

# **TITLE: Waves of words: mapping and modelling the history of Australia's Pacific ties**

## **AIMS AND BACKGROUND**

**Aims:** This project aims to rewrite the pre-history of Australia and its relations with its Pacific neighbours by developing and applying new technology-centred research methods to explore evidence from language change and relate it to anthropological evidence and the archaeological record. Pacific and Island Southeast Asian peoples have been involved in intense voyaging and social and cultural interaction for thousands of years. Australia did not sit alone in the middle of this busy highway; it was most likely an important part of the Pacific network, yet this perspective is absent from much of the linguistic and anthropological literature. This project aims to right this situation, in particular by focussing on the highly valuable but neglected evidence of language, using new approaches and new technologies.

### **The primary aim of this project is:**

**1. to understand Australia's prehistory as a networked agent in a complex web of Pacific interaction.** In order to achieve this, we aim to determine the extent and nature of early relationships between first peoples of Australia and the Pacific by bringing together, comparing, and modelling the linguistic, anthropological and archaeological evidence.

As a result of this research we also aim to:

2. discover what kinds of social configurations underlie different linguistic outcomes in language contact situations; and
3. improve our understanding of the relationship between language change and socio-cultural change and feed this back into improvements to linguistic and anthropological theory.
4. develop a toolkit of methods and software for integrating linguistic, anthropological and archaeological data that can be used in other regions as well.

Recent work on the relationships between Australia and its Pacific neighbours has begun to show how linguistic evidence can drive exciting new theories about the prehistory of the Pacific region, but this work so far has been fragmented.

The strongest published linguistic evidence of regular contact between Australia and its Pacific neighbours is for relationships between Sulawesi and Northern Australia (see e.g. Evans 1992; Macknight 2013; Urry & Walsh 1981). These relationships seem to have been due to trepanning activity by 'Macassans' from Sulawesi (probably speakers of various Austronesian languages, not only Makassarese). These were regular, intensive, and long-lasting (Clark & May eds 2013). A pidgin developed, cultural features such as ceremonial exchange practices diffused (Urry & Walsh 1981:99) and words were borrowed in ways that give us windows into the specific cultural interactions (e.g. Makassarese *barambarang* 'goods' > Yolgnu Matha *barrambarra* 'clothes' (Zorc 1986:45), suggesting trade in clothing, a semantic narrowing of the kind discussed in McConvell & Ponsonnet (2013:194-5)). Some words eventually spread a long way inland (Evans 1997:260; McConvell 1990:23).

Archaeological and anthropological evidence suggests a number of other paths of interaction between Australia and the Pacific, e.g. with early Austronesian movement through areas north and east of Australia (c.f. the introduction of the dingo), contacts in and across the Torres Strait (e.g. McNiven et al 2006), Papuan/Cape-York contact evidenced by canoe technology (cf. Balme 2013) and shell fishhooks suggesting wide networks of technological dissemination reaching as far south as Sydney (Attenbrow 2010), the details of which are archaeologically ambiguous but might be clarified by integration with linguistic data. There is intriguing recent evidence for contact between Norfolk Island and New South Wales (White et al. 2014). Strong linguistic and archaeological indications of contact between Vanuatu and northern Australia (particularly Lizard Island) from at least 2000 Before Present (BP) are also emerging (e.g. Tochilin et al. 2012; Sheppard et al. 2015).

Another, much later, Australia-Pacific relationship that is often mentioned in the literature but not yet fully understood is the spread of Pacific Pidgin English and its influence on the early development of Australian Kriols/creoles. Evidence shows that the pidgin used on ships and in early colonial encounters around the Pacific diffused via NSW Pidgin, then played a role in the development of Australian creole languages that are still widely used by tens of thousands of people today (Clark 1979; Baker 1993; Amery & Mühlhäusler 1996). South Sea Islanders were probably the vector for Pacific Pidgin in the origin of the Torres Strait creole known as *Broken* (Shnukal 1983). Pacific Pidgin English may have been a vector for some of the *Wanderwörter* (widely diffused words) we find shared among Pacific and Australian traditional languages, but not all. Some have undergone sound changes that suggest they have been present in the languages for much longer than several hundred years (e.g. Ross 1992:365), and others correlate with evidence from archaeology or plant genetics that sheds light on the approximate dating of the diffusion of the objects or plants the words refer to. Both the more ancient and the recent colonial relationships between Australia and the Pacific are therefore of interest, because the two must be untangled in the process of excavating the more ancient layers of linguistic and cultural diffusion.

Uncovering Australia's past Pacific ties, i.e. the steps to achieve aim 1, requires us to a. seek stronger evidence for those relationships that have been proposed but are not yet solidly proven (e.g. Vanuatu-Northern Australia), b. determine pathways of transmission across the wider Australia-Pacific region for words, structural linguistic features, and cultural practices that have thus far received attention only for smaller regions, and c. connect the disparate strands of evidence from linguistics, anthropology and archaeology into a coherent 'big picture'. The deeper and integrated study of these contact vectors could revolutionise our understanding of the region.

**Background:** Research on the history of a region through the perspective of linguistics has two major objects of study: the first is the genetic relationships of the language, i.e. how proto-languages split into daughter languages. For many of the Austronesian languages of the Pacific, some of the Papuan languages, and some Australian languages, this kind of work has already been carried out. The second major object of study— researching other kinds of relationship between languages of the region besides their genetic relatedness or lack thereof— must always build on this foundation, which is why Australia and the Pacific are now a prime site for such research.

When it comes to words, the kinds of words borrowed can tell us about the pathways of transmission (e.g. a hypothesis to test is that the borrowing of words for technology suggests trade relationships; the borrowing of religious or political words suggests longer or closer interactions between cultural groups, for example conquest, or intermarriage; borrowing of words for plants and animals often comes with migration, etc). Because sound change is regular, examination of the sound changes that have or have not affected the borrowed words, compared to sound changes that have affected the rest of the language, can indicate the time period at which words were borrowed, and therefore help us date periods of interaction. Examination of grammatical changes (or lack thereof) in the languages in question also help us to understand the duration and intensity of the contact (c.f. Ross 2013), since some kinds of linguistic structure appear to change more readily, while others tend only to change after very long intensive periods of contact such as pervasive multilingualism (Van Coetsem 1988: 25). We can therefore assume different kinds and durations of contact took place between two languages that only share a few words and little to no structural features, vs two languages that have converged to be grammatically almost identical (e.g. the 'metatypy' found in Takia and Waskia (Ross 2001)). To gain a full picture of the nature of interactions in a region, we therefore need to consider all of these elements together: word forms and their semantics, change in meaning, sound, and grammar.

Mapping of linguistic features ('typology', or 'what's where why' (Bickel 2007)) has a long history, and has experienced increased attention in recent years with advances in digital mapping and modelling (see e.g. the *World Atlas of Linguistic Structures Online* (Dryer & Haspelmath 2011) with its interactive typological maps and downloadable datasets). 'Diachronic typology', or 'what's where *when* and why' (Hendery 2012), is a newer and growing area of interest (e.g. Givón 2012; Fleischer et al. 2015; Evans 2016). Mapping of words, on the other hand, has most commonly been practised by dialectologists, and lexical typology or semantic typology is still underresearched (Koptjevskaja-Tamm 2016:4; McConvell & Ponsonnet 2013). Mapping of the diffusion of words through a region has even less frequently been attempted. Mapping of cultural practices or features is likewise rare (see discussion in Dousset In Press), and the rarer still is the layering of these with linguistic features or terms so as to better understand relationships between language and culture. However, in recent years members of this project team have shown how such an approach can illuminate a region's history. See, for example, the outputs from the *AustKin* projects (DP0878556 and DP120100632), which mapped kinship terms from Australian languages against social category systems and marriage rules to understand change in the linguistic and/or anthropological systems and how these changes reflect migration and/or contact (see e.g. Hendery & McConvell 2013; Keen 2013; Koch 2013; Koch 2012; McConvell 2012; McConvell & Dousset 2012; McConvell 2013; McConvell & Bowern 2011; Simpson 2013).

This project will extend this previous work by our CIs by (a) applying these proven methods to a wider region (beyond the Australian focus of the *AustKin* projects); and (b) to a broader set of semantic domains, selected to allow us to examine a number of possible vectors of contact and change (trade of technology, of agricultural products, transmission of spiritual concepts, political and religious practices, migration and marriage); (c) by including structural linguistic features as well as words; and (d) by exploring this data with a variety of data science and digital modeling methods rather than purely through maps. The key to bringing together so many theoretical and empirical layers is in the long humanities tradition of 'thick mapping' (Presner et al. 2014), while the means is provided by the expanding Digital Humanities community of 'Linked Pasts' (the construction of digital ecosystems of linked open data from history, geography, archaeology and cultural heritage, see Isaksen et al. (2014)), as well as new technological advances in interactive mapping, simulation, 3D modelling, and data visualisation, all of which are discussed further under 'project design and methods' below.

## INVESTIGATORS

**CI Hendery** (0.2 FTE in years 1 & 2, 0.3 FTE in year 3) is the lead investigator and will have overall responsibility for the direction of the project, in particular the *integration* of the technological, linguistic and anthropological elements. Within the linguistic research she will be responsible for the collation and synthesis of case studies of structural linguistic change. She will manage the RAs and supervise the PhD student. While she is a CI on a Linkage and Discovery grant, these will both be completed by the end of 2019, allowing significantly more time in the final year.

**CI McConvell** (0.2 FTE) will lead the lexical and semantic research, as well as providing anthropological expertise. This strand will draw on his work on kinship and social organisation, as well as extending his prior work on diffusion of plant terms and material culture. McConvell's commitments to two Discovery projects will be completed before this project begins and the Linkage grant on which he is co-CI with Hendery will be complete by the end of 2019.

**CI Simoff** (0.2 FTE) will be responsible for the conceptual development of the computational side of the project, advice on modelling methods, the appointment of RAs, and co-supervising the PhD student. His expertise in intelligent human-centred methods and algorithms for visual data mining and sense making, hypothesis formulation and visual explanation, including expertise in interactive visual analysis will be invaluable for the project, as will his experience in integration and synthesis of complex data of disparate kinds into holistic models. He will co-supervise the PhD student.

**PI Dousset** (0.2 FTE) is an anthropologist specialized in Australia and Island Melanesia, in particular Vanuatu. As former director of CREDO, Centre for Research and Documentation on Oceania, (the largest European centre for Pacific Studies) with responsibility for national technical platforms and databases for the social sciences and humanities in France, he is well placed to advise on both the technical and the anthropological sides of the project. He will contribute to the lexical and semantic research, and also his ongoing work on the history of Pacific political and social systems will provide new data on change in social and cultural systems. He will visit Australia each year via a DIA.

**Prof. Matthew Spriggs** (Laureate Fellow & Prof. of Archaeology, ANU) and **Assoc. Prof. Tim Denham** (archaeology, ANU) are unable to be CIs because of ARC rules on maximum number of concurrent grants, but will advise on relevant archaeological data and contribute to the research on diffusion of plant and agricultural vocabulary.

**The PhD student's** project will sit within simulation and/or 3D modelling part of the project. The specific subproject will depend on the interests and background of the student who takes up the role, but given the strong focus in SCEM at WSU on AI in simulation, cultural heritage simulation and modelling, and agent-based modelling, we anticipate that the project would sit within one of those themes.

**RAs** will be employed part time (1) to collect and collate relevant data and to assist with entering it into our project database and (2) for software development. This will include, for example, extending the existing database and mapping platform, extending existing VR data visualisation tools developed by Hendery with her 2016 Transdisciplinary Innovation Grant (TIG) grant, design and user testing of new visualisation metaphors and interfaces.

## PROJECT QUALITY AND INNOVATION

**Significance of the problem.** This project addresses four significant problems:

1. *That we have been misconceiving Australia's prehistory as more isolated than it most likely was.* In the 19th and most of the 20th centuries, Europeans saw the Pacific as a sea of static, isolated islands. Australia and the Pacific are multilingual regions whose languages and cultural practices still hold traces of significant historical interactions. Uncovering, collating and understanding these traces will allow us not only to better understand our region's past, but also to understand how long-term intercultural contact plays out. (See aim 1 above)
2. *That we do not yet understand the relationship between the different kinds of linguistic change and between each of these and social or cultural change.* How similar is change in language to change in cultural practices and in what circumstances are they linked with each other? Certain domains of language are thought to change more easily under linguistic contact (Van Coetsem 1988:25), but the *hows* and *whys* of this are not well understood. (Aims 2 & 3 above.)
3. *That the humanities and social sciences are increasingly adopting models and methods from the physical sciences before it is clear whether the underlying assumptions are appropriate to the nature of human language.* To what extent do language or cultural practice resemble deterministic systems (Montemurro et al. 2011; Loreto & Steels 2007) or biological organisms (Pagel et al. 2007; Beltran et al. 2009)? By including a range of different changes, we are thereby able to test more comprehensively what kinds of models and what assumptions or theories are the most useful for linguistics, giving us insight into the nature of language change itself. (See especially aims 3 & 4 above.)
4. *That most linguists and anthropologists are not in a position to take advantage of new modelling and data*

*visualisation technologies* as these are not yet accessible to those without strong computational backgrounds. While statistical modelling has enjoyed wide uptake in the social sciences (and to a lesser extent the humanities), more qualitative modelling methods such as mapping, simulation, data visualisation, and new possibilities in virtual and augmented reality have primarily been adopted for dissemination of research outputs (e.g. displaying findings on a map or in a data visualisation) rather than for exploratory modelling and hypothesis generation. One of the outcomes of this project will be to release methodologies and corresponding tools that can be adapted to other linguistic and anthropological data and allow other researchers to apply them to their own data for in-depth integral analysis and computational experiments. (Aim 4 above).

### ***Conceptual and theoretical frameworks***

This project will extend a typological linguistic and anthropological framework that has successfully supported two prior Discovery projects (DP0878556 and DP120100632) and combine this with digital modelling and mapping methods from CI Hendery's digital humanities research programme and CI Simoff's pioneering work in developing visual modelling methods for interdisciplinary problems. The theoretical framework will further draw from CIs Hendery and McConvell's ongoing work on language contact and change in Australia and the Pacific, and PI Dousset's research on social and political organisation and the ethnohistory of Australia and Melanesia.

The study of language change today takes place in (sub)disciplinary silos, with the different methods and theories for sound change, grammatical change and lexical or semantic change each supplying only part of the story of a given region's linguistic history. Approaches such as linguistic stratigraphy (where e.g. knowledge of how regular sound changes have affected a language or language family can be used to deduce approximate dates and orders of word borrowings) show how much more information can be gained by bringing these perspectives together.

For lexical and semantic analysis in the project we draw on the long tradition of the *Wörter und Sachen* approach, in which the history of an individual word is studied alongside the history of the material object that it denotes. McConvell developed linguistic stratigraphy (McConvell 1990) to provide absolute chronologies by importing archaeological and historical dates to layers of diffused items (generally material culture or plants). He has also shown how polysemies and semantic change in kinship terms can link to the rich anthropological work on the typology or kinship systems, and to known relationships with other aspects of culture. In this project we will also extend this approach to other social practices such as exchange, chiefly systems, and religious or spiritual practices, which have been the focus of PI Dousset's recent research programme. Alongside the geographical mapping of the terms of interest we will apply and extend François's semantic mapping approach (François 2008) to understand semantic change. For analysis of grammatical change we continue the diachronic typological approach pioneered by CI Hendery (2012), which among other influences draws on typological frameworks from Bickel (2007), frameworks and concepts for understanding the relationship between constructions as in Croft (2000), and an understanding of language contact and change heavily informed by Van Coetsem (2000) and Trudgill (2004).

A major focus of research today that investigates the patterns of cultural, social and linguistic change within whole regions is computational phylogenetics (see e.g. Jordan et al. (2009) for an example of this approach for the Pacific region). We situate this project as complementary to phylogenetic projects, in that we share similar goals in understanding the linguistic and anthropological history of the Pacific, but while their focus is more on quantitatively determining *what* was transmitted *when* to *whom*, we are primarily interested in *contextualising* the transmissions - *under what circumstances* did which linguistic and cultural features get taken up by whom, and *what might account for* the similarities and differences in such changes across the region?

### ***Project design and methods***

**Data sources and selection:** We will continue the successful strategy of our research team's other joint and individual projects (DP0878556, DP120100632, DP140102983, DP11010371 and LP160100192) of combining existing case studies and data from the published literature with detailed new investigations of selected items. Specific linguistic and anthropological focuses are:

- Systematic investigation and mapping of words and practices that have already been proposed to be indicative of contact between Australia and its Pacific neighbours, or that our own preliminary investigations have found likely to be so, for example, exchange practices, the words *tabu* and related terms including in-law kinship terms, the kinship term *tavi/tawi*, words for sweet potato, taro, yam, boab and kava, fishing and canoe technology (Aims 1, 3).
- Changes in structural features that have been well described for individual languages and (sub)families of the Pacific and Australia such as possessives, voice systems, alignment systems, word classes, relationships between verbs and objects, word order and verb serialisation. These items will provide linguistic contexts for the diffusion of vocabulary

to be situated against – whether the languages in question have undergone more thorough typological convergence, or whether they generally are more conservative, which would be evidence for less intensive contact. This strand of the project will also help us more fully model *the processes and mechanisms* of language change in the region. (This will especially contribute to aims 2 & 3).

The vocabulary and structural changes will be mapped against each other, but also against known historical networks of e.g. trade routes, migration routes, and intermarriage. Connections between the above domains will also be investigated, for example plant terms are used metaphorically in kinship in both Austronesian and (probably independently) in some Australian languages (c.f. McConvell 2015). (This will especially contribute to aim 3).

**Digital methods:** Just as layering and overlaying multiple sources of data on linguistic and cultural change in the Australia-Pacific region lead to a more nuanced understanding of the region and the relationships between those data, layering multiple ‘digital lenses’ is the only way to achieve a full understanding of the region’s complexities. The ‘lenses’ employed are three different kinds of modelling: maps are models that foreground place; simulations are models that foreground time; and 3D modelling in virtual and augmented reality offers a space in which to bring both time and place together. Each of these is outlined more fully below.

**Mapping (Years 1-2)** We will begin with ontology engineering, developing an ontology database (using eg. Stanford University’s Protégé) for structuring, storing and searching data, and a web-based interface for filtering and mapping results (Y1) and populate this (Y1-2). Word forms, meanings, and grammatical change will be mapped in both time and space, to see whether they cluster (aim 2), follow certain geographic or temporal patterns (aims 1, 2), and whether they correlate with social and cultural phenomena (aim 3). An example of correlated linguistic and non-linguistic data in this space is words for pottery and the archaeological finds themselves; words for parts of canoe, such as ‘outrigger’ and the existence of these kinds of canoes; or kinship terms and kinship systems. The distribution of these items, words, and their differing semantics (using various existing linguistics methods such as analysis of collocation and colexification) can be mapped alongside population movements and contacts, and more abstract changes in grammatical or phonological systems, in order better to understand the relationship between the different kinds of linguistic change and its relationship to language contact. It builds on Hendery’s 2016 Women’s Research Fellowship work, as well as prior work by Hendery & McConvell as part of the ARC Discovery projects DP0878556, DP120100632 and DP140102983.

**Simulation (Years 2-3):** The construction of theory-building models allows us to explore possible *explanations* for what we observe in our data (aim 2, 3), to operationalize theoretical claims about the mechanisms of language change and test whether they can produce the kinds of outcomes we find in our mapping results (aim 1). For example, some words have spread through the Pacific (*‘Wanderwörter’*) seemingly at a very different rate and in different distribution patterns than is followed by grammatical change, and this again differs from the patterns of sound changes, which are more often due to shared inheritance. Even within *Wanderwörter* themselves, there may be differences in the behaviour of words for abstract or spiritual items versus words for material culture if the spread of ideas follows different trajectories than the dissemination of technology (Haynie et al. 2014). Agent-based modelling is a way to explore how interactions between individual speakers (agents) can, over generations, lead to change at the system level (the language). This will help us understand how language change might relate to population movements and cultural change (aims 2, 3). This aspect of the project will tap into CI Simoff’s ‘curious negotiator’ agent-based technology and ‘believability’ formalisms developed in two of his previous ARC grants: DP0557168 and DP0879789. It also builds on joint work CI Hendery has been carrying out with Liam Magee, through a 2015 Western Sydney University internal Digital Humanities grant on simulating the effects of easily-navigable versus difficult terrain in linguistic change.

**3D modelling and Virtual Reality (Years 2-3):** This component will lead to improved understanding of the relationship between the geographic, genetic and historical dimensions of the data relating to the languages and region selected for the project (aim 1), as three dimensional interactive visualisations will allow researchers exploring the data to make connections that might otherwise be obscured. It builds on CI Simoff’s extensive experience with 3D visual data mining (Danish Research Council Grant No. 9900103) and CI Hendery’s TIG grant from the Centre of Excellence for Language Dynamics (2015-2016), which resulted in a museum exhibit at the Canberra Museum and Gallery Nov 2016-March 2017 of a Virtual Reality Pacific Language experience that explores questions of archiving practice, highlights the mismatch between the distribution of languages vs the distribution of linguistic attention, and multilingualism hotspots. The affordances of an extra dimension in data visualisation allows the uncluttered and more accurate representation of data that underlyingly has a large number of dimensions but has been reduced for the purposes of visualisation. This TIG grant focussed on comparing contemporary languages, so this component of the Discovery project will extend this research to include a time-dimension, something which the affordances of VR and AR lend themselves to especially well.

### ***Advancing knowledge***

This project will advance our understanding of the Australia- Pacific region, advance our understanding of linguistics and anthropology, and it will evaluate and develop cutting-edge technology in the humanities and social sciences, in particular advancing practices and platforms in Digital Humanities. It will re-write Australia's own pre-history, relocating Australia within trans-Pacific networks of material, cultural and linguistic exchange. It will especially advance our understanding of the history and anthropology of the Pacific in areas such as social organisation and political systems (for example transformations of kinship systems, chiefly systems, religious and spiritual practices such as sorcery, and material culture). It will advance our knowledge of Australia's first peoples' pre- and post-colonial links with our Pacific neighbours. In linguistics it will advance our understanding of the relationship between phonological change, grammatical change and lexical and semantic change by comparing the distribution of such changes and investigating regional patterns. It will contribute to advances in theories of language change by testing the appropriateness of these theories for modelling the observed patterns. In Digital Humanities we will advance the development of new visual analysis methods for integrated rich linguistic, anthropological and archaeological datasets. In computer science more generally advances will be made in discovering the specific complexities of extracting and managing humanities data. The collaboration between CI Simoff and the humanities-based CIs & PI will therefore lead to new knowledge generation.

### ***International collaboration***

International collaboration is built into this project formally and informally. PI Laurent Dousset is a professor of anthropology at EHESS (specifically CREDO) in Marseilles, former director of CREDO, and established the French national research unit (CNRS) for Digital Humanities. Dousset will travel to Australia for project meetings and intensive work sessions with the rest of the project team each year, and CI McConvell will spend time at CREDO in year two for further collaborative work. Alex François at CNRS in Paris has been involved in discussions around the development of this project, and we expect to continue these conversations throughout. Both François and Professor Eva Schultze-Berndt (Manchester) are participants in a workshop Hendery & McConvell are running mid 2017 on diffusion and change of meaning in cultural words, with a particular focus on *tabu*, which will feed into this proposed project.

### ***Dissemination of results***

Results will be disseminated through both traditional academic outputs (a joint monograph, journal articles, conferences see 'communication of results' below for more details) and also non-traditional research outputs such as an interactive mapping platform and adaptable digital tools for simulation, visualization and other modelling. Non-academic summaries of our research findings, as well as copies of publications (where allowed by the publisher) will be made available on our project website, and publications will also be disseminated through our individual research websites and the university repository, where this is permitted. We will also carry out a campaign of media engagement, submitting regular pieces about the project's findings to popular media outlets such as *The Conversation*.

## **FEASIBILITY**

The project's feasibility is enhanced both by the modular project design and by the team's track record of successful collaboration. The CIs have previously carried out successful ARC Discovery projects, in a variety of independent and joint configurations (see section F) and have published together in various configurations.

The project components (lexico-semantic, grammatical, anthropological, and the diverse technological modelling methods) are designed to have independent outcomes as well as to contribute to the global aims (1-3), which provides a built-in failsafe. In a project that relies on technological methods, the major risk factors are the feasibility, budget and scope of these technological elements. In this case the risk is mitigated by the fact that each of the technological methods builds on pilots CI Hendery has already carried out (see track record for details). The process of constructing the pilot softwares allowed more exact estimation of the costs and time that would be involved in extending them to the wider Pacific context and to more robust implementations.

The research environment at the lead institution and across the partner institutions is supportive, vibrant, and engaged in leading-edge research in the relevant disciplines. Western Sydney University is a strong research environment for interdisciplinary projects, in particular in Digital Humanities. The world-class Digital Humanities Research Group (DHRG) is involved in a number of projects and research programs relating to Australian First People's history, and this project will draw on the protocols and expertise from these and for an advisory board for this project. The DHRG also collaborates with two high-profile interdisciplinary institutes at WSU: the Institute for Culture and Society and the MARCS Institute for Brain, Behaviour and Development, which includes linguistics, engineering, music and psychology. These, as well as the School of Computing, Engineering and Mathematics, of which CI Simoff is Dean, host

leading researchers in linguistics, cultural studies, and computer science, as well as a stream of high profile international visitors. CIs Hendery and McConvell are also affiliates of the Centre of Excellence for Language Dynamics (COEDL), which gives access to an extensive international network of linguists, as well as to regular events and opportunities for collaboration with computer scientists and engineers in the technology stream of COEDL. This has already provided Hendery with training workshops on app development and data visualisation, run by international leaders, including Google engineers. The Partner Organisation CREDO is the largest centre for Pacific Studies in Europe. ANU where CI McConvell is based has the strongest Pacific research focus in Australia, with excellence in Pacific and Australian linguistics, anthropology and archaeology.

## **BENEFIT**

The proposed project will benefit Australia by increasing our understanding of Australia's pre- and post-contact history as a networked agent in a complex web of Pacific interaction. This will have benefits for curriculum, as well as public perception of Indigenous Australia. Today Australia faces challenges with Pacific immigration and integration, and follow-on economic effects of its smaller Pacific neighbours adopting new technologies, new political and economic systems. This project will help us to contextualise population movement in the Australia-Pacific region as a normal, long-term phenomenon not constrained by colonial nation-boundaries, with positive consequences for cultural exchange, which has flow-on effects for reconciliation and unity. The project will develop non text-based ways of representing information, which, with careful attention to developing visual metaphors that are cross-culturally appropriate, will produce outcomes that are of interest to diverse populations of Australia and the Pacific, including tourists visiting the region. Furthermore the project will raise Australia's profile in an area of research that is crucial to contemporary linguistics. Australia has developed a strong profile in linguistic research on the Pacific region, which this project will increase by integrating the linguistics with other Pacific disciplines, thus increasing its relevance and the reach of its dissemination.

The proposed project will be cost-effective because of it builds on a solid foundation of research that the ARC has already invested in. Data and frameworks will draw from Austkin I (DP0878556), Austkin II (DP120100632) McConvell's Dynamics of Social Organisation project (DP140102983), and Hendery's Palmerston Island project (DP11010371), as well as smaller projects and workshops. We anticipate finding further relevant data in our archival research as part of LP160100192.

## **COMMUNICATION OF RESULTS**

We will produce a monograph on *Language change and exchange in Australia and the Pacific*, in which we bring together the outcomes from the diverse strands of the project and discuss their implications for the history of the region. We would publish this with a high-profile publisher such as Oxford or Cambridge University Press. We also aim to publish in high-impact journals in linguistics and anthropology on empirical and theoretical findings, i.e. papers on the diffusion of the linguistics and anthropological features listed above; on how they together build up a clearer picture of the region; the role of Pacific Pidgin/NSW Pidgin in these diffusions; untangling pre- and post-colonial changes; the relationship between social, cultural and linguistic change; and on how each of mapping, simulation, and data visualisation assists our understanding of the above. We will target e.g. *Diachronica*, *Linguistic Typology*, *the Journal of Language Dynamics*, *the Journal of Pacific History*, *Current Anthropology*, *Oceania*, *Oceanic Linguistics*. A paper on how new computational methods enabled us to uncover previously unknown prehistory of Australia's relationship with the Pacific is also within the scope of subject matter that *Nature* or *Science* publishes. We intend to present our findings annually at the Australian Linguistics Society conferences, as well as at high profile relevant international conferences such as *Digital Humanities 2020*. The digital tools and mapping platform will be Non Traditional Research Outputs that provide a long-term resource for the field, including disciplines outside linguistics and anthropology. The website and media pieces will also be a visible public resource for researchers, participant communities and the general public. (Cf. 'dissemination of results' above.)

## **MANAGEMENT OF DATA**

Data will be managed using the infrastructure and tools of the Western Sydney University Research Data Repository. This means that the data will be secure, centralised, shared among our team, and backed up throughout the lifecycle of this project. At key stages and with the assistance of the WSU Library, we will be archiving raw and processed data through the use of the Research Data Catalogue. Archiving will provide a secure, permanent, citable home for relevant data, which is also searchable for possible reuse and future collaboration. Under the Open Access to Research Policy, data for completed research and data supporting publications will be made openly available via ResearchDirect.

## REFERENCES

- Amery, R. & Mühlhäusler, P.**, 1996. Pidgin English in New South Wales. In S. A. Wurm, P. Mühlhäusler, & D. T. Tryon, eds. *Atlas of languages of intercultural communication in the Pacific, Asia, and the Americas*. Berlin/New York: Mouton de Gruyter, 33–52. **Attenbrow, V.** 2010. Aboriginal fishing in Port Jackson, and the introduction of shell fish-hooks to coastal New South Wales, Australia. *The Nature History of Sydney*. Royal Zoological Society of NSW, Mosman, Australia, 16-34. **Baker, P.**, 1993. Australian influence on Melanesian Pidgin English. *Te Reo. Journal of the Linguistic Society of New Zealand*, 36, 3–67. **Balme, J.** 2013. Of boats and string: The maritime colonisation of Australia. *Quaternary International* 285: 68-75. **Beltran, F. S. et al.**, 2009. Forecasting a language shift based on cellular automata. *Journal of Artificial Societies and Social Sciences*, 12(3). **Bickel, B.** 2007. Typology in the 21st century: major developments. *Linguistic Typology*, 11, 239–251. **Clark, R.** 1979. In search of Beach-la-mar: Towards a history of Pacific Pidgin English. *Te Reo. Journal of the Linguistic Society of New Zealand*, 22, 3–64. Clark, M. & S. May (eds). 2013. *Macassan History and Heritage: Journeys, Encounters and Influences*. ANU E-Press. **Croft, W.** 2000. *Explaining language change: An evolutionary approach*, London: Longman. **Dousset, L.**, Systems in geography or geography of systems: Attempts to represent spatial distributions of Australian social organisation. In P. McConvell, H. Koch, & I. Keen, eds. *Australian Kinship through time*. Berghahn. **Dryer, M.S. & Haspelmath, M.** eds., 2011. *The world atlas of language structures*, Leipzig: Max Planck Digital Library. **Evans, N.**, 1992. Macassan loanwords in top end languages. *Australian Journal of Linguistics*, 12(1), 45–91. **Evans, N.**, 1997. Macassan loans and linguistic stratification in western Arnhem Land. In P. McConvell & N. Evans, eds. *Archaeology and linguistics: Aboriginal Australia in global perspective*. Melbourne: Oxford University Press, 237–260. **Evans, N.**, 2016. Typology and coevolutionary linguistics. *Linguistic Typology*. **Fleischer, J., Rieken, E. & Widmer, P.**, 2015. *Agreement from a diachronic perspective*, Berlin: Walter de Gruyter. **Francois, A.**, 2008. Semantic maps and the typology of colexification: Intertwining polysemous networks across languages. In M. Vanhove, ed. Amsterdam: John Benjamins Publishing Company, 163–216. **Givón, T.**, 2012. Toward a diachronic typology of relative clauses. In B. Comrie & Z. E. Fernández, eds. *Relative clauses in languages of Americas: a typological overview*. Amsterdam: John Benjamins, 3–25. **Haynie, H. et al.**, 2014. Wanderwörter in languages of the Americas and Australia. *Ampersand*, 1:1–18. **Hendery, R.**, 2012. *Relative Clauses in Time and Space: A case study in the methods of diachronic typology* Amsterdam, John Benjamins. **Hendery, R. & McConvell, P.**, 2013. Mama and papa in Indigenous Australia. In P. McConvell, I. Keen, & R. Hendery, eds. *Kinship systems: change and reconstruction*. Salt Lake City: University of Utah Press, 217–238. **Isaksen, L. et al.**, 2014. Pelagios and the emerging graph of ancient world data. In *WebSci 14: Proceedings of the 2014 ACM conference on web science*. ACM, 197–201. **Jordan, F.M. et al.**, 2009. Matrilineal residence is ancestral in Austronesian societies. *Proceedings of the Royal Society B: Biological Sciences*, 276(1664), p.1957. **Keen, I.**, 2013. The evolution of Yolngu and Ngarinyin kinship terminologies. In P. McConvell, I. Keen, & R. Hendery, eds. *Kinship systems: change and reconstruction*. Salt Lake City: University of Utah Press, 132–162. **Koch, H.**, 2012. Loanwords between the Arandic languages and their western neighbours: principles of identification and phonological adaptation. In L. Gawne & J. Vaughan, eds. *Selected Papers from the 44th Conference of the Australian Linguistic Society, 2013*. Melbourne: Australian Linguistic Society, 311–334. **Koch, H.**, 2013. The reconstruction of kinship terminology in the Arandic languages of Australia. In P. McConvell, I. Keen, & R. Hendery, eds. *Kinship systems: change and reconstruction*. Salt Lake City: University of Utah Press, 163–186. **Koptjevskaja-Tamm, M.**, 2016. The lexical typology of semantic shifts: an introduction. In P. Juvonen & M. Koptjevskaja-Tamm, eds. *The lexical typology of semantic shifts*. Berlin: De Gruyter Mouton, 1–21. **Loreto, V. & Steels, L.**, 2007. Social dynamics: Emergence of language. *Nature Physics*, 3(11), 758–760. **Macknight, C.C.**, 2013. Studying trepangers. In M. Clark & S. May (eds), *Macassan History and Heritage: Journeys, Encounters and Influences*. Canberra: ANU E-Press 19-40. **McConvell, P.**, 1990. The linguistic prehistory of Australia: Opportunities for dialogue with archaeology. *Australian Archaeology*, 31(1), 3-27. **McConvell, P.**, 2012. Omaha skewing in Australia: Overlays, dynamism and change. In T. R. Trautmann & P. M. Whiteley, eds. *Crow-Omaha: New light on a classic problem of kinship analysis*. Tucson: University of Arizona Press, 243–260. **McConvell, P.**, 2013. Proto-Pama-Nyungan kinship and the AustKin project: Reconstructing proto-terms for “mother’s father” and their transformations. In P. McConvell, I. Keen, & R. Hendery, eds. *Kinship systems: change and reconstruction*. Salt Lake City: University of Utah Press, 192–216. **McConvell, P.**, 2015. Language and culture history: the contribution of linguistic prehistory. In *The Routledge Handbook of Language and Culture*. London: Routledge, 218–220. **McConvell, P. & Bown, C.**, 2011. The prehistory and internal relationships of Australian languages. *Language and linguistics compass*, 5(1), 19–32. **McConvell, P. & Dousset, L.**, 2012. Tracking the dynamics of kinship and social category terms with AustKin II. In *Proceedings of the EACL 2012 Joint Workshop of LINGVIS & UNCLH*. Avignon: Society for Computational Linguistics, 98–107. **McConvell, P. & Ponsionnet, M.**, 2013. Results and prospects in the study of semantic change: A review of from polysemy to semantic change (2008). *Journal of Language Contact*, 6(1), 180–196. **McNiven, Ian J., et al.** 2006. Mask Cave: Red-Slipped Pottery and the Australian-Papuan Settlement of Zenadh Kes (Torres Strait). *Archaeology in Oceania*, vol. 41:2, 49–81. **Montemurro, M.A. et al.**, 2011. Universal Entropy of Word Ordering Across Linguistic Families. *PLoS ONE*, 6(5), p.e19875. **Pagel, M., Atkinson, Q.D. & Meade, A.**, 2007. Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature*, 449(7163), 717–720. **Presner, T.S., Shepard, D. & Kawano, Y.**, 2014. Thick mapping in the digital humanities. In T. S. Presner, D. Shepard, & Y. Kawano, eds. *HyperCities : thick mapping in the digital humanities*. 44–65. **Ross, M.**, 1992. The sources of Austronesian lexical items in Tok Pisin. In T. Dutton et al., eds. *The language game: Papers in memory of Donald C. Laycock*. Canberra: Pacific Linguistics, 361–384. **Ross, M.**, 2001. Contact-induced change in Oceanic languages in north-west Melanesia. In A. Y. Aikhenvald & R. M. W. Dixon, eds. *Areal diffusion and genetic inheritance: problems in comparative linguistics*. Oxford: OUP, 134–166. **Ross, M.**, 2013. Diagnosing Contact Processes from their Outcomes: The Importance of Life Stages. *Journal of Language Contact*, 6(1), 5–47. **Serzant, I.A. & Kulikov, L.** eds., 2013. *The diachronic typology of non-canonical subjects*, Amsterdam: John Benjamins. **Sheppard, P. et al.**, 2015. A Solomon sea interaction sphere? In C. Sand, S. Chiu, & N. Hogg, eds. *The Lapita Cultural Complex in time and space: expansion routes, chronologies and typologies*. Noumea: Institut d’archéologie de la Nouvelle-Calédonie et du Pacifique, 63–80. **Shnukal, A.**, 1983. Torres Strait Creole: the growth of a new Torres Strait language. *Aboriginal History*, 7(1/2), 173–185. **Simpson, J.**, 2013. Warumungu kinship over time. In P. McConvell, I. Keen, & R. Hendery, eds. *Kinship systems: change and reconstruction*. Salt Lake City: University of Utah Press, 239–254. **Tochilin, C. et al.**, 2012. Sourcing temper sands in ancient ceramics with U–Pb ages of detrital zircons: a southwest Pacific test case. *Journal of Archaeological Science*, 39(7), 2583–2591. **Trudgill, P.**, 2001. Contact and Simplification: Historical Baggage and Directionality in Linguistic Change. *Linguistic Typology*, 5(2-3), 371-374. **Trudgill, P.**, 2004. *New-dialect formation : the inevitability of colonial Englishes*, Edinburgh: Edinburgh University Press. **Urry, J. & Walsh, M.**, 1981. The Lost “Macassar Language” of Northern Australia. *Aboriginal History*, 5, p.91. **Van Coetsem, F.**, 1988. *Loan phonology and the two transfer types in language contact*, Dordrecht: Foris. **Van Coetsem, F.**, 2000. *A general and unified theory of the transmission process in language contact*, Heidelberg: Winter. **White, P. C. Reepmeyer & G. Clark.** A Norfolk Island basalt adze from coastal New South Wales. *Australian Archaeology* 79.1 (2014): 131-136. **Zorc, R.D.** 1986. *Yolngu-Matha dictionary*. Batchelor: School of Australian Linguistics.