Our Experience of Linguistic Mapping and Thoughts on its Future Direction

In preparation of the third edition of the
Atlas of the World's Languages in Danger of Disappearing

29–30 November 2007

UNESCO, Paris France

Ms Kay Dancey
Managing Cartographer
Cartographic Services
College of Asia and the Pacific
The Australian National University
Canberra, Australia

INTRODUCTION

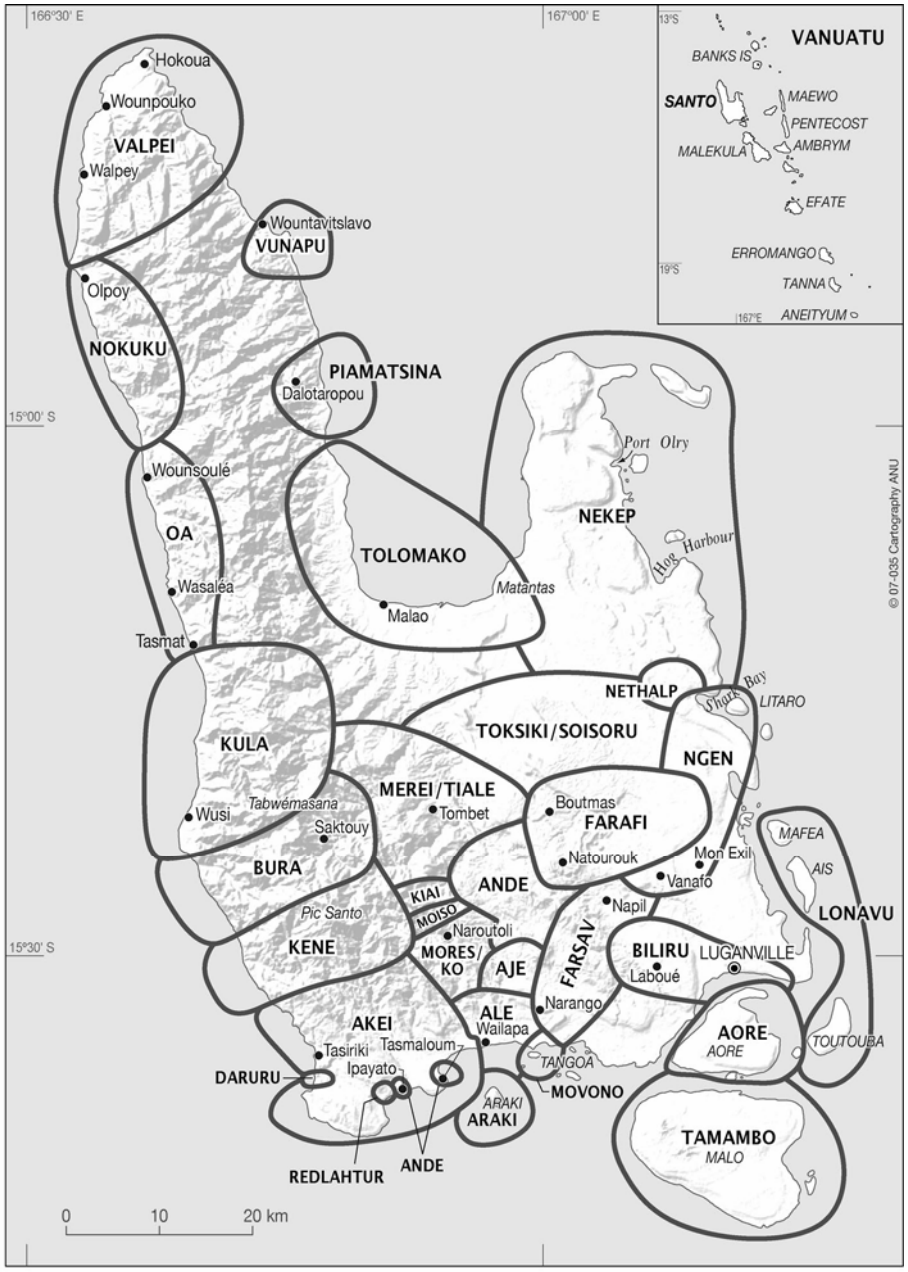
The Cartographic Services unit which has been in existence for over 50 years. The unit lies within the Research School of Pacific and Asian Studies, College of Asia and the Pacific, at The Australian National University. We are a staff of 4 full-time cartographers working to academics and graduate students in various disciplines including anthropology, archaeology, history, economics, strategic studies and linguistics.

OUR EXPERIENCE OF LINGUISTIC MAPPING

I joined the unit in 1989 and was aware of the extensive linguistic work of Professor Stephen Wurm in the School's Linguistic Department. At that time a cartographer worked full time on Stephen's linguistic mapping projects of the *Language Atlas of China,* the *Language Atlas of the Pacific Area* (eds. S.A. Wurm, S. Hattori) both large format A2 of 20 and 47 loose sheets respectively and the *Atlas of Languages of Inercultural Communication in the Pacific, Asia and the America* (eds S. Wurm, Peter Muhlhausler, D. Tryon, Mouton de Gruyter 1996) 151 A3 maps. Later the *Atlas of the World's Languages in Danger of Disappearing* (ed. S.A. Wurm) was also produced.

These first three atlases map areas at a relatively large scale. The detailed linguistic classifications are illustrated using hard borders, colours and stippling. The Disappearing Languages atlas has been mapped at a smaller scale (region, continent level) and used points to represent locations. Some lines were used, although their use was not always easy to interpret.

All but the 2nd edition of this atlas were produced by hand using traditional cartographic methods. All our mapping is now produced using ARC GIS and Adobe Illustrator.
In recent years our linguistic mapping has been both large and small scale.

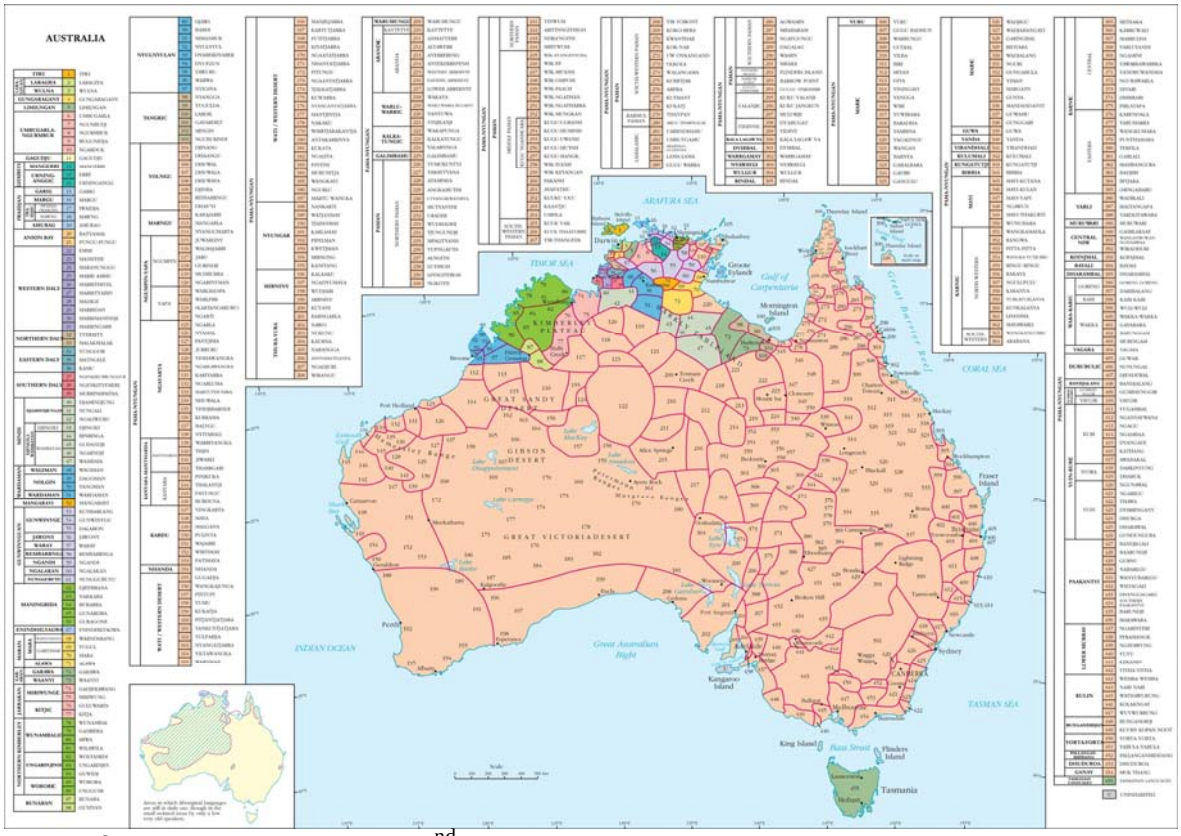


Santo Island languages

The large scale language map of Santo Island, Vanuatu was produced for a black and white journal. This was complied from the linguist's field work notes and sketches. The elevation was derived from the Shuttle Radar Topography Mission (SRTM) 3 arc-second data available from NASA's online site http://www2.jpl.nasa.gov/srtm/ . Hard borders are used to represent language names only–no language classification is shown.

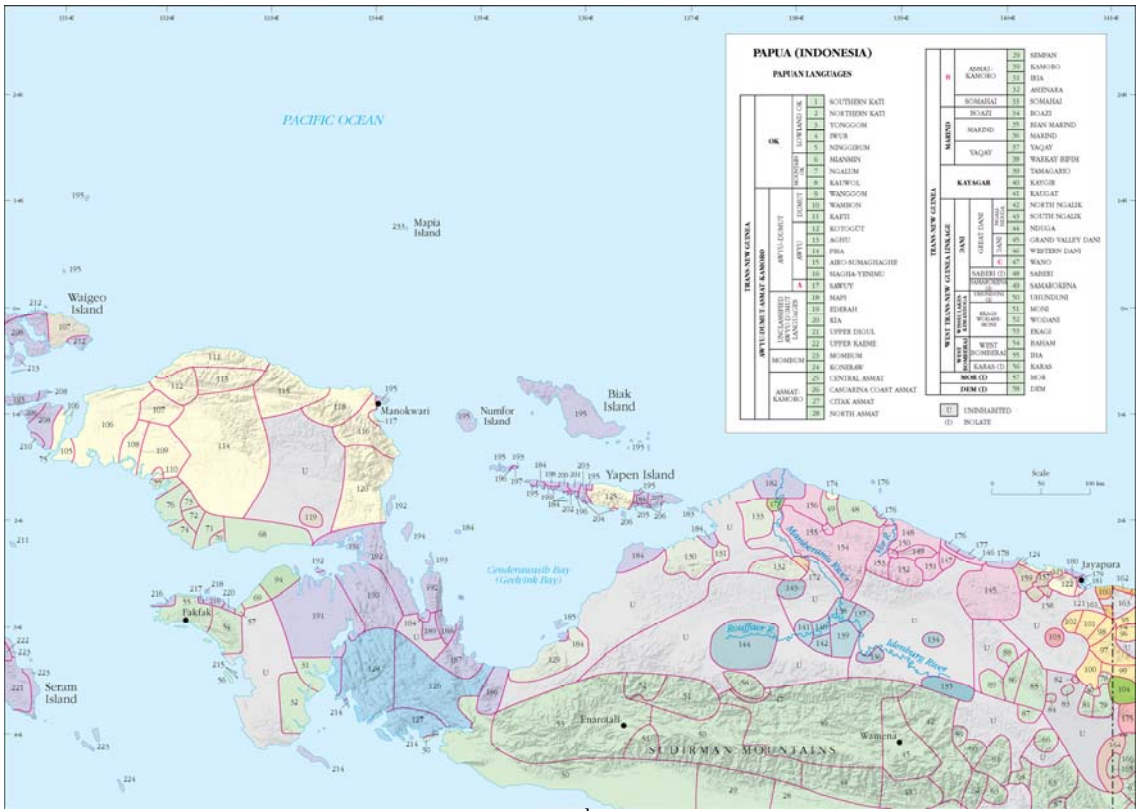
Encyclopaedia of Aboriginal Australia

This map from the *Encyclopaedia of Aboriginal Australia* (ed. D Horton), published in 1994 shows tribe/language areas using a fuzzy border. It does not illustrate any language classification. The colours are arbitrary and unrelated to each other. This map is appealing but of limited linguistic value. It does not correspond to precise needs. The information is based on the work of Norman Tindal, from the mid 1900s.



Atlas of the World's Languages 2nd edition

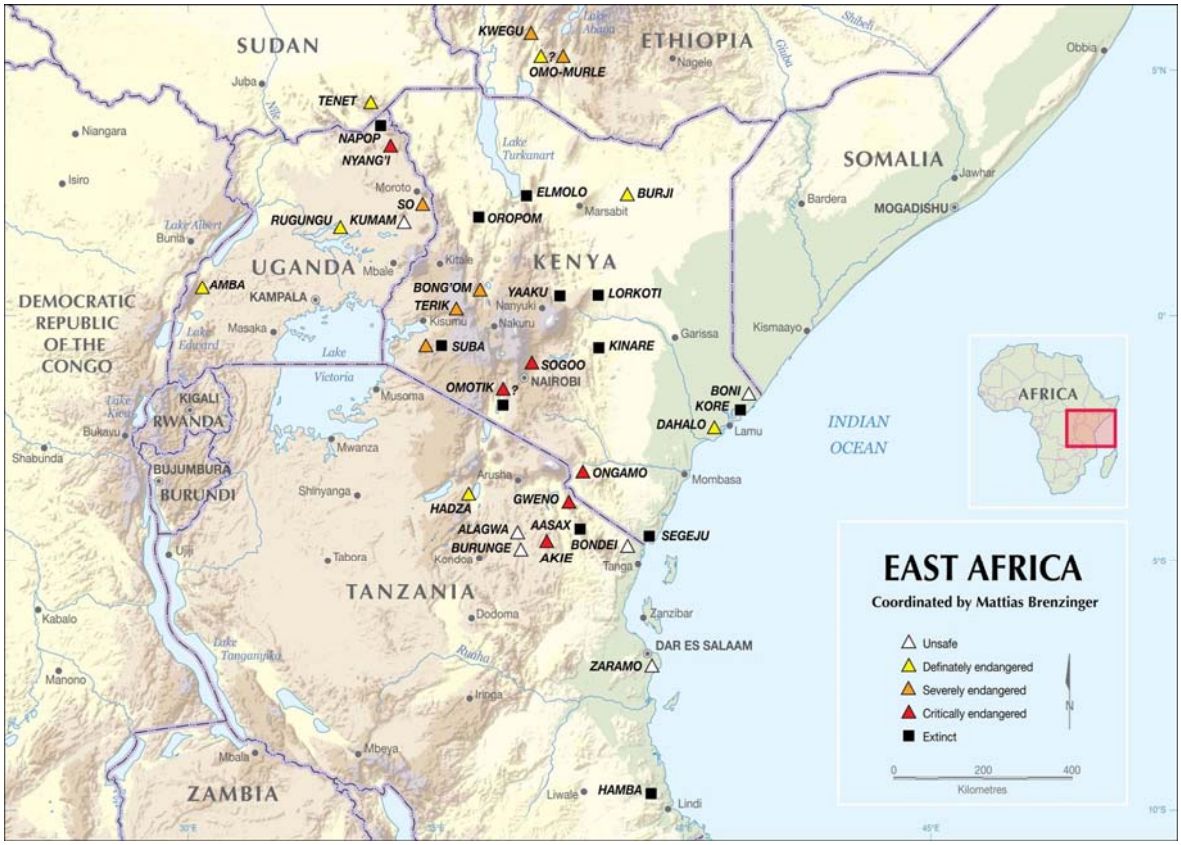
In comparison the *Atlas of the World's Languages* 2nd edition (eds. R.F. Asher, C. J. Moseley) shows more current information and extensive classification both on the map and in the key. Anthony Bright was responsible for updating the Australasia and Pacific maps in this 2nd edition.



Enhanced *Atlas of the World's Languages* 2nd edition

From the same atlas this slide shows the Papua map with hillshading as an example of topographic information that can be added without obscuring the linguistic information.

THOUGHTS ON THE 3RD EDITION OF THE ATLAS OF LANUAGES IN DANGER OF DISAPPEARING



This is a sample reworking of the Atlas's 2nd edition East Africa map.

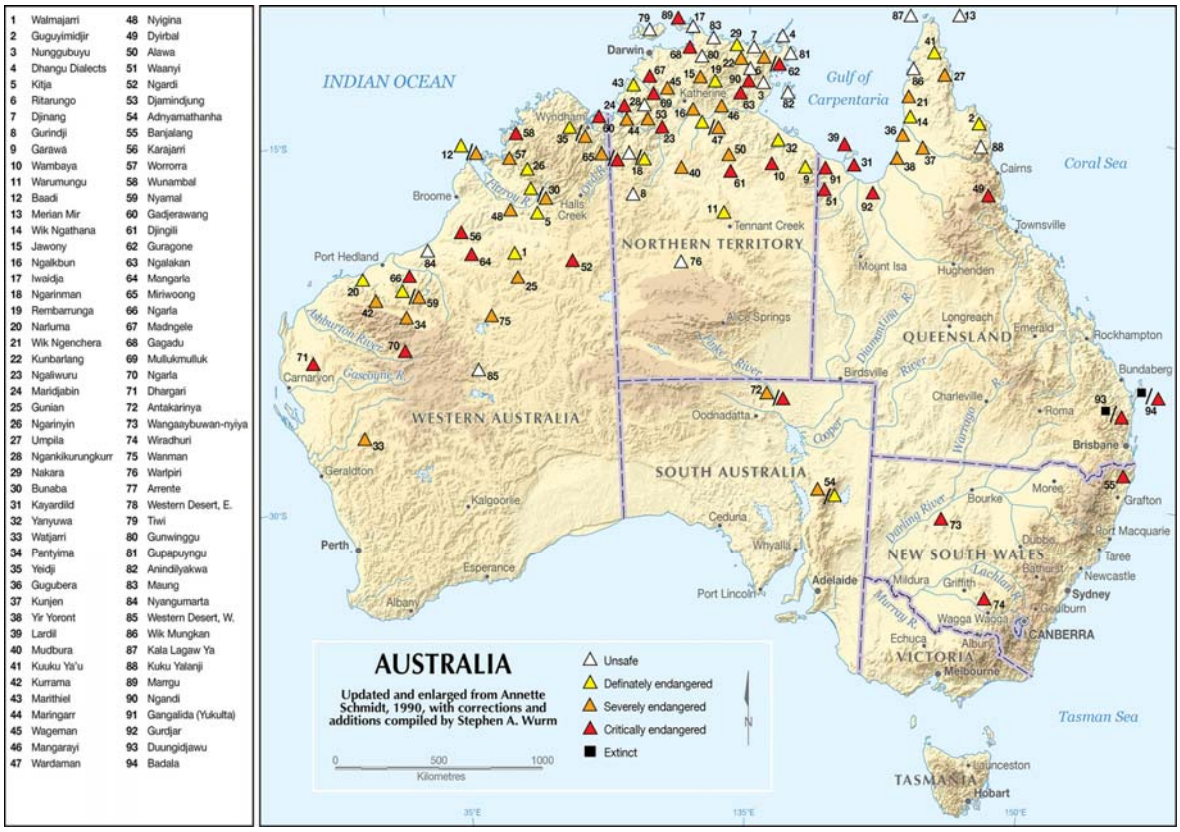
Data from public domain used in the production of this map include:

Elevation. SRTM–Shaded relief with hillshading to represent elevation. 7 colour tints or colour ramp and hillshading at 60% transparency.
Coastline. GSHHS_13– world coastline vectors. Generalised and smoothed
International borders. VMAP0
Lakes. VMAP0 (watercourse, internal waterbody)
Rivers. VMAP0, or digitise from georeferenced material,
Place names and locations. GEOnet Names Server (GNS) or Fallingrain Global Gazetteer (http://earth-info.nga.mil/gns/html/index.html http://www.fallingrain.com/world/).
Fonts: Optima bold, Georgia bold/regular, Helvetica regular.

Other datasets in use:

Australia 1M & 250K topo vector/raster. Geoscience Aust 2004.
Australia topographic, bathymetry raster grid. Geoscience Aust 2004.
AMBIS_maritime boundary vector. Geoscience Aust 2001.
(http://www.ga.gov.au/nmd/products/thematic/ambis.jsp)

Etopo2– 2 minute (4km) worldwide bathymetry/topography
Gtopo30–DEM 30 arc-second (1km). Centre for Earth Observation and Science 1996.
SRTM_NASA–DEM world 1 x 1degree tiles. 3 arc-second (~90m). NASA 2005.
SRTM_CGIAR–DEM world 5 x 5 degree tiles. 3 arc-second (~90m). Consortium for Spatial Information 2004.
VMAP0–topo vectors–transport, population, utilities, vegetation, hydrology, boundaries. National Imagery and Mapping Agency.
Indonesia 2000 census.



This is another mockup of a 3rd edition map (Aust) showing elevation using a monchrome colour ramp plus hillshading, in place of the tints in the previous East Africa example.

Issues in linguistic mapping and base information presentation.

• Define the audience of this *Atlas*. Can broaden the Atlas relevance and usage by adding data collection information, adding detail to the linguistic information and updating the linguistic data more regularly.

• Need to find balance between the amount of geographic information shown that is relevant to the linguistic data while at the same time enhancing the data.

• Elevation useful in finding correlations between geography and language endangerment. Less rivers needed when elevation shown.

• Show hierarchy in place names. Greyed to be less prominent but allows larger font size.

• Symbology as per the UNESCO online Africa map. Logical colour ramp, same symbol for all but extinct category, all language names in black.

• Note that any areas shown will be approximate unless detailed information (this side of river for X kms etc) or GPS is captured. GPS data of language (village) location would be very useful. A database of all language information needed.

• There is the need to be mindful of the potential use of hard language boundaries in land claim issues.

Additional features which would add value to the 3rd edition of the hardcopy *Atlas*.

Define the audience of this *Atlas*. Can broaden the Atlas relevance and usage by adding metadata on the date of data capture.

And detail to the linguistic information;
• Add the linguistic family classification of each language. Classify by broad family affiliation or subgroup as has been done in the online version of Africa? This dependent on the mapping scale.

• Update the linguistic data more regularly (annually online, every 3yrs in print?).

• Note the data, reliability and source of the linguistic data. One concern with present linguistic mapping is the currency of the data. The Australasian information in the *Encyclopedia of the World's Endangered Languages* (ed C. Moseley 2007) was collected in the mid 1980s. Updated data crucial when dealing with endangered languages.

• Statement regarding unawarded land.

• Show language names' variance (print/online version). Checklist of alternate language names.

FUTURE DIRECTIONS FOR ONLINE ATLAS

The online version of the Atlas offers a very exciting opportunity.

There are two approaches that we can see. An online product displaying maps with popup boxes of information on rollover and extended information on clicking. This approach would utilise software such as Flash.

The other alternative is the use of an internet mapping service accessing a GIS server (eg ARC IMS although it doesn't have to be an ESRI product). The browser becomes a remote ARC Map accessing the database of linguistic information. See example at http://seamless.usgs.gov/

This can become a very elegant and sophisticated tool.
The zoomable view extent would determine the data shown and also the data that is searchable from the database.

This approach allows more functionality to be added at a later time. It can be used for other mapping projects where data has been collected e.g. other intangible heritage elements such as dance, song.

The database sitting behind the browser interface allows you to do sophisticated analysis without the need by the user for dedicated software, as in the case of Flash interface, or plugins.

This approach would require more time initially but would be offer a much more sophisticated tool with the bonus of being easily updated since the maps area is generated by the server dynamically. The Flash approach requires maps for each anticipated query.

The summary information is dependent on the functionality offered to the user from the server. SQL access to the database would allow more variations of queries than could be anticipated and mapped using Flash.

There is the opportunity for the inclusion of audio with transcription. Crucial in the mapping of endangered languages.

SUMMARY

The *Atlas of Languages in Danger of Disappearing* hardcopy is a valuable resource. With updating of the data and enhancements of its presentation, the *Atlas's* usage and appeal would be increased.

An exciting area of linguistic mapping is the online application. An online version has the potential to offer large volumes of data which can more readily be updated. This would go some way to addressing the present concern in publication lag times of linguistic data. An interface to this data would allow custom queries. It is an exciting time to be mapping linguistics.