



GRESHAM COLLEGE
Founded 1597

Part Four - Language Documentation and Revitalisation: Where are we now? Transcript

Date: Thursday, 16 June 2011 - 12:00PM

Location: Barnard's Inn Hall

15 June 2011

Language, Documentation and Revitalisation: Where are we now?

Peter Austin

[Introduction in Gamilaraay]

The language you have just heard me speaking is a language that almost disappeared from the face of the Earth, and I shall tell you the story of that language in a few moments. First, I want to give you a little bit of background and a more global perspective, after which I shall zoom in on a couple of local instances, one of which will be the language you just heard.

So far in this symposium, we have looked at global and linguistic diversity. Tim Connell, Nicholas Ostler and Mark Turin have all outlined some of the threats to that diversity. Mark has given us a vision of one way of responding to these threats, and I want to tell you about some other responses, particularly in the area of linguistic research.

The second part of my discussion is going to look at language revitalisation and support. Mark mentioned revitalisation as an ongoing issue. I will outline some of the models that people have been thinking with, and then give you a couple of case studies and draw some conclusions.

In the world today, there are a handful of enormously huge languages, and a huge number of small languages. The top nine or ten languages in the world are spoken by 40% of the world's population. In fact, the top twenty languages are spoken by more than half the world's population. What you have is a tiny handful of languages - about twenty languages - that half the world speaks. 96% of the world's languages are maintained by only 4% of the world's population, and in fact, 50% of those have fewer than 10,000 speakers, and 25% of the world's languages have fewer than 1,000 speakers, so they are very small communities.

Furthermore, there has been a radical reduction in speaker numbers, particularly since the time of the Second World War, and increasing age profiles, which means that although they may be languages still spoken, it is only middle-aged or older people who continue to speak them. This was the situation, for example, in Welsh, up until the 1950s, where older people spoke Welsh but younger people were increasingly speaking English. The linguist Michael Krauss has argued that something like half of the world's languages - he says 90% - could disappear.

The distribution of the languages around the world is skewed, geographically. The bulk of language diversity is in Asia and Africa, while Europe is comparatively not very diverse at all.

We can also look at languages from the point of view of their diversity and how they are related to one another. Mark mentioned, for example, that Nepali belongs to the Indo-European language family; Thangmi belongs to the Tibeto-Burman language group - it is ultimately related to Tibetan and to Chinese. If we look at language families, we see that South America has 90 families, North America has about 45, and Papua New Guinea, an island of just three million people, has 35 different language families. There is close to 800 languages spoken on that one island that belong to 35 totally different language families. Again, you can see how skewed this distribution is. Languages can be categorised or typologised into several groups. There are languages that are safe, strong or viable, spoken by all age groups, actively supported and learnt by children. The population of such a language might be large, like English, with maybe 500 million speakers, or it may be small. A village in Vanuatu, with 300 people, has a strong, viable, safe language - in that village, everybody speaks the language and it is used on all occasions. In these situations, people are typically multilingual, rather than monolingual.

The second category is languages that are under social and economic pressure - usually from a bigger language - and spoken by a reducing population. Actually, in the bulk of cases, in Asia and Africa, it is not languages like French or English or German or Portuguese that are the problem. It is local, big languages, like Swahili in East Africa or Hausa or Yoruba or Wolof in West Africa, which are the languages that have the political and social dominance.

The third category is languages that are moribund, no longer being learnt by children, with fewer older speakers and little social function.

Of course there are also extinct languages, with no native speakers. Picene that used to be spoken on the Italian peninsula is such a language.

So, what can we do about this situation? What has been happening, inexorably and increasingly over the last 50 years, is that languages are moving from being strong to becoming endangered, to becoming moribund, and, in many cases, finally extinct.

There are several things that we can do. Mark's aphorism sums it up well: documentation, protection and support/revitalisation. This three imperatives should be followed in a respectful and collaborative manner, working together with members of language communities at a grassroots level.

It is really important for us to understand language patterns and language attitudes. Many communities have, because of the social and political and cultural dominance of their neighbours, developed a negative attitude towards their own language. They will say, for example, that it is just a dialect, you cannot write it down, it is not a 'real language' like Nepali or Chinese. We need to understand these kinds of attitudes in order to address the situation.

One of the things we can do is to provide reliable and comprehensible information. Many communities believe that their language has negative characteristics, that it is not strong because it cannot be written, but we can actually show them that that is not the case. You can easily write it down, whether you use Roman script or Devanagari script or Tibetan script and so on.

Another important role in our work is to inform relevant stakeholders, including international organisations,

governments and the general public. This is partly the reason for this lecture, information-sharing.

I will now discuss language documentation. What we do is collect and analyse linguistic, socio-linguistic and cultural data, including audio, video and text materials, with the goal of creating a corpus of material or data, representing different aspects of the language in use. We collect and analyse data about the social, cultural and political environment of the community, so that the language shift processes - whereby people change from one language to another - can be understood and interpreted. Why is it that communities are choosing not to speak one language and preferring to speak another to their children? We need to understand that, as part of the documentation, and to archive these materials and the associated metadata for current and future use.

Metadata is a fancy word that simply means data about data, which can mean information about when the recording was made, who was speaking, who was involved, what language was being used, and so on. We do this now, increasingly, in a digital format and framework.

Let me give you a couple of examples. Mark pointed out that there are several documentation projects going on around the world. The DoBeS project (<http://www.mpi.nl/DOBES>) - this stands for Dokumentation Bedrohter Sprachen, literally, the Documentation of Endangered Languages, funded by the Volkswagen Stiftung - so far has distributed €60 million to support a range of projects around the world, and has set up a major archive at the Max Planck Institute in Nijmegen in the Netherlands. They currently fund 50 teams of researchers around the world, documenting languages and cultures in a very wide variety of different community contexts. The archive that was set up in the Netherlands has developed a range of new software tools that allow us to document languages in ways that we have never been able to do before.

Here is a map, on the DoBeS website, showing these 50 projects. It has a global reach - they are doing research on every continent, except Antarctica, on the languages spoken all around the world.

I am associated with the Hans Rausing Endangered Languages Project at SOAS. You probably interact with Hans Rausing's legacy every day. Every time you open a cardboard box that contains liquid, you are dealing with Hans Rausing's invention, which is called Tetra Pak, the ability to put liquids inside boxes - such as milk or orange juice cartons. Hans Rausing is an exceptionally wealthy man, who gives away huge amounts of money to support humanities and other projects.

Arcadia Trust has funded a project at SOAS, with a benefaction of £20 million, to research languages around the world, and we currently distribute £1.5 million in research grants each year in three different types of grants. We have funded 230 teams of researchers around the world. The grant selection committee met last week, and is going to be funding another 40 projects of people who have applied for this work. We have a major digital archive at SOAS, and an academic programme, which I am responsible for, where we train Masters, PhD students, and postdoctoral researchers in language documentation. So far, we have trained 80 Masters' students and graduated about a dozen PhDs from the programme. We also publish books, newsletters, CD-ROMs and we have a website. We also publish a series called Language, Documentation and Description.

Let me give you an example of the work that some of our people do. A young man called Stuart McGill is currently finishing a postdoctoral fellowship; he has been working for the last six years in north-west Nigeria, documenting a language called Cicipu. This language was never recorded until he went to north-west Nigeria. It is one of 50 languages that belong to a close group. The other 49 have not yet been worked on, so this is ground-breaking new work. The outcome is that he has created a corpus of materials: a 2,000 item dictionary (with opportunities for the community to expand on it); 50gb of archive material; a grammar; an analysis of the language; and a website that includes some of his materials (<http://www.cicipu.org/>). He has done a primer for readers, translated stories, folk tales, and produced a keyboard that allows you to type in the language, because it needs extra characters than what is on the regular keyboard.

What is the kind of thing that we collect or record? Let me play you a video, from his collected corpus.

[Video plays]

This is part of a long conversation that these two people are having, talking about the fact that younger people these days are not speaking their language, Cicipu, but are in fact speaking Hausa, which is the language of their next-door neighbours. This group has an animistic religion. The Hausas are coming in with a jihad, the spread of Islam and their language, and basically forcing people to switch to speaking Hausa.

It is striking for us, as English speakers, to observe how these two people have no eye contact when speaking to one another. The lady is sitting there, and the man is in front of her, and they never look at each other, completely avoid any possibility of eye contact. This is because she is his mother-in-law, and in this community there is a taboo relationship between a man and his mother-in-law. They can speak, but they must never look face-to-face and must never have any eye contact. By just watching the video, you can see that the social relationships between them are represented by the way that they place themselves in the world - their body language, their interpretation. You will notice that she is moving her arms about when talking. He is quietly sitting there, not moving. So, by doing this kind of documentation, we can represent and show the social and cultural dimensions of language, language in context, the use of language to express social and cultural and other identities.

We use computer software in order to analyse this kind of material. You can do a transcription, using whatever alphabet, phonetic or otherwise, and also produce a translation. From this kind of material, you can produce, for instance, subtitle videos that are available to play. That is the kind of work that Stuart has been doing.

We also have an archive. At the moment, we have 70 deposits that are available, so if you would like to see more videos like that or Sami reindeer herders, they are all available and accessible through our website. We currently have something like seven terabytes of data - that is 7,000gb of material - available, and it is growing at a very

rapid rate.

Another dimension of our work is the dissemination of information about aspects of language diversity, both within communities and more broadly, with the idea that we should be raising awareness about endangered languages, through all available channels. One of the things we have been trying to do is use the arts and exhibitions to promote awareness. We have produced books and publications for general display, and we are using things like blogs and Facebook and YouTube, new media to engage younger people. We have been running in situ training courses in London, Tokyo and Ghana. This kind of outreach, involving talking to the general community, particularly to language-speakers, can be very challenging for academics, who would perhaps prefer to talk about how many angels there are on the head of a pin than try and engage with the general public, but hopefully we can do this and spread the word.

One of the things I have been doing is to bring together a group of scholars, including Nicholas Ostler, to publish a book called *1000 Languages - the Worldwide History of Living and Lost Tongues*. This has now been translated into Japanese, Spanish, Italian, French, German, Polish, Dutch, Estonian and, earlier this year, Icelandic. We have been doing exhibitions. We organised one called *Endangered Voices*, where we worked with a sound artist, and another one called *Living Languages*, and one called *Hearing Voices*. We are trying to work with the artistic community to get the message across- this is not a dry, academic field, this is something to really engage people in.

Every year, we have been running *Endangered Languages Week*. Earlier this year, some 600 people came to the event at SOAS, to see what it is that we are doing and how things are developing.

A group of students are working on London's *Language Landscape*, a project to map the diversity of London. So far, they have recorded 60 different languages and they have developed an interactive display using Google Maps that will be up on the web very soon.

The city of London, as you know, is an incredibly diverse place where many languages are spoken, including many endangered languages. We simply do not know how many languages are spoken in this city. Various surveys have been done, and estimates of 450 languages have been given, but that is certainly incorrect. For example, there is a language called *Talishi*, spoken on the border between Iran and Azerbaijan, which is impossible for us to do first-hand research on; as outside researchers, we simply cannot get to Iran. There is a man who lives in Edgware Road who speaks *Talishi* fluently, as a native language - a highly endangered language, spoken right here on the streets of London, but never identified because he would be classified as Iranian. A million Londoners speak a language other than English at home - one million people in this city. The top ten languages are French, Spanish, Polish, Hindi, Italian, Urdu, German, Russian, Bengali and Portuguese, but there are hundreds and hundreds of other languages, spoken right around us here today. Many of them are languages where people have moved here and are endangered in their homelands. Interestingly, there is a *Regional Language Network*, which is actually sponsored by businesses here in London, promoting and supporting languages.

We have also been working with local communities, helping to educate and train them, developing teacher training opportunities and supporting the communities in developing materials and the use of the language in a wide range of contexts.

This is all part of what could be called *language revitalisation*: the principles and practices for developing language use and reversing the shift that some communities seem to think is inevitable, increasing the numbers of speakers and the number of places and domains where languages can be used. Linguists around the world are involved in working together with communities to strengthen languages, develop orthographies, school curricular, materials, training and media.

There are several models for language revitalisation. One that is very prominent in North America is called the 'master-apprentice' programme, where a speaker of a language, usually an older person, works together with a younger 'apprentice'. The apprentice lives with the master and they speak the language as much as possible. The way that carpenters and electricians used to pass on their trade is now the model for passing on languages. The second model is called 'language nest'. This is a direct translation of the Maori term, *kohanga reo*. 'Kohanga' means nest; 'reo' is language. This was pioneered in New Zealand, and has been extended around the world, to Hawaii, to the Sami reindeer herders from the north of Scandinavia, and also to the Isle of Man. There is a language nest of *Manx*, where *Manx* languages are spoken.

A third possibility is *immersion schooling*, where children are taught languages and content - Mathematics, Science, Culture, Geography, History, Music - through the second language. At home, they might speak one language - perhaps English - and they go to school and are immersed in the second language. This is what has happened in the case of Welsh, the result being that, today, there are more children who speak Welsh than there have been for 100 years.

Bilingual schooling is teaching language and content in two languages - in the first language of the child, the mother tongue, and in the second language. This is the case in somewhere like Quebec, where people are taught in French and English. These kinds of models, unfortunately, tend to be transitional in nature - that is, the child enters the school system with their own first language and is introduced to the second one, and after several years, the first language disappears from the curriculum and only the second or dominant language is used. Finally, we have *language awareness*, which is teaching about the language - usually to learn some routines, some iconic expressions or usages - but not in order to achieve full fluency in the language.

Let me give you two case studies - first, involving Maori. In the 1970s, Maori people realised that the Maori language was contracting. It was a language that was increasingly being used just by older people, and in restricted contexts. The *marae*, which is the ceremonial house where various ceremonies and cultural activities are carried out, was a typical location for the speaking of Maori; however, English rather than Maori was spoken at home or down the pub. So they decided, in the 1970s, to introduce *kohanga reo*. They realised that the chain

of communication was being broken as parents were speaking English and not Maori to their children. They created pre-schools in which children were brought together with grandparents, and the single instruction was 'no English to be spoken inside these walls'. In the nest, the kids were only exposed to Maori, and when they went home after the pre-school, they spoke English to their parents and to their friends. As a result, we now have a whole generation of children who have grown up bilingually, speaking Maori and English. Maori was introduced at all levels of education – primary, secondary and tertiary – and the Government recognised it as a co-official language, along with English. The Maori developed terminology for objects like computer and iPad and so on, and the first university-educated students are now leaving the Maori education system. I recently met a person who teaches at the University of Technology in Auckland, New Zealand, whose daughter has just written her dissertation in Maori, the language. You can now go to the University of Technology and you can study all your subjects in Maori. The language has been revitalised, from a position in the 1970s where it was struggling, to a situation now where it is used in higher education and people are writing dissertations in the language. The first dissertation written fully in Hawaiian was defended at the University of Hawaii last year. Hawaiian was a literary language 100 years ago, but was under huge pressure from English, and is now coming back through a similar arrangement. At the Sami University College, in northern Norway, teaching is available in Sami languages. New contexts of use have become available. There is a full set of programmes on Maori television which you can now watch, and which can be viewed online. Maori is used on Facebook and many people are now 'tweeting' in Maori on Twitter. It is not the only language – there are dozens of endangered languages that are being used on Facebook and Twitter.

I shall finish by giving you another case study, from south-eastern Australia and centred on the language that I spoke in at the start of this lecture. In 1788, there were about 70 languages spoken in what is now New South Wales and Victoria, and we have some early, but really quite insufficient, documentation materials from the 18th and 19th Centuries. By 1900, most of the languages in this region were moribund, and by 1950, most were extinct or reduced to the last single speakers. Modern linguistic research really only began in the 1950s, and in the 1970s and '80s, we got documentation of the last remaining linguistic and cultural knowledge, and with the death of the last speakers, that knowledge has now disappeared.

This map of New South Wales shows you some of the languages. The language that I spoke to you in is called Gamilaraay, and it is from that shaded area in the north-west of New South Wales, about 350km from the city of Sydney.

In the 19th Century, the Gamilaraay were forced off their land by settlers, and we have some missionary materials, such as those of William Ridley from the 1850s. Unfortunately, he was not a trained linguist, so he was not able to hear subtle differences.

In the 1930s, some anthropologists visited the area and realised that it was the private language of a very few old people. In '55, a linguist called Stefan Wurm introduced the last fluent speaker, and we have his notes and a tape which he made on a wind-up clockwork tape recorder, with thirteen minutes of recordings.

I started work in this area and I was able to find people who could remember their parents and grandparents speaking the language, but all they could remember were a few words and expressions. They knew how to say things like water and fire and so on, but they could not put whole sentences together. One old lady remembered her grandmother calling out 'I am hungry – where's my bread?!' She did not know what the Gamilaraay word for bread was, but it was there somewhere.

Using the missionary materials and Wurm's documents, we were able to analyse these findings and combine them with some grammatical studies done with two half-speakers of the language encountered in the 1980s. The result is that we published a dictionary and uploaded it online. It is still there on the internet, launched in 1995.

In 1998, a revivalist movement began. They decided to start working together as a community and, in the year 2000, got approval to teach the language in school. Languages other than English are required of all primary school students, and the aboriginal people in this region said, instead of studying Indonesian or French, why can our kids not study our language? After receiving approval in 2000, we started producing materials – language CD-ROMs, school textbooks and CDs – and we started training teachers. We produced bilingual merchandise, like t-shirts: 'Let's talk Gamilaraay to one another.' A dictionary was created, in collaboration with linguists, school teachers, information technology specialists and local language and culture custodians. The state government introduced a language programme across the state in 2003. A language that was almost extinct has been brought back to a point where children are able to write it, speak it and use it in the classroom. In 2006, the University of Sydney started teaching language classes in this language, so that anyone could go and learn it and study it. There has been a huge cultural revival as well, with children being taught traditional dances and so on.

There are many challenges facing the world today in terms of linguistic and cultural diversity, both from a global perspective and more locally, within the city of London, within New Zealand, within Australia. Researchers and communities are responding to these challenges to understand language use, attitudes and ideologies, and to try and understand the dynamics of change. What I hope to have shown you is that efforts are going on around the world to document, revitalise and support threatened languages and to extend their domains of use, to move into new media areas and take advantage of the internet. These are efforts to foster the languages to help turn the tide of language shift that we have seen globally.

Everyone in this audience can do some good things to support us and to help us in our tasks. First, we all need to understand and appreciate language diversity – though I am sure I do not have to repeat this to people who

have come to a symposium on endangered languages! Language diversity is an amazing, wonderful thing and we should appreciate it.

We also need to combat the ideology of monolingualism, which has probably never existed anywhere in the world, and definitely not in the UK. Going back thousands of years, there was never one language ever spoken on these islands. There was Roman, there was Britannic, there was Pictish, and many, many more. English is not the only language of the United Kingdom - there are over 400 languages spoken here in London! There have always been multiple languages spoken in the United Kingdom. Please repeat this to your friends, repeat this to politicians, repeat this to Mr Cameron - tell the people who need to hear it. Please help to disseminate accurate information about the state of languages and cultures across the world, supporting communities in their struggles to maintain and develop them.

Thank you.

© Peter Austin, 2011