the verb is the final part of the sentence. So-called verb-final languages not only often have postpositions, but they also tend to have case marking.

Case marking is a grammatical system where you use elements like prefixes and suffixes on the noun to say something about its grammatical role. For example in Latin 'star' is stella if it is the subject (e.g. the star is bright), stellam if it is the object (e.g. I saw the star), and stellae if it is the genitive (e.g. the light of a star).

Presumably, verb-final languages do this for reasons of efficiency:

because the first two elements of the sentence are nouns, you provide more information about what is going on in the sentence much faster.

Likewise, there are languages where the verb is the first element of the sentence, such as Berber, Mayan and Salish. In these so-called verb-initial languages, the verbs tend to be have many prefixes and suffixes, such un- and -ed in un-pack-ed in English.

And in those languages the suffixes can even tell you something about who did what in the sentence, such as who is the subject and who is the object. They do this for similar reasons: if the verb is the first element you utter, you convey more information early in the sentence if this verb contains more information.

This in itself may seem rather basic, but when you correlate many of those traits, this allows you to construct theories about the workings of human language. And to do this, we need information about as many languages as possible.

Bird's-eye view of the world's languages at the click of a mouse

Thanks to the documentation of indigenous languages, the 21 st century has witnessed a veritable boom in the development of massive databases holding this type of information about the world's languages.

The World Atlas of Language Structures ([WALS](#)) is a large database of linguistic properties pertaining to sounds, grammar and words that were gathered from descriptive materials, such as reference grammars. It contains information on 192 linguistic features across some 2,680 languages, allowing us to take a world-wide view of language.



There are still children growing up as Reta speaking - but different mechanisms in society can make the language gradually disappear. Credit: Jeroen Willemsen

It also allows us to see whether the languages we speak or study are comparable to most other languages in the world based on a set of parameters. For instance, Danish word order is predominantly subject-verb-object (SVO).

A quick search in WALS reveals that 35.5 percent of the languages in the database also have this. On the other hand, only in 8.7 percent of them do sentences start with a verb, so languages like Berber, Mayan and Salish are more unusual in that respect.

That said, 41 percent of the languages surveyed in WALS have subject-object-verb order, like Latin and Turkish, so the Danish pattern is cross-linguistically not the most common pattern either.

WALS is of course not perfect: it necessarily simplifies facts about many languages, and also contains occasional misinterpretations. But it allows us to take a bird's-eye view of the world's languages at the click of a mouse, and the development of these types of tools are of crucial importance to linguistics as a science.

Languages tell us about human history

It is clear that documentation of indigenous languages is important for linguistic theory, especially since modern-day advantages allow us to conduct comparative studies on such a large scale.

However, it is not just linguistic theory for which the collection of cross-linguistic data is important. Another major goal for linguistics is to learn about human history in general.

Learning about what language our ancestors spoke, what words they used and how their language developed, can tell us a lot about their culture, history and prehistoric migrations. We can compare word forms in related extant languages and reconstruct the word forms of the proto-language in terms of how it sounded and what words were part of the vocabulary.

This can then tell us a lot about human history: we are often able to determine the urheimat or homeland of such a proto-language, from where people starting migrating and forming offshoot communities elsewhere, with decent accuracy.

And if all direct daughter languages that descend from it have a similar-

sounding word for, say, 'to milk', we know it is quite likely that they herded cattle, too.

Linguists can do a lot today – but data is lacking

As much as languages can teach us about [human history](#), ideally we combine linguistic research with research from other disciplines, like genetic studies and archaeology.

For instance, mainstream historical linguistics has proposed the Pontic-Caspian steppes to the north of the Black Sea as the Indo-European urheimat, and this is also corroborated by genetic studies and [archaeological studies](#).

This kind of interdisciplinary research has made its way into Denmark as well: in Copenhagen, Eske Willerslev and colleagues have made waves by triangulating genetic, archeological, and linguistic data in order to shed light on [prehistorical population movements in Europe](#).

Exciting as this research is, and as much as we are currently advancing, the truth of the matter is that data is severely lacking for languages in most parts of the world. Of all 7,000 languages in the world, only some 2,000 have been described in any kind of detail.

So while we can construct reasonable hypothesis about the history of Indo-European languages and their speakers, we simply lack the data for smaller and lesser-known language families.

Loss of languages, loss of knowledge

Yet another reason to document [indigenous languages](#) is what languages are actually used for by their speaker community. There is a wealth of

knowledge to be found in oral history, traditional poetry, myths and sagas, not to mention indigenous knowledge about flora and fauna.

For instance, one of us (Jeroen Willemsen) has recorded, transcribed and translated some 50,000 words worth of Reta stories, words and phrases, which includes a local genesis story, tales of tribal wars and migrations, highly symbolic ritual speech, and indigenous botanic knowledge. Had this not been done, there is a good chance this knowledge would have died with the eventual last Reta-speaker.

An African proverb states that when an old man dies, a library burns to the ground. The same can be said about when a language dies: every language makes sense of the world in its own way, and we lose a wealth of knowledge whenever the last speaker of a language passes away.

When a language is gone, it's gone

While linguists are currently working hard to gather as much information as possible, languages are dying out at an unprecedented rate. Unesco's [Atlas of the World's Languages in Danger](#) lists some 3,000 languages as endangered, 576 languages critically so.

And to make matters worse, if a language disappears, there are no remnants such as those that can be analysed by archaeologists and geneticists (e.g. human artefacts, carbon, DNA).

When a language is gone, it's gone. And sadly, thousands of the world's languages have either been documented poorly or not at all.

Language databases: The linguist's work tool

- Besides [WALS](#), other databases are also constantly being built in

order to cast new and exciting light on language as a phenomenon.

- [PHOIBLE](#), for instance, is a database for the sound systems of the world's languages (this is an area where Danish is bit more exotic - depending on the linguist you ask, Danish has the highest number of vowels in the world).
- [Glottolog](#) catalogues languages and language families including sources and references.
- The Atlas of Pidgin and Creole Language Structures ([APiCS](#)) is a [database](#) for linguistic traits of pidgin and creole languages.
- [D-PLACE](#) contains a wealth of linguistic, cultural, and environmental data.

This story is republished courtesy of [ScienceNordic](#), the trusted source for English-language science news from the Nordic countries. Read the original story [here](#).

Provided by ScienceNordic

Citation: Linguists need preservation of languages to study human language (2019, February 25) retrieved 26 April 2024 from <https://phys.org/news/2019-02-linguists-languages-human-language.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.