Morphology and Language History

EDITED BY
Claire Bowern
Bethwyn Evans
Luisa Miceli

JOHN BENJAMINS PUBLISHING COMPANY
MORPHOLOGY AND LANGUAGE HISTORY
## CONTENTS

Introduction ................................................................................................................1
Claire Bowern, Bethwyn Evans, Grace Koch and Luisa Miceli

### I. Genetic relatedness

1. Western Torres Strait language classification and development...............15
   Barry Alpher, Geoffrey O’Grady and Claire Bowern
2. The classification of Pinikura, Western Australia.................................31
   Peter K. Austin
3. Bound pronominals in the West Papuan languages ..........................43
   Mark Donohue
4. Alawa and its neighbours: Enigma variations 1 and 2.........................59
   Margaret Sharpe
5. Reconstructing pre-Warumungu pronominals ..................................71
   Jane Simpson

### II. Reconstruction

6. Splitting vs. lumping in morphological analysis: evidence from Greek ......91
   Avery Andrews
7. Pronominal accretions in Pama-Nyungan ...........................................99
   Paul Black
8. Associated eating and movement: further examination of Yuwaalaraay Gamilaraay verb suffixes .................................................................107
   John Giacon
9. The origin of conjugation markers in Australian languages ...............123
   Mark Harvey
10. Some remarks on negatives in Southeastern Australia .....................139
    Luise Hercus and Stephen Morey
11. *gʷes-, *(z)gʷes-, *(s)gʷesh₂-? The PIE root for ‘extinguish/go out’ ..........155
    Jay H. Jasanoff
12. The language of Central Australian Aboriginal songs ..........................167
    Grace Koch and Myfany Turpin
13. The origin of noun classes in Worrorran languages ...............................185
    William B. McGregor
14. Hittite duwān (parā) ..................................................................................201
    H. Craig Melchert
15. Morphological reconstruction and Australian languages ..........................211
   Luisa Miceli
16. Warlpiri verb roots in comparative perspective ....................................221
   David Nash
17. Oujiang Wu tones and acoustic reconstruction ....................................235
   Phil Rose
18. Issues in the morphological reconstruction of Proto-Mon-Khmer ..........251
   Paul Sidwell

III. Processes of change
19. Case selection Old and New Basque ...........................................................269
   Cathryn Donohue
20. Third person plural as a morphological zero: object marking in Marovo .................................................................281
   Bethwyn Evans
21. The morphological development of the perfect in Jersey Norman French .................................................................299
   Anthony J. Liddicoat and Timothy Jowan Curnow
22. Grand-daddy morphs: the importance of suffixes in reconstructing Pama-Nyungan kinship .....................................................313
   Patrick McConvell
23. Morphology of the eggs, and what it can tell us about Romanian nominal inflection ..........................................................329
   Kim Schulte
24. The refunctionalisation of first person plural inflection in Tiwi ...........341
   John Charles Smith
25. A chain vowel raising in the early history of Chinese ...........................349
   Xiaonong Zhu

Language Index ........................................................................................................355
Subject Index .............................................................................................................
CONTRIBUTORS’ ADDRESSES

Barry Alpher
3218 Wisconsin Avenue NW, Apt. B2
Washington, DC 20016
USA
barryalpher@yahoo.com

Avery Andrews
School of Language Studies
Australian National University
Canberra, ACT 0200
Australia
Avery.Andrews@anu.edu.au

Peter K. Austin
School of Oriental and African Studies
Thornhaugh Street, Russell Square
London, WC1H 0XG
United Kingdom
pa2@soas.ac.uk

Paul Black
Faculty of Education, Health and Science
Charles Darwin University
Darwin, NT 0909
Australia
paul.black@cdu.edu.au

Claire Bowern
Department of Linguistics
Rice University
6100 Main St, MS23
Houston, Texas, 77005
USA
bowern@rice.edu

Timothy Jowan Curnow
Research Centre for Languages and Cultures Education
School of International Studies
University of South Australia
GPO Box 2471
Adelaide SA 5001

Cathryn Donohue
Linguistics Program
Department of English/098
University of Nevada, Reno
Reno, NV 89557-0098
USA
donohue@unr.edu

Mark Donohue
School of Languages, Cultures and Linguistics
Monash University
Victoria 3800
Australia
mark@donohue.cc

Bethwyn Evans
School of Languages, Linguistics and Cultures
University of Manchester
Oxford Road
Manchester M13 9PL
United Kingdom
bethwyn.evans@manchester.ac.uk

John Giacon
School of Language Studies
Australian National University
Canberra, ACT 0200
Australia
jgiacon@ozemail.com.au

Luise Hercus
School of Language Studies
Australian National University
Canberra, ACT 0200
Australia
luise.hercus@anu.edu.au
CONTRIBUTORS’ ADDRESSES

Mark Harvey
School of Humanities and Social Science
The University of Newcastle
University Drive
Callaghan 2308 NSW
Australia
mark.harvey@newcastle.edu.au

Jay H. Jasanoff
Department of Linguistics
312 Boylston Hall
Harvard University
Cambridge, MA 02138
USA
jasanoff@fas.harvard.edu

Grace Koch
AIATSIS
GPO Box 553
Canberra, ACT 2601
Australia
Grace.Koch@aiatsis.gov.au

Anthony J. Liddicoat
Research Centre for Languages and Cultures Education
School of International Studies
University of South Australia
GPO Box 2471
Adelaide SA 5001
Australia
Tony.Liddicoat@unisa.edu.au

William B. McGregor
Department of Linguistics
Århus Universitet
Nordre Ringgade, bygning 1410, 3
DK - 8000 Århus C
Denmark
linwmg@hum.au.dk

Patrick McConvell
AIATSIS
GPO Box 553
Canberra, ACT 2601
Australia
patrick.mcconvell@aiatsis.gov.au

H. Craig Melchert
Department of Linguistics
3125 Campbell Hall, P.O. Box 951543
UCLA
Los Angeles, CA 90095-1543
melchert@humnet.ucla.edu

Luisa Miceli
Linguistics (M258)
School of Humanities
University of Western Australia
35 Stirling Hwy, CRAWLEY, WA 6009
Australia
lmiceli@cyllene.uwa.edu.au

Stephen Morey
Research Centre for Linguistic Typology
La Trobe University
Victoria 3086
Australia
s.morey@latrobe.edu.au

David Nash
School of Language Studies
Australian National University
Canberra, ACT 0200
Australia
David.Nash@anu.edu.au

Geoffrey O’Grady
Department of Linguistics
University of Victoria
PO Box 3045
Victoria BC V8W 3P4
Canada

Phil Rose
School of Language Studies
Australian National University
Canberra, ACT 0200
Australia
philip.rose@anu.edu.au
Kim Schulte  
Department of Modern Languages  
University of Exeter  
The Queen's Drive  
Exeter, EX4 4QH  
United Kingdom  
K.Schulte@exeter.ac.uk

Margaret Sharpe  
School of Behavioural, Cognitive and Social Sciences  
University of New England  
Armidale, NSW  
Australia  
msharpe@une.edu.au

Paul Sidwell  
Department of Linguistics  
Research School of Pacific and Asian Studies  
Australian National University  
Canberra, ACT 0200  
Australia  
paulsidwell@yahoo.com

Jane Simpson  
Department of Linguistics, F12  
University of Sydney, N.S.W. 2006  
Australia  
jane.simpson@usyd.edu.au

John Charles Smith  
St Catherine's College  
Oxford, OX1 3UJ, UK  
United Kingdom  
johncharles.smith@stcatz.ox.ac.uk

Myfany Turpin  
School of English, Media Studies & Art History  
Fourth Floor, Michie Building  
The University of Queensland  
St Lucia 4072  
myfturpin@netspace.net.au

Xiaonong Zhu  
Division of Humanities  
The Hong Kong University of Science and Technology  
Clear Water Bay, Kowloon  
Hong Kong  
hmxzhu@ust.hk
HAROLD KOCH
(Montréal, July 2007)
INTRODUCTION

CLAIRE BOWERN, BETHWYN EVANS, GRACE KOCH AND LUISA MICELI
Yale University, University of Manchester, Australian Institute of Aboriginal and Torres Strait Islander Studies and University of Western Australia

The present volume aims to make a contribution to codifying the methods and practices historical linguists use to recover linguistic history. We highlight the theoretical basis of processes of change and linguistic reconstruction, focussing predominantly on historical morphology.

Although there is a considerable body of literature dealing with these issues, it is difficult to find a single source whose findings are broadly cross-linguistic. The focus within historical linguistics is usually on Indo-European languages and this is understandable since it reflects the history of the field and the fact that its development has largely taken place within this language family. Nevertheless, Indo-European provides a very specific set of circumstances, as does any single language family, and while the theoretical and methodological issues encountered within these languages are highly relevant to historical and comparative research carried out in other parts of the world, work outside of Indo-European has also revealed different kinds of issues that need to be addressed. It is for this reason that the present volume includes papers on a wide range of language families: not only Indo-European, but also Austronesian, Sinitic, Mon-Khmer, Basque, and one Papuan language family, as well as a number of Australian language families. The volume is divided into three parts – genetic relatedness, reconstruction and processes of change. In the sections that follow we provide a brief discussion of each of these themes and the relevant papers.

The subject matter of this volume has been the life work of Harold Koch (see section 4 of this introduction for an account of Harold’s career by Grace Koch). We would like to honour the invaluable contribution that he has made to the discipline, both as a scholar and as a teacher, by dedicating this collection of papers to him.

1. Genetic Relatedness

Most of the papers in this part of the book deal with Australian languages. This is not surprising since genetic relationships of languages in this region have often proven to be difficult to recover. Establishing regular sound correspondences has not always been possible due to paucity of data, inaccurate transcriptions, absence of adequately diagnostic sound changes, or conflicting correspondence sets due to borrowing. As will be seen in the papers that follow,
Australianists often rely heavily on morphological similarities in making a genetic argument. The use of morphology as primary evidence in a genetic argument relies on the assumption that whenever the combination of similar morphemes with a shared paradigmatic organisation is observed, it is better explained in terms of historical connection than chance (see Evans 2003 amongst others), and that borrowing is also unlikely as morphological patterns — especially suppletive patterns — are supposedly resistant to borrowing (see Meillet 1925).

**ALPHER, O'GRADY and BOWERN** discuss evidence for the genetic classification of the Western Torres Strait language. They show that despite radical innovations in phonology, there is considerable evidence for its inclusion in the Pama-Nyungan language family. However, not all aspects of a language’s morphological system will be indicative of genetic relationships and M. DONOHUE’s paper discusses this very point with regard to West Papuan languages, demonstrating that in his case study bound pronominal elements are the most reliable indicators of genetic relationships. **SHARPE** discusses the implications of different levels of cognacy between lexicon, verb morphology and other morphology. She points out the problems for establishing the relationship of Alawa, Mara and Warndarang, despite a considerable degree of shared vocabulary and nominal morphology. **SIMPSON** investigates the genetic position of Warumungu by providing an internal reconstruction of its pronominal system that is then compared to that of neighbouring languages. Her investigation proves to be inconclusive, though she is able to identify the elements in the system that are most likely to be archaic. **AUSTIN** uses evidence from verb morphology and philology to investigate the possible closest relatives for Pinikura, a sparsely attested language of the Pilbara region in Western Australia.

Still on the topic of morphological similarities, in Part II of the book, **MICELI** discusses the validity of genetic arguments based solely on this type of evidence.

### 2. Reconstruction

A major part of investigating linguistic prehistory is reconstruction, namely hypotheses regarding the non-attested antecedent states of a language or group of languages. There is considerable variation in the ways in which reconstruction is viewed by historical linguists, in terms of both its role within the investigation of linguistic prehistory and the methods with which it is achieved.

For many scholars reconstruction is an integral part of the Comparative Method; the logical next step after establishing sound correspondences and cognate sets (Durie & Ross 1996, Nichols 1996, Rankin 2003). As such reconstruction is closely linked with demonstrating genetic relatedness, and can be used to strengthen the evidence supporting relationships already
established. For example, McGregor examines the Worrorran languages of northern Australia, showing how the reconstruction of both the forms and categories of noun classes further supports the family-status of this group of languages. McGregor’s reconstruction of the semantic restructuring of noun classes also highlights the need for methods of reconstruction that account for aspects of linguistic structure other than those which fall within the domain of the Comparative Method. The Comparative Method, strictly speaking, is limited to the lexico-phonological domain. It is a method that demonstrates cognacy amongst linguistic elements that are similar in terms of both form and meaning, where this pairing is arbitrary (Harrison 2003). Formal similarity is established through the recognition of regular sound correspondences, and reconstruction of the formal aspects of a proto-item is based on these correspondences and knowledge of sound change. The reconstruction of other aspects of linguistic prehistory requires techniques other than the Comparative Method (Harrison 2003). Methods of reconstruction used to complement the Comparative Method (or that sometimes stand in place of the Comparative Method) are rarely codified or made explicit, as highlighted by Miceli.

So what are the techniques and practices of historical linguists as they move reconstruction beyond the Comparative Method and towards “undoing” changes outside of the lexico-phonological domain? Primarily, as Koch (1996:222) describes for morphological reconstruction, reconstruction beyond the lexico-phonological domain is guided by knowledge of what constitutes plausible processes of change. In their papers, Jasanoff and Melchert demonstrate how traditional techniques of detailed etymological research can form the basis of reconstructions which also account for aspects of change other than phonological. Thus Jasanoff discusses the relationship between morphological explanations for change, morphological analysis as an input to reconstruction, and underlying phonological representations. The basis for discussion is the Ancient Greek root /extinguish, go out/ and he provides a good example of the need to consider morphological change and paradigm uniformity in their own right, and not simply as the basket in which all the exceptions to Neogrammarian sound change are put. Melchert discusses problems with the etymology of Hittite /duw/, which also touches on the role of paradigms (in this case, the reanalysis of a case-marked noun) and constructional semantic change.

In recent years, a major criticism of techniques of reconstruction is that their development within the Indo-European language family means they are not necessarily appropriate in other linguistic contexts where the basic processes of change may be different (Hale 2007:242-243). This is a challenge that has long been felt by those working with Australian languages because, as shown here by Miceli, these languages display a high degree of formal similarity between potential cognates, alongside an exceedingly low number of potentially cognate forms. Despite these problems, Harold Koch has always advocated the use of tried and tested methods of reconstruction aided by established models of
language change. Thus Koch (2003), for example, demonstrates the value of an etymological basis for morphological reconstruction with case studies from the Arandic languages of central Australia. Hercus and Morey similarly use etymological techniques to investigate the history of negative forms within southeastern Australian languages, demonstrating how such methods can form the basis of hypotheses regarding the patterns of inheritance and borrowing. Black’s detailed etymological study of pronouns in Pama-Nyungan languages adds to the understanding of the origins and development of those forms considered to be one of the strongest indicators of the genetic unity of Pama-Nyungan.

Alongside such traditional methods, Harold Koch promotes the need to codify the discipline’s collective knowledge of the ways in which languages can and do change. Thus Koch (1996) not only sets out a method for reconstructing morphology, he presents it alongside a typology of morphological change which must necessarily inform any reconstruction. Such a typology forms an important basis for Sidwell’s reconstruction of Proto Mon-Khmer morphology, based on ‘bottom-up’ methods of reconstruction in order to be independent of the Austic Hypothesis, which has influenced many discussions of Mon-Khmer morphology and its history. Giacon also uses knowledge of plausible kinds of morphological change to inform his reconstruction of verb specification morphemes in Yuwaalaraay-Gamilaraay, an Australian language of New South Wales, which in this case is the basis of grammatical development in a language revival situation.

For Harvey, too, typologies of change are invaluable. He examines the verbal ‘conjugation’ markers that are widely distributed across Australian languages with potentially cognate verb stems, and demonstrates that historically the system is best considered from the perspective of closed versus open paradigms and the different kinds of change associated with them. Nash is also concerned with closed paradigms, but in this case, with the reconstruction of small versus large closed verb systems in the Ngumpin-Yapa subgroup of Pama-Nyungan. Part of Nash’s argument concerns the evidence that reconstruction itself brings to bear on paths of change. That is, how much weight should we give to the reconstruction of changes that appear to conform to known pathways of grammaticalisation? (A similar question is the topic of Hendery 2007.)

A different facet of the contribution of reconstruction to theory is Andrews’ work on how results from historical linguistics should help us decide between competing synchronic theories of morphology. Andrews uses evidence from analogical restorative changes in the history of Greek verbal morphology to support an argument that at least some language learners (and innovators) prefer an analysis based on ‘splitting’ the verb into the smallest possible analytic units, even at the loss of a generalisable structure. C. Donohue also discusses the relationships between synchronic and diachronic explanations.
A crucial aspect of morphological reconstruction is the hypothesising on the kinds of linguistic features likely to be archaic and thus reconstructable. As Koch (1996:219) has noted, it is the synchronically irregular or anomalous features of a language that are often archaic. Such a principle has led to the hypothesis that certain specialised speech styles, which are characterised by linguistic features that are irregular in comparison to everyday speech, may thus reflect features reconstructable for the proto-language. G. Koch and Turpin challenge this from the perspective of Central Australian Aboriginal song language. Through the comparison of song language, everyday speech and Harold Koch's Proto-Arandic reconstructions, they demonstrate that while some archaic morphological features are retained in song language, processes such as phonological alteration due to metrical requirements and the multi-dialectal nature of song means that many aspects of song language are in fact innovative.

With different techniques of reconstruction, is it that all reconstructions represent the same kind of entity? All the papers within this volume take a realist approach to reconstruction. That is, the reconstructions presented have, as Harrison (2003:240) puts it, "the status of best approximations of antecedent historical states", rather than abstract representations of correspondences between attested language states. This is an issue taken up directly by Rose. Tonal systems are traditionally reconstructed in terms of an abstract system of contrasts on the basis of contrasts in attested languages. Using a novel method, Rose reconstructs the acoustic values of proto-tones for the Oujiang subgroup of the Wu dialects of Chinese. It is important, however, that the techniques and models of change which underlie any reconstruction are made explicit, because realist or not, it is only on this basis that their appropriateness as hypotheses for earlier systems can be judged.

3. Processes of change

Underlying any linguistic reconstruction is the assumption that certain changes have taken place in the languages' history, and as Koch (1996) notes, reconstruction needs to be based on knowledge of plausible and possible types of change. While morphological change has traditionally been viewed as isomorphic to analogical change (Hock 2003), Koch's (1996) typology of morphological change demonstrates the breadth of types of change that need to be considered in the histories of morphological systems. Analogy is clearly not the only process at work.

There are many claims in the literature about the universality of morphological changes and about the cross-linguistic applicability of reconstruction methods (Dressler 2003, Heine 2003, Joseph 1998, 2003). However, as already mentioned at the beginning of this introduction, it is quite rare that we find a set of studies which are truly cross-linguistic, and the aim of this book is to move beyond familiar data and clear cases.
Many of the papers in this volume explore processes of morphological change. Some, like Liddicoat and Curnow’s and Evans’, examine in detail cases of well-established types of change such as the process of levelling within paradigms or the creation of morphological zeroes providing a greater understanding of the mechanisms and motivations of change. Other authors, like Smith, C. Donohue and McConvell examine less well-described types of morphological change, not only providing detailed descriptions of change within particular languages, but also adding new insights to knowledge of types of morphological change. Smith uses data from the refunctionalisation of clusivity marking in Tiwi (an isolate in Australia) to argue for refining our conception of exaptation. C. Donohue looks at remapping of case marking in Basque four place predicates. McConvell’s focus is on reconstruction of kinship morphology in Pama-Nyungan, and he examines ways in which inflectional and derivational affixes have been absorbed into kin-term stems.

Zhu and Schulte’s papers both concern actualisation of change. Zhu discusses the processes of change which have resulted in the vowel chain-raising in the history of Chinese, while Schulte’s focus is on processes of paradigm restructuring, examining the dissolution of the Latin neuter plural category in Romance, particularly in relation to the isolated Romanian plural ouă ‘eggs’ and its genitive ouălelor.

4. Harold Koch: scholar and teacher

As mentioned, Harold Koch has long promoted the need for rigorous and codified methods of investigating language history. Beginning his own research on Indo-European languages, and in particular on Hittite, Harold has a deep knowledge and understanding of the theoretical and methodological issues which form the basis of the discipline. It is from this perspective that he began and has continued his historical and comparative research on Australian languages, and in particular the Arandic family. Harold’s skill as a teacher is largely responsible for the recent shift within Australian historical linguistics towards more traditional methods of comparative reconstruction (see Bowern & Koch 2004). Thus, just as studies of Indo-European languages have long dominated discussions of theories of language change and methods of comparative reconstruction, so studies of Australian languages are coming to be equally important within such discussions as Australianists rigorously apply traditional methods of historical linguistics, and present data which challenges some of the traditional notions within the discipline.

Working on Australian languages Harold has become not only an historical linguist, but also a fieldworker, and in the remainder of this introduction Grace Koch writes about this side of Harold’s career.

Having been inspired by Sally Hale’s article in Simpson et al (2001), I wanted to reflect upon some aspects of Harold’s career that some linguists may not know. Like Sally, I will begin by explaining how we first came to Australia, then
move on to some fieldwork experiences. I met Harold in Boston in the late
1960s. After over 5 years of learning to accept one another’s triumphs and
foibles, we married in 1972. So, the observations below come from first-hand
experience.

Harold received a PhD in Linguistics from Harvard in 1973, and, like many
other academics, was seeking a job in his field. Harold scanned the Harvard
bulletin boards as often as he could, looking for openings. What we saw then as
an exotic opportunity arose with the Canberra-based Australian National
University Linguistics Department headed by R.M.W. Dixon for an Indo-
Europeanist who would be willing to undertake fieldwork in either Australia or
the Pacific in order to document an indigenous language, and Harold applied.
Both of us had seen the film, Walkabout, and assumed that if he got the job, we
would be going to live in Canberra in the desert.

Bob Dixon looked favourably on Harold’s application and arranged a rather
unorthodox interview in Cambridge, Massachusetts at the Wursthaus, a German
restaurant in Harvard Square, to which I was also invited. Because the decision
would be a joint one, I needed to know what might be available for me, a music
educator and musicologist, and I was not shy in asking. Bob mentioned that the
ethnomusicologist Alice Moyle needed a research assistant at the Australian
Institute of Aboriginal Studies (A.I.A.S.) in Canberra to document music
recordings in their sound archive. This sounded ideal, especially as Harold had
been working in a shoe factory and I at Carl Fischer music publishers and retail
where conditions were minimal, with no sick days and a 45-minute lunch break.
After a few fits and starts, we decided to take up the challenge to move to
Australia for a two-year duration. Two years has stretched into thirty-two.
Harold has remained at the Australian National University and I, am still at the
(named) Australian Institute of Aboriginal and Torres Strait Islander Studies,
albeit now in another position.

Harold’s choice of Aboriginal languages to study had narrowed to Nyulnyul,
Bardi or Kaytetye. He opted for the latter. In 1974, it was our understanding that
he would have had to fly across Australia to Perth, hire a car and drive 2275
kilometres to Broome in order to contact the most knowledgeable Nyulnyul or
Bardi speakers. Our information may not have been correct, but it caused us to
go to Central Australia. Interestingly, over 20 years later, Claire Bowern chose to
study Bardi and was not faced with the same transportation difficulties.

Kaytetye had several advantages. Kaytetye speakers lived within a two or
three hours’ drive of either Tennant Creek or Alice Springs, both of which had a
fairly direct air service from Canberra. Also, Ken Hale had encouraged Harold to
consider working on Kaytetye because of its distinctive grammatical and
phonological features. Finally, Harold realised that in choosing Kaytetye, the
northernmost dialect of Arrernte, he might avoid the watchful and critical eye

1 I would like to extend most heartfelt thanks to David Nash, Jane Simpson, Jacque Lambert, and
Luise Hercus for editorial comments.
of Professor T.G.H. Strehlow, who had written extensively on Arrernte and had frightened away most linguists with an interest in that language. Strehlow had not included Kaytetye in the grammar he had written on Arrernte, so Harold felt on safe ground. Several years later Strehlow actually asked Harold to join the Strehlow Foundation.2

Our first field trip was in December 1974. I was still awaiting news on the position with Alice Moyle at the A.I.A.S. so was free to go. Not wishing to spend Christmas alone in Canberra, I decided to join Harold on this great adventure. We were to fly to Tennant Creek in the Northern Territory to pick up a vehicle stationed at the North Australia Research Unit from David Daffen, their field officer, then drive to Warrabri Aboriginal Settlement about two hours’ south. The flight stopped initially at Alice Springs, where we were met with a blast of sun and heat beyond anything I had ever experienced. After a brief wait, we reboarded and had a rollercoaster ride due to the savage thermals. I managed to keep lunch down unlike others on the flight, although I remember the stewardess trying to distract us by polite conversation as we struggled with our digestion, responding zombie-like if at all.

The heat in Alice Springs barely prepared us for Tennant Creek, where the temperature soared above the century. Dave Daffen met two people with red faces and queasy stomachs, and promptly informed us that academics were not really very much appreciated around there, speculating that we might be on a flight home soon. He had to do something with us, though, so he took us home and gave us a beer. The calming effect of the beer was offset by the fact that his wife had just bought a bullock from the butcher and was curing the raw meat on the dining room table. Shortly afterwards he took us to the Eldorado Hotel where our room had the air conditioning on full blast. I wanted to stay there indefinitely. After a nap, we headed out for dinner to the Boomerang Restaurant, watching the stick insects crawl up the walls. This entertainment foreshadowed one type of excitement I would experience later at Warrabri, as I lay on a vinyl couch, watching the life and death drama unfolding on the ceiling amongst the spiders, flies, stick insects and other creatures.

We spent a few days in Tennant Creek awaiting fit out of the vehicle. Some of the locals invited us to a picnic on the bone-dry Gosse River where they expected Harold to bog his four-wheel drive in the sand so that they could show him how to get out. They did not realise that they had a good old Waterloo County farm boy on their hands. Our friends got bogged doing wheelies, and Harold helped each get out even though the metal of the cars burned any hand that happened to touch it. He never did get stuck in the sand. The day was

---

2 Founded in 1978, the Strehlow Foundation was established to ensure the maintenance of Strehlow’s large collection of artefacts and documentation of Arrernte life and culture. In 2005, the Northern Territory Government passed an Act to create the Strehlow Research Centre, which contained a provision to work with Aboriginal communities to repatriate some of the objects according to their wishes.
around 45 degrees C so, after that adventure, we all sought shelter in the air-conditioning of the local seismic station.

When the vehicle was ready we headed south to Warrabri. Speakers of four language groups lived there, the Kaytetye (who were the traditional owners of the land) and Alyawarr and the Warumungu and the Warlpiri, who came from places farther north and had been transferred there in the 1950s from the old Phillip Creek Settlement near Tennant Creek. Each group camped in areas located in the general direction of their traditional territories, with Warumungu camping north, Warlpiri west, Kaytetye south and Alyawarr east. Interestingly enough, the name, Warrabri, was a combination of Warumungu and Warlpiri language names. Some people found the whole name too hard and just called it Warlpiri, which was especially upsetting for the traditional owners. The actual name of the place is Alekerenge, which is Kaytetye for “belonging to dog”, for this is Dog Dreaming territory. Lots of fierce and often scruffy camp dogs patrolled the area. Our house came with two large golden Labrador retrievers (appropriate to Dog Dreaming country) who lived beneath the house and often knocked against the floor as they rumbled around at odd hours.

The Superintendent at Warrabri either suggested or insisted that Harold work with Tracker Mick, who had done linguistic elicitation with Ken Hale. Mick worked as a yardsman, so his time needed to be negotiated with his supervisors. During their first elicitation session, after getting acquainted, Mick began to give vocabulary for body parts without even being prompted. After all, he had had practice working with the best! Harold found him a patient teacher, but the time limitation became a worry. A concentration of Kaytetye speakers lived at Neutral Junction Station to the south, so Harold began to forge connections with Tommy Thompson Kngwarreye, Peter Horsetailer Jabiyard and Alec Kapetye, who was featured in Chatwin (1986). They camped near the store on Neutral Junction, sometimes working as stockmen for the Petrick family, who lived on the station and were its leaseholders.

Harold made several other trips in the next two years working on grammar and vocabularies. Some of the Kaytetye women remembered that Harold had a spouse who was interested in music. In 1976, they planned their part of a Rain Dreaming ceremony, and the AIAS funded me to go to Barrow Creek and record and document the performance. This didn’t quite happen as planned, but to my joy I recorded significant parts of three women’s song series—Rain, Fire and Red Bank. At AIAS I had listened to and documented many field recordings of music and was thrilled to have the chance to record my own, putting into practice some of the documentation techniques taught to me by Alice Moyle.

Most of the Kaytetye women who chose to work with me were connected in some way to Harold’s teachers. Alec Kapetye’s wife, Daisy Akemarre, became my main teacher along with Katie Ampetyane (who had been married to Peter Horsetailer), and Mary Kngwarreye (Tommy Thompson’s sister). They gave me a Kaytetye “skin name” that categorised me as a “daughter” to Daisy, thus providing a point of reference for them to situate me within the complex social
structures of their community. Harold was disappointed that his Kaytetye “skin name” (Thangale) was fairly easy to pronounce. Mine, Kngwarreye, was the one he really wanted because that would have allowed him to wax eloquent on pre-stopped nasals. Lots of people, though, began calling Harold “Kaytetye” and referring to me as “Kaytetyarenge”, meaning “belonging to Kaytetye.” I have been trying to establish a separate identity ever since!

Harold met many of the pastoralists who ran stations where Kaytetye people were living because he needed to obtain their permission to talk to the residents; also, he had met them at the famous Barrow Creek Pub, where he stayed on several subsequent field trips. Obviously they, too, had their own names for Harold. David Nash tells me that when he first went to the Northern Territory in 1977, Harold told him to introduce himself to Slippery Harris, who managed McLaren Creek Station, and to tell him that ‘the Barrow Creek linguist’ had sent him. Perhaps the most interesting appellation came from a radio telephone operator at Warrabri. We needed to contact someone in Alice Springs, and requested help. The operator asked us our names, to which Harold replied “Harold and Grace Koch.” Harold’s Canadian accent may have put him off, so he identified us to the bemused recipient of the call as “Aroldan.”

Harold’s work in comparative linguistics has extended his formidable analytical skills to other Aboriginal languages. He takes great pride in his students and gets much pleasure introducing them to the Comparative Method and to the intricacies of Australian Aboriginal languages. For so much of his life he has put others first. I am very pleased that his students and colleagues have put this volume together in his honour.

References

3 The Kaytetye people have a complex series of eight “skin names” corresponding loosely to eight surnames. Everyone is related in some way to everyone else. I had the name, Kngwarreye, which meant that all other women or men with that name were my classificatory sisters or brothers.
4 Personal communication, David Nash. 28/09/2006.


PART I
GENETIC RELATEDNESS
1. Introduction

The language of the Western Torres Strait Islands (WT) has long been regarded as a member of the Pama-Nyungan (PN) family of Australian languages.\(^1\) It was Sidney Ray, a member of the Cambridge Expedition to Torres Straits, who first published explicit evidence from pronominal paradigms linking WT to an Aboriginal, and PN, language of the Australian mainland, Guugu-Yimithirr (1907: 267–8).\(^2\) Some 70 years later, Terry Klokeid and Ephraim Bani showed that like languages of the Australian mainland WT distinguished an apical series of stops (/t, d/) from a laminal one (/th, dh/), and they presented full pronominal paradigms in which there was a perfect correspondence between laminal stops in WT and in languages that had come to be regarded as Pama-Nyungan (O’Grady, Voegelin & Voegelin 1966). The pronoun paradigms (see below) constitute some of the most striking evidence for the PN membership of WT.

While we, among a number of scholars, consider it to be, in a sense, self-evident that WT belongs genetically with PN, much nonetheless remains to be demonstrated. We are currently engaged in assembling a list of etyma and in enumerating the sound changes that have characterised this language as it has evolved from proto-Pama-Nyungan (pPN). A number of sound correspondences are problematic, though we do not go into these data in detail here. The phylogenetic position of WT within PN constitutes a further set of problems. Again, we do not deal with these issues here, beyond repeating the observations of others that the languages of the nearest portions of the mainland do not appear to be closely related: that is, WT is not Paman.

It is thought that WT reflects heavy linguistic influence from Meryam, the language of the culturally similar people of the Eastern Torres Strait Islands. Meryam has been shown (e.g. Evans 2005: 255–256) to be related to languages of the Papua New Guinea mainland—to be a “Papuan” language of the Trans-Fly family (Eastern subgroup). Some features of WT seem consistent with a

---

\(^1\) For the genetic status of Pama-Nyungan, see Alpher (2004) and Evans (2005) and references. For the position of the Western Torres languages in this classification, see O’Grady, Voegelin and Voegelin (1966:25,55), Capell (1956), and Schmidt (1919). We thank the audiences of ICHL 2007 and LingAd (ALS) 2007 for feedback on a summary of the work presented here. Bowern’s work was funded in part by NSF-0643517.

\(^2\) Although note that Ray (1897:16) describes WT as a 'Papuo-Australian' language; that is, he classed it as Papuan, with considerable Australian features.
hypothesis that WT is a superstrate language, carried by migrants from the Australian mainland, with a “Papuan” substrate. But it is possible that some (or none) of the changes we discuss here have their origins in language contact or language shift. The details of WT’s shared history with Meryam remain to be studied.

The most abundantly documented WT dialects are Kala Lagaw Ya (KLY; of Mabuiag Island; Bani and Klokeid 1971; Ephraim Bani pers. comm.) and Kalaw Kawaw Ya (KKY; of Boigu, Dauan, and Saibai Islands; Kennedy (n.d.), and Ford & Ober (1991)). In this paper we deal mostly with data from these dialects. We consider some of the more striking correspondences between WT and other PN languages and their status as evidence for WT’s genetic affiliation.

2. PN-only features in WT

In order to demonstrate the classification of WT, we present features which are indicative of membership of this family. While we privilege grammatical and morphological features, we do not rely on these exclusively and our claims of morphological cognacy are accompanied by lexical reconstructions.

2.1 Pronouns

One very clear piece of evidence for WT’s membership in PN is the paradigm of the 1st-Sg pronoun (Table 1), with its four partially suppletive case-forms matching those of a number of other PN languages. It is so unlikely that suppletive forms of this type would be borrowed that we are confident in the assumption that they are inherited from pPN.

The nominative forms of most of the WT pronouns (Table 2) continue pPN originals in a fairly straightforward manner. Those that do not have apparent antecedents in the “Papuan” languages (See Evans 2005: 255, Table 1) or in the non-PN languages; rather, they appear to be developments within WT. The discrepancy in the initial consonants (n vs. ng, respectively) in the KLY and KKY
2nd-person pronouns is of interest here. As Table 2 shows, the KLY second-person pronouns all begin with n whereas their KKY counterparts begin with ng. This correspondence can be accounted for as an analogical adjustment by KKY to bring the “speech-act participant” pronouns under a single canonical form.\(^6\) However, the matter is not quite so simple, as there is some reason to believe that, at some pre-WT stage, the 2-Sg pronoun began with *ng.

<table>
<thead>
<tr>
<th>WT (KLY)</th>
<th>Diyari</th>
<th>Arabana-Wangkangurru</th>
<th>Ramindjeri</th>
<th>CE Arrente</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nom</td>
<td>ngay</td>
<td>nganha</td>
<td>ngape</td>
<td>ayenge</td>
</tr>
<tr>
<td>Erg</td>
<td>ngath</td>
<td>ngathu</td>
<td>ngathc</td>
<td>the</td>
</tr>
<tr>
<td>Acc</td>
<td>ngoena</td>
<td>nganha</td>
<td>ngan</td>
<td>ayenge</td>
</tr>
<tr>
<td>Dat</td>
<td>ngayka</td>
<td>ngakarni</td>
<td>anghirda</td>
<td>ayengna</td>
</tr>
</tbody>
</table>

Table 1: First-person singular pronouns in KLY and other PN languages

The 3rd-Pl forms continue pPN *cana (*dhana is a notational variant), and additionally show oblique forms based on thanama-, which is shared by a number of other Pama-Nyungan languages, including those of the Yolngu subgroup (cf. Yan-nhangu dhana ‘they’, dhanama ‘theirs’). The 3rd-Du forms are also widespread, although there is evidence both for *pala and for *pula.

The WT 3rd-Sg pronouns distinguish feminine from non-feminine gender (Bani 1987) with pronouns (see Table 2) that are clearly cognate with forms that evidence this distinction in a number of other PN languages. Alpher (1987:174), Bowern (1998:158–159), and others (probably including Dixon 1980:360) argue that the gender distinction is archaic. In the Karnic subgroup, the masculine singular pronoun is reconstructable to *nhulu in the ergative, with variable (and analogically formed) nominative, and accusative based on *nhi-. The feminine forms are all based on a stem of *nhan-. (See also Alpher 1987 for some comparative data).

We assume that the 2nd-Sg KKY forms beginning with ng- are analogical creations based on the first person, and that the forms in KLY beginning with n (< *ñ) are archaic.

2.2 Case inflection

A summary of WT case endings is given in Table 3. Not all forms are reconstructable, and the system itself is rather unusual in Pama-Nyungan terms: plural number forms part of the same morphological system as case, with all case distinctions neutralised in the plural.

\(^6\) The claim that the WT and pPN originals were *ni (pPN *ñi) and not *ngi is at present a matter of controversy, and the finding that pPN *ki > ci in most of the modern languages (see below) is clearly relevant to the question.

\(^7\) From the dative; cognate with *ngañca in the rest of Karnic. In the table, forms in bold continue archaic material.
Table 2: KLY and KKY pronouns

<table>
<thead>
<tr>
<th></th>
<th><strong>KLY</strong> (B&amp;K)</th>
<th><strong>KKY</strong> (K, F&amp;O)</th>
<th><strong>pPN</strong> (Alpher)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Sg</td>
<td>Nom ngay</td>
<td>ngay</td>
<td>*ngayu</td>
</tr>
<tr>
<td></td>
<td>Erg ngath</td>
<td>ngath</td>
<td>*ngacu</td>
</tr>
<tr>
<td></td>
<td>Acc ngoena</td>
<td>ngoena</td>
<td>*ngaña</td>
</tr>
<tr>
<td>1 ExDu</td>
<td>Nom ngalpa</td>
<td>ngalbe</td>
<td>*ngali</td>
</tr>
<tr>
<td></td>
<td>Poss ngalpun</td>
<td>ngalben</td>
<td></td>
</tr>
<tr>
<td>1 ExPl</td>
<td>Nom ngoey</td>
<td>ngoey</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Poss ngoelmun</td>
<td>ngoelmun</td>
<td></td>
</tr>
<tr>
<td>1 InDu</td>
<td>Nom ngaba</td>
<td>ngoeba</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Poss ngaban</td>
<td>ngoeban</td>
<td></td>
</tr>
<tr>
<td>1 InPl</td>
<td>Nom ngalbay</td>
<td>ngalpa</td>
<td>*ngali (1 InDu)</td>
</tr>
<tr>
<td></td>
<td>Poss ngalbayn</td>
<td>ngalpan</td>
<td></td>
</tr>
<tr>
<td>2 Sg</td>
<td>Nom ni</td>
<td>ngi</td>
<td>*ñun or *ñín</td>
</tr>
<tr>
<td></td>
<td>Erg nidh</td>
<td>nijdh</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Acc nín</td>
<td>nín</td>
<td></td>
</tr>
<tr>
<td>2 Du</td>
<td>Nom nipel</td>
<td>njipel</td>
<td>*ñupal(a)/ñipal(a)</td>
</tr>
<tr>
<td></td>
<td>Poss nippet</td>
<td>njipen</td>
<td></td>
</tr>
<tr>
<td>2 Pl</td>
<td>Nom nitha</td>
<td>njitha</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Poss nithamun</td>
<td>njithamun</td>
<td></td>
</tr>
<tr>
<td>3 non-F Sg</td>
<td>Nom nuy</td>
<td>nuy</td>
<td>*ñu</td>
</tr>
<tr>
<td></td>
<td>Erg nuydh</td>
<td>nuydh</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Acc nuyn</td>
<td>nuyn</td>
<td></td>
</tr>
<tr>
<td></td>
<td>other⁸</td>
<td>nubi- (poss nungu)</td>
<td>nube-</td>
</tr>
<tr>
<td>3 F Sg</td>
<td>Nom na</td>
<td>na</td>
<td>*ña</td>
</tr>
<tr>
<td></td>
<td>Erg nadh</td>
<td>nadh</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Acc nan</td>
<td>nan</td>
<td></td>
</tr>
<tr>
<td></td>
<td>other nabi- (poss nanu)</td>
<td>nabe-</td>
<td></td>
</tr>
<tr>
<td>3 Du</td>
<td>Nom palay</td>
<td>palay</td>
<td>*pula, *pala (?)</td>
</tr>
<tr>
<td></td>
<td>Poss palaymun</td>
<td>palamun</td>
<td></td>
</tr>
<tr>
<td>3 Pl</td>
<td>Nom thana</td>
<td>thana</td>
<td>*cana</td>
</tr>
<tr>
<td></td>
<td>Acc thanamun</td>
<td>thanamun</td>
<td></td>
</tr>
</tbody>
</table>

The WT ablative ending *-ngu (all nouns and pronouns) has cognates of the form *-ngu in a number of PN languages, and the cognates occur over a wide area. Examples include Dyirbal *-ngum (ablative) Yandruwandha *-ngura, Punthamara⁹ -ungu, and Proto-Yolngu *-nguru.¹⁰ The pPN dative ending *-ku continues in WT either in -w genitive (common nouns except A-declension; cf. also the 1st-Sg

⁸ The “other” stem is used for cases other than those mentioned above.
⁹ The initial -u- of the suffix in Punthamara is a relic of the case marker’s status as a former article (Bowern 1998:56–57). In Wangkumara and Punthamara, many cases are reflexes of inflected pronouns which functioned as articles when suffixed to nouns.
¹⁰ Both *-ngu and *-nguru are reconstructed for pPN.
masculine genitive /ngaw/ or in -ka dative (all nouns and pronouns); we are unable at present to justify one of these hypotheses over the other.

<table>
<thead>
<tr>
<th></th>
<th>'foodstuff'</th>
<th>'grandchild'</th>
<th>'woman'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abs</td>
<td>ayza</td>
<td>ngep</td>
<td>yørpkaz</td>
</tr>
<tr>
<td>Erg/Ins</td>
<td>ayzapun</td>
<td>ngépan</td>
<td>yørpekœzin</td>
</tr>
<tr>
<td>Poss</td>
<td>ayzaw</td>
<td>ngépan</td>
<td>yørpekowziw</td>
</tr>
<tr>
<td>Dat</td>
<td>ayzapu</td>
<td>ngépaw</td>
<td>yørpekœzipa</td>
</tr>
<tr>
<td>Abl</td>
<td>ayzangu</td>
<td>ngépangu</td>
<td>yørnekœzingu</td>
</tr>
<tr>
<td>Loc</td>
<td>ayzanu</td>
<td>ngépanu</td>
<td>yørkkœzinu</td>
</tr>
<tr>
<td>Com</td>
<td>ayzaya</td>
<td>ngépiya</td>
<td>yørpekœziya</td>
</tr>
<tr>
<td>Imit</td>
<td>ayzadh</td>
<td>ngépadh</td>
<td>yørpekœzidh</td>
</tr>
<tr>
<td>Pl</td>
<td>ayzapul</td>
<td>ngépal</td>
<td>yørpekœzil</td>
</tr>
</tbody>
</table>

Table 3: Select Western Torres case paradigms (after F&O 1991:136)

Relics of other pPN morphology include the -l in /'.kubil/ (plural /'.kubilal/) ‘nighttime’ (added to a base *kubi, as recorded in /'.kubikub/ ‘black’), which continues the pPN locative ending *-la. The locative is commonly used in PN languages to derive time adverbials; compare Bidyara (Breen 1972: 61–2) /'.gunda/ ‘dark, nighttime’, /'.gundan nga/ ‘in the nighttime, at night’ and the more locational /'.bala/ ‘leg’, /'.bala nga/ ‘on the leg’. PN *-la (and the corresponding ergative ending *-lu) is otherwise apparently lost in WT. This in itself is not very surprising, since elsewhere in the family we have evidence for the stranding of allomorphs precisely in contexts such as this. Compare, for example, Yolngu /'.yalala/ ‘later’ (continuing the pPN demonstrative *yala ‘that’) with the more usual Yolngu locatives in *-nga.

The ergative –dh, occurring on pronouns, very probably continues the *-ñcu ergative alternant that is reconstructable for pPN (and to some earlier more inclusive grouping as well).

2.3 Verbal tense-aspect-mood

WT verbs are inflected in two conjugations, termed N and Ø (B&K for KLY) or transitive and intransitive (F&O for KKY), respectively. Tense-endings with affinities to the rest of PN (see Alpher 1990) include:

(1) -ngul Yesterday Past (-dhinngul with Dual subject, -ayngul with Ø Conj

Singular subject): .ngula/ is widely attested as a free-standing adverbial in senses like ‘and then’.13

---

11 Ayza is probably a compound of ay ‘food’ (< *mayi) and za ‘thing’.
12 Possibly cognate with Yir-Yoront kolpkol+h ‘black’, although we know of no other cognate sets showing a correspondence of WT b with lp in another PN language.
13 It is unknown as a verb-inflactional ending in mainland PN, but note that WT distinguishes other tenses, like ‘Remote Past’ and ‘Last Night’, that are not in the inventory of the typical mainland PN language.
-nu Near Past (N Conj): cf. WT bananu ‘speared, pierced’ and Warlpiri pakanu ‘hit (Past)’; note that the -u might have been independently added in these two languages (Alpher 2004:100 fn 8). Western Desert also has -rnu as a past punctual marker in the L conjugation.

-ma Near Past (Ø Conj): as in mulima ‘talked’; with this compare Nonpast forms in -m(V) such as Warlpiri nyinami and Koko-Bera nyinem ~ nyinvm.

-n Present (with N Conj and with Ø Conj Nonsingular; in KKY Present Perfective with Transitive Conj Singular and all Nonsingulars): as in banan ‘spears’. 14

-k Future (-ayk with N Conj non-Dual; -ka elsewhere; not in KKY): as in mulika ‘will talk’; with this compare Purposive forms in -ku such as Watjarri nyinaku (~ nyinawu) ‘will sit’.

2.4 Voice marking of verbs

There is further evidence for the genetic affiliation of WT in voice marking. WT transitive verbs typically have an intransitive counterpart (belonging to the intransitive conjugation). An example is given in Table 4 below. The latter is used in an antipassive construction with an intensive sense, as in ‘drank all the water’ (B&K, Comrie (1981)). Uninflected verbs are typically (but not invariably) disyllabic, and in a great many cases the second vowel, which we term the THEMATIC VOWEL, is -a- for transitive verbs and -i- for intransitive verbs. 15 So for example in KKY there are pairs of verbs as displayed in Table 4.

<table>
<thead>
<tr>
<th>Transitive</th>
<th>Intransitive (antipassive; intensive)</th>
</tr>
</thead>
<tbody>
<tr>
<td>patha- ‘chop’</td>
<td>patha/-i ‘to chop’</td>
</tr>
<tr>
<td>paga- ‘spear, pierce [etc]’</td>
<td>paga- ‘spear, pierce [etc]’</td>
</tr>
</tbody>
</table>

Table 4: Examples KKY transitive/intransitive verb pairs

The correspondence of this distinction in form and function to distinctions made in a number of other PN languages is striking. 16 For example, the regular valency-decreasing (typically, reflexive) marking in Kuuk-Thaayorre (western Cape York Peninsula; Gaby 2006) is replacement of the thematic vowel with -e.

---

14 This tense-form can apparently have an Immediate Past sense: KKY Nadh sanan koewsa pathan ‘she squashed a fruit with her foot’ (Kennedy n.d.: 41). For the non-past sense of this form compare Nonpast tense forms in -m in Warumungu (pakkanta ‘spears’; Heath & Simpson 1982) and Western Desert (patjarni ‘bites’; KKY pathan ‘chops’).

15 Kennedy transcribes the latter as -e, and not -i.

16 Note that the transitivity values for the -a- and -i- thematic vowels described below are the opposite of those discussed briefly in Alpher (2004: 99–100); both sets of values are probably ancient in PN. It is imaginable that both of these originate in a single suffix *-i that was a transitivising form. If the association of intransitivisation and of -i- (not necessarily intransitivising) with intensiveness proves to be other than a happenstance one, then perhaps Arabana-Wangkanguru (with a number of -a- intransitives with transitive counterparts in -i-, such as thika- ‘return’, thik- ‘take back’) can be taken as continuing a prototypical arrangement.
(Table 5) to form a verb of the second (typically intransitive) conjugation. 17
(Verbs are cited as Nonpast, Past.)

<table>
<thead>
<tr>
<th>Active</th>
<th>Reflexive</th>
</tr>
</thead>
<tbody>
<tr>
<td>katpr, katparr ‘grasp something’</td>
<td>katpe, katpe(y)r ‘grasp oneself’</td>
</tr>
<tr>
<td>pathr, patharr ‘bite something’</td>
<td>pathe, pathe(y)r ‘bite oneself’</td>
</tr>
</tbody>
</table>

Table 5: Kuuk-Thaayorre reflexive marking

The Thura-Yura (South Australia) languages Adnyamathanha and Wirangu show a similar process, apparently moderately productive. For Adnyamathanha, Schebeck (1974: 16–7) describes an intransitivising suffix -i-, as in annga- ‘to lift up’, annga-i- ‘to get up, arise’. 18 McEntee and McKenzie (1992: 6), who represent this root as arrnga-, cite alternative pronunciations for the intransitivised form: arrnngi-i- and arrnngi. 19 In Wirangu (Hercus 1999: 97–8, 101–2, 109–11) a change of thematic vowel from -a- to -i- occurs with both transitive and intransitive verbs and indicates continuous or repeated action: warna- ‘drop something [vt], fall [vi]’, warni- ‘throw to ground [vt], come down (rain) [vi]’; barlda- ‘stab, spear, split’, barldi- ‘punch repeatedly, chop up’. 20 Like this in sense is the Adnyamathanha pair (McEntee & McKenzie 1992: 1) aka+rrhi- ‘to burp’, aka+rrhi-i- ‘to repeat (meal), to suffer indigestion’. In the neighbouring language Arabana-Wangkangurru (Hercus 1989 and 1994:146-147), a few transitive a-stems have (transitive) counterparts in –i indicating an action that is more

---

17 Since this -e can replace the thematic -i of verbs that have it in their active forms, since the historical origin of the /i/e/a distinction in Kuuk-Thaayorre is not yet clear, and since the past tense-form of the reflexive is often recorded as -eyr (as well as -er), it is not clear how -e is to be accounted for etymologically.

18 This -i- is held to be in an allomorphic relationship (under conditioning as yet undetermined) with -rri-; a suffix possibly identical to this latter is represented as -r(i)- (Schebeck 1974: 17, 23). One example in McEntee and McKenzie suggests that -i- and -ri- are not quite synonymous: nharla- ‘to pour (liquid), uproot, push over’, nharla-i- ‘to flow (as running water)’, nharla-ri- ‘to pour down, stream down (as flood)’. Note that for Adnyamathanha and other languages in which a tap contrasts with a trill, we write (standardising, following Austin 1990: 179) r for a glide, rr for a tap, and rrh for a trill.

19 However, they cite no such truncated alternants for other intransitivised examples in their dictionary, such as /una-/una-i- ‘to come undone’ (una- ‘to pull down, dismantle, take apart’), irnba- ‘to get together, join up (meet up), mob up (people)’ (irnba- ‘to add, join’). McEntee (pers. comm.) says that sequences written “a-i” and “i-i” are pronounced with a weak high front glide between the vowels, respectively [a’i] and [i’i]; Schebeck does not comment on this point. In Adnyamathanha, i-i and u-i are also written. In synchronic studies, a number of scholars have postulated underlying sequences like a-i in which the second vowel replaces the first in pronunciation. An example is Heath’s (1980) account of Ritharrngu, in which he analyses the paradigm of a verb like wanga- ‘speak’ as follows: Pres wanga <wanga-Ø>, F wanganha, F wangi <wanga-i>. (Whether a tense-marking process in some languages has a common origin with a voice-marking process in others is a question that remains for future work.)

20 This -i-, in Hercus’s analysis (109–11), occurs also preceding the reciprocal suffix -rri-, since “the reciprocal implies that an action is double at the least”(109); -rri- also marks the reflexive and (101–2) derives stative and inceptive verbs from adjectives. Simpson and Hercus (2004: 205) reconstruct a proto-Thura-Yura reflexive-reciprocal *-rrhi-.

In Ngawun (a Mayi language of Queensland; Breen 1981a: 55), the suffix –ila (which we suggest is –la added to an –i– stem) makes a verb reflexive (puñça ‘to kill’, puñcila ‘hit oneself’; thaya ‘chop’, thayila ‘chop oneself’). There are also verb/noun pairs in a/i (kaca ‘tell a lie’, kaci ‘a lie’; maya ‘to speak’, mayi ‘speech’).21 Our suggestion that this verb stem alternation in Ngawun attests the same phenomenon as the examples discussed above is corroborated by a cognate verb in Yidiny (bunja and bunji, which see below; with this compare Ngawun puñça and puñci).

Furthermore, several languages attest apparently one-off transitive-intransitive pairs derived by thematic-vowel replacement. In Kugu-Uwanh (spoken just to the north of Kuuk-Thaayorre; Smith and Johnson 1999), there is wañja ‘to hang something up’, wañji ‘to be hanging’ (no conjugational difference); likewise in Yidiny (Dixon 1991: 255–6): bunja– ‘hit to inflict injury; kill’ (transitive, but in the predominantly intransitive Conj N; all dialects), bunji– ‘bump into, collide with; burst, explode’ (intransitive, but in the predominantly transitive Conj I; Tablelands and Coastal dialects).22 Finally, in at least one language, Diyari, an intransitive verb unmatched by a transitive one, paki– ‘to burst’, corresponds in this way to the widely attested *paka- ‘(to dig, pierce)

We find these examples sufficient to justify a conclusion that the transitive vs. intransitive conjugation thematic vowel alternation in WT is archaic, and from a pPN source.23 This alternation in WT is yet another strong argument for its inclusion in PN, since morphological alternation of this type is rarely borrowed. Note especially the parallel reflexes of pPN *paka- ‘dig, pierce’ and *punca ‘hit (to kill)’ listed above.

It is tempting to account for the replacement of thematic a by i as the product of a conflation of a with *-yi, attested as a reflexivising (or otherwise valency-decreasing) suffix in various Australian languages including a few PN ones. Note in this connection however that actual synchronic or comparative evidence for such a conflation in the relevant languages is almost nonexistent, and that efforts like those of Dixon (2002:321-2; esp. Table 7.6) to relate forms like -yi and -ci to a prototypical *dharri are apparently posited as one-off changes. And note especially the fact from Yidiny (see fn 22 above) that changes to thematic -i and suffixial -yi (here clearly from *-ci) have contrastive value.

---

21 We thank Mary Laughren for pointing out these examples to us.
22 Note that Yidiny regularly intransitivises verbs with -ji- (antipassive, reflexive, etc.; Dixon 1977: 217–9), as in bahad ‘speared’, bahagiñu ‘speared self’ (etc.), in a suffixation process that seems to be distinct from the thematic-vowel alternation illustrated above. Cognate with Yidiny -ji- is (via lenition) Djabugay -yi- (Hale 1976: 238 & b: 324-6; Patz 1991); the protoform *-ci- is not a probable source of the WT antipassive thematic -i-.
23 No mainland language that attests this alternation is geographically close to WT.
2.5 Number marking in verbs

WT exhibits an agreement system in which stems inflect for the number of the absolutive argument. The marking is an increment to the stem of the form -ma- in the dual and -mary- in the plural in KKW. KLY has -ma- and -uma- respectively. An example of the dual is given below:

(2) Ngath puy patha-ma-n. (cf. singular pathan)
   I-ERG tree chop.down-DL-PF.
   “I chopped down two trees.” (Comrie 1981:17)

There are parallels for this too in other PN languages. For example, in Marrgany (Breen 1981b: 319-320) the verb stem increment -ma- marks plurality of objects (and causative on intransitive verbs), and in Parnkala (South Australia; Keith 1998:31 citing Schürmann 1844:17-20) -ma- marks dual subjects.

It is unlikely that the same function would be independently innovated with the same morphology.

3. Phonological history

In this section we discuss the basis for phonological reconstructions given above, give lexical cognate data to support reconstructions, and conclude with discussion of problematic (multiple) correspondences.

3.1 Preliminaries: inventories of phonemes

Table 6 gives the WT phoneme inventory. It is somewhat unusual for a Pama-Nyungan language in the absence of retroflexion, in the presence of a voicing contrast in stops, and in the fricatives /s/ and /z/.

<table>
<thead>
<tr>
<th>p</th>
<th>t</th>
<th>s</th>
<th>k</th>
<th>b</th>
<th>d</th>
<th>z</th>
<th>g</th>
<th>m</th>
<th>n</th>
<th>ng</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t</td>
<td></td>
<td></td>
<td>d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>l</td>
</tr>
<tr>
<td>u</td>
<td>i</td>
<td></td>
<td></td>
<td></td>
<td>e</td>
<td>æ</td>
<td>o</td>
<td>a</td>
<td>w</td>
<td>r</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>y</td>
</tr>
</tbody>
</table>

Table 6: WT Phonemes

In KLY there is a vowel-length distinction (B&K 1972; long vowels are written double); in KKY there is apparently no such distinction, although Kennedy’s description (n.d.: 14) is equivocal in this regard. Kennedy (n.d.:109) claims that length is not contrastive, while in another place (p14) presenting pairs of forms in which vowels apparently contrast as tense and lax.

We thank Mary Laughren for suggesting we look into this as a possibly archaic feature of WT. Keith (1999) surveyed verbal number marking; here we focus specifically on putative cognates to WT forms only.
3.2 Phonological developments in WT

While it is clear that there are many lexical stems in WT which are ‘cognate’ in some way with forms in other PN languages, much of the phonological development of WT remains unknown. There has been considerable language contact in the region, and it is not always possible to decide whether multiple correspondences are retentions, innovations, or the result of multiple loans. Here we use the term ‘descent’ with caution, since we recognise that many forms which are reconstructable to pPN may not descend directly into WT, and the forms may therefore not be strictly cognate. Because of this, the lexical evidence for genetic relationship is less convincing than the morphological data.

3.2.1 Initial dropping. One prominent sound change in many languages of the Cape York Peninsula (south of WT) is the loss of initial consonants (see e.g. Alpher 1972). We have some evidence for this change in WT too. Initial k is regularly lost if the initial vowel of the word was long; cf. aga-, age- ‘carry’ < *ka(:)ngka-, aba- ‘cover’ < *ka:mpa-, and ima- ‘see’, perhaps < *ki:mV- (for the distribution of these pPN terms and evidence for their reconstruction, see Alpher 2004). Initial *yi also exhibits initial dropping (although variation between yi and i in closely related languages is common in Australia). Examples include ika ‘laughter’ < *yika and ipika, ipkazil (KLY), ipkaz, ipkæzil (KKY) ‘woman, female’ < *yipī. Another potential example is ay ‘food’ < *mayi, however note that initial nasals do not usually drop; cf *ñ, which descends as n, and *m in muma- ‘squeeze, hug’ (KKY) < *muma- and mugu ‘anthing’ (KLY, KKY) < *mungku).

Initial dropping is also found in many (although not all) WT kinship terms. Examples with initial dropping include athe ‘grandfather’ < *ngaci, aka ‘grandmother’ < *paka, ama ‘mother’ [syn. apu] < *ngama, alay ‘husband’ < *ka:la², ipi ‘wife’ *yipi (covered by the change mentioned above), imi, imil ‘brother/sister in law’. All of the WT kinship terms that are clearly pPN in age have lost their initial consonant; compare Blevins’ (2001:485–6) argument that the propensity to be used utterance-initially is often a conditioning factor for initial-dropping. She mentions kinship terms as having this propensity and cites (486) the Arabana pair ama ‘mother’, ngama ‘milk’, both continuing pPN *ngama.²⁶

3.2.2 The voicing contrast. Since pPN is not reconstructed with a voicing contrast, we must account for the presence of such a contrast in WT as an innovation (or

²⁵ For the final -y of this word, cf. McConvell (this volume) and Yolngu gali/galay.
²⁶ Initial-dropping does not appear in thathi ‘Fa’, bab(a) ‘dad’, babath, babthal ‘opposite sex sibling’, babbeth ‘grandfather’, thawi & thawian ‘brother in law’. It is possible that some of these terms are loans from a neighbouring initial-dropping Pama-Nyungan language; for example, Urudhi has athi ‘mother’s father’ (cf. WT athe) and amii ‘mother’ (cf. ama).
seriously revise reconstructions). There is no evidence from elsewhere in PN that voicing is old; in Yolngu languages, for example, a phonemic voicing can be shown to have entered the language primarily through extensive borrowing of lexical items; Yolngu etyma with Pama-Nyungan etymologies always appear with voiceless stops; the voiced stops can be shown to appear either in loans or through secondary morphological processes (Alpher & Bowern 2005). No such solution is available at this stage for WT, however. There are reconstructable etyma where the reflexes of the initial consonant are voiced, while in others, the initial is voiceless:

(3) a. Voiced: badh ‘sore’ < *paci
   Voiceless: patha-/e- ‘chop (etc)’ (KKY) < *paca-; pirupirru ‘rainbow bird’ (KLY, KKY) < *prippirru; palay ‘they Du’ < *pula or *pala
b. Voiced: dhang ‘tooth’ < *canga, dhuugu ‘driftwood; raft’ ??< *cuku
   Voiceless: thala-/e- (vt/vi) ‘eat’ *cala-; thœra, thœral ‘ridge; reef; shin; back [usu. of fish]’ (< *carra ?); thœray ~ thara- ~ thare- ‘appear; place, be placed; stand; ...’ (< *carra-); thanur- ‘stand’ < *cana-; thana ‘they’ (< *cana).
c. Voiced: gam, gamul ‘colour, hue’ (& gamu, gamul ‘body’) < *kamu; guyar ‘stingray’ (< *kuya); garka ‘man’ < *karrka; gœyar ‘sun’ < *kaya; gururu ‘peaceful dove’ < *kuru+
   Voiceless: kun, kunal ‘hind, stern, back’ < *kuna; kosar ‘two’ < *kuc’arra; kuma ‘shit’ < *kuma; klak, kælak ‘spear’ < * kalaka; kisay, kisayil ‘moon, month’ (KLY, KKY) <(loan?) *kica (contrast sib ‘liver’ < *cipa)

We see the same problems in correspondences for medial consonants that are found initially. (See below for consonant clusters.) Pama-Nyungan *p continues as p in ipi(ka) ‘woman, wife’ (KLY, KKY) < *yipi; however it appears as b in sib ‘liver’ < *cipa and ibay ‘grate, scrape’ < *kipa-. *c descends as th in patha-/e- ‘chop (etc)’ (KKY) < *paca- and gath, gathal ‘shallows, reef’ < *kaca+, but as dh in badh ‘sore’ < *paci and idhan ‘bite’ < *cica-. It has the further reflex s in kisay, kisayil ‘moon’ < *kica and kosar (~ ukasar) ‘two’ < *kuc’arra.

3.2.3 Delaminalisation. Proto-Pama-Nyungan *n̥ is regularly reflected in WT as n: naka ‘situated here’ < *ñaaka, nagi-ka (KLY) ‘look at’ < *ñaaka, and ngaena < *ngañaì.

3.2.4 Nasals. The nasals have their expected reflexes: *m descends as m, *n and *n̥ as n,27 and *ng as ng. The two rhotics fall together; compare thara-/e- ‘stand’ (vt/vi) (KKY) < *carra- with the near-minimal pair dhœrang(+alay) ‘tooth’ (KKY; syn. dhang) < *carra.

---

27 An exception is miya ‘who’, assuming it is a reflex of *miña ‘what, meat’.
3.2.5 Clusters. Homorganic nasal-stop clusters simplify, and the resulting stop is voiced:

\[(4)\]
aba- ‘cover’ < *ka:m-; mugu ‘ant-bed’ < *mungku; nag ‘how’ < *ñangkV; naguy ‘down there’ < *ñangku; wadhawadhagumiya ‘everywhere, all around’ < *wañca

Heterorganic clusters involving *rr are retained; compare ngursi ‘snot’ < *ngurrci and dhardhi ‘tie’ (KLY,KKY) < pNE *currca-am, as well as garrpath < *karrpa-. Evidence is too slight at this stage for other clusters, but consider kunakana ‘strong’, possibly from *kunka ‘alive’.

3.3 Changes in vowels
Although KLY has a vowel length distinction, it does not appear to continue a length distinction inherited from pPN. That is, the distinction has been reinnovated. The individual pPN vowels continue as follows.

\[(5)\]
*a > a in thana ‘they (Pl)’ < *cana; thanur- ‘sit, live, stay’ < *cana; gath, gathal ‘shallows, reef’ < *kaca; kakur, kakural ‘egg, testicle’ < *kaku; gam, gamul ‘colour, hue, body’ < *kamu; garrka(zi), garkazil ‘man, male’ < *karrkay; garrpath- ‘gather, collect’ < *karrpa-; mari ‘spirit (of living person), ghost’ < *mari; matha ‘one’ < *maca; ay ‘food’ < *mayi; ath ‘grandfather’ < *ngaci < *ngaki; ngath ‘I’ < *ngacu/a; ama ‘breast’ < *ngama; ngamal ‘cumulus cloud’ < *ngamal; ngap ‘grandchild’ < *ngapa; ngaru ‘must, have to’ < *ngarru; ngay ‘I’ < *ngayu; nag ‘how’ < *ña(+ngkV); naka ‘situated here’ < *ña; nag-ka (KLY) ‘look at’ < *ña; naguy ‘down there (F)’ < *ñangku; badh ‘sore’ < *paci; palay ‘they (Du)’ < *pala (< *pula); thala-/e-/vt ‘eat’ < *cala; wadhawadhagumiya ‘everywhere’ < *wañca; wardh ‘different, questionable’ < *warrca; wati ‘bad’ < *wati; waw (Q particle) < *wawu

There are also reflexes of *a in œ, in fewer items. WT œ has several sources, since it is apparently partly conditioned both by vowel harmony and also grammatically. Finally, *a appears to continue as e in gethaw(a) ‘small digging stick’ (if < *kacin), as well as in *ñipala > *nupal > nipel.

---

28 For example, verbs in KKY have participles in œ but tensed forms with a, which implies at least in part a grammatical distribution to the vowel in the modern language, whatever its historical source (see also the following footnote).
THE WESTERN TORRES STRAIT LANGUAGE

(6) *a: > ø in dhœrang(-alay) 'tooth' (KKY; syn. dhang) < *ca:ra
*a: ø in ngaña 'me' < *ngaña; thœra 'ridge, reef, shin' < ? *carra 'thigh';
gœnga 'sun' (KLY, KKY) < *kayka; waydha/-e- (vt/vi) 'put' (KKY) < (?)
wœydha- (although if so y in WT is unexplained)29

Long vowels are reflected as short (most examples involve *a:), except in dhang(a) 'tooth, edge' (KLY, KKY) < *canga; however the principles which govern the secondary length innovation are not well understood. Examples are given below.

(7) *u: > u in kuku, kukul 'bang, noise, thump' < *ku:ku
*a: > a in alay 'husband' < *ka:la+y, aba- 'cover, hide something' < *ka:mpa-;
lawnga 'not' < *lawa; ma-y 'take, give, bring' < *ma-; nga+ 'who(m)'
< *nga:ni; wati 'bad' < *wati

The vowels *i and *u descend unchanged as i and u (except kœrkak, kœrkakal 'throat' < *kurkka). Examples are given below.

(8) *i > i: idhan 'bite' < *cica--; siti 'there' < *ciñcu; sib, sibal 'liver' < *cipa;
kisay(il) 'moon' < *kica; ibay 'grate, scrape' < *kipa--; iwi 'mosquito' < *kiwi;
mina 'good, true' < *miña; miyay 'what, which' < *miña; pirpuru 'rainbow
bird' < *pirrupirru; ipi(ka) 'woman, wife' < *yipi
*u > u in kuki(y) 'northwest wind' < *kukiya; kuma 'faeces' < *kuma; kun(a),
kunal 'rear end, in reverse direction' < *kuna; kunakana 'strong, hard,
powerful' < *kunka; kubi(kub)(+) 'charcoal, black, night' < *kula;
gururu, gururul 'peaceful dove' < *kuru--; guyar, guy(ar)a'l 'stingray' < *kuya;
muma- 'squeeze, hug' < *muma--; muçu 'anthill' < *mungku; ngu:ki water' <
*nguku;i; ngul 'yesterday' < *ngula; ngursi 'snot' < *ngurrci; ngu:projection, point' < *ngurru; yu: 'lie down' < *yu

Finally, two cognates where *u apparently descends as o are thonar 'time,
occasion' < *cunu, and kosar (~ ukasar) 'two' < *kuc(+)arra.30

The above examples all involved initial correspondences. There is limited
evidence for correspondences in V2. The evidence is partly limited because V2 is

29 A number of forms have alternations between a and ø. Examples include patha/-e- 'chop' ~
patha/-e- and nagi/e- ~ nagay 'look at' (KKY) < *ña:ka. Ultimately, these will prove to be of
importance in the solution of the problem of the origins of ø.

30 We have been unable to determine whether the variation in this word is dialectal, speaker-
based, or simply a data error. Obviously the source of this variation affects the certainty with
which we claim this as an example of a possible instance of *u > o.
lost in many disyllabic WT words. In (9) we list the reconstructions which provide attestation of descent of V₂, while (10) lists those that suggest V₂ loss.

(9) *a > a in thana 'they' < *cana, alay 'husband' < *calaa, gæygæa < *kayka 'sun', kisay 'moon' < *kica, kuma 'excrement' < *kuma, kuna (KLY) 'back' < *kuna, kærkak 'throat' < *kurrka, guyar 'stingray' < *kuya 'fish', miyay 'what' < *miña, ama 'mother' < *ngama, naka 'situated here' < *naka
*i > i in iwi 'mosquito' < *kiwi, mari 'spirit' < *mari, ngu(leri)ki 'water' < *nguki, ngursi 'snot' < *ngurcci, wati 'bad' < *wati, ipika 'woman' < *yipi
*u > u in kuku 'noise' < *ku:ku 'speech', gururu 'peaceful dove' < *kurlu+ 'dove', mu(eri)gu 'anthill' < *mungku, ngaru 'must' < *ngarru, pirupiru < 'rainbow bird' < *pirrupirru

(10) kun (KKY) 'back' < *kuna, gath 'reef' < *kaca+ 'coral', sib 'liver' < *cipa (< *kipa), ay 'food' < *mayi, ngath '1.ERG' < *ngacu, ngay '1.NOM' < *ngayu, ngan 'who' < *ngani/a, ngul 'yesterday past' < *ngula 'then', ngu(eri)J 'point' < *ngurrur 'nose', badh 'sore' < *paci, kam 'colour' < *kamu

4. Conclusions

Over the last hundred years, the literature has suggested several classifications of the WT language: 'Papuan', 'Australian', 'Australo-Papuan' and 'Papuo-Australian'. Each of these classifications implies a different history for the language speakers, and each is in principle demonstrable or falsifiable through detailed linguistic comparison. To make claims of mixed language descent and substrate influence without the detailed work required, however, trivialises the notion of comparative linguistics.

WT clearly shows reflexes of pPN lexemes, pronouns, case marking and verb morphology. Reconstruction work is obviously at an early stage, but the number of correspondences (as well as shared irregularities) unambiguously point to a genetic relationship with PN languages. Cognate lexical items, nominal inflection, verb morphology and pronouns all indicate that WT is PN, albeit at considerable time depth in a complex contact area. The alternative – that WT is a Papuan language with PN substrate elements – would require us to assume the borrowing of irregular morphology in multiple distinct areas of grammar. As regards the 'Papuan'-substrate hypothesis, we wish to put into consideration the possibility that the case for substrate influence has been overstated.
References


Ray, Sidney and Haddon, A.C. 1897. A study of the languages of the Torres Straits. Dublin: Dublin University Press.


Schürmann, C.W. 1844. A vocabulary of the Parnkalla language spoken by the natives inhabiting the western shores of Spencer’s Gulf. Adelaide: George Dehane.


THE CLASSIFICATION OF PINIKURA, WESTERN AUSTRALIA

PETER K. AUSTIN
School of Oriental and African Studies

1. Introduction

Among the languages traditionally spoken in the region between the Gascoyne and Ashburton Rivers in the north-west of Western Australia, Pinikura is something of an enigma. Although we have data on most of the languages, and have been able to classify them as members of three genetic subgroups, namely Kanyara, Mantharta and Ngayarta (Austin 1981, 1988, 1989, 1996, Dench 1995, O’Grady et al 1966, Thieberger 1993), materials on Pinikura have been rather harder to come by. This has not prevented various scholars referring to the language and classifying it into one or the other genetic group. In this paper I review some of the discussions about the language and take into account some new data that became available as a result of fieldwork in the region.

2. Names and Locations


The first explicit published location for Pinikura is given by Brown (1912:144) as “on the north side of the Ashburton River about Duck Creek”. Following Tindale 1940, Capell 1963 lists a rather more elaborated location for what he calls language “W7 Binigura” as follows:

---

1 My fieldwork on Kanyara and Mantharta languages has been supported at various times by University of Western Australia Department of Anthropology, AIATSIS, La Trobe University School of Humanities, and the Australian Research Council. I am grateful to Alan Dench, G.N. O’Grady and T.J. Klokeid for sharing data and ideas with me. My greatest debt is to the many native speakers who assisted my research, especially Jack Butler, Dolly Butler, Hamish Cameron, Maudie Dowton, Annie Eagles and Helen Hayes. I am delighted to be able to offer this paper in honour of Harold Koch, who has been an inspiration in the application of proper historical linguistics methods to the study of Australian Aboriginal languages over the past 30 years.

2 On the two maps in the back of Oates and Oates 1970, for example, Pinikura appears as a blank for “grammar/phonology” information and “vocabulary”.

Location: Duck Creek; south to Ashburton River; north east to Hamersley Range; headwaters of Robe and Cane Rivers. Information: No work appears to have been done as yet.

Oates and Oates (1970:55) give the following details about “Binigura”, which they also number W7 (AIATSIS number W34) and classify as a Ngayarda language, following O’Grady, Voegelin and Voegelin 1966:

Ashburton, Cane and Robe Rivers area, WA. There is still no information on this language, except that Brandenstein has made some brief recordings.

Oates and Oates (1970:65) list a separate language “Pinikurra”, which they number 12Wr and which they place in a somewhat different location:

Nanutarra, Oater Junction, Wyloo area, WA. Is the inland neighbour of Noala, 50.8 and Parduna, 51.1b; a Pilbara dialect being studied by Brandenstein with the help of an informant in Onslow. It is grammatically more akin to Ngarluma, 50.5b than to its neighbours. See Binigura, 50.7a.

Oates (1973:71) gives the information on “Binigura” as

Ashburton River between Wyloo and Nanutarra Stations and north of Cane and Robe Rivers (from von B’s map). Derived from pini ‘to go’ and gura ‘having’ (commonly used in geographical names). Von Brandenstein has surveyed and found the language to be closer grammatically to Ngarluma than to neighbouring languages. Some key words are different to neighbouring languages as pini-/punī- ‘to go’ and -daa instead of -na for accusative.

Tindale (1974) gives the following location, which is south of the areas indicated in the other sources:

On Ashburton River between Mount Price and Kooline; south to Wannery Creek; north to near Mount Amy, Urandy, and the beginning of the uplands of Duck Creek; east to lower headwaters of Hardye River on western boundary of Ashburton Downs. Kulanji Pool, 10 miles (15km.) upstream from Kooline, was an important ceremonial place. Another important water was Minbun, the pool at Glen Ferrie. 116°15’E x 22°40’S

Thieberger (1993:131) reports:

Location: Duck Creek, south to Ashburton River, north-east to Hamersley Range, headwaters of the Robe and Cane Rivers; Austin (1983) situates the language south of the Ashburton River.

According to speakers of neighbouring languages who I interviewed between 1978 and 1994, the country owned by Pinikura speaking groups took in the upper Ashburton from the Henry River junction upstream to about Hardye junction and Mount McGrath. In my database of place names of the Gascoyne Ashburton region there are 23 identifiable names associated with the Pinikura language owning group. These are listed in Appendix 1.

There is one further source of data on place names and locations that contains information on Pinikura. In the papers of A.R. Radcliffe-Brown at the
University of Sydney Fisher Library archive there is a card file listing the names
of local groups from the Gascoyne-Ashburton region. The material in this card
file was probably collected by Radcliffe-Brown during his expedition to the lock
hospital\(^3\) on Bernier Island west of Carnarvon in 1911, or in his later excursions
inland. The materials have never been published, but they include data on 19
groups identified as Pinikura, including one sketch map of the region near the
junction of the Ashburton and Hardey Rivers that shows the location of a
number of named pools. Radcliffe-Brown’s data is presented in Appendix 2
below. This material confirms the general upper Ashburton location suggested
by my consultants and the sources quoted above. Note that Radcliffe-Brown’s
cards include a couple of words for totems in language however these are
insufficient to identify which language group Pinikura might belong to.\(^4\)

3. Language materials

Although we have a fair idea of the location of the country associated with
Pinikura speaking local groups, there is almost nothing available about the
language. C.G. von Brandenstein collected a small amount of data in the 1960’s,
but there is little in his data that distinguishes it from the Thalanyji spoken by
his consultant as his first language, apart from the use of the verb \(\text{pini-}\) meaning
‘to go’ (the corresponding verb root in the Kanyara languages is \(\text{puni-}\) and
Mantharta languages \(\text{yana-}\) (with suppletive \(\text{puna}\) in the present tense)). The
name Pinikura is also unusual for this region in being potentially analysable as
\(\text{pini-ku-ra} ‘\text{to go-PRESENT-INHABITANT’}\).

Oates (1973:71) reports a reclassification of the Pilbara languages by von
Brandenstein in which Pinikura is placed in the “Coastal Ngayarda Subgroup”
along with “Nhuwala, Djiwarli, Gurama, Jindjibarni, Ngarluma, Gariera and
Ngarla”. She says that it is an “A(CTive) V(ERBAL) C(ONCEPT) language”, meaning
that it has nominative-accusative case marking morphology, rather than a split-
ergative system. Unfortunately, she presents no data to support this analysis
and neither has von Brandenstein published anything to support this

\(^3\) Two hospitals, one on Bernier Island for men and one on Dorre Island for women,
were established in 1908 ostensibly for sufferers of syphilis who were rounded up by a
policeman, Constable Grey, throughout the Gascoyne-Ashburton region and
transported there in neck chains via Carnarvon. They included speakers of a wide range
of languages (see Grant Watson 1946, Bates 1985).

\(^4\) There is a native title claim group Puutu Kurnti Kurrama Pinikura (PKKP) that
includes Pinikura descendants – for further information see http://www.yamatji.
org.au/about/claimlist.htm. The location of the area which this group claims can be
found on the map http://www.yamatji.org.au/about/Pilbara_NTCA_Schedule.pdf
reclassification. Wordick 1982 also mentions Pinikura as a member of the Ngayarta subgroup, however he too presents no data or discussion.5

I have been able to collect a little material from two people who identified themselves as speakers of Pinikura. I have 10 pages of notes collected from Molly Butler in Onslow on 23rd November 1978 (Field Notebook 5), and a further 9 pages of notes collected from Maudie Dowton at Onslow on 3rd November 1994 (Field Notebook 16). Molly Butler was the daughter of Topsy Ashburton, who was identified by everyone I met as a Pinikura — I attempted to meet and talk with Topsy Ashburton at Nanutarra Station in 1978, however I was not successful in doing so due to her shyness. Molly Butler had been married to my main Jiwarli consultant, Jack Butler, and identified herself as Pinikura, though she mostly spoke Thalanyji. I was able to interview her once only in Onslow in November 1978. Maudie Dowton was introduced to me in November 1994 by Alan Dench as a speaker of Kurrama, however in the course of that meeting she pointed out that her mother was Pinikura and that she spoke that language also. I was able to work with her on just one occasion and collected some basic grammatical and lexical information. Follow up work was done by Penny Lee, then employed by Wangka Maya Aboriginal Language Research Centre in Onslow – I have not seen any of the materials that she collected. The Wangka Maya website (www.wangkamaya.org) mentions Dowton and Mackay 1994, which I have also not seen.

The materials collected from Molly Butler and Maudie Dowton are relatively similar and both sets of data show many striking similarities with Thalanyji, including forms of pronouns, case inflections (clearly of a split-ergative type), verb inflections, and vocabulary. On the basis of this (admittedly rather limited) data Pinikura should probably be classified into the Kanyara subgroup along with Payungu, Purduna and Thalanyji.6

Interestingly, the materials provided by both speakers diverge from Thalanyji identically in a number of important ways, and are suggestive of a borrowing relationship with the neighbouring Mantharta languages. The verb ‘to go’ is given by both speakers as pini- with a present tense form piniku, compared to Thalanyji punin (cf. Payungu punima, Purduna puniyi), as in (1).7

---

5 Classification of Pinikura as “Ngayarta” is repeated also in Ruhlen 1987, and on the Rosetta Project website (http://www.rosettaproject.org/archive/pnv).
6 All the Pinikura data is available in Toolbox format from the author. It is planned to publish this material on the SOAS Linguistics Department website during 2008.
7 The transcription adopted here follows usual Australian conventions. Each example is followed by a citation giving the speaker (M B = Molly Butler, MD = Maudie Dowton) and the location in my fieldnote books. Abbreviations are acc: accusative case, allat: allative case, emph: emphatic, imper: imperative mood, incl: inclusive, loc: locative case, nom: nominative case, pl: plural, pres: present tense, prop: proprietive, purpSS: purposive same-subject.
(1) Ngatha pini-ku ngurnta-yi-ya.
   1SG.NOM go-PRES lie-PURPSS-EMPH
   “I am going to sleep.” [MBPAN5p99s1]

(2) Wanthalu nyinta kuwarti pini-ku
    where.ALLAT 2SG.NOM now go-PRES
    “Where are you going now?” [MDPAN16p21s5]

The verb ‘to sit’ is given by both speakers as kumpa- with a present tense forms kumpaku, compared to Thalanyji nyinayin (cf. Payungu nyinama, Purduna nyinai). This is particularly interesting because the neighbouring Mantharta languages (Jiwarli, Thiiin, Warriyangka) spoken south-west of Pinikura also have a root kumpa- for ‘to sit’ and show no sign of the common Pama-Nyungan root nyina-. It is possible that there has been borrowing between Pinikura and Mantharta here, but the direction and source is unclear (there is no present tense inflection -ku in Mantharta). Examples are:

(3) Ngatha kumpa-ku karla-ngka murna-ngka.
    1SG.NOM sit-PRES fire-LOC near-LOC
    “I sit close to the fire” [MBPAN5p92s4]

(4) Nganhurru kumpa-ku nyamu.
    1PLINCL.NOM sit-PRES here-LOC
    “We sit here.” [MDPAN16p20s1]

Note that a number of intransitive verbs appear in the data with a -ku present tense inflection, including marrkarri-ku ‘wait’, martawirri-ku ‘bleed’, thanta-ku ‘come’, ngurnta-ku ‘sleep’, ngurlungkarri-ku ‘be afraid’, pukarri-ku ‘become bad’, malyarrarri-ku ‘be sick’, and the reciprocal mirlija-lparri-ku ‘hit one another’. Note also that three of these verbs, ngurnta-, thanta- and pini- take a purposive-same subject inflection -yi, compared with the corresponding Thalanyji inflection -thu (but cf. Mantharta -yi as purposive-same subject for the corresponding conjugation to which cognates of these verbs belong).

The propitieve (‘having’) nominal affix is given by both speakers as -jaka, which is identical to the Jiwarli and Thiiin forms and quite distinct from Thalanyji -wartu (note also in Warriyangka and Tharrkari, the other Mantharta languages, this affix is -parri). There appears to have been borrowing between Jiwarli and Pinikura here. Examples are:

(5) Mantu-jaka nyinta.
    meat-PROP 2SG.NOM
    “Do you have any meat?” [MBPAN5p98s1]
(6) *Nyinta mirrura-jaka kupuju-jaka.*
2SG.NOM one-PROP child-PROP
“You have one child.” [MDPAN16p21s4]

In the data supplied by Molly Butler the imperative inflection is -*ma* on intransitive verbs and -*nma* on transitive verbs. This is essentially the pattern we find in Mantharta languages also, and differs strikingly from Thalanyji, where the intransitive conjugation takes -ya or zero, and the mainly transitive conjugation takes -n. Notice that the Mantharta languages have innovated a phonological restriction that all words must end in a vowel; they add -*ma* to all uninflected roots that would otherwise end in a nasal. If this affix were historically added to imperatives ending in -n and then transferred by analogy to the zero ending imperatives, then we would find the Pinikura and Mantharta pattern. Unfortunately, the data from Maudie Dowton contains no imperative examples. Some from Molly Butler are:

(7) *Kumpa-ma ngatha nyinta-nha mirlija-ru*
sit-IMPER 1SG.NOM 2SG-ACC hit-LEST
“Sit down or I’ll hit you!” [MBPAN5p96s2]

(8) *Ngulha-ya paja-nma. Wartija-nma-ya*
nothing-EMPH eat-IMPER throw-IMPER-EMPH
“Don’t eat it! Throw it away!” [MBPAN5p97s2]

4. Conclusion

The fragmentary data available on Pinikura allows us to more or less accurately locate the traditional homeland of Pinikura speakers and to identify a number of place names as Pinikura. The available brief fieldnotes on the language from two speakers (who are now deceased) suggest that Pinikura is related to Thalanyji as a Kanyara language, however there are a number of intriguing differences that point to possible borrowing to or from the Mantharta languages. Unfortunately, due to a lack of speakers it is unclear if we will ever know more about this intriguing language.

References


Austin, Peter. 1992a. A dictionary of Payungu, Western Australia. Melbourne: Department of Linguistics, La Trobe University.

Austin, Peter. 1992b. A dictionary of Thalanyji, Western Australia. Melbourne: Department of Linguistics, La Trobe University.

Austin, Peter. 1994a. A reference grammar of the Mantharta languages. Ms, La Trobe University, Melbourne.

Austin, Peter. 1994b. A reference grammar of the Kanyara languages. Ms, La Trobe University, Melbourne.


Brown, A. R. 1912. “The distribution of Native Tribes in part of Western Australia”. Man 75.143–146.


Dench, Alan C. 1982. ”Kin terms and pronouns in the Panyjima language of northwest Western Australia”. Anthropological Forum 5.305–120.


### Appendix 1 : Pinikura place names

In my database of place names of the Gascoyne Ashburton region there are 23 identifiable names associated with the Pinikura language owning group. These are:

<table>
<thead>
<tr>
<th>Place name:</th>
<th>English name:</th>
<th>Tribe:</th>
<th>Map reference:</th>
<th>Source:</th>
</tr>
</thead>
<tbody>
<tr>
<td>kajiriwari</td>
<td>Mount Florrie</td>
<td>Pinikura</td>
<td>Wyloo LV8860</td>
<td>JB PAN7p132</td>
</tr>
<tr>
<td>karlahuntu</td>
<td>Pinikura</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Place name:</th>
<th>English name:</th>
<th>Tribe:</th>
<th>Map reference:</th>
<th>Source:</th>
</tr>
</thead>
<tbody>
<tr>
<td>kalkuru</td>
<td>Calgra Bore</td>
<td>Pinikura</td>
<td>Wyloo 60493, Wyloo MV9459</td>
<td>JB PAN7p134</td>
</tr>
<tr>
<td>karrkarti</td>
<td>Duck Creek Gorge</td>
<td>Pinikura</td>
<td>Wyloo MA5012</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Place name:</th>
<th>English name:</th>
<th>Tribe:</th>
<th>Map reference:</th>
<th>Source:</th>
</tr>
</thead>
<tbody>
<tr>
<td>kartajarri</td>
<td>Duck Creek</td>
<td>Pinikura</td>
<td>Wyloo LA9202</td>
<td>JB PAN7p133</td>
</tr>
<tr>
<td>kurara*</td>
<td>Curara Claypan</td>
<td>Pinikura</td>
<td>Wyloo MV0594</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Place name:</th>
<th>English name:</th>
<th>Tribe:</th>
<th>Map reference:</th>
<th>Source:</th>
</tr>
</thead>
<tbody>
<tr>
<td>mamurpa</td>
<td>Hooley camp</td>
<td>Purduna</td>
<td>Wyloo LV6672</td>
<td>JB PAN7p135, Radcliffe-Brown “mamurba”</td>
</tr>
<tr>
<td>mantari*</td>
<td>Mundarie pool</td>
<td>Pinikura</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

8 Near Ashburton River, north of Hardey junction, used to be a store there.
9 On Henry River, Radcliffe-Brown lists this place as Pinikura also and locates it between pirtitha and mamurpa on the Henry River.
10 Border of Pinikura and Yankurangku country.
PLACE NAME: martangu
English name: Barliyunnu pool
Tribe: Pinikura
Group number: 67
Map reference: Wyloo LV9898
Source: JB PAN7p137, Radcliffe-Brown

PLACE NAME: palijangu
English name: Barliyunnu pool
Tribe: Pinikura
Group number: 67
Map reference: Wyloo LV9898
Source: JB PAN7p133, HH PAN15p19, Radcliffe-Brown “palid’angu”

PLACE NAME: pantima
English name: Mount Edith
Tribe: Pinikura
Group number: 67
Map reference: Wyloo MV0997
Source: JB PAN7p139

PLACE NAME: parlu
English name: Mount Edith
Tribe: Pinikura
Group number: 67
Map reference: Wyloo MV0997
Source: JB PAN7p139

PLACE NAME: partarawari
English name: Mount Stuart
Tribe: Pinikura
Group number: 67
Map reference: Wyloo MA0320
Source: JB PAN7p138

PLACE NAME: pipijuku
English name: Mount McGrath
Tribe: Pinikura
Group number: 67
Map reference: Wyloo MV2290
Source: JB PAN7p134

PLACE NAME: pipijuku
English name: Mount McGrath
Tribe: Pinikura
Group number: 67
Map reference: Wyloo MV2290
Source: JB PAN7p134

PLACE NAME: pululu
English name: Boolaloo
Tribe: Pinikura
Group number: 67
Map reference: Wyloo LA8402
Source: JB PAN7p133, Radcliffe-Brown “bululu”

PLACE NAME: thinhu
English name: Urandy Station
Tribe: Pinikura
Group number: 67
Map reference: Wyloo LA8402
Source: JB PAN7p133, Radcliffe-Brown “bululu”

Well, eight miles south-east up East Henry River from pırkirlirniny.
Seven miles from Curara claypan back towards Duck Creek, two miles from Duck Creek.
Tabletop hill south of Duck Creek Gorge, upstream from junction of Duck Creek and Ashburton River.
East from Mount Stuart Station right on Duck Creek Gorge.
Junction of Henry and Ashburton Rivers.
Appendix 2: Pinikura local groups

Radcliffe-Brown’s card file of local groups from the Gascoyne-Ashburton region contains the following materials for groups identified as Pinikura:17

<table>
<thead>
<tr>
<th>Local Group</th>
<th>Totem class</th>
<th>Totem</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>wal’ida</td>
<td>kajartu</td>
<td>bungurdi</td>
<td>Mt Amy, includes nogunmara, walzainmara, bilarabuga, walgadazara</td>
</tr>
</tbody>
</table>

16 Ten or fifteen miles north-west of Mount Stuart Station.
17 In Radcliffe-Brown’s transcription system <z> represents the lamino-dental stop, often realised phonetically as a voiced lamino-dental fricative in intervocalic position in Kanyara and Mantharta languages. Speakers of Mantharta and Kanyara languages did not have a four-term section system such as that found among Ngayarta language groups to their north but rather a system of totem classes. For each totem class there are male and female cover terms, and one or more totem species which people belonging to that class are prohibited from eating. There are also thalu (Radcliffe-Brown’s <talu> ‘totem centre’) which are places where increase ceremonies are held for particular totem species, eg pungkurti thalu ‘kangaroo totem centre’ in Thalanyji. I have normalised Radcliffe-Brown’s spelling of totem class names to the corresponding forms in the conservative Kanyara and Mantharta languages (Thalanyji and Jiwarli) – see Austin 1994a, b for further details of the totemic naming system.
CLASSIFICATION OF PINIKURA

kurara  kajartu  d'igura fish
kalgalgarra  wiyaji  ?ngurawari  at junction of Hardey-Ashburton
mardulu  wiyaji  kul’iri
wongadamuga  wiyaji  wongada
walgubadu  kajartu  bungurdi
balid’angu  wariyarra  emu
bululu  wiyaji  warida eaglehawk
bangud’a  wiyaji  kigili
pidan  wariyarra  emu  miran is the totem centre
n’agun  kajartu  rain
pidiza  kajartu  bungudi  at junction of Henry River and Ashburton River. Pools belonging to the group are on map 5. bungudi talu at ngurin-ngurin.
wadura  wariyarra  emu  on Henry River between pidiza and mamurba (baingula)
yirbira  wariyarra  emu
mardangu  walirri  kalban’black opossum
  t’udungad’i honey
yirda  kajartu  d’igura
  kalban’ totem centre is Nanutara station. The country included mardangu rockhole north of Ashburton, ngand’ungarina, kgiangulera etc.
yaribidi or yirgu  (?kajartu)  ngurawari
kandang?  kajartu?  d’igura  see yirda
talu at kandang, at a stone near a pool creek off Ashburton River below kurara. Territory includes pools in Ashburton River above Hardey junction; mazan’i, mid’almid’ala, yirda, kandang or kandangu (order upstream) - also two claypans d’undal’a to west of river and kurara or kurarda to east. At madand’i there is a wanamangura (mythical snake). At kandang is the totem centre of d’igura
BOUND PRONOMINALS IN THE WEST PAPUAN LANGUAGES

MARK DONOHUE
Monash University

1. Introduction

The West Papuan languages are a genetic grouping of languages found in the east of Indonesia. Languages belonging to this family are spoken in the north of Halmahera and in the western Bird’s Head at the far western edge of New Guinea. As originally established (e.g., Cowan 1953, 1957, 1958, and later 1960, summarised in Capell 1975), the family also included the languages of the eastern and southern Bird’s Head, those of central Yapen island to the east, and the non-Austronesian languages of Timor, Alor and Pantar in the south of Indonesia. Voorhoeve (1975a, b), upon a more careful examination of the data, concluded that the Timor-Alor-Pantar languages were members of the Trans New Guinea (TNG) family, and that the Yapen languages belonged to the East Geelvink Bay family, a grouping that included the languages of the interior of East Geelvink (= Cenderawasih) Bay.¹

In this paper I shall demonstrate that, based on the evidence of bound pronominal morphemes, the languages of Yapen have to be treated as members of the West Papuan family, as originally suggested (see Nichols 1996 for methodological considerations). The evidence for this claim comes from an examination of the bound pronominal prefixes that are a feature of this part of the north coast of New Guinea, including these languages. The reason for using bound morphology is simply that it is more likely to reflect older pronominal elements than younger ones. New Guinea has long been regarded as an area in which independent pronominals are subject to widespread borrowing (e.g., Foley 1986) or widespread restructuring, and so these independent words are less reliable than bound forms as indicators of genetic relationships.³ Bound pronominal elements have been compared because this is the only type of morphology consistently found across all the languages in the survey. The locations of the languages and language groups in question are shown in Map 1.

¹ I would like to thank this volume’s editors and referees for valuable comments on this paper.
² There are both Austronesian and non-Austronesian languages on Yapen island. In this paper shall use the term ‘Yapen language’ only to refer to the non-Austronesian languages.
³ For instance, in the languages examined for this study the free pronominals and bound pronominals show remarkable phonological agreement in the North Halmahera languages, and an almost total lack of correspondence in the Yapen languages, making the free forms nearly useless for comparative purposes.
In the remainder of this paper I shall examine bound pronominal elements from the languages of Yapen, Halmahera and the western Bird’s Head, making local reconstructions and then integrating the results. We shall see that there is extraordinarily little evidence for the inclusion of the Yapen languages in a Geelvink Bay unit, but that they fit well into a revised West Papuan family.

Map 1: Halmahera, the Bird’s Head and Yapen

2. Yapen

As can be seen in Map 1, Yapen island is found in the middle of Cenderawasih Bay in western New Guinea. While the west and east of the island are home to Austronesian languages, the centre of the island contains two Non-Austronesian languages, Yawa and Saweru. While Saweru is a small language spoken on a single 4.5km long island off the south coast of Yapen, there are many varieties of Yawa spoken by upwards of 6,000 people. These two languages are, as we shall see closely related. These languages distinguish three different bound pronominal sets, which can be described as ‘nominative’ (or ‘subject’), ‘genitive’ = ‘accusative’ (or ‘object’), and ‘dative’. Since the dative affixes are transparently derived from the fusion of the genitive prefixes with the dative postposition/enclitic, they shall not be considered further here. The nominative and genitive agreement markers are part of a system of semantic alignment (Donohue & Wichmann 2008; also referred to as ‘split-intransitive’,
'stative/active', and other terms), such that in addition to being used to mark transitive subjects and objects can also be used to encode the distinctions found between semantically distinct predicates; some of the possibilities are shown in (1) - (3) for Saweru, and similar data could be presented for Yawa as well (see Jones 1986a) (glosses follow Donohue 2001). In (1) we can see the basic marking of As and Ps in bivalent predicates, while (2) shows that the A-marking clitics from (1a) are also used to mark some agentive Ss, and that the P-marking prefixes from (1b) are also used to mark some non-agentive Ss. (3) demonstrates the fact that a full description of agreement in Saweru involves further complications, in that not all As are marked with the proclitics, and not all Ps with the prefixes (see Donohue 2001, 2004 for further discussion).4

Saweru bivalent

(1) a. Mo=na-ba-i.  
3SG.F.NOM=2SG.GEN-hit-TNS  
“She hit you.”

b. No=ra-ba-i.  
2SG.NOM=3SG.F.GEN-hit-TNS  
“You hit her.”

monovalent

(2) a. Mo=rayan-i.  
3SG.F.NOM=swim-TNS  
“She swam.”

b. Ra=teson-i.  
3SG.F.GEN-diarrhoea-TNS  
“She has diarrhoea.”

non-accusative bivalent

(3) a. Mo=komi-nai.  
3SG.F.NOM=search.for-2SG.DAT  
“She searched for you.”

b. Ra=meme-nai.  
3SG.F.GEN-think-2SG.DAT  
“She thinks of you.”

The basic bound pronominals for Sarawandori Yawa (the variety reported in Jones 1986a, from the south-west of the language area), Yapanani Yawa (the most south-eastern Yawa variety, which shows considerable reduction in contrasts) and Saweru are shown in Table 1, together with some non-controversial reconstructions for proto-Yapen. The correspondences, shown here, such as /s:Ø/ and /p:ø/ are regularly attested in lexical items as well as in these affixes.

Examining the reconstructions in Table 1, we can clearly analyse various morphological formatives. The reconstructions in Table 1 can in most cases be split segment-by-segment into the separate morphemes, or morpheme formatives, described in Table 2. Table 3 presents the pronominal roots that are

4 In addition to the conventions of the Leipzig glossing rules, the following abbreviations have been used: AUG: augmented, EX: exclusive, F: feminine, IN: inclusive, M: masculine, MIN: minimal, NH: non-human, P: Proto-, TAP: Timor-Alor-Pantar (family?), TNG: Trans New Guinea (family), TNS: tense, WBH: Western Bird’s Head (family), WP: West Papuan (family).
found in addition to these formatives. Note that this analysis depends on the assumption that m is found in the formation of many prefixes, but is semantically empty.

<table>
<thead>
<tr>
<th>Formative</th>
<th>Appears with</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>o</td>
<td>(all 'subject' forms)</td>
<td>NOMINATIVE</td>
</tr>
<tr>
<td>ns</td>
<td>*(w)ans, *(r)ans, *(w)ans</td>
<td>NSG.LOCAL.ACCUSATIVE</td>
</tr>
<tr>
<td>r</td>
<td>*(r)imo, *(r)amo, *(r)ans</td>
<td>1NSG</td>
</tr>
<tr>
<td>w</td>
<td>*(w)amo, *(w)ans, *(w)ans</td>
<td>2/3PL</td>
</tr>
<tr>
<td>i/y</td>
<td>*(r)imo, *(r)ins, *(r)imo, *(r)ins</td>
<td>DUAL</td>
</tr>
<tr>
<td>p</td>
<td>*(r)imo, *(r)ins, *(r)imo, *(r)ins</td>
<td>2NSG.NOM</td>
</tr>
<tr>
<td>a</td>
<td>*(r)amo, *(r)amo, *(w)amo, *(w)ans, *(w)ans</td>
<td>PL.LOCAL</td>
</tr>
<tr>
<td>&quot;m&quot;</td>
<td>(many)</td>
<td>–</td>
</tr>
</tbody>
</table>

Note: ‘Local’ refers to local person: first or second person.

Table 1: Yawen agreement prefixes

<table>
<thead>
<tr>
<th>Formative</th>
<th>Appears with</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>o</td>
<td>(all 'subject' forms)</td>
<td>NOMINATIVE</td>
</tr>
<tr>
<td>ns</td>
<td>*(w)ans, *(r)ans, *(w)ans</td>
<td>NSG.LOCAL.ACCUSATIVE</td>
</tr>
<tr>
<td>r</td>
<td>*(r)imo, *(r)amo, *(r)ans</td>
<td>1NSG</td>
</tr>
<tr>
<td>w</td>
<td>*(w)amo, *(w)ans, *(w)ans</td>
<td>2/3PL</td>
</tr>
<tr>
<td>i/y</td>
<td>*(r)imo, *(r)ins, *(r)imo, *(r)ins</td>
<td>DUAL</td>
</tr>
<tr>
<td>p</td>
<td>*(r)imo, *(r)ins, *(r)imo, *(r)ins</td>
<td>2NSG.NOM</td>
</tr>
<tr>
<td>a</td>
<td>*(r)amo, *(r)amo, *(w)amo, *(w)ans, *(w)ans</td>
<td>PL.LOCAL</td>
</tr>
<tr>
<td>&quot;m&quot;</td>
<td>(many)</td>
<td>–</td>
</tr>
</tbody>
</table>

As we shall see later, the reconstruction of these pronominal forms is a strong argument against the theory that links Yawa (and by implication Saweru) with the languages of the eastern interior of Geelvink Bay.

Examining the forms shown in Table 1, we can apply principles of internal reconstruction to arrive a more regularised pre-proto-Yawen stage, shown in Table 4. Many of the uncertainties in the reconstructions from Table 1 have been removed here, though, as we shall see in Table 10, the irregularities do not reconstruct to a higher level, and may represent multiple innovations within the Yawen group.
Table 4: Pre-proto-Yapen pronominal prefixes

<table>
<thead>
<tr>
<th></th>
<th>NOM:</th>
<th>SG</th>
<th>DU</th>
<th>PL</th>
<th>ACC:</th>
<th>SG</th>
<th>DU</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>*jo</td>
<td>*rimo</td>
<td>*ramo</td>
<td></td>
<td>*in</td>
<td>*rans</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>*wamo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>*no</td>
<td>*ipo</td>
<td>*wapo</td>
<td></td>
<td>*n</td>
<td>*ins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3M</td>
<td>*po</td>
<td>*yo</td>
<td>*wo</td>
<td></td>
<td>*a</td>
<td>*y</td>
<td>*w</td>
<td></td>
</tr>
<tr>
<td>3F</td>
<td>*mo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*r</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. North Halmahera data

The languages of North Halmahera have been described by a range of authors, and have received comparative attention by van der Veen (1915) and Voorhoeve (1988). The morphological alignment of the majority of these languages is very similar to that shown in (1) and (2) for Saweru, showing a semantically-aligned contrast (Holton 2008 discusses this system in detail). In some languages the ‘object’ prefix has become specialised as a possessive agreement prefix (see the use of the same prefixes in Yapen languages for both genitive and accusative functions).

Voorhoeve reconstructed the pronominal forms shown in Table 5. Table 6 presents data from four languages of North Halmahera (drawn from Voorhoeve’s work, and also van Staden 2000, Visser and Voorhoeve 1987, Wimbish 1991), and Table 7 presents reconstructions for Halmahera based on the data in this table. There is very little disagreement between Voorhoeve’s reconstructions and mine, except that I give more weight to the irregular West Makian forms, since West Makian is assumed to be a first-order subgroup in the North Halmahera family (Voorhoeve 1988). I reconstruct two sets of prefixes rather than Voorhoeve’s single set, though it is quite easy to isolate an -i formative that indicates the possessor-object-nonagent function, just as an -o ‘nominative’ formant is readily identified in the singular pronominals (see Wimbish 1991). The other main differences between my reconstructions and Voorhoeve’s are simply the result of more deliberation on my part, since I wish to preserve any uncertainties for higher-level comparison. Many of the irregular developments from the forms in Table 7 to the primary data in Table 6 can be accounted for by analogy; the Tidore no-1PLEX prefix has acquired the o vowel that characterises the nominative prefixes. Similarly, the lack of a reflex of *r for the 1SG OBJECT/NON-AGENT form in Pagu and West Makian is the result of two separate instances of analogic levelling with the 1SG SUBJECT form (in West Makian) or the developing less specified 1/2NONAGENT form that has spread in Pagu.

5 While the comparisons in the Yapen datasets in section 2 involve sound correspondences that are testably regular between the two languages, the lack of as-yet uncovered lexical cognates means that comparison between areas is not yet supported by the presence of similar sound correspondences in the lexicon.
Table 5: Voorhoeve’s proto-North Halmahera pronominal prefixes

<table>
<thead>
<tr>
<th></th>
<th>Tidore ‘subject’ ‘(poss’r)’</th>
<th>Sahu ‘subject’ ‘object’</th>
<th>Pagu ‘agent’ ‘nonagent’</th>
<th>West Makian ‘subject’ ‘(poss’r)’</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>to- ri-</td>
<td>to- ri-</td>
<td>to- ni-</td>
<td>tV- ti</td>
</tr>
<tr>
<td>2SG</td>
<td>no- na-</td>
<td>no- ni-</td>
<td>no- ni-</td>
<td>nV- ni</td>
</tr>
<tr>
<td>3SG.M</td>
<td>wo- i- ‘w’</td>
<td>o- u-</td>
<td>wo- wi-</td>
<td>i- mV</td>
</tr>
<tr>
<td>3SG.F</td>
<td>mo- mi-</td>
<td>mo- mi-</td>
<td>mo- mi-</td>
<td>i- mV</td>
</tr>
<tr>
<td>3SG.N</td>
<td>yo- ma-</td>
<td>yo- ma-</td>
<td>yo- /i- (Ø)</td>
<td>i- dV</td>
</tr>
</tbody>
</table>

Table 6: Halmahera agreement prefixes (representative)

<table>
<thead>
<tr>
<th></th>
<th>‘subject’ ‘object’</th>
<th>‘subject’ ‘object’</th>
<th>‘subject’ ‘object’</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>‘to’ ‘r’ (‘n’)</td>
<td>‘mi’</td>
<td></td>
</tr>
<tr>
<td>2SG</td>
<td>‘no’ ‘ni’</td>
<td>‘mi’</td>
<td></td>
</tr>
<tr>
<td>3SG.M</td>
<td>‘wo’ ‘wi’</td>
<td>‘mi’</td>
<td></td>
</tr>
<tr>
<td>3SG.F</td>
<td>‘mo’ ‘mi’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3SG.NH</td>
<td>‘yo’ ‘Ca’</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7: Proto North Halmahera agreement prefixes

4. Bird’s Head

The languages of the Bird’s Head are more problematic for historical linguists than those of Yapen or Halmahera, probably reflecting the influence of various unrelated languages in the area. While the languages of the north-west of the peninsula are, as I shall show, related to the Yapen and Halmahera languages, those of the east (Meyah, Sough, Hatam and Mansim) and the north-east (Mpur) are not so obviously related to these groups. In Table 8 the three languages on the left, Tehit, Moi and Mai Brat, are the languages of interest for this study.

The Bird’s Head languages show fewer forms, with only one set of nominative (‘subject’) prefixes, and, as can be seen, show much more erratic correspondences. Table 9 presents reconstructions for putative proto-language ancestral to Tehit, Moi and Mai Brat (data from Flaszy 1991, Hesse 1983, Stokhof & Flaszy 1985, Brown 1999, Dol 1999, Reesink 1999, 2002).
BOUND PRONOMINALS IN THE WEST PAPUAN LANGUAGES

Table 8: Bird’s Head agreement prefixes

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>*t</td>
<td>*m/*p</td>
</tr>
<tr>
<td>12</td>
<td>*p</td>
<td>*[p/f]</td>
</tr>
<tr>
<td>2</td>
<td>*n</td>
<td>*n</td>
</tr>
<tr>
<td>3M</td>
<td>*w/*p</td>
<td>*[C/y]</td>
</tr>
<tr>
<td>3F</td>
<td>*m</td>
<td></td>
</tr>
</tbody>
</table>

Table 9: Tentative proto-Western Bird’s Head reconstructions

In addition to the languages shown in Table 8 there are also a number of languages of the Southern Bird’s Head which do not, following Voorhoeve (1975), appear to be related to the more northern languages at all (see also De Vries 1998), and cannot be considered part of a genetic grouping that includes the northern languages.

5. Combined picture

Putting together the reconstructions seen earlier in Tables 4, 7 and 9, we can reconstruct the Proto-West Papuan pronominal forms shown in Table 10.

<table>
<thead>
<tr>
<th></th>
<th>A-type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1SG</td>
</tr>
<tr>
<td>PWP</td>
<td>*to/*tio</td>
</tr>
<tr>
<td>P Yapen</td>
<td>*fo</td>
</tr>
<tr>
<td>[DUAL: P-]</td>
<td>*ri</td>
</tr>
<tr>
<td>Halmahera</td>
<td>*to</td>
</tr>
<tr>
<td>PWBH</td>
<td>*t</td>
</tr>
</tbody>
</table>
Table 10: Proto-West Papuan reconstructions and their descendants

The following changes are assumed to have taken place between proto-West Papuan and (pre-)proto-Yapen:

- the 1SG A-form *tio became *fo, presumably via **tjio and **tjoo; the other singular A-forms are retained without change
- while the plural A-forms *mo and *po are carried through into Yapen in their original functions, *yo shifts to a dual function, pulled into this new specialisation through its similarities to the proto-Yapen *i/y dual formative.
- in a similar elaboration of NSG pronominal forms, *-r- develops as a first person nonsingular formative, derived from a generalisation of the 1PL P-form.
- *-a- appears as a first or second person plural formative.
- *mo and *po are also used with dual functions, distinguished from the plural form by the addition of *i (thus resulting in *rimo, *ipo), and the development of the proto-West Papuan *w(o) '12PL' as a non-first person plural formative (yielding the plural forms *ramo, *wamo and *wapo).
- a replacement 3PL form *wo- is innovated, based on this plural formative and the nominative o.
- for the P forms, the 1SG and 2SG are retained, with the loss of the vowel in the 2SG form presumably as a device to maximise contrast with the 1SG.
- a new morpheme *-ns- develops as a marker of (local, nonsingular) accusative, presumably because the old *-i- accusative declined as a result of the development of *-i- as a dual formative.
- the same formatives that we saw developed in the A-forms are also used in the P-forms, but with the *-ns- formative, and not an of the nominative forms.
- the old 3PL P-form shifted to mark (feminine) singular, stripped of the now obsolete accusative formative.
- a new, essentially morphologically unmarked, 3SG.M P-form develops.
Similarly, the bound pronominal system of North Halmahera can be derived with the following set of changes:

- the 3SG.M A-form lenites from *po > *wo; the other singular forms survive unchanged (potentially *tio > *to), and the 3PL prefix extends its range to mark 3SG.NONHUMAN as well as being retained as a marker of 3PL.

- The plural forms are irregularly distinguished by using a plural formative *-i-, possibly related to the Yapen dual *-i-, and if so reflecting an aspect of the morphology of proto-West Papuan.

- The plural forms are irregularly distinguished by using a plural formative *-i-, possibly related to the Yapen dual *-i-, and if so reflecting an aspect of the morphology of proto-West Papuan.

- The 1PL.EX A-form reflects this plural formative, while the 1PL.IN is either replaced by the unspecified ‘pronominal’ formative *m (see discussion surrounding Table 2), or else is replaced by the proto-West Papuan 2PL form.

- the 2PL *po lenites to *fo and in some cases to *wo, and additionally extends reference to 1PL.IN (=12PL) in most North Halmaheran languages. This ambiguity allows an innovated form *ni to extend into the 2PL position.

The development of *ni- as an A-form probably postdates the breakup of proto-North Halmahera, since West Makian does not reflect *ni- in the 2PL, though we already see *ni in proto-North Halmahera.

The loss of distinction between SG and PL, and so the spread of the 2SG *n into the plural, is initiated by the development of the singular nonhuman forms that are formally identical to the 3PL.

- the a ‘1PL.EX’ form in West Makian suggests that the reconstructed *-a- for proto-Yapen should be reconstructed to a higher level as well (data below from the 1PL.IN P-form, and from the related Bird’s Head languages, provide further support for this hypothesis). In West Makian this form replaced the inherited *mo, while the underspecified *m form became used for the 1PL.IN.

- the *di ‘3PL’ form presumably reflects crossover from the P-forms.

- the P-forms show the wide spread of the *-i- ‘accusative’ formative, and the reconstruction of many P-forms based on the A-forms. In the PL the Halmahera A-forms dominate, with only the *na- 1PL.IN being distinct; the *-a- has been discussed above, and the *-n- probably reflects the consonant of the West Papuan 1SG.P or the North Halmahera 2PL.A form.

- the *in ‘1SG.P’ loses this function; the Pagu forms probably reflect a spread of the above innovative 1PL.P form, marked for accusative or plural with *-i-. West Makian shows a 1SG.P based on the A-form, while the other languages reflect *r.
When we examine the changes required from proto-West Papuan in Table 10 to the reconstructions for proto-Western Bird’s Head in Table 9, we find remarkably few inclusive changes in the forms themselves, though the collapse of the two sets into one makes for a complex modern picture.

- the 2PL *po has replaced the 1PL.IN form; in some daughter languages the reflex of *po has extended to the 1PLEX as well.
- the West Papuan 2PL has been replaced by a form with n, presumably formed by analogy from the n in the 2SG. The reconstructions for all of the West Bird’s Head prefixes are single segments, but there is some unusual behaviour associated with the 2PL (and 1PL) prefixes. (4), showing inflectional paradigms from Mai Brat (Brown 1999) demonstrate, from left to right, the use of epenthetic schwas with C-initial roots; the uncomplicated prefixation of inflection on V-initial roots; and the irregular behaviour of 1PL and 2PL forms with a-initial roots or VG-initial roots. As can be seen, the 1PL and 2PL prefixes appear to ‘absorb’ an initial low vowel, so that *p-aws > p-ws (absorption) > [pu:s] (resyllabification). This might reflect an earlier vowel position in the syllable template, differentiating *n- ‘2SG’ from *nV- ‘2PL.’ Recall that proto-Yapen showed a 1PL/2PL formative of the shape *-a- (Table 2).

<table>
<thead>
<tr>
<th></th>
<th>-C</th>
<th>-V</th>
<th>-a</th>
<th>-aw</th>
<th>-ay</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>tāpo</td>
<td>tisi</td>
<td>tamo</td>
<td>taws</td>
<td>taym</td>
</tr>
<tr>
<td>2SG</td>
<td>nāpo</td>
<td>nisi</td>
<td>namo</td>
<td>naws</td>
<td>naym</td>
</tr>
<tr>
<td>3SG.M</td>
<td>yāpo</td>
<td>yisi</td>
<td>yamo</td>
<td>yaws</td>
<td>yaym</td>
</tr>
<tr>
<td>1PL</td>
<td>pāpo</td>
<td>pisi</td>
<td>pāmo</td>
<td>pu:s</td>
<td>pi:m</td>
</tr>
<tr>
<td>2PL</td>
<td>nāpo</td>
<td>nisi</td>
<td>nāmo</td>
<td>nus</td>
<td>nīm</td>
</tr>
<tr>
<td>3PL / 3F</td>
<td>māpo</td>
<td>misi</td>
<td>mamo</td>
<td>maws</td>
<td>maym</td>
</tr>
</tbody>
</table>

- there is considerable reanalysis of the 3PL position, none of it reconstructable to a proto-West Bird’s Head stage, but involving extension of either the unmarked *m- form, or extension of the (semantically unmarked) 3SG.M form. In Mai Brat we can see the opposite, the replacement of the 3SG.M with the 3PL y-.

Based on the development of the proto West Papuan 2PL form as a marker of 1PL.IN and the subsequent innovation of *n(i) as a 2PL marker we might suppose a closer connection between the Bird’s Head languages and the Halmahera

---

Data from other Western Bird’s Head languages are, of course, desirable in order to confirm whether this can be taken as an argument for the other languages in the Bird’s Head as well as for Mai Brat.
languages. The fact that the Halmahera *ni 2PL postdates the breakup of these languages, however, means that this cannot be used to group the Halmahera and the Bird’s Head languages together. Similarly, the fact that West Makian does not reflect *po as a 1PL.IN (or EX) form means that this cannot be used to subgroup the two groups together. The Yapen languages potentially share the development of the *-i- DUAL with the Halmahera *-i- PLURAL, but this is the only possibly shared innovation, and we will have to leave subgrouping for further consideration.

In sum, the pronominal data strongly supports the relatedness of the three language groups discussed here, though we have not seen any strong evidence for any of these two groups being more closely related to each other than to the other one. This implies that the Yapen languages do belong to a (revised) West Papuan family.

6. **Revising the ‘West Papuan Phylum’: the Geelvink Bay hypothesis**

We should consider the alternative hypothesis, that the Yapen languages are members of the Geelvink Bay phylum, related to the languages of the Waropen coast hinterland. This classification was proposed by Voorhoeve (1975b), based solely on lexical comparisons; Voorhoeve (1975b: 875-876) lists the lexical forms presented in Table 7 (Baropasi and Bauri [= Bauzi] were also compared, but only Tarunggare forms were given for comparison; in this table Yava = Yawa).

<table>
<thead>
<tr>
<th>Yava</th>
<th>Tarunggare</th>
<th>Yava</th>
<th>Tarunggare</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘come’ nde, re</td>
<td>nere</td>
<td>‘tail’</td>
<td>ateva</td>
</tr>
<tr>
<td>‘eat’ rai(s)</td>
<td>ghayo</td>
<td>‘water’</td>
<td>waro</td>
</tr>
<tr>
<td>‘fly’ bariri</td>
<td>bunana</td>
<td>‘wind’</td>
<td>bwa</td>
</tr>
<tr>
<td>‘foot’ najo</td>
<td>nal</td>
<td>‘you (sg)’</td>
<td>ein</td>
</tr>
<tr>
<td>‘give’ ra</td>
<td>nore</td>
<td>‘you (pl)’</td>
<td>wea</td>
</tr>
<tr>
<td>‘I’ nei, rei</td>
<td>ei</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 11: Voorhoeve’s Yapen: Geelvink Bay correspondences

Examining these lexical items, the following cautions apply: Yawa rai(s) is morphologically ra- ‘3SG.F.Gen/Acc’ plus i ‘eat (transitive)’ (I cannot explain the bracketed ‘(s)’, except to note that it is not heard in Yawa or Saweru these days). The ‘probable cognate’ is then reduced to an i ‘correspondence’. Yapen bariri: Tarunggare bunana evidences only one out of six phonemes as cognate, unless an r correspondence can be established, and so this remains tentative. Yapen najo is morphologically na- 2SG.Gen plus yo ‘foot, leg’, yielding no correspondences at all with Tarunggare, and only a single vowel with Baropasi naro. A single consonant defines Yawa -ra and Tarunggare nore ‘give’, and the putative rr correspondence would nullify the value of ‘fly’ earlier. However, Saweru awe ‘give’ means that -ra cannot be reconstructed for proto-Yapen in any case. Yapen inei : Tarunggare ei relies solely on vocalic correspondences, as does ‘you (singular)’ and ‘you (plural)’. Yawa obar: Tarunggare bwa is possible,
and would confirm the b:b seen in ‘fly’. The set Yapen re: Tarunggare nere ‘come’, and ateva: Tarunggare atapara, in which a- is the 3SG.NFEM.GEN prefix for the Yapen forms, seem plausible, as is Yawa karu: Tarunggare waro, but these are then only three possible cognates, with no regular sound correspondences beyond two very tentative b:b sets, not enough to base a genetic relationship on.

We can examine the pronominal data from the one East Geelvink Bay family language for which we have adequate data, Bauzi (Briley 1997). The pronominal bases are shown in Table 12. Significantly, there is no verbal agreement, a point of structural disagreement.

<table>
<thead>
<tr>
<th>1SG</th>
<th>e-</th>
<th>1PL</th>
<th>i-</th>
</tr>
</thead>
<tbody>
<tr>
<td>2SG</td>
<td>o-</td>
<td>2PL</td>
<td>u-</td>
</tr>
<tr>
<td>3SG</td>
<td>a-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 12: Bauzi pronominal forms

Establishing any sort of correspondence between these monovocalic free pronominal forms and the bound morphology discussed in section 4 would be speculative, and establishing a connection between the forms seen in Table 12, and those in Table 10, would be highly speculative. In short, there do not appear to be any compelling reasons to accept a Geelvink Bay phylum hypothesis that includes the Yapen languages, just as there are no reasons to exclude the Yapen languages from a West Papuan group.

7 The rest of the old West Papuan Phylum: connections with Timor-Alor-Pantar languages?

The West Papuan Phylum, as summarised in Capell (1975), included not only the Yapen languages which were removed by Voorhoeve, but also the languages of the eastern Bird’s Head, the southern Bird’s Head and of the Timor-Alor-Pantar (TAP) region.

The eastern Bird’s Head languages appear to be a diverse group of languages with typological similarities, but little in the way of formal resemblance. Reesink (2002) discusses the difficulties of grouping these languages, and while he arrives at no firm conclusions (see Donohue 2005), it is clear that they cannot be unproblematically grouped with any of the languages described here, certainly not on the basis of pronominal evidence (see the right-hand side of Table 8).

7 There is no 3PL pronominal form in Bauzi, and all the pronominal bases in this table must be suffixed for case in order to be used in grammatical clauses; case-inflected pronouns such as these are not a feature of the other languages we have been considering. The case choices are -ho ergative, -m absolutive, and -ba dative, as well as some pragmatically, and not syntactically, determined choices. Notice that, in addition to the fact of case-marking on free pronouns itself, the alignment marked in Bauzi matches neither the nominative-accusative pattern of the Bird’s Head, nor the stative-active systems of Halmahera and Yapen.
The southern Bird’s Head languages, such as Inanwatan, have been shown to share little if any structural or lexical material with the Bird’s Head languages that have been considered here. Voorhoeve suggests that they should be classified as members of the wide-faring TNG phylum, and De Vries (1998) suggests, on the basis of unusual structural features, that they are most closely related to the Marind languages of the far south coast.

The languages of the Timor-Alor-Pantar region have also been grouped as TNG languages (Stokhof 1975), albeit with very different structural profiles (see Stokhof 1987, Steinhauer 1995, Donohue 1996). These languages show many lexical (and bound pronominal) similarities with the ‘average TNG’ languages of mainland New Guinea, but also show some elements that cannot be traced to current TNG reconstructions. Table 13 presents reconstructions of the TAP bound pronominals (Donohue & Schapper 2007) and of the TNG pronoun system. Given the similarities in the first person forms, much of the rest of the TAP material can be related to the TNG forms, if one assumes a swap of second and third person; even the association of plural with high front vowels is present, and (not shown here) the dual *-li formative (Suter 1997) from mainland New Guinea also appears in Timor languages such as Buna’ (Berthe 1961).

The TAP 12AUG form cannot be related to TNG reconstructions, but is similar to the Bomberai form mbi, which has exclusive, not inclusive, meaning. The fact that many lexemes from the TAP languages have correspondences in the Bomberai peninsula, and not elsewhere in New Guinea, makes this pronominal connection enticing. There is no evidence of the Bomberai languages having prenasalised correspondences for plain voiceless stops, however, and no explanation for the inclusive/exclusive cross-over. An equally plausible history for TAP *pi matches the Bird’s Head and Halmahera story, in which a 2PL form acquired 12PL reference. But this would require us to assume a West Papuan influence – is there any independent evidence of this? The Bomberai peninsula is located immediately south of the Bird’s Head, making some early influence plausible. When we examine the rest of the TAP pronominal reconstructions, we find that the 12MIN form (‘you and I’) is also not related to any proto-TNG reconstructions. It can, however, plausibly be related to the West Papuan *to ‘1SG’, with the addition of a second person referent as well. This does not appear to argue for a genetic connection with the West Papuan languages, but a possible borrowing relationship. Table 14 presents representative lexical data from the TAP languages (from Donohue 2007a), the Bomberai languages (principally Iha), and the languages of the West Papuan family. While some words show apparent cognacy over a wide area, the Bomberai correspondences are clearly closer for the TAP languages than for any of the other columns. In the Bird’s Head we find some forms apparently cognate between the (western) Bird’s Head languages discussed here and the TAP languages, in which case the lexeme is also shared with the Halmahera group, and some isolated lexemes in
the unrelated languages of the eastern Bird’s Head (such as Hatam) or the Southern Bird’s Head. This suggests that we can legitimately discuss a contact area which might have given rise to borrowings between a West Papuan source, previously spoken over a wider area than is currently the case, and a pre-TAP language in the Bomberai/South Bird’s Head area (see also Donohue 2007b, 2007c).

---

<table>
<thead>
<tr>
<th>Timor-Alor-Pantar</th>
<th>Trans New Guinea</th>
<th>Iha (Bomberai)</th>
</tr>
</thead>
<tbody>
<tr>
<td>minimal</td>
<td>augmented</td>
<td></td>
</tr>
<tr>
<td>1 *n</td>
<td>*ni</td>
<td>*na *ni *nu</td>
</tr>
<tr>
<td>12 *ta</td>
<td>*pi</td>
<td>in</td>
</tr>
<tr>
<td>2 *Ø / *a-</td>
<td>*i</td>
<td>*ga *gi *ja</td>
</tr>
<tr>
<td>3 *g</td>
<td>*gi</td>
<td>[y]a [y]i</td>
</tr>
</tbody>
</table>

Table 13: Three pronominal sets reconstructions

\[
\begin{array}{|l|l|l|l|l|l|}
\hline
& TAP & Bomberai & Bird’s Head & Halmahera & Yapen \\
\hline
\text{‘tree’} & *a\text{Da} & adop & ara & *gota & - \\
\text{‘water’} & *yar & kar(y)a & *kala & *gala & karo (strait) \\
\text{‘hand’} & *tan & tan & (Hatam: nda) & - & - \\
\text{‘dog’} & *yapal & yamb(l)a[l/r] & - & - & - \\
\text{‘sit’} & *mi & mese & - & - & - \\
\text{‘stomach’} & *atok & tokar & (SBH: atoko) & - & karawe (faeces) \\
\hline
\end{array}
\]

Table 14: Possible correspondences between TAP languages and WP languages

We have seen how the evidence from bound pronouns clearly argues for the inclusion of the Yapen languages into a re-assessed West Papuan phylum, but that the other members of the family, removed by Voorhoeve in 1975, must remain outside its bounds, though within the scope of plausible (and detectable) historical influence. A considered evaluation of the bound pronominal evidence, based on the well-established principles of the comparative method and supported by a careful investigation of the morphophonological peculiarities of the languages concerned, allows us to establish with confidence the new boundaries of the West Papuan family.

References


1. Introduction

This paper grew out of, or should I say, shrank from a paper originally given at an International Conference on Historical Linguistics some years back to a minimal-sized audience. It is appropriate that the enigmatic relationships of Alawa and its neighbours be further explored, and that a start be made in a volume in honour of Harold Koch. Many a time I have stayed with Harold and Grace, and been stimulated by the contact and conversations on both historical linguistics, Aboriginal and other music, and archiving. I was also invited by them — and took up the offer — to house-sit for them while they and I were on study leave and I was beginning to focus more strongly on historical linguistics.

The similarities and differences in Alawa and nearby languages put me in mind of Elgar’s *Enigma Variations*. Elgar composed variations on a theme, but the underlying theme was never made overt. Alawa and its neighbours considered here have a lot of similarities, but the connections between them are enigmatic.

The three non-Pama-Nyungan languages Alawa, Mara and Warndarang have been generally considered to form a subgroup (Sharpe 1972, Heath 1978). On the basis of shared features with Alawa, Merlan (1989) has argued for Mangarayi being added to this group. However Baker (2006) has called into question the evidence for such a group, or at least whether these languages belong in one family, drawing attention to the differences between Alawa and the other languages in finite verbs. I will summarise what we can say at the present time, before I focus on two small aspects which can be handled in a short article.

- Structurally all these languages are very similar, in that syntax is similar and translation from one to another mostly involves a linear replacement of morphemes (extract one, drop in another). The structures for marking of tense, aspect, mood and case categories are very similar; however, *the forms can be quite different*. One should bear in mind that it was common for traditionally-living people in this area to be multilingual; these similarities could be partially a result of multilingualism.

- All have closely similar systems of an invariant coverb linked to a finite verb.

- The languages share certain phonological processes, such as ‘hardening’ of initial semivowels or addition of a stop to an initial vowel of a finite verb.
when the coverb ends in a stop or nasal. Probably historically a consonant was lenited when ‘unprotected’ in such contexts.

- Somewhere near half of the coverbs in these languages are of the form CVC; shapes CVCVC, CVCV and CVCC account for most of the other forms. Many coverbs are cognate.
- All have genders: Alawa has just two (masculine and feminine), Mara and Mangarayi three (masculine, feminine and neuter), and Heath established that three extra genders in Warndarang were diffused from Nunggubuyu (Heath 1978).
- All mark dual with a prefix having /r/, and plural with a prefix having /l/, generally contrasting with other languages of the area.
- Many pronominal prefixes on verbs seem clearly related and use a similar system, and all show a similar ranking of first and second person in prefixes indicating one of these acting on the other.
- Alawa and Mangarayi share an adjectival ending -mayin which is reflected in Mara’s adjectival ending -min.
- If we set aside probable borrowing, common cognate vocabulary would suggest that the most recent connection is between Mara and Warndarang, with Alawa related at a longer time depth, and Mangarayi connected at even greater time depth. See Table 1, which shows cognacy in a maximum of 157 common words.¹

<table>
<thead>
<tr>
<th></th>
<th>Warndarang</th>
<th>Mara</th>
<th>Alawa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mara</td>
<td>71%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alawa</td>
<td>43%</td>
<td>54%</td>
<td></td>
</tr>
<tr>
<td>Mangarayi</td>
<td>16%</td>
<td>17%</td>
<td>22%</td>
</tr>
</tbody>
</table>

Table 1: Cognate percentages

- A number of the coverbs (and other words) in Alawa and Mara show evidence of plausible shifts in meaning consistent with family relationship.
- A number of (masculine) nouns in Alawa that begin with prenasalised stops correspond to nouns in Mara with the neuter gender prefix n- and with stems that begin with prenasalised stops in Warndarang.
- However, no-one has yet unravelled how the finite verb forms are related.

Baker (2006:1) takes this last point as “proof that these languages are, in fact, unrelated at the family level”. He has a valid point. Although one of the

¹ Distant or possible cognates are counted as halves, clear cognates as full units. While the cognate percentage is low between Mangarayi and the other languages, it should be noted that clear cognates between Mangarayi and Alawa are practically the same phonetically, any phonemic difference being due to the recognition of five vowels in Mangarayi as opposed to Alawa’s four. Mara and Warndarang only have three vowels: the fourth in Alawa is shown from my research to be a recent development from an original three-vowel system which would underlie all four languages at an earlier time.
finite verbs seems to me to be related in the four languages, neither I nor anyone else has yet come up with any plausible account of how the various forms developed, let alone any relationship between other stems and their inflections. That does not rule out as yet unrecognised relationships. I recall in early 1960 or 1961 hearing the late Stephen Wurm commenting that the only common feature of New Guinea Highland languages was their enormous complexity. Within a few years after that date, the relationships between them had been established. I would hope that in the near future collaboration between different researchers might crack the problem of the apparent unrelatedness of the finite verbs as well as other tantalising similarities and differences in these languages and others.

Certainly more is now known about other languages of the area than was known when the Sharpe (1972), Heath (1978) and Merlan (1989) publications appeared. Now Baker has a better knowledge than I do of other languages of the area which show shared features with some of these languages, and his questioning is a call to reconsider all the evidence.

It is worthwhile therefore to look at what seems to be shared and what that can tell us, as well as the problematic areas which cast doubt on the relatively close relationship between these languages compared to others of the area or elsewhere. While Baker may be right in saying these languages are not from one linguistic family, the most plausible hypothesis at the present time is that they are related at a considerable time depth, even if a number of their features may have other sources. To me, with my fair knowledge of Alawa, and experience in comparing texts in the different languages, it is clear that in syntax and in how affixes are organised, Alawa, Mara and Warndarang are similar. Mangarayi is not so similar, except for many near identical forms, especially in adjectives. But the system is still the same. In one session by eliciting Mangarayi from an Alawa man through the medium of Alawa, I found it easy to get data on the language2.

In this paper I will focus on two of what seems to be the simplest of the correspondences between the four named languages, that is the relationship which appears to exist between the gender and number prefixes for substantives, and on words and stems which begin with homorganic nasal-stop sequences in one or other language.

2 It is worth noting here as an aside, that when I was in search of a language to study and permission to stay in a community where the language was in use, Arthur Capell was intrigued by the intonation patterns of both Mangarayi and Alawa, wondering if they were tonal languages. They are not, but both have very interesting intonation patterns. Permission to work with Mangarayi was at the time denied by the cattle station owner (tired of researchers he was), and I began on Alawa. Later, when the Mangarayi at Djembere (now Djilkminggan) obtained their own land, Merlan was able to live with them and learn the language.
2. **General comments on these languages**

Mangarayi, Alawa, extinct and unstudied Yugul, Mara and Warndarang were spoken in contiguous areas in the eastern part of the Northern Territory. It is claimed by Aboriginal people of the area that Yugul, spoken along the Roper River near Ngukurr, was 'like' Alawa and Mara, but as we only have a few words, and no grammatical material, this language will not be considered here. Warndarang is now extinct, and those who know Alawa and/or Mara have for many decades been using Kriol more regularly.

It is estimated that traditionally these languages had at most 150-200 speakers each. Territory was owned by exogamous patriclans, and lineages numbering about 40 persons owned various territories. Clan affiliation was of greater importance in defining a person’s social identity than language (Heath 1978:12). In traditional times many of the people would have been multilingual; my Alawa teacher, Heath’s and my Warndarang teacher and many others in the 1960s could converse fluently in more than one traditional language. They also claimed to know whether a commonly used form in Alawa was ‘really’ Alawa or Mara, etc.

The relationship between the non-Pama-Nyungan (NPN) or prefixing languages of North Australia is generally considered to be extremely old, predating the spread of the Pama-Nyungan (PN) languages across 80% of Australia. Nonetheless, there are some widely attested Common Australian stems which are shared by both groups of languages, including not only nga- 1st person singular, but some, mostly monosyllabic, verb stems (either as coverbs or finite verbs in languages which have these). A later study will look at this feature in these languages.

Exogamy and other social ties allowed interaction between and multilingualism among speakers of different patriclans, whether they were PN or NPN. Common sign languages covered much wider territories than the spoken languages, and even today short conversations are carried on by sign. Many flora and fauna names and bird names have wide distributions and sometimes near identical forms across languages that are only remotely related, with scattered geographical distribution over a thousand or more kilometres. (Such words have not been included in my cognate estimation.) Under these conditions, diffusion of many vocabulary items is not surprising, and Heath (1978) established that morphemes smaller than whole words (three noun class prefixes, and the ablative marker) were diffused from Nunggubuyu to Warndarang (both NPN but very distantly related), and between another pair of languages, one PN and one NPN. We therefore cannot rule out diffusion of other features, even though many speakers claim to know which language a particular form belongs to. The near-identical systems in these four languages could also be attributed to multilingualism.

The patterns of relationship between these and other languages is therefore not easy to untangle. On the basis of the construction of the verb phrases and
some morphology, including noun class prefixes (Heath 1978, Merlan 1989) and
pronominal forms (Heath 1978), Sharpe, Heath and Merlan all independently
considered these languages to share a closer relationship with each other than
with other languages. However Merlan states that although Mangarayi shares
evidently archaic features of (semi-productive and frozen) derivational verbal
morphology with Warndarang and Mara, it shares many similarities (in verbal
inflectional categories and in some cases even their morphological expression)
with another subgroup of NPN languages (Merlan 1982:xiii).

3. Enigma variation 1: Gender and Number Prefixes for Substantives

Alawa has two genders, Mara and Mangarayi have three, and Warndarang six
genders. In Alawa, feminine (f) is used for the sun, female humans (and
sometimes for female introduced animals) and adjectives applying to them, and
is marked by a prefix \textit{an}- (absolutive) or \textit{ar}- (oblique) (‘sun’ lacks the prefix). All
other substantives are masculine (m), which is unmarked except in some kin
terms: these have a masculine prefix \textit{na}- (abs) or \textit{a}- (obl). The oblique form
occurring alone signals ergative/instrumental; in combination with suffixes it
signals other cases such as genitive/purposive, locative, allative and ablative.
Substantives unmarked for gender take an ergative case suffix; suffixation for
other oblique cases is identical for all singular substantives. Duals and plurals of
some gender-marked kin terms have the feminine prefix, whatever the natural
gender of the persons, and are \textit{suffixed} rather than prefixed to mark dual or
plural, as in (1).

(1) \textit{an-biŋa-ŋa-yili}
F-‘son’-1SG-PL
“my (brother’s) children”

Mara and Mangarayi have three genders, masculine, feminine and neuter;
al\ll four languages (including Warndarang) use masculine for male humans and
feminine for female humans. In Mara a wide range of animals, natural objects,
ceremonies and some internal body parts are masculine; a few female marsupial
terms are feminine, while most body parts, a few topographic terms, most
abstract nouns, and \textit{n-guguru} ‘sun’ are neuter. In Mangarayi male humans,
higher animals and mythological beings are masculine, feminine is for female
humans and female mythological beings, and other nouns are neuter. In
Mangarayi, masculine and feminine classes have nominative-accusative case
marking, while the neuter class has ergative-absolutive; Merlan (1989:xiii)
considers that, at least for the feminine, the nominative-accusative marking is a

\footnote{In this paper I am using underlining to mark retroflexed consonants, including the rhotic
continuant. \textit{r} is used for the flapped/trilled rhotic. The digraph \textit{ng} is used for the velar nasal, \textit{n,g}
marks the sequence of alveolar nasal and velar stop.}
development from an earlier ergative-absolutive pattern. Warndarang has masculine and feminine for male and female humans respectively, the same gender prefixes for a range of other items, and three other class prefixes for other items, which Heath established were borrowed from Nunggubuyu. Merlan has established the connection between the Mangarayi and Alawa feminine prefix, and between some non-productive forms in some Alawa demonstratives and the neuter gender prefix in Mangarayi. When the free standing pronouns of Alawa are examined, together with one verb prefix (nganu 3f subject acting on 3f object) and the third feminine singular object verb suffixes (-ngVr(V)) they confirm Merlan’s suggestion that most Alawa affixes for 3sgf have lost an initial /ng/, or replaced it with {w}, a morphophoneme which presents itself as /g/ when ‘hardened’, and zero when not. The substantive gender and number affixes in the four languages under consideration are shown in Table 2.

<table>
<thead>
<tr>
<th></th>
<th>Mangarayi</th>
<th>Alawa</th>
<th>Mara</th>
<th>Warnd. noun</th>
<th>Warnd. dem</th>
<th>Suggested proto-form</th>
</tr>
</thead>
<tbody>
<tr>
<td>m abs</td>
<td>na-(nom),</td>
<td>0</td>
<td>na-</td>
<td>na-</td>
<td>*na-</td>
<td>*na-</td>
</tr>
<tr>
<td>m obl</td>
<td>na-, na-</td>
<td>na-</td>
<td>na-</td>
<td>*na-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f nom</td>
<td>nga-</td>
<td>nga-</td>
<td>nga-</td>
<td>*nga-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f acc</td>
<td>nga-</td>
<td>nga-</td>
<td>nga-</td>
<td>*nga-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f abs</td>
<td>nga-</td>
<td>nga-</td>
<td>nga-</td>
<td>*nga-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f obl</td>
<td>nga-</td>
<td>nga-</td>
<td>nga-</td>
<td>* nga-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>n abs</td>
<td>nga-</td>
<td>nga-</td>
<td>nga-</td>
<td>* nga-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>n obl</td>
<td>nga-</td>
<td>nga-</td>
<td>nga-</td>
<td>* nga-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>du</td>
<td>-yəran/-wər-</td>
<td>wur-</td>
<td>yir-</td>
<td>*yir-</td>
<td>*wil(ə), *yil(ə)-</td>
<td></td>
</tr>
<tr>
<td>pl</td>
<td>-yə,ələ/-wə,ələ</td>
<td>yil-</td>
<td>wul-</td>
<td>*wil(ə), *yil(ə)-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Evidence in Table 2 and contained elsewhere in pronouns and verb pronominal prefixes suggests there either were two proto-forms from which the dual and plural markings were derived, or one to mark a ‘large’ plural and another to mark a ‘small’ plural (paucal). Alawa has taken the purported latter and lost the former except in suffixes for 3rd dual and plural in the direferential

---

1 ō in allative, ablative
2 na- erg, obl na-; ō- allative, ablative
3 There are other noun class prefixes, borrowed form Nunggubuyu (Heath 1978)

Table 2: Noun Class and Number prefixes

---

4 Use of the masculine prefix is rare with non-human nouns, but seems to occur chiefly with place names; the feminine prefix occurs with a fair number of fauna terms and for explicitly female kangaroo terms.
verbs; Mangarayi has alternate forms based on both these, Warndarang uses yili- for paucal, but only Alawa uses the yil- form in verbs. Symbols y, and w, stand for morphophonemes which ‘harden’ from semivowels to stops contiguous to a nasal or stop.

All four languages have taken /r/ which marks plural in many other languages of the area and used it for the dual, except where, in direferential verb prefixes, no dual/plural distinction is made. Some other NPN languages suffix various forms to the /r/ plural to form the dual, whereas these languages mark the plural with a lateral. Alawa, Mara and Warndarang are consistent in using the alveolar lateral. Merlan (1989) makes a good case for Mangarayi having lost the /*l/ marking of plural under the importing of *-la-(with retroflexed lateral) to mark nominative, etc. She suggests this -la is the same form that persists in Alawa nula ‘he’, ngandula ‘she’, yirula/yurula ‘they two’ and yilula/yulula ‘they’.

Merlan has argued that most Alawa feminine prefixes have lost an initial velar nasal, which conclusion I also reached. The supporting evidence is from two independent third person singular pronouns and one direferential verb prefix, which have an initial /ng/ in Alawa: ngandula ‘she’, ngadu ‘her (possessive)’, and nganu- 3fs/3fo), Merlan also argued that Alawa lost a pre-existing neuter word class, marked in the protolanguage by *ña-. The occurrence of a feminine form with initial /ng/ is more widespread, but Merlan argued (1989), the 3f forms in the four languages under consideration can be shown to share features not seen in other languages.

Merlan also argued for invariant *ña- as proto-form for all languages for the masculine prefix, and its specialisation to mark absolute in proto-Mangarayi, proto-Alawa and in Warndarang, and to mark oblique in Mara. Later, when Mangarayi moved to a nominative-accusative system in nouns, na- was also used for nominative (as well as – variably – for oblique). While in this group only Alawa marks the distinction between masculine and feminine in 3s pronominal affixes in the verbs, all four languages show these distinctions in the independent pronouns.

4. Enigma variation 2: Word Initial Prenasalised Stops

Alawa as spoken in the 1960s had many words which began with prenasalised stops, and from the patterns of the language these functioned as single phonemes. In slow repetitions, my language teacher Barnabas Roberts, a literate and astute man, also always made a syllable break before a word medial homorganic nasal-stop ‘cluster’. While Warndarang had some stems in common with Alawa that began with prenasalised stops, they carried a gender prefix so that words did not begin with prenasalised stops (with one exception). Heath (1980:9, 1981:9) noted that there was some evidence for prenasalised stops in Mara and Warndarang, but they did not occur word initially as in Alawa. He
concluded that “[c]onsequently we have a borderline case and it is not productive to waste ink and mental energy worrying about whether mb is a unit or a cluster” (Heath 1981:9).

However Heath includes in his Mara dictionary a number of words that do begin with prenasalised stops. There is also one word in his Warndarang dictionary listed with a word initial prenasalised stop. These are included in the list below which shows the stems we have record of with prenasalised stops in at least one of the three languages Alawa, Mara and Warndarang and which in general have cognates in at least one more language of those three.

<table>
<thead>
<tr>
<th>Al, Ma mbagar</th>
<th>honeycomb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al mbaďiriňa</td>
<td>ghost; Ma mbardirňa ghost</td>
</tr>
<tr>
<td>Al mbangara</td>
<td>clear sky (clouds gone, open country); Ma n-bangara (-mbangara) (cloudless) sky</td>
</tr>
<tr>
<td>Al mbįlal</td>
<td>waterlily leaf; Ma n-bįlal (-mbįlal) water lily leaf (dual wur-mbįlal); Wa (wu-)mbįlal water lily leaf</td>
</tr>
<tr>
<td>Al mbul-nala</td>
<td>big mob go away; Ma mbul to load up, bring (many people or things) together</td>
</tr>
<tr>
<td>Al mbulwul</td>
<td>sweat, skin group; Ma n-buwul (-mbuwul) semimoiety, skin</td>
</tr>
</tbody>
</table>

| Al ndagi⁵, Wa (wu-)ndaqi | nape of neck |
| Al ndaway | hole in tree (as of witchetty); Ma n-daway (obl na-yaway), Wa (wu-)ndaway word, language |
| Al dula, Ma n-dula | upper leg, thigh; Wa (wu-)ndula leg |
| Al dun | arise, set out; Ma ndun get up (and set out); Wa dun to get up (and set off), depart |
| Al nduru | shin or lower leg; Ma n-duru leg (only 1 attestation); Wa (wu-)nduru shin (from knee to ankle) |
| Al jagų- | saliva, salivate; Ma jagų saliva, foam on waves; Wa (wu-)ňjagų saliva |
| Al jagiri | red waterlily; Ma jagiri, n-jagiri, Wa (ma-)ňjagirų dark red waterlily fruit |
| Al ńjalųyų | young man just starting beard; Ma ńjalųyų adolescent boy (pl wul-ńjalųyų or wul-ńjalųyų); Wa (mu-)ńjalųyų man |
| Al janur, yenur | head cold; Ma janur light head cold; Wa (wu-)ńjanur cold (disease) |
| Al ńjariňiyi one, ńjari one(?) , ńjari ńjari only a little, one; Ma ńjari many much, sometimes redup ńjari-ńjari; Wa ńjari do repeatedly or continuously, ńjari many, much (possible cognates) |

⁵ Heath in his dictionaries of Mara and Warndarang has written in the retroflexion on initial alveolars, whereas Merlan and I have not. In examples quoted here I have written in this initial retroflexion. However the gender prefix n- in Mara is not retroflexed. Heath has also written what I assume to be a homorganic nasal-stop sequence (or phoneme) as nj in both Mara and Warndarang. I assume this to be the ň sequence or phoneme, as it is in cognate Alawa words.
Al ñjawanda chin, beard; Ma n-jawanda whisker, beard, jaw, chin; Wa (ra-)jawanda beard
Al ñjeyelng tongue; Ma n-jiyil tongue; Wa (wu-)ñjįyilng tongue
Al ñjigu̍r tail of animal; Ma n-ju̍gur tail of dog, fish etc.; Wa (wu-)ñjįgur tail
Al, Ma ñjįjan totem, dreaming
Al ñji̍l̍i bank of river, jili-naļ̍a go along side of
Al ji̍wa widower, an.ji̍wa, ñjiwar widow; Ma ñjiwa charcoal, bereaved spouse (widow, widower); Wa (wu-)ñjiwa charcoal
Al, Ma ji̍wal saltwater tree (Pemphis acidula); Ma ñjiwal shrub sp. w small fruits in early dry season (distinct from jiwal)

Al nggabul-nerneni, Ma, Wa gabul cooked, ripe, ready to eat
Al nggagar, Wa (ra-)nggagar River Oak (Casuarina cunninghamiana)
Al nggawil gorge, space between two sides, gawil-naļ̍a go through; Ma nggawil to go through, n-gawil pass, gap
Wa (ma-)nggawujā rough, irregularly shaped waterlily root
Ma nggilingillnga galah; Al ailinggailni̍n̍ galah
Ma nggumir shrub w edible white fruits; Wa (ra/-wu-)nggumir shrub w juicy black fruits
Ma gurya excrement, n-guriya, n-gurya anus etc.; Wa (ma/-wu-)nggurya excrement

The next list shows words which have word initial prenasalised stops in Alawa or Mara which occur with plain stops in Mangarayi.

Al mbalbalbi, Mg balbalbi slow(ly)
Ma mbal̍na opening, crack, space between two large masses; Mg bal̍na open space
Al njalmar fin, also pectoral fin; Mg jalmar tail of fish
Al ngagu̍ spear without wire, either half made or with sharpened point; Mg gagu̍ spear made out of bamboo
Al ngudaru hill, stone, rock, coins, hailstones, ngudaru-enu; Mg gudaru horn
Al ngulur eye, point, seed, battery of torch, etc.; Mg gulur- ma- be sorry for

Amongst the comparatively few fluent speakers today, most of Alawa’s prenasalisation has been lost, although some persists, particularly after gender or number prefixes. In most instances in Mara, and in all but one Warndarang stem, the prenasalisation only occurred following a vowel final prefix. (The Warndarang exception is ḳįąiqi many, much.) If the suggested relationship put forward in the past by Heath, Merlan and Sharpe is supported, we must look for the development of these subsequent to the separation of proto-Mangarayi from any possible proto-language for the other three, or the more plausible reduction of pre-existing homorganic nasal-stops to stops in proto-Mangarayi.

Despite the slow fading of the word initial prenasalised stops in Alawa,
speakers of Alawa at the main Alawa settlement (Hodgson Downs) were insisting on the form ŋjeyeling for ‘tongue’ some years back, retaining the prenasalised stop. Possibly significantly, one of these is a daughter of the last Warndarang speaker, who was Heath’s (and my) informant.

What else can be said about words with initial prenasalised stops in Alawa?

- Most (but intriguingly not all) are nouns or adjectives or verbs derived from nouns.
- Just under half are body part terms, none are bird names (one of Mara’s words is a bird name).
- A few relate to plants, only one seems to be primarily a coverb.

At one stage of analysis I noted that a number of the Alawa words beginning with prenasalised stops were cognate with Mara words with the neuter prefix, and thought perhaps the Alawa prenasalised stops resulted from assimilation of this prefix to the following consonant. But as the corresponding Warndarang stems had prenasalised stops, this hypothesis had to be rejected. The stem-initial prenasalised stops in Warndarang are ‘protected’ (one might assume) by a preceding borrowed gender prefix (wu-, ga-, ma-). However, if Merlan’s hypothesis that ŋa- in Alawa ŋagul ‘other’ is a frozen relic of the neuter prefix, the neuter prefix may well have been *ŋ- or *ŋa-. Mara seems also to have shifted */ŋ/ to /n/ in some pronominal prefixes. Another history of words or stems which have prenasalised stops must be looked for, and it would suggest that any proto-language contributing (directly or indirectly) to these four languages must have had word initial prenasalised stops. Looking forward, Baker (2007) has postulated, giving evidence, that ‘monogestural’ items, that is, homorganic nasal-stop sequences in Australian languages function as syllable onsets.

5. **A forward look to future research on more Enigma Variations**

Pursuing my analogy of Enigma Variations, we have the following further list of variations:

- I can present the relationships between pronominal affixes on finite verbs on these four languages, and working with others perhaps produce a better account than the one I gave some years ago.
- A lot more can be done in collecting evidence of cognates, not only in these languages but also in others of the area, including PN languages.
- Perhaps we can formulate patterns behind the limited evidence of sound shifts in these languages and others of the area.

---

6 It is nggilinggilinga ‘galah’, cf Al. alinggaligñin.
• On the whole (except for the ablative), nominal case inflection in the four languages under consideration does not show obvious relationship.
• I have begun work compiling intriguing relationships between some coverb forms and meanings in these languages, and calculating cognate percentages in these.
• More work can be done on forms (nouns, coverbs, finite verbs and others) which seem to trace to Common Australian forms.
• And lastly, and most challenging, we can work to establish the prehistory of the various finite verb stems and their inflections.

The famous *Enigma Variations* of Edward Elgar numbered 14. Do we have any advance on the nine I have listed?

**References**

Sharpe, Margaret C. 1972. *Alawa Phonology and Grammar*. Canberra: AIAS.
1. Introduction

The genetic position of Warumungu (spoken in the Northern Territory of Australia) is uncertain. The speakers’ traditional country is surrounded by the countries of speakers of languages from four different language families. To the north are two non-Paman-Nyungan languages, Jingilu and Wambaya, from the Mirndi family. The other neighbours are all Pama-Nyungan. On the west are Ngumbin-Yapa languages; Warlpiri and Warlmanpa are Yapa languages, while Mudburra is Ngumbin. To the south and south-east are Arandic languages, Kaytetye and Alyawarr. To the east is Wakaya, a Warluwaric language.

Pronouns are an obvious place to start looking for relationships between Warumungu and its neighbours. At first glance the Warumungu pronoun system diverges markedly from its neighbours, both in syntactic structure and in morphological shape (Hale, 1959, Heath, 1977, Simpson, 1990). Having both free and bound pronouns (‘double pronoun systems’) is a characteristic of Warumungu’s neighbours, except the Arandic languages, (although the free pronouns in Warlmanpa are reduced to just two forms for first and second person regardless of number). Warumungu has one system of pronouns, which are restricted to second position (1), or when prominent, first position (2). They all start with the vowel /a/, and may end with a future tense inflection -a, or an optative inflection -aparn. The restricted pronouns express subjects of intransitive, transitive or ditransitive verbs, objects, reflexive objects and Dative arguments and adjuncts. Third person singular subject and absolutive object pronouns have no overt forms. Object, reflexive and Dative pronouns never appear separately from subject pronouns, as in (2). The order is normally S-O.²

¹ This paper is inspired by Harold Koch’s work on Kaytetye, on reconstructing Arandic, and on reconstructing Pama-Nyungan, and it’s a great pleasure to dedicate it to him. I thank him, Mark Harvey, Mary Laughren, Pat McConvell, David Nash, the editors and an anonymous referee for helpful comments.

(1a)  \( \text{wartarnp}(a)= \text{arni} \) \( \text{apina} / \)  
always  \( =1S \) was.walking  
“I walked about all the time.”

(b)  \( \text{Apirrkarl}= \text{arna}! \)  
go.away.FUT=1S.FUT  
“I’ll go later.”

(2)  \( \text{ajulajji}=\text{ngini } \text{ warliniyinjina} / \)  
333S.1O=disc.marker were.chasing  
“(When) they chased me.”

In (2) \( \text{ajulajji} \) (3rd plural subject 1st singular object) appears to be made up of \( \text{ajul} \) (333S) and \( \text{ajji} \) (1O), but in isolation 333S is \( \text{ajjul} \), not \( \text{ajul} \) (i.e. with an intransitive verb, or a null third person singular object), and 1O is \( \text{ajju} \), not \( \text{ajji} \), (i.e. with a null third person singular subject).

While the restricted pronouns are disyllabic, with at least a weak stress on the first syllable, and so fit the template for minimum word, their restricted position and the combined structure of S and O makes them more like the bound pronouns in double pronoun systems than like free pronouns, (Mushin & Simpson, submitted). Warumungu also has possessive pronouns which act like nominals in taking case-endings and in not being positionally restricted.

The Warumungu restricted pronouns show considerable regularity in structure, (Heath, 1977). It is likely that they have undergone paradigmatic reshaping. But there are some irregularities, especially in the non-singular forms and in the combined Subject-Object forms. An obvious hypothesis is that pre-Warumungu had two sets of pronouns, free and bound, and then the free pronouns were lost. In this paper I present reconstructions of pronouns at an earlier stage, “pre-Warumungu”. These internal reconstructions are then compared with pronouns in neighbouring languages, and to reconstructions. The comparisons do not result in a conclusive assignment of Warumungu to any of the neighbouring language families, but they do reveal likely connections with, and influence from, Arandic languages and Wakaya.

First I provide some background on phonotactic, morphological, and phonological processes affecting analysis of Warumungu pronouns. Then I analyse the Warumungu pronouns, first the singular pronouns and then the non-singular pronouns. Each analysis begins with conservative internal reconstruction. I then compare the results of the internal reconstruction with synchronic and reconstructed forms for pronouns in neighbouring languages.

2. Contemporary Warumungu pronouns

Factors to consider in reconstructing the pronouns include what grammatical functions and features the pronouns express, and what order the formatives appear in. Warumungu pronouns mostly appear in the order S-O:. However, in
some non-singular forms the order S-O is reversed, and a suffix -ngkki appears, acting as an inversion marker (Heath, 1977), as shown in (3). I shall suggest that -ngkki may be an old Ergative marker.

(3) a. ankkul 1st plural exclusive Subject (111S),
    b. ajjul 3rd plural Subject (333S)
    c. ajurnu 3rd plural Object with zero 3rd singular Subject (3S-333O)
    d. ankuljarni (111S-333O)
    e. ankuljarningkki 3(3)3S-1(1)1O 3.

Apart from the possessive pronouns, the positionally restricted Warumungu pronouns fall into three types: Subject forms (subjects of intransitive verbs and of verbs with third person singular objects) (3a,b, 4a-c), Object forms (direct or indirect objects of verbs with third person singular subjects) (3c, 4de-f), and combined Subject-Object forms for verbs with subjects and objects that are not third person singular (3d-e).

Subject pronominal forms are all disyllabic, as are about half of the Warumungu object pronominal forms with third person singular subjects. The remaining S-O forms consist of a disyllabic stem with an ending -ngkkV, -kkV (4e), -V or -rnV, where V is /i/ or /u/, subject to vowel harmony.

(4) a. ankkul (111S (-3O))  d. ankku (3S - 111O)
    b. amppul (22S (-3O))  e. ampukku (3S - 22O)
    c. arni (1S (-3O))  f. angku (3S - 20)

The combined Subject-Object forms show several irregularities which are significant for reconstructing pre-Warumungu pronouns. Owing to dual neutralisation (Hale, 1973), if both the subject and the object are non-singular, the distinction between dual and plural is neutralised to a non-singular, 1(1)1, 1(2)2, 2(2)2, and 3(3)3. In such complexes the base forms of S and O resemble the plural forms and not the dual forms. Thus (5a,d) show the first exclusive Subject forms (plural and dual). (5b, e) show the third plural and dual Object forms. (5c) shows the form used for a first non-singular Subject operating on a third non-singular Object. The forms have more in common with the plural counterparts than with the dual counterparts.

(5) a. ankkul (111S)  d. ajjil (11S)
    b. ajurnu (3S-333O)  e. apulu (3S-33O)
    c. ankuljarni (1(1)1S-3(3)3O i.e = non-sing. Subject, non-sing. Object

3 In fact, owing to dual neutralisation to be discussed later, ankuljarni and ankuljarningkki represent non-singular, i.e. both dual and plural subjects and objects. I write this as 1(1)1S-3(3)3O and 3(3)3S-1(1)1O.
Modern Warumungu pronouns show two properties which require consideration in reconstruction. The first is that all pronouns start with /a/. This is pronounced [a] when stressed (as in word-initial position) and as schwa or /a/ when weakly stressed (as when cliticised to a preceding word⁴). In Warumungu, initial /a/ is found mostly on closed class items (pronouns, function words and two verb roots), on loanwords from Arandic languages, and on some nouns and preverbs in variation with initial /wa/ or /ya/. So, in internal reconstruction of pronouns, we must consider for each pronoun whether the synchronic presence of /a/ is due to paradigmatic levelling and insertion of /a/ (or change of vowel to /a/), or whether the /a/ is reconstructable to an ancestral form which underwent a phonological process, such as initial consonant loss.

The second is that all non-singular Subject pronouns end with /l/ if there is no overt pronominal object. Again, we must consider whether the synchronic presence of /l/ is reconstructable to an ancestral form, or whether it results from paradigm levelling, or from reanalysis of a bound morpheme /l/.

2.1 Synchronic phonological processes affecting forms

Reconstruction also requires considering the effects of synchronic phonological and morpho-phonological processes.

Vowel elision rule. Across the board, if two vowels end up adjacent, the first vowel is deleted. Since all pronouns start with /a/, the final vowel of the first morph in a Subject-Object form is indeterminate: \(aCV_1+aCV_2\) is realised as \(aCaCV_2\).

Apical dissimilation rule. A suffix beginning with an apico-alveolar consonant is realised as retroflex when attached to stems with a medial apico-alveolar consonant, e.g. the future suffix -l (alveolar lateral) is realised on nyirri- ‘put’ with a retroflex lateral. This alternation is not represented orthographically when roots and suffixes are clearly synchronically isolatable, as \(nyirril\) ‘will put’. It is important when reconstructing combinations of subject-object pronominal affixes with sequences of apical consonants. For example, \(anyuljarni\) (1(2)2S on 3(3)3O) contains a morpheme boundary between \(anyul\) (1(2)2S) and \(-jarni\) (3(3)3O). The juxtaposition creates a sequence \(l.. rni\). The original form of the second morph could thus have been \(-jani\), not \(-jarni\).

The next two rules are morphologically conditioned processes: a fortis/lenis alternation of medial stop consonants and vowel assimilation in suffixes (Heath, 1977), which are linked in that they apply to the same set of suffixes in nouns.

Fortis/lenis alternation. Orthographically, CC represents the fortis, and C represents the lenis, regardless of whether the realisation of the fortis consonant is long or short. The alternation does not apply in morphemes of

---

⁴ Usually the stresses on the clitics are subordinated to those of the preceding word but sometimes the initial /a/ is stressed. This seems to depend on the stresses assigned to the preceding word, but rhythm in Warumungu has not been investigated.
more than two syllables. In disyllabic nouns, the medial consonant is usually fortis (long and voiceless), *kumppu* 'big'. When a case suffix is attached, the consonant is lenis (short and voiced/unaspirated), *kumpu-ku* 'big-DAT'. The few irregular disyllabic nouns with lenis medial consonants can mostly be explained as results of long vowels or reduplications.

If Warumungu pronouns derive from free pronouns, and if free pronouns are a kind of nominal, then we would expect disyllabic pronouns with medial stop consonants to have fortis medial stops, and to show lenition when a suffix is added. A disyllabic pronoun with a medial stop consonant which does not show fortition, (e.g. *aku* (3sg Dative) contrasting with *makku* (meat)), needs explanation.

The initial consonants of case suffixes themselves are short and voiced/unaspirated when added to such two-syllable nouns (written here C despite the fact that they are short	extsuperscript{3}), but short and voiceless/aspirated when added to other nouns (written CC). The most obvious manifestation of the fortis/lenis alternation occurs in nouns, but the fortis/lenis alternation occurs regularly in one of the main verb conjugations. In other parts of speech, including pronouns, more irregularity is involved. An example which fits the regular pattern is the pair *ajjil* (11S), disyllabic with fortis medial CC, *aji-nginyi* (11 possessive), lenis medial C.

**Vowel assimilation.** This applies most predictably with the same set of suffixes involved in the fortis-lenis alternation. The vowel of the suffix is determined by the final vowel of the stem: *kumpu-ku* (big-DAT), *karti-ki* (man-DAT), *kana-ka* (digging.stick-DAT). Some vowel assimilation is found in parts of speech other than nouns, but usually only with high vowels. High vowel alternations are found among pronoun forms: *amppul* (22S), *ampu-kku* (22O), *ajjil* (11S), *aji-kki* (11O).

**Word-final /i/.** Several words ending in /i/ end in /a/ when followed by case suffixes: *watti* 'far' *wattangara* 'from far'. Correspondences exist between word-final /i/ in Warumungu and word-final /a/ in other languages, *Nangali*, *Warlpiri Nangala* (subsection name). These suggest that /i/ may not always be stable.

### 3. Reconstructing pre-Warumungu pronouns

I discuss singular pronouns, dual pronouns and third person plural pronouns. First and second person plural pronouns are mentioned in passing, but their reconstruction requires more space. Modern pronouns are in the Appendix.

---

	extsuperscript{3} Speakers tend not to write the difference between fortis and lenis consonants initially in suffixes.
3.1 Singular pronouns

3.1.1 First person singular. First person singular pronouns are the only Warumungu pronouns where the Subject and Object stems are different.

<table>
<thead>
<tr>
<th>internal reconstruction</th>
<th>subject of IV, or with 30</th>
<th>subject with non-30</th>
<th>subject with 222O</th>
<th>subject with reflexive O</th>
</tr>
</thead>
<tbody>
<tr>
<td>*=(a)rn</td>
<td>=arni</td>
<td>=arn.a.</td>
<td>=arnturrrku</td>
<td>=arnajju</td>
</tr>
</tbody>
</table>

Table 1: First person singular subject

There is no evidence to decide whether the initial /a/ is original or not (hence *(a)rn). No final vowel is reconstructed, as the arnt in arnturrrku suggests a stage where there was originally no final vowel. The /i/ only appears in one form, as subject of intransitive verb or transitive verb with third singular object. The other forms are all complex, with a linking vowel /a/. The periods in arn.a.- indicate that in a form such as arnajuru (1S-333O) it is unclear as to whether the ‘a’ is a realisation of the final vowel of arni (1S) or the initial vowel of ajuru (333O).

I propose therefore the stages *arn - arnV. The trigger for the addition of V could be to create a minimal two-syllabic word, and it could have arisen from a situation similar to that in modern Arandic languages, where most final vowels are schwa. Word-finally, the schwa could be reanalysed as /i/, while word-medially the ambiguity of analysis of the linking /a/ makes it unclear what happens.

The apical stop ‘t’ in *arnturrrku (which is phonetically retroflex following a retroflex nasal) is either due to an apical stop which appears with other second plural objects, or is an old Ergative marker, since the modern allomorph of the Ergative following a nominal stem ending in a retroflex nasal is rtV.

<table>
<thead>
<tr>
<th>internal reconstruction</th>
<th>object (3S)</th>
<th>object (2S)</th>
<th>object (2/3nsgS)</th>
<th>possessive</th>
<th>subject with reflexive O</th>
</tr>
</thead>
<tbody>
<tr>
<td>*=aju</td>
<td>=aju</td>
<td>=ang-aju</td>
<td>=akul-aji</td>
<td>gijinyi</td>
<td>=arnajju</td>
</tr>
<tr>
<td>*=aji</td>
<td>=ang-aji</td>
<td>=aju-aji</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: First person singular object

The fortis consonant jj is circumstantial evidence for reconstructing a disyllabic form with an initial vowel, since fortition operates on the medial consonant of disyllables. No obvious phonological reason exists for the alternation of /u/ and /i/ as the final vowel of the Object. The only common

---

6 The /a/ may belong to the subject pronoun or may belong to the following object pronoun.
7 In modern Warumungu this allomorphy applies to disyllabic stems, not monosyllabic stems.
8 A reviewer has pointed out that case suffixes may also have initial fortis consonants, and so *-ju could be reconstructed. But this would not explain the difference between aju (1O) and aku (3Dat). If both are monosyllabic suffixes, both would be expected to have fortis CC (or both lenis C, depending on what they attach to).
factor appears to be the Subject’s number: /u/ if it is singular, /i/ if it is not. We will see this pattern elsewhere. The alternation between -aju with null subject and -anjju in anganjju (2S-1O) probably reflects the effect of the medial velar nasal; there is sporadic nasal dissimilation and nasal assimilation of clusters in modern Warumungu.

The possessive form is always formed on the basis of the object stem. The ending -(in)i appears throughout the pronoun paradigm - in all duals and 122 and 333 plural (albeit with a -ngi- augment in all duals and 122 and 333 plural). The -(in)i form does not affect fortition: *aju 1O, aijniyi 1possessive, while the -ngi augment does. The fact that the final vowel of the stem changes to /i/, aijniyi, may be evidence for the /i/ being part of the suffix. Counterparts to this ending are found in Arandic languages: 1sg possessive Alyawarr atye-nh, atjinha [a’cinya], and Kaytetye atye-ye-ngge [a’cinya]. The -ye- of Kaytetye thus resembles the supposed ‘i’ of the Warumungu, even though the following nasals are different.

Table 3 compares the reconstructed pre-Warumungu form with reconstructions for neighbouring Pama-Nyungan languages, and some synchronic free and bound pronouns forms.9

<table>
<thead>
<tr>
<th>preWaru</th>
<th>Kaytetye</th>
<th>pArandic</th>
<th>Wakaya</th>
<th>pNgarna</th>
<th>pYapa</th>
<th>pPN</th>
</tr>
</thead>
<tbody>
<tr>
<td>*(a)rn</td>
<td>atye(A),</td>
<td>*aye(A),</td>
<td>*arn</td>
<td>*ngarna</td>
<td>*=rna</td>
<td>*ngathu</td>
</tr>
<tr>
<td></td>
<td>ayenge</td>
<td>(S)</td>
<td></td>
<td></td>
<td></td>
<td>(Erg),</td>
</tr>
<tr>
<td></td>
<td>(ajye-</td>
<td>*aje-</td>
<td></td>
<td></td>
<td></td>
<td>*ngayi</td>
</tr>
<tr>
<td></td>
<td>rne 1nsg</td>
<td>nge)</td>
<td></td>
<td></td>
<td></td>
<td>(Nom)</td>
</tr>
<tr>
<td>*=aju,</td>
<td>atye(nge)</td>
<td>atye(nge)</td>
<td>anh</td>
<td>pWarluwarric</td>
<td>*=ja(Loc,Acc)</td>
<td>*ngatyu</td>
</tr>
<tr>
<td>*=aji</td>
<td>(Dat)</td>
<td>(nge)</td>
<td></td>
<td></td>
<td></td>
<td>(Dat)</td>
</tr>
</tbody>
</table>

Table 3: Local comparison for reconstructed first person singular subject and object (Shaded cells show correspondences). Proto-Ngarna and proto-Yapa share 1sgS correspondences with pre-Warumungu *=arn, and proto-Arandic shows a related form *=erne as an augment. This is an areal difference from the pPN first person singular forms. The lack of final vowel in the pre-Warumungu corresponds most closely with Wakaya, and with the schwa vowel in Arandic.


10 The alternation -ji~ju in Yapa languages is phonologically conditioned by preceding high vowels.
Proto-Arancid and proto-Yapa share 1sgO correspondences with pre-Warumungu *=aju, *=aji, and presumably derive from the pPN Dative *=ngatyu. The pre-Warumungu forms in *=aju, *=aji correspond with the Arandic forms in having the initial vowel, but with proto-Yapa in lacking the augment –nge.

With respect to our earlier question of initial consonant dropping versus /a/ insertion, proto-Ngarna *=ngarna suggests initial velar nasal dropping for the *(a)rn stem, and pPN *=ngatyu suggests the same for the *=aju stem.

3.1.2 Second person singular. 2sg pronouns form the Object based on the Subject stem.

<table>
<thead>
<tr>
<th>Second person</th>
<th>internal reconstruction</th>
<th>subject with 3O</th>
<th>subject with non-3O</th>
<th>subject with reflexive O</th>
</tr>
</thead>
<tbody>
<tr>
<td>*=ng</td>
<td>~angi</td>
<td>~ang-a</td>
<td>~anqurnu</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Second person singular subjects

<table>
<thead>
<tr>
<th>Second person</th>
<th>internal reconstruction</th>
<th>object (3S)</th>
<th>object (1S)</th>
<th>obj, (1plS)</th>
<th>object (3nsgS)</th>
<th>object (1duS)</th>
<th>possessive</th>
</tr>
</thead>
<tbody>
<tr>
<td>*=ng+ku (sgS)</td>
<td>~ang.ku</td>
<td>~arn-ang.kku</td>
<td>~X-urn.kku</td>
<td>~X-irn.kki</td>
<td>angkinyi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*=N+ku (nsgS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*=N+ki (nsgS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Second person singular objects

(N stands for an apical nasal whose place of articulation is indeterminate because of possible apical dissimilation, X stands for a subject form.)

Two stems are reconstructed for 2sg, a stem with an initial velar nasal found in the Subject and singular Objects, and a stem with an initial apical nasal found in the combined forms with non-singular Subjects. In these latter forms, the presence of morpheme-final /l/ (e.g. ajilirnkkii) means that apical dissimilation may be operating, and so only an apical N can be reconstructed. Vowel harmony determines the high vowels in the combined forms X-urn.kku, X-irn.kki.

Subject *=ng without initial or final vowel is reconstructed on the basis of the *=ngku Object form. The +ku in 2O and 3O is presumably the same as the Dative –ku in the nominal case-marking system (and note the parallel with pPN Dative ngatyu as the basis for the 1sgO form). So the form ang+ku suggests that at an earlier stage there was no final vowel i, just as there was not in the first person singular *=rn. The lack of fortis consonant in angku and aku suggests there was no initial vowel. The possessive builds on the *=ngku Object form, but shows the same change of vowel to /i/ seen with 1sg possessive.

11 ajilirnkkii (11S-2O), ankulurnkku (111S-2O), awałurnkku (33S-2O), ajulurnkku (333S-2O)
Table 6a: Local comparison – reconstructed 2nd person singular, apical nasal stem

<table>
<thead>
<tr>
<th></th>
<th>pre-War</th>
<th>Kaytetye</th>
<th>pArandic</th>
<th>Wakaya</th>
<th>pNgarna</th>
<th>pYapa</th>
<th>pPN</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>*=ng</td>
<td>nge (S)</td>
<td>*(un).nge</td>
<td>-(y)imb</td>
<td>*(yinda)</td>
<td>*=n</td>
<td>*nyun</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>*(unte)</td>
<td></td>
<td></td>
<td></td>
<td>(Nom)</td>
</tr>
<tr>
<td>20</td>
<td>*=ng+ku</td>
<td>ngke.nge</td>
<td>ungke(nge)</td>
<td>/engke(nge)</td>
<td></td>
<td>*=ngku</td>
<td></td>
</tr>
<tr>
<td></td>
<td>*=N+ku</td>
<td></td>
<td>Pre-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Arandic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>*=N+ki</td>
<td></td>
<td>*nyunku</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>*=ng(V)-</td>
<td>ngkeyenge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pos</td>
<td>iny1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The *=n stem is found in Yapa languages and Arandic languages, and is reconstructable to pPN.

Table 6b: Local comparison for reconstructed 2nd person singular – velar nasal stem

<table>
<thead>
<tr>
<th></th>
<th>pre-War</th>
<th>Kaytetye</th>
<th>proto-</th>
<th>Wakaya</th>
<th>pNgarna</th>
<th>pYapa</th>
<th>pPN</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>*=ng</td>
<td>nge (S)</td>
<td>*(un).nge</td>
<td>-(y)imb</td>
<td>*(yinda)</td>
<td>*=n</td>
<td>*nyun</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>*(unte)</td>
<td></td>
<td></td>
<td></td>
<td>(Nom)</td>
</tr>
<tr>
<td>20</td>
<td>*=ng+ku</td>
<td>ngke.nge</td>
<td>ungke(nge)</td>
<td>/engke(nge)</td>
<td></td>
<td>*=ngku</td>
<td></td>
</tr>
<tr>
<td></td>
<td>*=N+ku</td>
<td></td>
<td>Pre-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Arandic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>*=N+ki</td>
<td></td>
<td>*nyunku</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>*=ng(V)-</td>
<td>ngkeyenge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pos</td>
<td>iny1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The 2S *=ng stem is closest to Arandic languages. 20 *=ng+ku corresponds to Arandic and Yapa languages. However, Koch reconstructs pre-Arandic *nyunku via assimilation in place, and so ultimately from pPN. At an earlier stage in Warumungu, it is possible that the original alveolar nasal of *=n+ku assimilated to *=ng+ku, except following an apical /l/ when the alveolar had become retroflex through apical dissimilation.

With respect to the question of initial consonant dropping versus /a/ insertion, the reconstructions suggest /a/ insertion.

3.1.3 Third person singular. Third person singular is normally unmarked, except for a Dative form which corresponds to the -ku marking Object in the second singular, and to the Dative found in the nominal case system. The third person Subject pronoun ama has the same positional restrictions as the restricted pronouns, but does not enter into morphological composition with other forms, and hence is not treated here.
Table 7: Third person singular forms

The lenis consonant /k/ suggests there was no initial vowel. The form *=ku is identical to one allomorph of the Dative case on nominals. But, unlike the nominal Dative which undergoes vowel assimilation when following /a/ (kana-ka ‘for the digging-stick’), the final vowel is not determined by vowel assimilation to the /a/. However, when attached in compound pronouns, there is vowel assimilation anyulukku (122S-3Dat.Obj), ayilikki (125-3Dat.Obj). This suggests that the /a/ is added to ensure that aku has the two syllables necessary to be a minimum word.

The possessive builds on the Dative form, but undergoes a change of vowel to /i/. The reflexive may also build on the Dative form if the /w/ is assumed to result from an otherwise unattested lenition of /k/ to /w/.

Table 8: Local comparison for reconstructed third singular pronouns

The *=ku corresponds to the proto-Arandic form, and, more distantly to the Wakaya form, and the proto-Pama-Nyungan form. However, with respect to initial consonant dropping versus /a/ insertion, the reconstructions suggest /a/ insertion.

3.2 Non-singular pronouns

3.2.1 Dual pronouns. Dual pronouns show some properties not found in other pronouns. They are more transparent than plural pronouns, since, because of dual neutralisation, there are no complex object forms for duals. Some at least may be more recent additions to the Warumungu pronoun system.

Table 9: Dual pronouns
The final /l/ appears on plural pronouns too, and is a non-singular subject marker. /l/ is part of the stem for ayil, and awul, since it appears in the reflexive and possessive of these forms. Since it does not appear in the possessive for ajil and amppul and is in variation for the reflexive, I assume that it is not part of their stems, but rather is a non-singular subject marker.

The fortis/lenis alternation in ajil and amppul is circumstantial evidence for the presence of the initial /a/ vowel. There is no synchronic motivation for the /w/ ~ /p/ alternation in 33S and 33O, although lenition is apparent in the 3Dative aku, reflexive awurnu. The fact that apulu has no fortis counterpart is not evidence for lack of initial vowel because it is a three syllable form (although possibly derived through addition of an object marker /u/), and trisyllabic stems are not subject to the fortis-lenis alternation.

It is likely that the -kku, -kki on the 11O and 22O forms, the -ngkki of the 12O form and the -u of the 33O form are all old object or Dative markers. It is not clear to why -kku, -kki and -ngkki are fortis (short voiceless); their nominal counterparts would be lenis following a regular disyllabic stem. I have found no Warumungu-internal reason for the absence of the proposed stem final /l/ in =ayingkki, and in its plural counterpart =anyul, =anyungkku, or for the use of the =ngkkV rather than =kkV. (There is no other evidence for assimilation of the lateral /l/ to /k/, and /lk/ is a possible cluster in present-day Warumungu).

Considering the fact that the singular possessive forms are built on the object forms, it is arguable that the *-(i)ngi augment in all four dual possessive forms is an equivalent of the proposed object marker *-ku, -ngkki and -u of the 12O form. Since it causes the final vowel of the stem to become /i/, it is tempting to relate it to the Kaytetye -yenge possessive form.

<table>
<thead>
<tr>
<th>pre-War</th>
<th>Kaytetye</th>
<th>Pre-Arandic</th>
<th>Wakaya</th>
<th>pNgarna</th>
<th>pYapa</th>
<th>pPN</th>
</tr>
</thead>
<tbody>
<tr>
<td>12S</td>
<td>*=ayil</td>
<td>ayle-</td>
<td>*=ngali</td>
<td>ngali,</td>
<td>*=ngali</td>
<td>*=li</td>
</tr>
<tr>
<td></td>
<td>*=ajil</td>
<td>ayle.rne</td>
<td>*=ngali.rna,</td>
<td>ngali.r,</td>
<td>ngatharra</td>
<td>*=ngali</td>
</tr>
<tr>
<td>11O</td>
<td>*=ajil</td>
<td>ayle.rne</td>
<td>*=ngali.rna,</td>
<td>ngali.r,</td>
<td>ngatharra</td>
<td>*=ngali</td>
</tr>
<tr>
<td>22S</td>
<td>*=ampu+</td>
<td>mpwele</td>
<td>*=ngumpVLV</td>
<td>yibul,</td>
<td>*nimba</td>
<td>*=n-pVL</td>
</tr>
<tr>
<td></td>
<td>*=ampu+</td>
<td>mpwele</td>
<td>*=ngumpVLV</td>
<td>yibul,</td>
<td>*nimba</td>
<td>*=n-pVL</td>
</tr>
<tr>
<td>33S</td>
<td>*(a)wul</td>
<td>elwe</td>
<td>*=pula</td>
<td>yawul,</td>
<td>*(ya)bula ?</td>
<td>*=pVL</td>
</tr>
<tr>
<td></td>
<td>*=pul+</td>
<td>elwe</td>
<td>*=pula</td>
<td>yawul,</td>
<td>*(ya)bula ?</td>
<td>*=pVL</td>
</tr>
</tbody>
</table>

Table 10: Local comparison for reconstructed dual pronouns

12 Cf. Warlmanpa, Gurindji =ja
Looking first at the inclusive/exclusive distinction, it seems that the inclusive form is shared across all the languages and proto-PamaNyungan. But the Warumungu form looks like a broken form of the Kaytetye prepalatalised form first dual form ayle (which is a regular descendant from pPN *ngali) and is the base of inclusive and exclusive, and so the Warumungu form is possibly a borrowing from Kaytetye. Kaytetye also provides interesting possibilities for the otherwise unexplained -ngkki, -kki forms of the object. In Kaytetye, the form of the object marker on pronouns has a three-way split depending on same or different kinship group, and same or different generation group. The form ayle has variously the suffixes -we, -wake, -wanthe, (1 dual inclusive), and -ngke, -lake, -kanthe (1 dual exclusive). It is possible that the Warumungu alternation is a remnant of such a system. If the Warumungu form is a borrowing, then the question of initial consonant dropping versus /a/ insertion is irrelevant, since the Kaytetye initial vowel [e] is easily converted to /a/ in Warumungu.

Unlike Kaytetye and other Arandic languages, the Warumungu exclusive ajiil is not transparently related to the inclusive ayil. The /j/ of the exclusive could be related to the /j/ found in the Proto-Yapa dual suffix -jarra, which is hypothesised as the source of the Warlmanpa and Gurindji inclusive =ja, similar to the proto-Ngarna form *ngatharra. Warumungu, like proto-Ngarna, Warlmanpa and Gurindji, has two apparently distinct stems. The pre-Warumungu inclusive *=aji+ may be related to a form like *-ja, via *=aj. This would require the same addition of final /i/ found with first and second person singular.

The second person dual forms show strong resemblances across all the languages. However, one striking shared property between Warumungu and Kaytetye is the absence of the /l/ in the Object form. As Table 11 shows, this feature is not shared with Kaytetye’s Arandic relative Alyawarr. Warumungu reflexive ampurnu and possessive ampinginyi resemble Kaytetye mpewe-wenhe and mpewe-yenge.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Reflexive</th>
<th>Object</th>
<th>Possessive</th>
<th>Dative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alyawarr</td>
<td>mpwel</td>
<td>mpwel-ke-nh</td>
<td>mpwele-nh</td>
<td>mpwele-w/ mpwele-k</td>
</tr>
<tr>
<td>Kaytetye</td>
<td>mpwele</td>
<td>mpweye-wenhe</td>
<td>mpwele</td>
<td>mpwele-yenge</td>
</tr>
</tbody>
</table>

Table 11: Second person dual pronouns in Kaytetye and Alyawarr

The third person dual forms also show strong resemblances across all the languages. While the form of the Warumungu subject pronoun is identical to that of Wakaya, a striking shared property between Warumungu and Kaytetye is the presence of the final /u/ vowel as the object marker, as in pre-Warumungu
*±pul-u, and Kaytetye elwewe. Considering the pPN form *pula, Table 10 shows, as a referee has pointed out, evidence for insertion of initial /a/ in Warumungu.

The similarities between the dual pronouns across the languages especially between Warumungu, Wakaya and Kaytetye, the similarity of the possessive form to that of Kaytetye, the relative transparency of the Warumungu dual forms and the fact that they do not appear with non-singular objects, all suggest that some of the Warumungu dual forms may have developed more recently, perhaps even through borrowing from its neighbours.

3.2.2 Third person plural pronouns. Finally, we turn to the third person plural pronouns, because the reconstructions of these illustrate several points that have arisen earlier, most notably the presence of two stems, the effect of apical dissimilation, and the correspondence of final /i/ to /a/ elsewhere.

<table>
<thead>
<tr>
<th>internal reconstruction</th>
<th>subject</th>
<th>subject (1msgO)</th>
<th>reflexive</th>
<th>possessive</th>
</tr>
</thead>
<tbody>
<tr>
<td>3335 *±aju-l</td>
<td>=ajul</td>
<td>=anyuljarningkki</td>
<td>=ajurnu,</td>
<td>ajurnunginyi</td>
</tr>
<tr>
<td>-jaNi-ngkki</td>
<td></td>
<td>=ankuljarningkki</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>internal reconstruction</th>
<th>object (3S)</th>
<th>object (1S)</th>
<th>object</th>
<th>object</th>
</tr>
</thead>
<tbody>
<tr>
<td>3330 *=aju+rnu</td>
<td>=ajurnu</td>
<td>=arnajurnu</td>
<td>=ajuljarni</td>
<td>ankuljarni</td>
</tr>
<tr>
<td>(sgS)</td>
<td></td>
<td>(3(3)S)</td>
<td>(11(1)S)</td>
<td></td>
</tr>
<tr>
<td>*-jaNi (msgS)</td>
<td>=akuljarni</td>
<td>=anyuljarni</td>
<td>(2(2)S)</td>
<td>(12(2)S)</td>
</tr>
</tbody>
</table>

Table 12: Third person plural subject and object pronouns

The third person plural has two stems for S and O: *=aju and *-jaNi (the indeterminacy of the apical is due to the presence of /l/ in the preceding morphemes). The stem *=aju is reconstructed because of the variable absence of the /l/ in the third person reflexive and possessive forms, and because the fortis /j/ in ajjul is circumstantial evidence for the presence of the initial vowel. This stem is used for subjects, and for objects when the subject is singular.

The stem -jaNi is used for objects with non-singular subjects, or for inversion subjects with first or second person non-singular objects. As an inversion subject -jaNi has a suffix -ngkki added. Synchronically this acts as an inversion marker, but historically it may have been an Ergative marker, since it is identical in form (except for the fortis consonant) with the modern Ergative allomorph on disyllabic vowel-final words, -ngkV (the final vowel is determined by vowel assimilation). As to why *-jaNi, and no other pronoun, has Ergative, this is not unexpected given the principle that in a split system Ergative case is more likely to be assigned to third person than to first or second person, and more likely to be assigned to a plural than to a singular.

The augment -ngi in the third plural possessive is that found in the dual pronouns and the first inclusive plural possessive anyinginyi (cf anyul 122S).
The 333 reflexive form has two variants, depending on whether the stem is analysed as ending in /l/ or not. The 333O (with singular subject) and 333 possessive forms are the same as the reflexive form. This suggests that perhaps the reflexive form itself is a generalisation of the object form. I have no explanation for this collapse, other than to suggest that the system was undergoing reshaping.

<table>
<thead>
<tr>
<th>pre-War</th>
<th>Kaytetye</th>
<th>Pre-Arandic</th>
<th>Wakaya</th>
<th>pNgarna</th>
<th>pYapa</th>
<th>pPN</th>
</tr>
</thead>
<tbody>
<tr>
<td>333S</td>
<td>*aju</td>
<td>atangke</td>
<td>yal, =al</td>
<td>*yalu</td>
<td>*-lu</td>
<td>*tyana</td>
</tr>
<tr>
<td></td>
<td>+l</td>
<td>atak</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>333S</td>
<td>*=jaNi+ngki</td>
<td>atanthe</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>333O</td>
<td>*=aju+rnu</td>
<td>atewanthe</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>333O</td>
<td>*=jaNi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 13: Local comparison for reconstructed third plural pronouns

The third person stem *=jaNi has correspondences in final /a/ in Pre-Arandic, proto-Yapa and proto-Pama-Nyungan *tyana. As in Arandic languages and Yapa languages, it can be used for O, but it can only be used for S in an inversion context. It may have undergone the same change that we saw with arni, angi and possibly ajjil, and so the final /i/ is not original.

The source of the *aju stem is less clear. However, the /l/ form which is generalised in Warumungu still further as the /l/ on all non-singular subjects may relate to the *=lu third person plural of proto-Ngarna and proto-Yapa. In proto-Ngumbin-Yapa this is generalised to first exclusive plural, and in many of the daughter languages to second plural as well. The use and behaviour of *=lu as a plural subject clitic is argued to be an innovation in proto-Ngumbin-Yapa (McConvell and Laughren, 2004). The Warumungu /l/ could be an extension to duals as well. Factors promoting this would have been the presence of /l/ in the right position in duals reconstructable to pPN *pula and *nhumpVlV, and in the borrowed Kaytetye ayle.

4. Conclusion

The singular, dual and plural pronouns show some similarities and differences in development. Similarities include:

- The addition of final /l/ to all non-singular pronouns. This /l/ may derive from the plural subject (V)lu found in proto-Ngarna and proto-Yapa.
- The appearance of initial /a/.
- The process whereby final *a is realised as /i/, possibly via a stage where the /a/ is unstressed (1S and 2S, 11S/O, 333S/O).
**(V)rn(V) -> *(V)rn(o) -> arni  1S
*ng -> *ng(o) -> angi  2S
*jaN(o) -> *jaNi  333S/O
? *(V)(o) -> *(a)ji +l/ki  12S/O

- The formation of possessives by adding -(i)nyi and, in the dual and two plural pronouns, an augment (i)ngi as well: -(i)nginyi. These augments most closely resemble forms in Arandic languages, and may reflect a later development. More remote are the Western Desert object form -nya, and Nyangumarta object -(i)nyi.
- The formation of reflexives by adding -(V)rnV to the subject form

Several pronouns have differences between object forms depending on whether the subject is singular or not, summed up in Table 14.13

<table>
<thead>
<tr>
<th></th>
<th>singular S</th>
<th>non-singular S</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>-aju</td>
<td>X-aji</td>
</tr>
<tr>
<td>20</td>
<td>-ng-ku</td>
<td>-N-kV</td>
</tr>
<tr>
<td>333</td>
<td>aju+rmu</td>
<td>jaNi</td>
</tr>
<tr>
<td>inversion</td>
<td>-ngki</td>
<td></td>
</tr>
</tbody>
</table>

Table 14: Contrasting object forms with singular and non-singular subjects

The shape of the modern Warumungu pronoun system is determined by the interplay between positional restrictions which resemble those of bound pronoun systems, the requirement for pronouns to be minimal words, that is, disyllabic forms, and paradigmatic reshaping. Some of the pronouns descend from forms which were mostly likely original bound pronouns. Creating minimal words has led to a phonological change - in the singular pronouns this has required adding /a/ to some, e.g. *ku -> aku 3sgDat. The initial /a/ in other pronouns may have resulted from initial consonant dropping, for example ajju (10). Paradigm levelling is probably responsible for the insertion of initial /a/ in some words, such as apulu (330).

It is possible that an absence of bound pronouns for certain numbers and persons have led to the development of new forms for some of these. This is suggested by the regularity of the dual and inclusive forms, the restriction of dual object forms to those with singular subjects only, and the strong resemblances between the dual forms and those in Kaytetye and Wakaya.

13 Similar differences are found with 1110 and 2220.
References


Hale, Kenneth L. 1959. *Warumungu notes*. Ms, copy held at AIATSIS, Canberra.


Laughren, Mary. in prep. First and second person singular pronouns: The shared past of Nyungic and Warluwarric languages. Ms, University of Queensland, Brisbane.


Appendix 1: Pronouns

This excludes extra suffixes for dative or benefactive or future, or optative.

<table>
<thead>
<tr>
<th></th>
<th>Subject</th>
<th>Object (3sg Subject)</th>
<th>Dative (3sg Subject)</th>
<th>Reflexive</th>
<th>Possessive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>arni</td>
<td>ajju</td>
<td>ajju</td>
<td>arnajju</td>
<td>ajiinyi</td>
</tr>
<tr>
<td>2</td>
<td>angi</td>
<td>angku</td>
<td>angku</td>
<td>angurnu</td>
<td>angkinyi</td>
</tr>
<tr>
<td>3</td>
<td>[ama] or NULL</td>
<td>NULL</td>
<td>aku</td>
<td>awurnu</td>
<td>akiinyi</td>
</tr>
</tbody>
</table>

Table 1: Singular forms
<table>
<thead>
<tr>
<th>Subject</th>
<th>Object (includes Dative)</th>
<th>Reflexive</th>
<th>Possessive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg.S-&gt;sg.O</td>
<td>1st person object</td>
<td>2nd person object</td>
<td>3rd person object</td>
</tr>
<tr>
<td>1sg.S-&gt;du.O</td>
<td>arnangkku</td>
<td>arni</td>
<td>arn</td>
</tr>
<tr>
<td>1sg.S-&gt;pl.O</td>
<td>arnturrrkku</td>
<td>arnajurnu-arnaju</td>
<td>arnu</td>
</tr>
<tr>
<td>1du.exc.-S-&gt;sg.O</td>
<td>ajilirnkki, ~ajilinkki</td>
<td>ajil</td>
<td></td>
</tr>
<tr>
<td>1du.inc.S-&gt;sg.O</td>
<td>ayil</td>
<td>ayil</td>
<td></td>
</tr>
<tr>
<td>1pl.exc.S-&gt;sg.O</td>
<td>ankulurnkku</td>
<td>ankul</td>
<td></td>
</tr>
<tr>
<td>1pl.inc.S-&gt;sg.O</td>
<td>anyul</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1nsg.exc.S-&gt;nsg.O</td>
<td>alkurntukku, alktukku</td>
<td>anyuljarni</td>
<td></td>
</tr>
<tr>
<td>1nsg.inc.S-&gt;nsg.O</td>
<td>anyuljarni</td>
<td>anyuljarni</td>
<td></td>
</tr>
<tr>
<td>2sg.S-&gt;sg.O</td>
<td>angajju, anganijju</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2sg.S-&gt;du.O</td>
<td>angajikki (t exc. du O)</td>
<td>angapulu</td>
<td></td>
</tr>
<tr>
<td>2sg.S-&gt;pl.O</td>
<td>angankku (t exc. pl O)</td>
<td>angajurnu,</td>
<td>angajunu</td>
</tr>
<tr>
<td>2du.S-&gt;sg.O</td>
<td>ampulajji</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2pl.S-&gt;sg.O</td>
<td>akulajji</td>
<td>a(rr)kulkan</td>
<td></td>
</tr>
<tr>
<td>2nsg.S-&gt;nsg.O</td>
<td>a(rr)kularnkki (1exc.nsg.O)</td>
<td>akuljarni</td>
<td></td>
</tr>
<tr>
<td>3sg.S-&gt;sg.O</td>
<td>ajuu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3sg.S-&gt;du.O</td>
<td>ajikki (1exc.du O)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3sg.S-&gt;du.O</td>
<td>ayingkki (1inc.du O)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3sg.S-&gt;pl.O</td>
<td>anyukku (1exc.pl O)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3sg.S-&gt;pl.O</td>
<td>anyukku (1inc.pl O)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3du.S-&gt;sg.O</td>
<td>awulajji</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3pl.S-&gt;sg.O</td>
<td>ajulajji</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3nsg.S-&gt;nsg.O</td>
<td>ankuljarningkki (1exc.pl O)</td>
<td>ajurntulku, ajultuku</td>
<td></td>
</tr>
<tr>
<td>3nsg.S-&gt;nsg.O</td>
<td>anyuljarningkki (1inc.pl O)</td>
<td>ajuljarni</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Combined forms
PART II
RECONSTRUCTION
SPLITTING VS. LUMPING IN MORPHOLOGICAL ANALYSIS
EVIDENCE FROM GREEK*

AYER ANDREWS
Australian National University

1. Introduction

A standing question in morphological analysis is whether forms should be split up into the smallest units for which any kind of case can be made, at the expense of more complex rules governing how these units can be combined ('splitting'), or whether they should be divided into larger, fewer, and possibly easier to manage units ('lumping'). Here I will adduce some evidence from the history of Greek verb inflection that suggests that fairly aggressive splitting is the correct approach. The two phenomena I will consider are the history of an /s/ formative that arguably meant '2nd person singular' for at least two millennia (and probably just means '2nd person' now), and a /t~sth/ contrast that emerged to code a distinction between active and medio-passive voice in certain forms in Greek verbal paradigms. Since these are both small formatives, for which the evidence was at times not overwhelming, and which appeared in forms that would be quite complex in analyses that recognised them, evidence that they really existed as independent units is evidence that splitting should be favored over lumping.

The issue of splitting versus lumping holds across a wide variety of contemporary morphological theories. In paradigm-based approaches such as Anderson (1992) or Stump (2001), it corresponds to using more rules making smaller changes as opposed to fewer making larger ones. In Distributive Morphology (Noyer (1997) and much additional literature), it would correspond to more vocabulary items each spelling out fewer features, as opposed to fewer spelling out more.

An issue I won’t consider here is exactly what kind of theory of morphology learning would favor splitting over lumping, or vice-versa. Working out explicit methods for deteremring segmentation is hard (Manning 1998), and depends partly on the logically prior question of what kinds of segmentation we should accept as correct in the first place. It seems to me that the rather fine splittings argued for below are sufficiently noteworthy to deserve corroboration from as many directions as possible.

---

*I am indebted to two anonymous referees for comments, to Peter Svenonius for discussion and encouragement, and to Harold for book loans and reading suggestions. All errors remain my own.
2. /s/ means ‘you’

2.1 The Appeal of Lumping

The primary and secondary medio-passive verb desinences in Homeric Greek were:

<table>
<thead>
<tr>
<th></th>
<th>primary</th>
<th>secondary</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>-mai</td>
<td>-mE:n</td>
</tr>
<tr>
<td>2sg</td>
<td>-(s)ai</td>
<td>(s)o</td>
</tr>
<tr>
<td>3sg</td>
<td>-tai</td>
<td>-to</td>
</tr>
<tr>
<td>2dl</td>
<td>-sthon</td>
<td>-sthon</td>
</tr>
<tr>
<td>3dl</td>
<td>-sthon</td>
<td>-sthE:n</td>
</tr>
<tr>
<td>1pl</td>
<td>-metha</td>
<td>-metha</td>
</tr>
<tr>
<td>2pl</td>
<td>-sthe</td>
<td>-sthe</td>
</tr>
<tr>
<td>3pl</td>
<td>-ntai</td>
<td>-nto</td>
</tr>
</tbody>
</table>

Table 1

In the 2nd person singular row we see an optional /s/, which was always present after consonants (in the perfective indicative, for example, where there were no mood-markers or thematic vowels), but usually absent after vowels, always absent after the thematic vowel, or optative or subjective mood markers, and sometimes absent and sometimes present after the stem-vowels of athematic verbs including perfects (for example /dunasai/ ‘you are able’, but /dizdE:ai/, not */dizdE:sai/ ‘you seek’; /memnE:(s)ai/ ‘you remember’.

A splitter would recognise an /s/ formative meaning ‘2sg’ in these paradigms, but a lumper might challenge this on various grounds. First, there isn’t any segmentation of the desinences that holds for all the rows. In 1-3sg and 3pl we seem to have a person-number complex followed by a primary/secondary marker (which could be treated as irregularly spelled out as /E:n/ in the secondary 1sg), which can plausibly be regarded as ‘tense’, while in the other forms, this breaks down. The nature and history of the active singular forms is too unclear for them to reveal much, but in the dual and plural we have:

<table>
<thead>
<tr>
<th></th>
<th>primary</th>
<th>secondary</th>
</tr>
</thead>
<tbody>
<tr>
<td>2dl</td>
<td>-ton</td>
<td>-ton</td>
</tr>
<tr>
<td>3dl</td>
<td>-ton</td>
<td>-tE:n</td>
</tr>
<tr>
<td>1pl</td>
<td>-men</td>
<td>-men</td>
</tr>
<tr>
<td>2pl</td>
<td>-te</td>
<td>-te</td>
</tr>
<tr>
<td>3pl</td>
<td>-si</td>
<td>-(sa/e)n</td>
</tr>
</tbody>
</table>

Table 2

---

1 The primary endings were used in subjunctive and present-tense indicative forms, the secondary in optative and past-tense indicative forms. The desinences appear after the stem, mood-markers and various other material. In the form-citations, I use ‘E:’ and ‘O:’ to represent the original long vowels rendered as eta and omega in standard Greek orthography.
The 1st and 3rd plural forms are also hard to segment in a simple and useful way, but in the other forms we can recognise /t/ vs. /st/ as a voice-marker, with the rest of the ending as a combined person-number-tense marker. The net effect is that if we want to recognise the /s/ as a formative, we’re going to need some reasonably complex morphotactic rules to get the expressions of the various relevant features into the correct positions in the paradigm.

The other problem is that /s/ with this meaning is not all that widespread in the medio-passive. Most verb forms have a thematic vowel and/or mood-marker, and with these the /s/ is always missing. In some pre-Homeric earlier stage of the language, the /s/ would have been missing more often, since it disappeared due to a highly regular sequence of sound changes converting prevocalic /s/ to /h/, and then deleting the /h/ intervocalically.

The other place where /s/ appears meaning /2sg/ is in active paradigms. The secondary singular desinences are /n/, /s/, Ø, clearly original; in the primary we have Ø/mi, s/stha/Ø, Ø/si (alternates given in order of frequency), with /s/ by far predominant but not original (it should have been deleted by the sound changes, and is indeed absent with some verbs, such as ‘be’ and ‘go’, but has been restored with all other forms. Although /s/ meaning ‘2sg’ is quite widespread in the active, a lumper could plausibly contend that this is irrelevant to the mediopassive, because there isn’t any common scheme of segmentation that works cleanly across both kinds of paradigms.

Therefore, somebody using an explicit morphological theory might well find it quite plausible to simply have a rule adding each desinence to the appropriate stem, rather than trying to have individual rules placing /m/, /s/, /ai/, etc. in the appropriate places.

2.2 Why it would be wrong

In spite of its appeal, the historical development of the language seems to indicate that a lumping-style analysis would be wrong. We’ve already seen some indication of the problem: the apparent restoration of the /s/ in certain forms such as /dunasai/, /memnE:(s)ai/ in Homeric Greek. Although Homeric Greek doesn’t appear to be a real spoken language, but rather an artificial Kunstsprache combining the features of a number of dialects, it seems plausible to take its treatment of the /s/ in 2sg mediopassive forms as characteristic of the state of the language for a period after the loss of intervocalic /s/.

We can’t be entirely sure that the /s/ wasn’t retained in some of these positions, but in later stages of the language, we find /s/ occurring successively in all of the positions from which the original /s/-deletion sound-change erased it (cf. Siehler 1995:476). So in Classical Attic, the /s/ is restored in all athematic forms (which are imperfective and perfective (aorist) stems of a small number of common irregular verbs, and perfect mediopassive forms, which aren’t much used), but remains absent from forms with a thematic vowel or mood-marker. Then in New Testament (NT) Greek, it spread to the ‘contract verbs’ with stem-
vowel /a/, and later into all verbs, where it is still found in Modern Greek. Here are some representative paradigms (duals omitted due to marginality in Hometic and Attic, and absence from Modern):²

<table>
<thead>
<tr>
<th></th>
<th>Homeric</th>
<th>Attic/NT</th>
<th>Modern</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>erkhomai</td>
<td>erkhomai</td>
<td>erxome</td>
</tr>
<tr>
<td>2sg</td>
<td>erkheai</td>
<td>erkhei/erkhE:i</td>
<td>erxes</td>
</tr>
<tr>
<td>3sg</td>
<td>erkhetai</td>
<td>erkhetai</td>
<td>erxete</td>
</tr>
<tr>
<td>1pl</td>
<td>erkhometha</td>
<td>erkhometha</td>
<td>erxomaste</td>
</tr>
<tr>
<td>2pl</td>
<td>erkhesthe</td>
<td>erkhesthe</td>
<td>erxeste</td>
</tr>
<tr>
<td>3pl</td>
<td>erkhontai</td>
<td>erkhontai</td>
<td>erxunte</td>
</tr>
</tbody>
</table>

Table 3

Summing up, we see that an /s/ formative which at one time appeared quite widely in various active forms, but only in few places in the mediopassives, due to a sound-change which had erased it, restored itself into those positions. This has always been attributed to the operation of ‘analogy’; the point here is that the analogically driven spread is only explicable if a reasonable number of speakers are analysing the subsequence in question as a formative meaning ‘2sg’ in the positions where it occurs. That is, ‘spitting’ these forms is justified even when the paradigms don’t get a uniform structure, and the postulated formatives fail to appear in many places where they might be expected. It appears to be justified even for dialects similar to Homeric; even more so for Attic and NT, in which the /s/ is appearing in more forms.

An objection (raised by an anonymous reviewer) is that perhaps a lumping analysis could explain the facts anyway, on the basis that after the loss of intervocalic /s/, there would be two different formatives meaning 2sg, /ai/ and /sai/. Then, for some reason, the latter begins to spread. The problem with this is the lack of any plausible reason. The spread of /s/ from a rather marginal corner of the morphology back into its full range of original positions is a steady trend holding over more than a millenium. It is possible that it ‘just happened’, but it would seem worthwhile to try to find possible reasons for it, and a preference for splitting over lumping would appear to be a good candidate.

2.3 Modern Greek

Although /s/ has been preserved in Modern Greek, its meaning has arguably altered somewhat, with the singular property dropping out. In Ancient Greek, the 1st and 2nd person plural pronouns were /hE:meis/ and /hu:meis/ while in

² In Attic, the result of /s/-deletion undergoes vowel-contraction, with /E:i/ the expected outcome, /ei/ (whose phonetic implementation would have been a closer vowel then /E:i/, details unclear) peculiar to this particular form.
Modern they are /emis/ and /esis/ (all nominative case). The accusative/genitive plural forms are furthermore:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>GenSg</td>
<td>mu</td>
<td>su</td>
</tr>
<tr>
<td>AccSg</td>
<td>(e)me(na)</td>
<td>(e)se(na)</td>
</tr>
<tr>
<td>Acc/GenPl</td>
<td>(e)mas</td>
<td>(e)sas</td>
</tr>
</tbody>
</table>

Table 4

It seems justified to identify /m/ as a first person marker and /s/ as a second person marker, both indifferent to number.

3. \( /t/ \) vs \( /st^h/ \)

Our second case is already evident from the contrast between the dual and plural forms of Table 1 and Table 2. Many of the active forms of Table 2 have ‘t’ where the corresponding medio-passive ones of Table 1 have \( /st^h/ \). This is an innovative rather than original feature of the paradigms. The reconstructed PIE forms for the active dual and 2pl have the \( /t/ \) (Siehler 1995:455-6; some minor simplifications and orthographic changes here):

<table>
<thead>
<tr>
<th></th>
<th>primary</th>
<th>secondary</th>
</tr>
</thead>
<tbody>
<tr>
<td>1dl</td>
<td>-wos</td>
<td>-we</td>
</tr>
<tr>
<td>2dl</td>
<td>-t(H_1)es</td>
<td>-tom</td>
</tr>
<tr>
<td>3dl</td>
<td>-tes</td>
<td>-tam</td>
</tr>
<tr>
<td>1pl</td>
<td>-mos</td>
<td>-me</td>
</tr>
<tr>
<td>2pl</td>
<td>-te</td>
<td>-te</td>
</tr>
<tr>
<td>3pl</td>
<td>-nti</td>
<td>-nt/(e):r</td>
</tr>
</tbody>
</table>

Table 5

Many of the mediopassive forms are ‘more than ordinarily’ conjectural (as marked by parentheses), but no trace of an \( /t-/st^h/ \) alternation is present, and in fact the origin of the 2pl \( /st^h/ \) is unclear (Siehler:277ff):

<table>
<thead>
<tr>
<th></th>
<th>primary</th>
<th>secondary</th>
</tr>
</thead>
<tbody>
<tr>
<td>1dl</td>
<td>(-wosdh(H_+_+))</td>
<td>(-wedh(H_+))</td>
</tr>
<tr>
<td>2dl</td>
<td>(-Hto(H_))</td>
<td>(-te(H_))</td>
</tr>
<tr>
<td>3dl</td>
<td>(-Hte)</td>
<td>(-te)</td>
</tr>
<tr>
<td>1pl</td>
<td>-mosdh(H_2)</td>
<td>-medh(H_2)</td>
</tr>
<tr>
<td>2pl</td>
<td>-dhwo</td>
<td>-dhwo</td>
</tr>
<tr>
<td>3pl</td>
<td>-(e):ror/ntor</td>
<td>-(e):ro/nto</td>
</tr>
</tbody>
</table>

Table 6
It is taken as obvious that the dual mediopassive forms are formed by analogy with the active ones; the most plausible historical sequence being that it accidentally becomes possible to segment the 2pl forms as /t-st⁸/+/e/, with the /t-st⁸/ contrast then spreading to the 2/3du forms. And indeed further, to the third person, dual and plural imperatives, where /st⁸/ displaces original /t/. The relevant Greek forms are (2sg excluded):

<table>
<thead>
<tr>
<th></th>
<th>active</th>
<th>mediopassive</th>
</tr>
</thead>
<tbody>
<tr>
<td>3sg</td>
<td>-to:</td>
<td>-st⁸:o:</td>
</tr>
<tr>
<td>2dl</td>
<td>-ton</td>
<td>-st⁸:on</td>
</tr>
<tr>
<td>3dl</td>
<td>-ton</td>
<td>-st⁸:on</td>
</tr>
<tr>
<td>2pl</td>
<td>-te</td>
<td>-st⁸:e</td>
</tr>
<tr>
<td>3pl</td>
<td>-nton</td>
<td>-st⁸:on</td>
</tr>
</tbody>
</table>

Table 7

The available PIE forms on the other hand show no trace of a the /t-st⁸/ alternation (Siehler:601):

<table>
<thead>
<tr>
<th></th>
<th>active</th>
<th>mediopassive</th>
</tr>
</thead>
<tbody>
<tr>
<td>3sg</td>
<td>-tu</td>
<td>-to</td>
</tr>
<tr>
<td>2pl</td>
<td>-te</td>
<td>-dhwo</td>
</tr>
<tr>
<td>3pl</td>
<td>-(e)nu</td>
<td>-nto</td>
</tr>
</tbody>
</table>

Table 8

So what we have here is a further case of the spread of a rather ‘small’ position in the morphological form of the verb, containing only two possible occupants.

4. Conclusion

Historical linguists have of course known about examples of analogy such as these for a very long time. But drawing the consequences for morphological analysis does not seem to have been a prominent concern of the the generative literature. For example I find no discussion of how refined a morphological analysis ought to be in the recent, comprehensive morphological framework of Stump (2001), or earlier proposals such as Anderson (1992) or Lieber (1992). One of the conceptual difficulties of the theory of inflectional morphology is that, because of the finite nature of the systems, the standard assumption that the members of a speech community all acquire the same grammar is even weaker than it usually is. We cannot therefore assume that these or any other analogical change phenomena show that all learners of a language with rich morphology are avid splitters. But we can conclude that enough of them are to have a significant effect on the historical development of such languages.
References
1. *Introduction*

One thing that has long helped it seem obvious that many of Pama-Nyungan languages constituted a genetic unity is the similarity in their pronominal systems. This similarity has led to attempts to reconstruct pronominal paradigms that have given us a general idea about what some of the ancestral forms must have been like (e.g. Dixon 1980:334-362, Blake 1988), but which still leave many matters of details unresolved (see e.g. Alpher 2004:123-125). Even within subgroups the reconstruction of pronominal paradigms is not necessarily straightforward; for one example see Simpson and Hercus (2004:199-201).

A basic idea of what Pama-Nyungan nominative (or unmarked) pronouns tend to look like can be had from looking at Table 1, which shows forms reconstructed more than thirty years ago by Hale (e.g. 1976:56-58) on the basis of the so-called Paman languages of Cape York Peninsula.

In reconstructing the first person plural inclusive as *ngampul(a) 'we (1pi)' Hale (1976:57) cited reflexes only in more northern languages, and he reconstructed an alternant *ngampa to accommodate data from Middle Paman languages. In view of its marginal nature it is perhaps unsurprising that this reconstruction was subsequently ignored by Dixon (1980:351-356). However, apparent cognates can be found elsewhere in Australia, such as Yanyuwa ngambala 'we (1pi)'.

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>dual</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>first person exclusive</td>
<td>*ngayu 'I'</td>
<td>*ngali 'you and I'</td>
<td>*ngana 'we'</td>
</tr>
<tr>
<td>first person inclusive</td>
<td>-</td>
<td>*ngampa(a) 'we'</td>
<td></td>
</tr>
<tr>
<td>second person</td>
<td>*nyuntu 'you'</td>
<td>*nyupula 'you two'</td>
<td>*nyarra 'you lot'</td>
</tr>
<tr>
<td>third person</td>
<td>*nyulu 'he/she'</td>
<td>*pula 'they two'</td>
<td>*tyana 'they'</td>
</tr>
</tbody>
</table>
element -nge and -re had been added to certain pronouns in Arandic. In the present paper I will suggest that there may actually be some basis for reconstructing such additions of shape *pa and *kV (where V is an uncertain vowel) in some earlier protolanguage, whether or not in Proto-Pama-Nyungan itself. Note that there seems to be no basis for reconstructing a contrast in voicing in Pama-Nyungan, and I have simply followed Hale and some others in using symbols for voiceless consonants in such reconstructions (thus p and k rather than b and g). The rough locations of the languages I will be discussing is shown on Map 1.

Map 1: Languages discussed in the paper

2. Evidence for a syllable *pa

Evidence for adding a syllable *pa to pronouns can be found in both the Kurtjar language of southwestern Cape York Peninsula and the Uradhi language near the tip of the Cape. In Kurtjar, all forms of third person pronouns end with a bilabial fricative bh that follow any other affixes. The forms in Table 2 are from Black and Gilbert (1996:16).

The roots of these pronouns are more or less what one would expect in view of the sound changes Kurtjar has undergone (see Black 1980). These include a lenition of initial stops that can be seen in third person dual wila-bh 'they two' < *pula-pa and in third person plural dhana-bh 'they lot' (where dh represents a voiced dental fricative) < *tyana-pa, and also the loss of the initial consonants and the following vowel in third person singular laa-bh 'he/she' < *nyulu-pa, allowing that vowel quality is problematic.

That the -bh might be inherited from a syllable *pa in an earlier protolanguage is suggested by the fact that a syllable bha can optionally be added to many nominative pronouns in Uradhi (Crowley 1983:354-355). Crowley (1983:352) noted that it was not clear what determined the choice between the longer and shorter variants, except that slower, more deliberate speech almost
always used the longer forms. Table 3 shows the nominative forms of Uradhi compared with Kurtjar forms and with reconstructed forms based on Hale (1976) with the possible addition of "pa.

<table>
<thead>
<tr>
<th>case</th>
<th>singular 'he, she' etc.</th>
<th>dual 'they two' etc.</th>
<th>plural 'they' etc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>nominative</td>
<td>laa-bh</td>
<td>wila-bh</td>
<td>dhana-bh</td>
</tr>
<tr>
<td>accusative</td>
<td>nhaa-bh</td>
<td>wilnga-nha-bh</td>
<td>dhalnga-nha-bh</td>
</tr>
<tr>
<td>possessive, ablative</td>
<td>nga-ninga-bh</td>
<td>wilnga-ninga-bh</td>
<td>dhalgi-ninga-bh</td>
</tr>
<tr>
<td>purposive, allative, dative, locative</td>
<td>nga-nta-bh</td>
<td>wilnga-nta-bh</td>
<td>dhalngi-nta-bh</td>
</tr>
<tr>
<td>possessive</td>
<td>nga-ntika-bh</td>
<td>wilnga-ntika-bh</td>
<td>dhalngi-ntika-bh</td>
</tr>
</tbody>
</table>

Table 2: Third person pronouns in Kurtjar

<table>
<thead>
<tr>
<th></th>
<th>reconstructed</th>
<th>Uradhi</th>
<th>Kurtjar</th>
</tr>
</thead>
<tbody>
<tr>
<td>1s 'I'</td>
<td>*ngayu(pa)</td>
<td>ayyu(bha)</td>
<td>ngaay</td>
</tr>
<tr>
<td>1di 'we two'</td>
<td>*ngampula</td>
<td>ampu(la)</td>
<td>ntighal</td>
</tr>
<tr>
<td>1pi 'we lot'</td>
<td>(~ 1d)</td>
<td>ntighal</td>
<td></td>
</tr>
<tr>
<td>1de 'you and I'</td>
<td>*ngali(pa)</td>
<td>ali(bha)</td>
<td>ngaal</td>
</tr>
<tr>
<td>1pe 'you, i and others'</td>
<td>*ngana(pa)</td>
<td>ana(bha)</td>
<td>ngaan</td>
</tr>
<tr>
<td>2s 'you'</td>
<td>*nyuntu(pa)</td>
<td>antu(bha)</td>
<td>aant</td>
</tr>
<tr>
<td>2d 'you two'</td>
<td>*nyupula</td>
<td>iplu(la)</td>
<td>wual</td>
</tr>
<tr>
<td>2p 'you lot'</td>
<td>*nyurra</td>
<td>(~ 2d)</td>
<td>oerr</td>
</tr>
<tr>
<td>3s 'he/she'</td>
<td>*nyulu(pa)</td>
<td>ulu(bha)</td>
<td>laab</td>
</tr>
<tr>
<td>3d 'they two'</td>
<td>*pula(pa)</td>
<td>ula(bha)</td>
<td>wilab</td>
</tr>
<tr>
<td>3p 'they lot'</td>
<td>*tyana(pa)</td>
<td>(~ 3d)</td>
<td>dhanab</td>
</tr>
</tbody>
</table>

Table 3: Nominative pronouns in Uradhi and Kurtjar, with reconstructed equivalents

Note that two of the Uradhi forms can add la rather than bha, but that this simply reflects parts of the normally reconstructed roots, namely *ngumpula in the case of first person dual inclusive ampu(la) 'we' and *nyupula in the case of second person plural ipu(la) 'you lot'.

As noted earlier, in Kurtjar the bh also occurs on oblique forms of the third person pronouns. In Uradhi the only oblique forms to take bha are the accusatives of the only two irregular pronouns, the first and second person singular. The other Uradhi accusatives are formed by simply adding the suffix -nha to the nominatives stems, although in the Atampaya variety of Urathi this -nha can optionally be preceded by nga to make longer variants. This syllable nga is also a non-variant part of other obliques in Atampaya, if not other Uradhi varieties, and it is also found in many Kurjjar obliques, as can be seen in Table 4.

Various other evidence for a possible reflex of *pa in pronouns can perhaps be found elsewhere in Australia. Along the Queensland coast to the south of Cape York Peninsula, Giramay (Dixon 1972:50), Wargamay (Dixon 1980:329) and Nyawaygi (Dixon 1983:464) all make a distinction between intransitive and transitive subject forms in the first and second person singular, i.e. intransitive ngayba and transitive ngadya 'I' and intransitive nginba and transitive nginda 'you' in all three languages. This may suggest that earlier *pa may have marked
intransitive subjects, but aside from problems in the phonological detail (namely the lack of a vowel before that \(\text{ba}\)), this would not explain why it is found in oblique forms in both Kurtjar and Uradhi. Dixon (1980:343) himself suggested that the \(\text{ba}\) in the east coast languages was just a “dummy syllable added in order to achieve a disyllabic word”, but this also would not account for the Kurtjar and Uradhi forms, which appear to reflect disyllabic stems.

Table 4: Accusative pronouns in Uradhi and Kurtjar, with reconstructed equivalents

Some evidence for a pronominal *-pa suffix can be found even in non-Pama-Nyungan languages, in such first person singular forms as Gunibidji (or Ndjębbana) \(\text{ngaayapa}\) and perhaps Maung \(\text{ngabi} \, \text{‘I’}\) (cited in Dixon 1980:351).

3. Evidence for a syllable *\(k\)V

There is somewhat more problematic evidence suggesting the reconstruction of a syllable *\(k\)\(\text{V}\), ending in an uncertain vowel \(\text{V}\). Some of this evidence is from the poorly attested Nganyaywana language of New South Wales. In showing how Nganyaywana had lost initial consonants and sometimes the following vowels, Crowley (1976:38-39) proposed that \(\text{ga}\) was added to roots that would otherwise have become monosyllabic, as in the case of \(*\text{ngaya} > *\text{ngaya-ka} > \text{yaga} \, \text{‘I’}\). As Black (2007:260) notes, however, all five forms that add \(\text{ga}\) are pronouns, where the \(\text{ga}\) could perhaps represent an affix. Table 5 shows my retranscription of the Nganyaywana pronouns and their protoforms given by Crowley.

Table 5: Nganyaywana pronouns (Crowley 1976:38-39)

Note that the initial \(d\) of the dual forms represents the most common Nganyaywana reflex of earlier *\(l\). This raises questions about the source of the \(la\)
in the exclusive first person plural (and perhaps the first person dual as well); I will return to this briefly in the following section. My only concern here is the fact that some of the pronouns seem to contain a reflex of what could go back to earlier *kV. The reason *V represents an uncertain vowel is because Crowley took most Nganyaywana forms to end in the vowel a, raising the question of whether final vowels remained contrastive and thus can provide a basis for reconstructing vowel quality; see Black (2007:256-257).

Other possible reflexes of a reconstructed *kV can be seen in two nominative pronouns in Mpkwithi (Crowley 1981:170), an initial dropping Northern Paman variety spoken just to the southwest of Uradhi in northern Cape York Peninsula. These pronouns are the first person dual inclusive lægi ‘you and I’ < *ngali-kV and the second person plural re:geh ‘you lot’ < *nyira-kV. Vowel changes in Mpbkwithi again make it difficult to reconstruct the quality of the final vowel.

Pronouns in Koko-Bera (indigenously Kok-Kapér), of western Cape York Peninsula, appear to end in a w reflex of the earlier *kV, a sound change that can also be seen word finally in kutéw ‘dog’ < *kutaka and in such purposive forms as may-éw ‘for food’ < *maya-ku. The addition of w alone is found only the nominative third person pronouns yélew ‘he/she’ < *nyulu-kV and púlew ‘they two’ < *pula-kV. Other pronouns that do not involve second person (even as inclusive first person) end in -ntew (where unaccented e in polysyllables represents a schwa) in the nominative, accusative, and dative (but not possessive) cases; see Table 6.

<table>
<thead>
<tr>
<th>person</th>
<th>nominative</th>
<th>accusative</th>
<th>dative</th>
</tr>
</thead>
<tbody>
<tr>
<td>1s ‘I’ etc.</td>
<td>ngá-ntew</td>
<td>ngénye-lew</td>
<td>ngánye-ntew</td>
</tr>
<tr>
<td>1de ‘we two’ etc.</td>
<td>ngalínye-ntew</td>
<td>ngalénye-ntew</td>
<td>ngaléngke-ntew</td>
</tr>
<tr>
<td>1pe ‘we lot’ etc.</td>
<td>nganthé-ntew</td>
<td>nganténe-ntew</td>
<td>nganténgke-ntew</td>
</tr>
<tr>
<td>1di ‘you and I’ etc.</td>
<td>ngel</td>
<td>ngaléngk</td>
<td>ngaléngk</td>
</tr>
<tr>
<td>1pi ‘you, I and others’</td>
<td>ngen</td>
<td>ngenéngk</td>
<td>ngenéngk</td>
</tr>
<tr>
<td>2s ‘you’ etc.</td>
<td>yen</td>
<td>yintéw</td>
<td>yintéw</td>
</tr>
<tr>
<td>2d ‘you two’ etc.</td>
<td>yupél</td>
<td>yulpéngk</td>
<td>yulpéngk</td>
</tr>
<tr>
<td>2p ‘you lot’ etc.</td>
<td>yurr</td>
<td>yurréngk</td>
<td>yurréngk</td>
</tr>
<tr>
<td>3s ‘he/she’ etc.</td>
<td>yèle-w</td>
<td>yungónye-ntew</td>
<td>yungóngke-ntew</td>
</tr>
<tr>
<td>3d ‘they two’ etc.</td>
<td>púlé-w</td>
<td>pulénye-ntew</td>
<td>puléngke-ntew</td>
</tr>
<tr>
<td>3p ‘they lot’ etc.</td>
<td>thá-ntew</td>
<td>thenínye-ntew</td>
<td>thenéngke-ntew</td>
</tr>
</tbody>
</table>

Table 6: Koko-Bera pronouns (author’s own data)

One may wonder whether the w could perhaps represent a lenited reflex of *p rather than *k, although the lenition of *p to w is otherwise unestablished for Koko-Bera. Aside from that lack of supporting evidence, there are two reasons for favouring *k over *p. One is the possibility that the Koko-Bera addition -ntew could perhaps reflect an earlier affix *-ntVkV that is also reflected in a Kurtjar suffix for the genitive oblique, i.e. for a possessive form in the ergative or other oblique case. While Black and Gilbert (1996:14) give this suffix as a -tik (with
schwa-like \( i \), it always occurs after \( n \) in pronouns and demonstratives, and it appears as \( -\text{ingintak} \) after nouns. The second reason it that pronouns in the closely related Kok-Papónk actually show a \( k \) reflex of an earlier final syllable \(*kV\), if always in the combination \( nk \) or \( ngk \); see Table 7 for the nominative pronouns in Kok-Papónk as compared with Koko-Bera.

<table>
<thead>
<tr>
<th>Kok-Papónk</th>
<th>Koko-Bera</th>
</tr>
</thead>
<tbody>
<tr>
<td>1s 'I'</td>
<td>nga-nk</td>
</tr>
<tr>
<td>1de 'we two'</td>
<td>ngašənye-nk</td>
</tr>
<tr>
<td>1pe 'we lot'</td>
<td>ngašəñē-nqk</td>
</tr>
<tr>
<td>1di 'you and I' etc.</td>
<td>ngašə-ngk (~ ngašə-ngən)</td>
</tr>
<tr>
<td>1pi 'you, I and others'</td>
<td>ngaen</td>
</tr>
<tr>
<td>2s 'you'</td>
<td>yon</td>
</tr>
<tr>
<td>2d 'you two'</td>
<td>yipəl ~ yipəl</td>
</tr>
<tr>
<td>2p 'you lot'</td>
<td>yuri, yuri-ngk</td>
</tr>
<tr>
<td>3s 'he/she'</td>
<td>nhul-ngk</td>
</tr>
<tr>
<td>3d 'they two'</td>
<td>pule-ngk</td>
</tr>
<tr>
<td>3p 'they lot'</td>
<td>thə-ntew</td>
</tr>
</tbody>
</table>

Table 7: Kok-Papónk and Koko-Bera nominative pronouns (author’s own data)

The Kok-Papónk data is problematic is several ways. Firstly, since the main Kok-Papónk informant was far more fluent in Koko-Bera, possibly the \( ngk \) could simply be a Koko-Bera mispronunciation of the nonhomorganic cluster \( nk \) attested in the first person singular, since \( nk \) is generally lacking in Koko-Bera. Secondly, vowels in such forms as \( yen \) 'you' may not be authentic, but rather based on the informant’s awareness that Kok-Papónk often has \( o \) where Koko-Bera has \( e \), as in \( yen \) 'you'. Evidence for such ‘ethnoreconstruction’ has been presented at length by Black (2005), and it may also explain why the second dual was attested as \( yipəl \) 'you two', although the alternative \( yipəl \) could alternatively just be a contamination from Koko-Bera. Thirdly, the fact that \( -ngk \) was attested in two second person or inclusive pronouns, unlike the \( -ntew \) in Koko-Bera, could again be due to ethnoreconstruction: the informant may have added it to these pronouns simply because she was aware that it occurred on others.

Further evidence for a \(*-kV\) suffix may be found in Yir-Yoront, to the north of Koko-Bera, Alpher (2004:125) notes that such nominative pronouns as \( nguəy 'I' \) could not have continued as disyllables “without (in at least one of them and quite probably in all) [the loss of] a third syllable of the form \(*(ng)kV\)”. The evidence is thus ambiguous for a suffix of shape \(*-kV\) or \(*-ngkV\). The latter could perhaps have been the ergative suffix \(*-ngku\), since the pronouns in question are all nominative and since occasional Australian languages have ergative marking on pronouns (Blake 1987:21), whether or not using this particular suffix. Note that the Kok-Papónk suffixes are less amenable to such an explanation because the first person suffix is clearly \( -nk \) rather than \( -ngk \), but in any case it is unclear whether the syllable that was lost in Yir-Yoront contained a nasal or not.
4. Other added syllables

In the Kok-Nar language between Kurtjar and Koko-Bera, most non-second person pronouns either end in, or have variants ending in, the velar nasal ng: e.g. ngaling 'we two', nganang 'we lot', nhulang 'he/she', pil or pilang 'they', and thanang 'they lot' (Breen 1976:248). This may remind us of the addition of -nge in Arandic (Koch 2004:142), but alternatively the final ng in Kok-Nar could well reflect earlier *l, as it does it such words as ngamáng ‘big’ < *ngamal (Black 2004:258). One may be tempted to relate this to the la that was added to one or two Nganyaywana pronouns (see earlier), but this also seems to be found on some Nganyaywana nouns, where it could perhaps reflect the ergative suffix *-lu (considering that final a may not be contrastive in Nganyaywana); see Black (2007:261). While we also saw an optional syllable la in ‘long’ forms of Uradhi pronouns, recall that this seemed to reflect part of the original root.

5. Conclusion

While we have had a general idea about the development of pronouns in Pama-Nyungan languages for some thirty years, many details remain to be worked out. From the present paper it can be seen that this may even include the reconstruction of such additional syllables as *pa and *kV that have reflexes in the pronouns of several languages of Cape York Peninsula and (for *kV) in the Nganyaywana language of New South Wales.

References


ASSOCIATED EATING AND MOVEMENT
FURTHER EXAMINATION OF YUWAALARAY GAMILARAAY VERB SUFFIXES

JOHN GIACON
Australian National University

1. Introduction

Harold Koch’s seminal paper on ‘Associated Motion’ (1984:164) made a major contribution to the study of verb specification in Australian languages. He discussed the concept of associated motion as a verbal category outside the more traditional categories of tense, aspect and mood. This paper examines two areas of verb specification in Yuwaalaraay and Gamilaraay (YG). The first is the continuous aspect suffixes which also encode further information about movement or non-movement. Similar suffixes are reasonably common in Australian languages. The second area is the ‘Associated eating’ category which seems limited to what Austin (1997) calls the Central New South Wales languages, a distinct group including Yuwaalaraay, Gamilaraay, Wiradjuri, Wangaaybuwan and Wayilwan. The last two are also known as Ngjuryambaa. This suffix is well outside the TAM categories.

1.1 Background - Yuwaalaraay-Gamilaraay

YG are languages from north central New South Wales, sometimes regarded as dialects of one language as they share approximately 70% of their lexicon (Austin et al. 1980). Current handed down knowledge of YG is limited to words and a few short phrases. The majority of the linguistic information available for YG is from the Yuwaalaraay, and a significant part of that comes from tapes made in the 1970s, with Arthur Dodd and Fred Reece as informants. Williams used earlier material and these tapes for her ‘Grammar of Yuwaalaraay’ (1980), which significantly expanded the available analysis of Yuwaalaraay. This grammar is the main source used for Gamilaraay revival.

The context of work on YG is language revival, where linguistics has a role beyond description. Descriptive linguistics can stop at what is known about a language from the sources. In language revival the task of linguistics includes providing a grammar which enables people to use the language. It will include all information that can be obtained from the historical sources, and also go beyond that to provide a ‘reconstructed’ grammar – a well based estimation of what else the language would have done. Information about related languages, the proto-language and new linguistic knowledge are the major sources for this ‘reconstructed’ grammar. This ‘grammatical development’ parallels the lexical

---

1 Harold Koch is supervising John Giacon’s PhD at the Australian National University.
development common in language revival. This paper re-examines the historical sources and takes some small steps in grammatical development, mainly based on the analysis of cognate morphemes in Wanggaaybuwan found in Donaldson (1980). The major historical sources for YG are listed in Austin (in press) and include Ridley (1875) and Mathews (1902, 1903) - short grammars which are quite limited by recent standards.

1.2 Yuwaalaraay Gamilaraay verb morphology

This section includes background information on the basic YG verb paradigm, and a brief discussion on some suffixes which form complex verb stems. There is no simple present tense, rather present tense is shown by multiple-morpheme suffixes, as discussed in section 2.2. YG has four verb classes. There are two large classes, \( l \) and \( y \), and two smaller classes, \( rr \) and \( ng \). The \( l \) class is mainly transitive, and the \( y \) class mainly intransitive. Apart from \( l \) class verbs, the past tense suffix is \(-nyi\) after ‘i’ final stems, \(-nhi\) elsewhere. Examples of inflected simple verbs (root + one suffix) are provided in Table 1.

<table>
<thead>
<tr>
<th>Verb class</th>
<th>Inflection/gloss</th>
<th>NON-Past</th>
<th>PAST</th>
<th>IMPerative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>( l )</td>
<td>( y )</td>
<td>( ng )</td>
</tr>
<tr>
<td>Inflection/gloss</td>
<td>‘eat’</td>
<td>( dha-li )</td>
<td>( gubi-y )</td>
<td>( gaa-gi )</td>
</tr>
<tr>
<td>NON-Past</td>
<td>( dha-y )</td>
<td>( gubi-nyi )</td>
<td>( gaa-nhi )</td>
<td>( wuu-nhi )</td>
</tr>
<tr>
<td>IMPerative</td>
<td>( dha-la )</td>
<td>( gubi-ya )</td>
<td>( gaa-nqa )</td>
<td>( wuu-na )</td>
</tr>
</tbody>
</table>

Table 1: Examples of simple YG verb inflections

As well as the simple inflections shown above there are a number of more complex inflections that involve one or more suffixes between the main root and the final inflection. Williams’ (1980:70-72) list of suffixes includes: ‘continuous’ (2), time (4), causative (2), reciprocal, reflexive, benefactive, completive, and ‘back’. The time suffixes have glosses such as ‘near past, near future, distant past’ and ‘all day’. They are prominent in Ridley and Mathews, but much rarer in 20\textsuperscript{th} century sources. Some suffixes do not fit into the standard TAM categories, for instance the completive suffix, \(-aba-li\) ‘all’, (Ash et al. 2003:314; Williams 1980:80) operates ergatively, indicating all of the Subject with intransitive verbs and all of the Object with transitive verbs.

2. Continuous Aspect and Motion

2.1 Associated Motion and Aspect+Motion

It is important to distinguish between Associated motion and aspectual suffixes which also contain information about motion.\(^2\) Koch (1984:26) points out that in

\(^2\) Thanks to the anonymous referees for a number of suggestions and comments. One of them highlighted this distinction and pointed out the Wilkins reference. Special Thanks to Jo Caffery, Jane Simpson, Steve Morelli and Luise Hercus for the comments on drafts of this paper.
Central Australia ‘Associated Motion’ is an areal feature, and in Kaytej, (Kaytetye) “[v]erbs that indicate any kind of activity may also be specified for various kinds of motion associated with the activity.” Such Kaytetye verbs have a complex verb stem containing the main root and a stem-forming suffix and the semantic content of the suffix can vary. It can include prior motion (before the action of the main verb), subsequent motion and concurrent motion. Similar suffixes are found in Adnyamathanha (Tunbridge 1988) and Arrernte/Aranda Wilkins (1989:272). An example is the suffix irtne (REVERSIVE: do coming back) (Wilkins 1989:272) given below.

(1) Nthenhe annanthere arlkw-irtne-tyenhe Thursday-nge?
Where 2pl(A) eat-REVERS-NPC Thursday-ABL
“Where will you eat on your way back on Thursday?”

A number of languages have suffixes which contain information about both aspect and motion. In Arrernte ‘the fillers of the continuous aspect slot’ are -rle.ne ‘do verb action continuously’ or -rle.pe ‘do verb action continuously while moving along’, or similar forms which also encode plural subject (Wilkins 1989:252). These are remarkably similar in use to the YG suffixes to be discussed next. Austin (1989:50, 52) points out similar suffixes in Yandruwandha (-durrha- continuing action while in motion) and Wangkumara (-manja- progressive aspect as well as concurrent motion).

2.2 YG Continuous Aspect + Motion

Each YG verb class has two suffixes which can indicate ‘continuous’ aspect. One is glossed REGular and always contains the short vowel /a/. The other is glossed MOVing and always contains the long vowel /aa/. The suffixes form y class verb stems which have the added possibility of a present tense final suffix ‘-nha’, in addition to the final inflections to be found on other verb stems. The table below shows the form of the continuous suffixes in YG and (2) lists verbs incorporating some of the suffixes. For y and ng class roots the MOVing suffix begins with ‘y’ after root final ‘i’ and with ‘w’ elsewhere. These two suffixes never co-occur. Their position in the complex stem is always immediately before the final suffix.

<table>
<thead>
<tr>
<th>Verb class</th>
<th>MOVing suffix</th>
<th>REGular suffix</th>
<th>Final suffixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>l class</td>
<td>-l-aa-y</td>
<td>-lda-y</td>
<td>-y, -nha, -nih, -ya</td>
</tr>
<tr>
<td>y class</td>
<td>-w/y-aa-y</td>
<td>-y-la-y</td>
<td>-y, -nha, -nih, -ya</td>
</tr>
<tr>
<td>ng class</td>
<td>-w/y-aa-y</td>
<td>-gi-la-y</td>
<td>-y, -nha, -nih, -ya</td>
</tr>
<tr>
<td>rr class</td>
<td>-rr-aa-y</td>
<td>-dha-y</td>
<td>-yi, -nha, -nih, -ya</td>
</tr>
</tbody>
</table>

Table 2: YG continuous suffixes

(2) banaga-waa-nhi, dha-lda-ya, willa-y-la-nha
run-MOV-PAST eat-REG-IMP sit-REG-PRES
“was/were running” “Keep eating!” “am/is/are sitting”
The interpretation of the suffixes depends on the semantic class of the verb root they are attached to. The important distinctions are between active and stative verbs, with active verbs further divided into those which usually involve linear motion and those which do not. When continuous suffixes are attached to the YG copula (gi-gi) the interpretation is as for stative verbs. Donaldson (1980:74) defines active verbs as “verbs denoting actions controlled by those who perform them, such as baga-l ’dig’, bibuwa-y ’run’” and stative verbs as “verbs [that] denote physiological or emotional states which are not controlled by those who experience them, except in so far as they expose themselves to their causes, or avoid them. Examples are girrambi-l ’sick, in pain’, birrabi-l ’hungry’, dharrambi-l ’fond’, walindja-l ’lonely’”. I will now discuss the interpretation of continuous suffixes on active verbs, then on stative verbs, then the use of one suffix to indicate habitual aspect and finally discuss some examples which do not fit the current analysis.

2.2.1 Continuous suffixes on active verbs. Examples (3) to (10) illustrate the differences between the suffixes when they are attached to active verb roots. The MOVing suffix shows linear motion, and the REGular suffix indicates absence of linear motion. Examples (3) to (6) show the MOVing and REGular suffixes attached to the motion verbs baa-y ’hop, jump’, and yanaa-y ’walk/go’. In (3) and (5) the MOVing suffix shows linear motion: baa-waa-nha ’the kangaroo hopping along’, and yanaa-waa-nha ’walking’. In (4) the REGular suffix indicates non-linear motion, i.e. the movement is up and down and in one place, baa-y-la-nha ’the fish is ‘jumping’. Example (6) also illustrates non-linear motion (walking about). The nature of the movement in (6) is further emphasised by the locative question word, minyaaya ‘where at’, which contrasts with the allative question word, minyaarru ‘where to’.

(3) bandaarr nhama baa-waa-nha
grey.kangaroo that hop-MOV-PRES
“The kangaroo is hopping along.” Tape 3219A

(4) guduu baa-y-la-nhi ngiyarra, ganuu-ga
fish hop-CM-REG-PAST there canoe-LOC
“The fish were jumping in (the bottom of) the canoe.” Tape 8184

(5) giirr ngali yanaa-waa-nha
true 1duSA go-MOV-PRES
“We are walking.” Tape 3996A

(6) minyaaya-nda yanaa-y-la-nhi
where.at-2sgSA go-REG-PRES
“Where was you walking about?” Tape 1853
Examples (7) and (8) are parts of one sentence and show the MOVing and REGular suffixes attached to an active, non-motion verb, *gula-li* 'bark'. The MOVing suffix is attached to the verb when the dog is both running and barking. The REGular suffix is used when the dog is stationary and barking. Example (9) also shows the REGular suffix on a non-motion verb and (10) shows MOVing suffixes on motion and non-motion verbs.

(7) *maadhaay dhaay banaga-waa-nha, gula-laa-nha*
dog to.here run-MOV-PRES bark-MOV-PRES
“The dog is running this way and he’s barking at something.” Tape 2439A

(8) *ngaarrma nguu gula-laa-nha,*
there 3sgSA bark-REG-PRES
“(and now) he is barking (at the goanna up the tree).”

(9) *minya nginda gimbi-lda-nha*
what 2sgSA do-REG-PRES
“What are you doing?” Tape 3996A

(10) *giirr nhama bandaarr dhuu-rraa-nha,*
true that grey.kangaroo crawl-MOV-PRES
 *buunhu nhama dha-laa-nha*
grass that eat-MOV-PRES
“That kangaroo is crawling along, it’s eating that grass.” Tape 2833B

2.2.2 Continuous suffixes on stative verbs. With stative verbs or the copula the REGular suffix indicates present continuous aspect, with no change, for instance no change in the emotion being felt in (11) and (12), and the MOVing suffix indicates inchoative aspect, as in (13) and (14). Inchoative aspect can be understood as metaphorical motion, moving from one state of being to another.

(11) *walindja-lda-nha ngaya*
be.lonely-REG-PRES 1sgSA
“I am lonely.” Tape 5130

(12) *giirr nhama garigari gi-gi-la-nhi*
true that frightened be-REG-PAST
“(The boy) was really frightened.” Tape 5131

(13) *maadhaay ngay balu-waa-nha*
dog 1sgGEN die-MOV-PRES
“My dog is dying.” Tape 1852B
(14) giirr ngarran gi-yaa-nha
    true dawn be-MOV-PRES
    “It’s getting daylight.” Tape 5052

2.2.3 Regular continuous suffix and habitual aspect. The REGular suffix can also indicate habitual aspect as shown below with a motion verb, yanaa-y-la-nha, in example (15), with an active, non-motion verb, dha-lda-nhi, in example (16), with stative verbs, dhanduwi-y-la-nha, example (17) and bundaa-gi-la-nha (15) and on the copula, gi-gi-la-nha, in example (18).

(15) birralii bundaa-gi-la-nha, waala maayu yanaa-y-la-nha
    child fall-CM-REG-PRES, can’t well walk-REG-PRES
    “The child falls over all the time, it can’t walk (properly).” Tape 5131

(16) giirr ngiyani bandaarr dha-lda-nhi
    true 1plSA grey.kangaroo eat-REG-PAST
    “We used to eat kangaroo.” Tape 3996A

(17) giirr ngaya maayu dhanduwi-y-la-nha
    true 1sgSA well sleep-REG-PRES
    “I always sleep well.” Tape 1989A

(18) waal-bala ngaya garigari gi-gi-la-nha
    NEG-CONT 1sgSA afraid be-REG-PRES
    “I don’t get frightened.” Tape 3998B

2.2.4 Summary of YG Continuous Suffix rules. The specification of linear or non-motion by verb suffixes is quite limited in YG and is linked to use of one of the continuous suffixes. Complex stems that incorporate the MOVing suffix and a non-motion active verb specify two actions, that included in the main stem, and linear movement. The rules for use of the continuous suffix in YG are listed below and exemplified in the Table 3.

- the MOVing suffix indicates continuous aspect and linear motion when used with an active verb (motion or non-motion)
- the MOVing suffix, when used on stative verbs and the copula has an inchoative interpretation
- the REGular suffix can have a habitual meaning or indicate continuous aspect
- when used to show continuous aspect on active verbs it specifies that there is no linear motion
- when used to indicate continuous aspect on stative verbs and the copula it indicates a steady state - there is no change happening
the MOVing suffix occurs most commonly on motion verbs and the REGular suffix on non-motion verbs. The use of a MOVing suffix with a non-motion verb or of a REGular suffix with a motion verb is marked.

<table>
<thead>
<tr>
<th>Suffixes</th>
<th>MOVing</th>
<th>REGular</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verb type</td>
<td>(non-moving)</td>
<td>(habitual)</td>
</tr>
<tr>
<td>Active verbs</td>
<td>dha-laa-nha</td>
<td>dha-lda-nha</td>
</tr>
<tr>
<td>non-movement verb</td>
<td>is eating</td>
<td>eats (habitual)</td>
</tr>
<tr>
<td>movement verb (hop)</td>
<td>baa-y-la-nha</td>
<td>baa-y-la-nha</td>
</tr>
<tr>
<td>Stative verbs</td>
<td>dhanduwi-y-la-nha</td>
<td>dhanduwi-y-la-nha</td>
</tr>
<tr>
<td>is hopping</td>
<td>dhanduwi-y-la-nha</td>
<td>dhanduwi-y-la-nha</td>
</tr>
<tr>
<td>is sleeping</td>
<td>dhanduwi-y-la-nha</td>
<td>dhanduwi-y-la-nha</td>
</tr>
</tbody>
</table>

Table 3: Use of the continuous suffixes in YG (marked uses are shown in bold)

2.2.5 The growing understanding of the continuous suffixes. The understanding of YG associated movement has steadily developed. Mathews (1902:142) states that “[t]here are forms of the verb to express ‘beating going along the road’” (my emphasis) but gives no suffix form or example verb. Williams (1980:70) recognised the different forms of the continuous suffixes, but did not find any differences in meaning. The description of these suffixes above is largely the one included in Ash et al (2003:305) but with some modifications, particularly regarding the MOVing suffix, following a close study of Donaldson’s (1980:190-191) description of its Wangaaybuwan cognate -wa-y. The distinction between verb types in the current analysis is also more detailed. However there are exceptions to the above rules, as shown in examples (19) to (21). It may be that ongoing examination of the sources may lead to more detailed rules, it may be that these examples are actually errors by the informants, or that we will never have a good understanding or the grammar involved. The rules above are the ones of currently taught YG.

Example (19) contains the MOVing suffix and active verb, but there is no indication of linear movement associated with the eating. More common are examples such as (20) and (21), with a MOVing suffix on a stative verb but the translations do not include an inchoative meaning. In (20) the informant was asked to translate: ‘He felt sick and couldn’t leave the camp.’ There is repetition and unsureness in his reply.

(19) giirr-nqa ganugu dhinggaa dha-laa-nha
true-NOW 3plA meat eat-MOV-PRES
“They are having a good feed now.” Tape 2436A

(20) bamba-nqa-nha dhaala-waa-nha, dhinggaa nguu dha-ndaay
hard-then-that be.sick-MOV-PRES meat 3sgSA eat-REL 2436A
Arthur Dodd, when asked to translate ‘it (the lizard) was very sleepy’ (when I killed it), uses:

(21) *dhanduwi-yaa-nhi*  
    sleep-MOV-PAST  
    Tape 8184

2.2.6 ‘Back’. There is another YG verb suffix, -uwi-y, which often occurs with movement verbs and the verb ‘give’. It adds the meaning ‘back’, as in ‘give it back’ or ‘come back’. This suffix does not indicate action additional to that of the main verb, rather it specifies the direction of the main action. Examples (22) and (23) have -uwi-y as the only pre-final suffix, while (24) shows -uwi-y preceding a continuous suffix. Example (24) shows that the ‘back’ suffix is a separate category from the continuous suffixes, since they co-occur here. It also shows that the ‘back’ suffix precedes the continuous suffixes. Donaldson (1980:193) details Wangaaybuwan cognates of the suffix.

(22) *wuu-rr-uwi-ya nhama ngay bilaarr!*  
    give-CM-BACK-IMP that 1sgGEN spear  
    “Give my spear back to me!” Tape 5056

(23) *giirr ngaya-laa yanaa-w-uwi-y*  
    true 1sgSA-then go-CM-BACK-FUT  
    “I will come back. (to see you).” Tape 5055

(24) *gaa-g-uwi-yaa-nha*  
    take-CM-BACK-MOV-PRES  
    “is taking back” Tape 5056

3. *Associated eating/drinking/smoking*

A further category of verb specification in YG is what I call here ‘associated eating’. The suffix has the form -DHa-y, and is clearly related to the verb dha-li ‘eat’. In the Yuwaalaraay sources this suffix conveys the meaning that there is ‘eating’ (and in one example ‘smoking’) associated with the action or state indicated by the main root. A suffix with this meaning is much rarer cross linguistically than ‘associated motion’ suffixes. Cognates of the YG suffix, also of the form -DHa-y, are found in both Wangaaybuwan and Wayilwan. The detailed description of the suffix in Donaldson (1980:175-6), provides many clues to understanding the YG suffix. I first summarise the Wangaaybuwan analysis, then give some Yuwaalaraay examples and then develop some Yuwaalaraay rules. A brief discussion about the ordering of YG verb suffixes, and also about

---

3 Realised -dha-y after verb stems ending in /a/ or /u/, -dja-y after stem final /i/.
homophones suffixes in the Central NSW group and also in some other languages follows.

3.1 Associated eating in Wangaaybuwan

Donaldson (1980:175-6) states that in Wangaaybuwan the 'associated eating' suffix indicates that an event occurs in association with eating and/or drinking. When -DHa-y is attached to verbs indicating position, (e.g. sit) the eating or drinking is concurrent with the event referred to by the verb root. When (it) is attached to other, active, verbs, (e.g. lean over, cook) it indicates that the action is undertaken in order to eat or drink. When -DHa-y is attached to a stative verb or ga-l 'be' in a nominal predicate construction, it indicates that the state results from eating or drinking (e.g. 'choke from eating'). (my emphases).

-DHa-y is obligatory whenever some NP argument in the sentence has to be interpreted as being ingested. e.g. (25) (Donaldson’s 6-53).

(25) badhaambhaan-dhi+ni balu-nh-dha-y-guwa-nhi
    poison-CIRC=3ABS+VIS die-EAT-CM-PITY-PAST
    “It was because of eating poison that this poor fellow died.” (6-53)

Sentence (26) shows the Wangaaybuwan suffix having to do with drinking.

(26) ngadhu nginu-ga gaanh-dha-nha
    I+NOM you+OBL-LOC carry-DRINK-PRES
    “I am bringing (some) for you to drink.” (6-55)

3.2 Associated eating in Yuwaalaraay Gamilaraay

The suffix is not discussed in the Gamilaraay sources, but may occur in the story of Emu and Brogla. (Austin 1975:12, 13) in the verb stem ngudhaRu 'feed'. This stem may be monomorphemic, or a compound of ngu- ('give' in Wayilwan and Wangaaybuwan, so perhaps in a version of Gamilaraay) followed by -dha- 'eat' and Ru (unknown meaning). The suffix has only been found once in early descriptions of Yuwaalaraay. Williams (1980:74) discusses this occurrence, noting that Mathews (Notebook 3:62) includes a sentence, example (27), which incorporates the suffix. Williams did not find evidence in the tapes to support Mathews’ example, however quite a few instances have since been found where -DHa-y is attached to a verb root and the translation includes 'eating', or an associated concept, or at least eating is implied. This shows the value of having time to build on Williams’ work. The complex stem which incorporates -DHa-y

---

4 There is no published information about the suffix in other Central NSW languages. I have transcribed Wayilwan tapes, where the suffix is common, but it has not been systematically studied. I have not found it in a preliminary search of the Wiradjuri of Gunther (1892) or Hale (1968 [1846]).
can refer to two subjects - one which is the subject to the main verb and one which is the subject of 'eating', as shown in examples (30) and (31), where the subjects of 'give' and 'eat' are different, and in (36). The transitivity of the complex stem is determined by that of the main root.

Below are a number of the Yuwaalaraay examples of the suffix. Apart from (27) they are from the tapes, with translations provided by the informants. Often the Wangaaybuwan semantics also apply to YG, but not always. The meaning 'after eating' does not occur in the Wangaaybuwan analysis, where the suffix can only mean 'concurrent eating; in order to eat; as a result of eating'. The meaning, 'after eating' occurs in (27) from the late 1900's and in (28), from a tape made in the 1970s.

(27) buma-dha-y
hit-EAT-FUT
“(Will) beat after eating.” Mathews’ Notebook

(28) bamba ganunga yulu-dha-nhi,
hard 3plSO dance-EAT-PAST
“They all danced after the meal.” Tape 3220A

Both (27) and (28) have the suffix attached to an active verb, as do (29) to (31). In these YG examples the suffix is translated as 'after eating' or not at all. The activities of the main verb ('carry, give') can be more or less explicitly related to 'eating', often with a purposive meaning. In (32) the purposive meaning of -DHa-y is made explicit. With a copula (33) or stative verb (34), the suffix is not explicitly translated. The verbs describe the result of eating.

(29) ngaarri-bala ngaya gi-yaa-nha gaa-dha-waa-y
there-CONT 1sgSA be-MOV-PRES take-EAT-MOV-FUT
(going.to)
“I’ll take this meat across (the river).” Tape 3217A

(30) waal ngiyani-luu wuu-dha-y-la-nha
NEG 1plSA-ALL give-EAT-REG-PRES
“We are not going to give (them) any.” Tape 8187
(This sentence follows a statement that some people had asked for food.)

Arthur Dodd translates: 'They were fat children and she fed them well.' with:
ASSOCIATED EATING AND MOVEMENT

(31) a. giirr ngaama ganungu wamu gi-dja-nhi,
    really that 3plGEN fat be-EAT-PAST        Tape 8185
b. ngambaa-gu maayu dhuwarr ganungu wuu-dha-nhi
    mother-ERG well bread 3plGEN give-EAT-PAST

(32) wayamaa-gu wuu-dha-y-la-nha birralii-gal-gu,
    old.man-ERG give-EAT-CM-REG-PRES child-PLUR-DAT
    ngaarrrma dhinawan
    there emu
    “The old man is giving the children emu to eat.”       Tape 3217A

(33) giirr-nga-bala ngay mubal gi-dja-nhi
    really-now-CONT 1sgGEN stomach be-EAT-PAST
    “My stomach is full now.”                  Tape 2440A

(34) nhama maadhaay-u dha-li, waal nhama balu-dha-y
    that dog-ERG eat-FUT NEG that die-EAT-FUT
    “The dog will eat it but won’t die.”      Tape 2833B

Examples (35) and (36) illustrate quite different meanings of the suffix. Example (35) is the only one found where the suffix is associated with ‘smoking’, and (36), a comment on a goanna being cooked, is the only instance where the subject of the main verb is inanimate.

(35) ngaama nguu gagil dhaygal gi-dja-nhi,
    that good bad head get-EAT-PAST
    “[He was smoking that tobacco] and he got a headache out of it.”  Tape 5131

(36) ngaama gaba buwi-dja-nhi, mangun.gaali
    that good smell-EAT-PAST tree.goanna
    “The goanna smelt good.”                 Tape 8185

3.2.1 Historical and reconstructed rules for YG -DHa-y ‘Associated Eating’. It is relatively easy to list what is found in the YG sources, and a ‘historical grammar’ could be limited to organising that material. However a reconstructed YG grammar aims to provide a full set of rules, in this instance for the suffix -Dha-y. Such a grammar could take notice of the Wangaaybuwan occurrences, the rare YG occurrences and could also ask the question ‘What are other likely uses of

---

5 After a version of this paper was presented Harold Koch noted that a number of Australian languages semantically linked eating, drinking and smoking and that this suffix might follow that pattern. The ‘smoking’ example was found on a subsequent re-examination of the tape transcripts.
the suffix which have not been captured in the historical material?'. I here
develop a conservative reconstructed grammar which does not explore that last
question.

In Wangaaybuwan the suffix is found with meaning ‘drinking’, and on
position verbs, but similar examples are not found in YG. This is more likely to
reflect the limitations of the historical material rather than traditional YG.
Given that eat/drink/smoke are semantically related in some Australian
languages, and that the suffix is associated with smoking, I would expect that
the suffix did refer to ‘drinking’ in YG. It also seems likely that the suffix
occurred on YG position verbs. There are some uses found in YG which are not
found in Wangaaybuwan – for instance with active, non-position, verbs, in
Wangaaybuwan the suffix only indicates purpose to do with eating/drinking,
whereas in YG it indicates ‘after eating’ or purpose - with the purpose often
being implicit, rather than stated.

The question arises as to what is the meaning in YG when the suffix is used
with an active, non-position, verb such as ‘hit’. The answer is context and the
semantics of the main verb. The Mathews example (27) could easily be part of a
longer sentence, such as bumadhay (birralii), ‘hit (the child) after eating’. The
‘after’ interpretation fits here. If however the longer sentence were bumadhay
bandaarr (bandaarr ‘kangaroo’) it would be translated with a purposive meaning
‘kill the kangaroo (in order to eat)’. It is hard to imagine a purposive translation
for (28) - ‘dance after eating’ is much more likely than ‘dance for the purpose of
eating.’ A ‘different subject’ meaning is implied by a main verbs such as ‘give’.
Both different subject and same subject interpretations are possible with ‘cook’
and ‘bring’, and a ‘same subject’ interpretation is the only likely one with verbs
such as ‘sit’ and ‘be ill’ and ‘die’. These rules for -DHa-y in YG are the same as the
Wangaaybuwan rules, with the additional meanings: ‘after eating’ and
‘smoking’.

The table below shows the main well understood or easily specified
semantics of the suffix in Wangaaybuwan and YG. For YG it contains both rules
based on the YG sources and rules (marked#) which copy the Wangaaybuwan
rule in the absence of any YG evidence. Given the very close similarity between
the suffixes in other aspects, the assumption is that these occurred in YG but
have not been recorded in the historical sources. It is also assumed that the use
of the YG suffix is obligatory in the same circumstances that the Wangaaybuwan
suffix is obligatory. The YG evidence points to this but is not conclusive.

A full set of rules for the suffix will also include other material. One example
is the ordering and co-occurrence with other suffixes. Examples (30) and (32)
show the EAT preceding the continuous suffixes. No examples have been found
of the EAT and BACK suffixes co-occuring, but this is more likely to reflect the
limitations of the sources rather than the intrinsic qualities of the language. The
order of the EAT and BACK suffixes is not certain and a topic for further
investigation and determination.
### 3.3 Historical development of the suffix.

While this paper is mainly about the semantics of the suffix, some preliminary comments will be made about the source of the suffix, and cognate forms. It is clearly derived from the verb *dha-li* ‘eat’, and has been found only in the Central NSW languages, so is likely to be a genetic feature. There are other suffixes which are formally identical. In YG the Regular continuous suffix for rr class verbs is –*dha-y*. Two of the common rr class verbs are *wuu-rrri* ‘give’ and *dhu-rrri* ‘spear’, the ‘eat’ suffix is found on these verbs and it has presumably been re-analysed as the continuous suffix. Wangaaybuwan and Wangkumara have formally identical suffixes with a benefactive meaning, suffix –*DHa-y* ‘reflexive focus’ (Donaldson 1980:177), and *thayi* ‘do for oneself’ (Breen 2004:164). It is easy to imagine the ‘Associated eating’ suffix being reinterpreted with a benefactive meaning. This could easily happen with examples such as (29). As a free form *thayi* is ‘eat’, but none of the Wangkumara examples given are related to eating. A Muruwari suffix may be related, but at this stage the link is more speculative. (Oates 1988:186) has a morpheme –*tha* glossed ‘Object focus (OBJ)’ However the use of it in the examples given is very similar to ‘Associated eating.’ Oates’ 5.405 is reproduced here as (37). Her other example of the suffix, 5.406, includes the verb *kaa-tha* ‘take’ in the sentence: ‘Go and get water and have a drink.’ Both occurrences of –*tha* are open to an ‘Associated eating’ interpretation. It may be that an ongoing investigation will find examples of related morphemes in other languages.

---

6 Muruwari borders Yuwaalaraay to the North-west, and Wangkumara is further to the north-west.

7 The difference between ‘t’ and ‘d’ is merely one of the spelling system adopted by a particular language.
(37) Ti-tjana maa-tha-ga tapa-ku
that put-OBJ-IMP supper-DAT
“Put the billy on for supper.”

4. Conclusion

The paper has reviewed and refined the description of two categories of verb specification in Yuwaalaraay Gamilaraay. After revisiting the historical sources and comparing YG with Wangaybuwan, a well described and closely related language, I have suggested that for the continuous suffixes key distinctions are related to the semantics of the main verb root. The meaning of the ‘associated eating’ suffix depends on the semantics of the main verb root and on context. There is scope for further work in YG grammar, both more historical study and grammar development. This will result in a better understanding of the traditional language and provide a sound base for language revival.

References

THE ORIGIN OF CONJUGATION MARKERS
IN AUSTRALIAN LANGUAGES

MARK HARVEY

The University of Newcastle

In this paper, I compare two approaches to the reconstruction of verbal paradigms in Australian languages. The immediate point at issue is illustrated by the Guugu Yimidhirr conjugational system. The relevant data is marked in bold.

<table>
<thead>
<tr>
<th>RR</th>
<th>L</th>
<th>VOWEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>ngalpu 'close'</td>
<td>karrpa 'hold, grab'</td>
<td>thatha 'go'</td>
</tr>
<tr>
<td>Purposive</td>
<td>ngalpu-nhu</td>
<td>karrpa-nhu</td>
</tr>
<tr>
<td>Counterfactual</td>
<td>ngalpu-nta</td>
<td>karrpa-nta</td>
</tr>
<tr>
<td>Subordinate 2</td>
<td>ngalpu-nhun</td>
<td>karrpa-nhun</td>
</tr>
<tr>
<td>Non-Past</td>
<td>ngalpu-rr</td>
<td>karrpa-l</td>
</tr>
<tr>
<td>Imperative</td>
<td>ngalpu-rr-u</td>
<td>karrpa-l-a</td>
</tr>
<tr>
<td>Past Negative</td>
<td>ngalpu-rr-nta-nta</td>
<td>karrpa-al-muku</td>
</tr>
<tr>
<td>Cautionary</td>
<td>ngalpu-rr-nta-nta</td>
<td>karrpa-ya</td>
</tr>
<tr>
<td>Anticipatory</td>
<td>ngalpu-rr-nta-nta</td>
<td>karrpa-yiku</td>
</tr>
<tr>
<td>Past</td>
<td>ngalpu-rr-nta-nta</td>
<td>karrpa-y</td>
</tr>
<tr>
<td>Subordinate 1</td>
<td>ngalpu-rr-nta-nta</td>
<td>karrpa-ya</td>
</tr>
<tr>
<td>Precautionary</td>
<td>ngalpu-rr-nta-nta</td>
<td>karrpa-ya</td>
</tr>
<tr>
<td>Members</td>
<td>about 50</td>
<td>about 150</td>
</tr>
</tbody>
</table>

Table 1: Guugu Yimidhirr - northern coastal Queensland (Haviland 1983:81)

The material in bold cannot be straightforwardly assigned to either the root or the tense/aspect/mood suffixes. This is obvious in terms of the surface forms, if no allowance is made for any phonological processes. However, even if allowance is made for phonological processes, straightforward assignment remains problematic.
We may consider the RR and L conjugations. One alternative analysis of these conjugations is that the root in these paradigms is consonant-final. Thus ‘close’ would be ngalpurr, and ‘hold’ karrpal. Whenever this consonant-final root took a consonant-initial suffix, a cluster would result, at least theoretically. Consonant clusters are commonly subject to assimilation and deletion, which would then produce the actual forms.

This analysis cannot account for the Guugu Yimidhirr forms systematically. We may consider the Anticipatory category, where the suffix is evidently -yiku.

<table>
<thead>
<tr>
<th>Root</th>
<th>Anticipatory</th>
<th>Attested form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ngalpurr</td>
<td>-yiku</td>
<td>ngalpurr-iku</td>
</tr>
<tr>
<td>karrpal</td>
<td>-yiku</td>
<td>karrpa-yiku</td>
</tr>
</tbody>
</table>

Table 2

There does not appear to be any systematic explanation for the deletion of the /y/ in the putative /rr+y/ cluster, but the deletion of the /l/ in the putative /l+y/ cluster. The consonant-final analysis is inconsistent, not only within the verbal paradigms, but also inconsistent when the verbal paradigms are compared with nominal paradigms. In the 1970s, the oldest and most fluent speakers of Guugu Yimidhirr had the following patterns for nominals (Haviland 1979:41-45).

<table>
<thead>
<tr>
<th>Root</th>
<th>Ergative</th>
<th>Attested Form(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘thunder’</td>
<td>wulungkurr</td>
<td>-nta wulungu-nta wulungkurr-nta</td>
</tr>
<tr>
<td>‘hand’</td>
<td>mangal</td>
<td>-nta manga-rnta</td>
</tr>
</tbody>
</table>

Table 3

Following the consonant-final hypothesis for verb roots, the forms involving the Counterfactual suffix –nta are set out in Table 4.

<table>
<thead>
<tr>
<th>Root</th>
<th>Anticipatory</th>
<th>Attested form</th>
<th>Not attested</th>
</tr>
</thead>
<tbody>
<tr>
<td>ngalpurr</td>
<td>-nta</td>
<td>ngalpu-nta</td>
<td>*ngalpurr-nta</td>
</tr>
<tr>
<td>karrpal</td>
<td>-nta</td>
<td>karrpa-nta</td>
<td>*karrpa-rnta</td>
</tr>
</tbody>
</table>

Table 4

The historical question is the origin of these segments which cannot synchronically be assigned to either the root or the suffix. Under one hypothesis, these segments were originally the final segments of verbal roots. Various historical changes have led to the loss of these final segments in some forms, with the result that the analysis of forms where the final segments survive is synchronically problematic.

The other hypothesis is that these segments were originally suffixes, and that the combination of these suffixes and the root served as a stem for further suffixation in some parts of verbal paradigms. In other words, some suffixes were inner suffixes and attached directly to the root. Other suffixes were outer suffixes and attached to stems which consisted of the root and an inner suffix. Various historical changes have obscured this root vs stem distinction.
I will show that the evidence supports the second hypothesis and not the first. The choice between these two hypotheses is important to Australian historical linguistics generally. The issue of unassignable segments arises in many languages, distributed across the continent, as illustrated in Tables 5-7.

<table>
<thead>
<tr>
<th>-YA Conjugation</th>
<th>-YA Conjugation</th>
<th>-LA Conjugation</th>
</tr>
</thead>
<tbody>
<tr>
<td>wangka 'say, tell'</td>
<td>warni 'fall'</td>
<td>paja 'bite'</td>
</tr>
<tr>
<td>wangka-nyja</td>
<td>warni-nyja</td>
<td>paja-nta</td>
</tr>
<tr>
<td>wangka-ku</td>
<td>warni-ku</td>
<td>paja-l-ku</td>
</tr>
<tr>
<td>wangka</td>
<td>warni</td>
<td>paja-n</td>
</tr>
<tr>
<td>wangka-ma</td>
<td>warni-ma</td>
<td>paja-n-ma</td>
</tr>
<tr>
<td>wangka-manyja</td>
<td>warni-manyja</td>
<td>paja-n-manyja</td>
</tr>
<tr>
<td>wangka-ya</td>
<td>warni-ya</td>
<td>paja-l-a</td>
</tr>
<tr>
<td>wangka-nya</td>
<td>warni-nya</td>
<td>paja-na</td>
</tr>
</tbody>
</table>

Irregular Verbs

<table>
<thead>
<tr>
<th>Participle</th>
<th>Purpose</th>
<th>Perfect Imp</th>
<th>Impf Imp</th>
<th>Present</th>
<th>Future</th>
<th>Past</th>
</tr>
</thead>
<tbody>
<tr>
<td>nyi-nya-nya</td>
<td>pi-nya-nya</td>
<td>pi-nya-ku - pi-nya-ku-rlu</td>
<td>pi-nya-ku</td>
<td>pi-nya</td>
<td>pi-nya-ku ~ pi-nya-ku-rlu</td>
<td>pi-nya-nya ~ pi-nya</td>
</tr>
<tr>
<td>nyi-nya-nya-nya</td>
<td>pi-nya</td>
<td>pi-nya-ku</td>
<td>pi-nya-ku</td>
<td>pi-nya-nya</td>
<td>pi-nya-ku ~ pi-nya-ku-rlu</td>
<td>pi-nya-nya-nya</td>
</tr>
</tbody>
</table>

Table 5: Watjarri - southern Western Australia (Douglas 1981:230-231)

<table>
<thead>
<tr>
<th>Regular</th>
<th>Irregular</th>
</tr>
</thead>
<tbody>
<tr>
<td>paana 'arrive'</td>
<td>nyaa 'see'</td>
</tr>
<tr>
<td>paana-ng</td>
<td>nyaa-wang</td>
</tr>
<tr>
<td>Present</td>
<td>Purposive</td>
</tr>
<tr>
<td>paana-y</td>
<td>nyaa-yaki - nyaa-kay</td>
</tr>
<tr>
<td>paana-y-ku</td>
<td>nyaa-kay-ku</td>
</tr>
<tr>
<td>paana</td>
<td>nyaa-(y)ka - nyaa-yaki</td>
</tr>
<tr>
<td>~ paana-la</td>
<td>~ paana-ka</td>
</tr>
<tr>
<td>Future</td>
<td>paana-w</td>
</tr>
</tbody>
</table>

Table 6: Gumbaynggirr - northern coastal New South Wales (Eades 1979:301)
Dixon (1980:408-409) proposes the following synchronic analysis of the conjugations set out in Table 1 and Tables 5-7.

The most efficient way to describe the verbal morphology of just about every modern Australian language outside the prefixing area is to set up a structure

ROOT (+CONJUGATION MARKER) + INFLECTION

specifying that the conjugation marker morpheme (which is always a single consonant – y, ng, m, n, l or rr) is included only before certain inflections, and then only for certain conjugations. All roots, set up under this scheme, end in a vowel.

As comparison of Table 1 and Tables 5-7 shows, cognate verb roots and their attendant conjugation markers are widely distributed across Australia. Dixon (2002:215-216) proposes a number of verbal reconstructions of this type, on the basis of verbal paradigms from across the continent.

(1) N conjugation: *ya-n 'go', *pa-n 'fall', *thaa-n ~ *ta-n 'copulate with', *thu-n 'put, tell, say', *thu-n 'swear at, be angry, scold', *nyii-n ~ *ni-n 'sit', *thaa-n ~ *ta-n 'stand', *nguu-n (*wu-n) 'lie'

M conjugation: *pu-m 'hit', *ngaa-m 'hear, understand', *la-/ra-/ta-/ya-m 'spear, throw', *wal-m 'get up'

NG conjugation: *nhaa-ng ~ *na-ng 'see', *nyu-/yu-/ngu-/wu-ng 'give', *ngu-ng 'eat', *kaa-ng 'take, hold, carry', *nuu-ng ~ *nu-ng 'smell', *lu-/ru-/tu-/yu-ng 'cry', *waa-ng 'follow'

NY conjugation: *maa-ny 'hold, get, take'

L conjugation: *ma-l 'do, make, tell, say', *nga-l 'eat', *hta-l 'eat, consume', *kunpa-l 'cut', *paka-l 'pierce, dig, spear, etc', *nyma-l 'hold, pinch, squeeze', *thuupa-l 'spit', *yaa-l 'speak', *nuu-attha-l 'kiss', *matha-l 'chew, bite, suck, eat'

RR conjugation: *patha-rr 'bite, eat, drink, smoke'

Y conjugation: *tharrpa-y 'enter, dive', *wanta-y 'fall', *kampa-y 'burn, cook'

Vowel conjugation: *yung(k)a-Ø 'swim', *wula-Ø 'die, disappear', *luwa-Ø 'hit with a missile'

Dixon (2002:187-193) provides an overview showing that synchronically the degree of openness of directly inflecting verbal paradigms varies considerably.
In many languages, the entire system is itself closed. In nearly all systems, regardless of whether the overall system is closed or not, there is a considerable disparity between the membership of various conjugations. In nearly all languages, there are one or two large conjugations, and a number of very small conjugations.

The grouping of small closed paradigms with large open paradigms, under the label 'conjugation' is misleading. Both synchronically and diachronically, these two classes of paradigms show distinct patterns. It would be descriptively more adequate to restrict the term 'conjugation' to large paradigms, and analyse the small closed 'conjugations' as a set of irregular verbs.

There is evidence that the status of these reconstructed 'conjugations' varied along the closed – open continuum. This evidence is summarised in (2), based on Dixon (2002:227-228).

(2)  
   a. The M class is always a small closed class with a few monosyllabic members.  
   b. The NG class is generally a small closed class. It is somewhat larger in the Yolngu languages, but the verbal system as a whole is closed in the Yolngu languages (Morphy 1983:65). Consequently, there is no language in which the NG class is open.
   c. The RR class is a smallish class (6–50) members. In no language synchronically is it an open class.
   d. The L, N, Y and vowel classes tend to be larger. The L class is an open class in a number of languages.

Based on this evidence, it is reasonable to reconstruct the M and NG classes as closed classes, and the L class as an open class. It seems likely that the N, Y and vowel classes were also open classes, given their larger membership. I reconstruct the RR class as open on the basis of those languages, such as Guugu Yimidhirr, where it has 50 members. This is too large a membership to have derived from an originally closed class.

Dixon (1980:409) proposes the following hypothesis for the diachronic development of conjugation markers.

We suggest that proto-Australian verbs had the following structure:

\[
\text{ROOT} \ast \text{INFLECTION}
\]

Roots could end just in a vowel or in y, ng, m, n, l or rr [the 'conjugation marker']. Verbal inflections were simply added to the root. We are positing a simple agglutinative-type structure, where the morpheme boundary between root (or stem) and affix would have been consistently and clearly recognizable.

Dixon (2002:213-214) proposes the following inflections.

\[\text{Inflections}\]

1 Dixon (2002:215) restates this hypothesis.
The following types of changes are required in the development from the posited proto-structure of $\text{ROOT} + \text{-ka/-nhu} \sim \text{-nyu/-ku}$ to the attested reflexes (Dixon 1980:413).²

(3) [i] Assimilation e.g. -n + ka > -nta for imperative in Walmatjari and Warlpiri
[ii] Blending e.g. -y + ku > -ju for future in Warlpiri
[iii] Deletion e.g. -y + ka > -ya for imperative in Warlpiri and -rr + ka > -rra for imperative in Walmatjari
[iv] Lenition e.g. -ng + ka > -ka > -wa and -n + ka > -nta > -rra for imperative in Western Desert
[v] Epenthesis e.g. -rr + lany > -rralany for customary in Walmatjari and -n + thi > -nay for past in Guugu Yimidhirr

Alpher (1990:167-169) argues that these required changes are arbitrary and unsupported by the general historical phonology of the relevant languages. He proposes that fully inflected forms in the various daughter languages should be compared with one another, in order to reconstruct fully inflected forms in a proto-language. Further, he proposes that there would have been no stage at which the reconstructed verbal system would have been uniformly analyzable into a straightforward agglutinative $\text{ROOT} + \text{INFLECTION}$ pattern.

Dixon (2002:222-223) responds to Alpher’s critique by arguing that reconstruction of verbal paradigms cannot be determined solely by the regularity of proposed sound changes. I agree with Dixon that it cannot be the sole criterion. However, regularity of sound change must be the default criterion. Any departures from this default criterion must follow from the operation of other principles which are independently motivated. Dixon does not provide any such independently motivated criteria.

One important criterion is productivity. It is well established that productive systems show different patterns of change to non-productive systems (Hock 2003:441). If a system involves both small closed unproductive paradigms and large open productive paradigms, it is predicted that the two types of paradigms will show different patterns of change.

Over time, the structure of the members of small closed paradigms shifts towards the structure of simplex word forms. Phonological structures which indicate internal morphological divisions tend to be eliminated. This elimination proceeds on an individual basis, and not on a regular basis. In other words, token effects can play a major diachronic role in small closed paradigms. Type effects may also play a role.

² Dixon (2002:224-225) restates these changes, but in a less explicit form.
By contrast, the members of large open paradigms do not shift towards simplex word structures on an individual basis. Rather, phonological structures indicative of internal morphological divisions tend to be maintained, precisely because the paradigm is open. Further, if there is an elimination of some such structures and a shift towards simplex word structure, then it proceeds on a paradigmatic basis and not an individual basis. In other words, token effects do not play a diachronic role in large open paradigms. Only type effects will have a diachronic role.

As we have seen, there is evidence for variation in the status of reconstructed verbal paradigms on the closed – open continuum. Consequently, there is no reason to posit that verbal paradigms will have had a uniform history, either in source or trajectory. Rather, different paradigms will have had distinct histories, depending upon their degree of openness over time. This paper does not examine the full range of differences that may be predicted. It examines only one – the use of substantively inflected verbal forms as stems for other verbal forms.

I begin by providing some further data – in this case from the Non-Pama-Nyungan languages of northern Australia.

<table>
<thead>
<tr>
<th>Root</th>
<th>Non-Past</th>
<th>Past Imperf</th>
<th>Past Perf</th>
</tr>
</thead>
<tbody>
<tr>
<td>'cry'</td>
<td>ru</td>
<td>ru-n</td>
<td>run-iny</td>
</tr>
<tr>
<td>'die'</td>
<td>tho(wi)</td>
<td>tho(wi)-n</td>
<td>tho(wi)-ny</td>
</tr>
<tr>
<td>'eat'</td>
<td>ngu - ngo</td>
<td>ngu-n</td>
<td>ngun-iny</td>
</tr>
<tr>
<td>'follow'</td>
<td>wa</td>
<td>wa-n</td>
<td>wan-iny</td>
</tr>
<tr>
<td>'give'</td>
<td>wo</td>
<td>wo-n</td>
<td>won-iny</td>
</tr>
<tr>
<td>'hear'</td>
<td>nga</td>
<td>nga-n</td>
<td>ngar-iny</td>
</tr>
<tr>
<td>'hit'</td>
<td>pu ~ po</td>
<td>pu-n</td>
<td>pun-iny</td>
</tr>
<tr>
<td>'inchoative'</td>
<td>-me</td>
<td>-me-n</td>
<td>men-iny</td>
</tr>
<tr>
<td>'inchoative'</td>
<td>-thi</td>
<td>-thi-n</td>
<td>thin-iny</td>
</tr>
<tr>
<td>'reciprocal'</td>
<td>-nthi ~ -nyci</td>
<td>-nthi-n ~ -nyci-n</td>
<td>-nthi-ny ~</td>
</tr>
<tr>
<td>'see'</td>
<td>na</td>
<td>na-n</td>
<td>nan-iny</td>
</tr>
<tr>
<td>'spear'</td>
<td>ra ~ re</td>
<td>re-n</td>
<td>ren-iny</td>
</tr>
<tr>
<td>'take'</td>
<td>ka</td>
<td>ka-n</td>
<td>kan-iny</td>
</tr>
<tr>
<td>'chop/crush'</td>
<td>tho</td>
<td>tho-ng</td>
<td>thong-iny</td>
</tr>
<tr>
<td>'get'</td>
<td>ma</td>
<td>ma-ng</td>
<td>manq-iny</td>
</tr>
<tr>
<td>'punch'</td>
<td>patca</td>
<td>patca-ng</td>
<td>patca-ng</td>
</tr>
<tr>
<td>'tell off'</td>
<td>thu</td>
<td>thu-ng</td>
<td>thunq-iny</td>
</tr>
<tr>
<td>'do, say'</td>
<td>ma</td>
<td>ma-r</td>
<td>mar-ay</td>
</tr>
<tr>
<td>'lie'</td>
<td>yo - yu</td>
<td>yu - yo-ngen</td>
<td>yo-y</td>
</tr>
<tr>
<td>'sit'</td>
<td>ni</td>
<td>ni</td>
<td>ni-ny</td>
</tr>
<tr>
<td>'be standing'</td>
<td>thi</td>
<td>thi</td>
<td>thi-ny</td>
</tr>
<tr>
<td>'stand up'</td>
<td>tha</td>
<td>tha-ngen</td>
<td>tha-ny</td>
</tr>
</tbody>
</table>

Table 9: Proto-Gunwinyguan (Alpher, Evans & Harvey 2003:345)

Alpher, Evans and Harvey (2003:347-349) make the following points about the Proto-Gunwinyguan paradigms, their descendents in the various Gunwinyguan languages, and their relationship to forms in Pama-Nyungan languages.

(4)  a. The reflexes of the ‘conjugation markers’ cannot be so analysed for PGN. They are simply Tense/Aspect/Mood suffixes. This is also true for all daughter languages.

b. The reflexes of the ‘conjugation markers’ are unpredictably distributed across the TAM categories. Most appear in the Past Perfective, e.g. *po-m ‘hit-PP’ – Dixon (2002:216) *pu-m ‘hit’. However, some appear in the Non-Past, e.g. *ma-r ‘say/do-NP’ – Dixon (ibid) *ma-l ‘do, make, tell, say’

c. The PGN Non-Past generally serves as the stem for the Past Imperfective.

The final point is the issue of central concern in this paper. Alpher, Evans and Harvey (ibid) draw out its implications “Moreover as it appears that in Proto-Gunwinyguan the Past Imperfective consisted of the Non-Past + *i/any, it is not difficult to conceive of changes that would cause the Non-Past inflections *-n, *-ng, *-r to be reanalyzed as conjugation markers.”

In this paper, I will provide evidence that this type of structural change – the use of substantively inflected verb forms as stems – is indeed the source of the ‘conjugation markers’. In response to this hypothesis Dixon (2002:223-224) states:

Alpher, Evans and Harvey … suggest that verbal paradigms with ‘conjugation markers’ … developed out of a system similar to that in Gunwinygu … by analogical extension of some consonants (e.g. –m in bo-m ‘hit-Past.Completed’, -ng in wo-ng ‘give-Past.Completed’) to occur between root and inflections for all TAM choices. This would involve analogical changes of unusual power. Some explanation is then needed for why ‘conjugation markers’ are found just in some ‘Pama-Nyungan’ languages but not in others, and why – in those languages in which they do occur – they occur at certain root-affix boundaries and not at others.

Dixon’s response is misdirected in one central point. The change mechanism proposed by Alpher, Evans and Harvey is not analogical, in any usual sense of this term. Analogical change is change where forms in one paradigm are changed to conform to the pattern of forms in another paradigm. Thus bring ~ brought is changed to bring ~ brang ~ brung on the basis of sing/ring ~ sang/rang ~ sung/rung. Alpher, Evans and Harvey propose change by re-analysis.

(5) \[ \text{[[verb root + TAM suffix]}_{stem} + \text{suffix A /} + \text{suffix B /} + \text{suffix C]}_{word} > \]
\[ \text{[[verb root + conjugation marker]}_{stem} + \text{suffix A /} + \text{suffix B /} + \text{suffix C]}_{word} \]
I will show that the answers to the other issues raised by Dixon may straightforwardly be accommodated within the line of analysis in (5).

As a starting point, we may consider the use of the Non-Past as a stem for the Past Imperfective in Proto-Gunwinyguan, and in a number of daughter Gunwinyguan languages. This appears counter-intuitive, at first sight, as there is a Past Perfective form. It might be predicted that this should serve as the stem for the Past Imperfective. However, this pattern, where a Non-Past form serves as the stem for a Past form, is independently attested elsewhere in Australia. In Guugu Yimidhirr, the Non-Past serves as the stem for the Past Negative in the two major conjugations – the L and RR conjugations (Table 1).

The Non-Past also serves as a stem for irrealis forms. Baker (2004) shows that the Gunwinyguan languages, Ngalakgan and Rembarrnga have innovated new Future and Irrealis forms using the inherited Non-Past as a stem. The Non-Past also commonly serves as the stem for various derived non-finite verb forms. For example, in Warlpiri, the Non-Past serves as the stem for the Infinitive and the Participle.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-past (a) (=prt)</strong></td>
<td>wangka-Ø 'speak'</td>
<td>paka-rni 'hit'</td>
</tr>
<tr>
<td><strong>Non-past (b)</strong></td>
<td>wangka-mi</td>
<td></td>
</tr>
<tr>
<td><strong>past</strong></td>
<td>wangka-ja</td>
<td>paka-rnu</td>
</tr>
<tr>
<td><strong>future</strong></td>
<td>wangka-ji</td>
<td>paka-ka</td>
</tr>
<tr>
<td><strong>imperative</strong></td>
<td>wangka-ya</td>
<td>paka-ka</td>
</tr>
<tr>
<td><strong>irrealis</strong></td>
<td>wangka-ya-rla</td>
<td>paka-ka-rla</td>
</tr>
<tr>
<td><strong>infinitive</strong></td>
<td>wangka-Ø-nya-</td>
<td>paka-nya-</td>
</tr>
<tr>
<td><strong>participle = non-past(a)</strong></td>
<td>wangka-Ø-nja-</td>
<td>paka-nya-</td>
</tr>
<tr>
<td><strong>Number of members:</strong></td>
<td>± 60</td>
<td>± 85</td>
</tr>
</tbody>
</table>

Table 10a: Warlpiri (Laughren p.c)

<table>
<thead>
<tr>
<th></th>
<th>3a</th>
<th>3b</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-past (a) (=prt)</strong></td>
<td>kanja-Ø 'take'</td>
<td>nga-rni 'eat'</td>
<td>ma-ni 'get'</td>
<td></td>
</tr>
<tr>
<td><strong>Non-past (b)</strong></td>
<td>kanja-nya</td>
<td>ka-nju</td>
<td>nga-rnu</td>
<td>ma-nu</td>
</tr>
<tr>
<td><strong>past</strong></td>
<td>kanja-nku</td>
<td>ka-nga-nya</td>
<td>nga-rnu-ka</td>
<td>ma-nku-ka</td>
</tr>
<tr>
<td><strong>future</strong></td>
<td>kanja-nka</td>
<td>ka-ngka</td>
<td>nga-nja</td>
<td>ma-nja</td>
</tr>
<tr>
<td><strong>imperative</strong></td>
<td>kanja-nka-rla</td>
<td>ka-ngka-rla</td>
<td>nga-rla</td>
<td>ma-rla</td>
</tr>
<tr>
<td><strong>irrealis</strong></td>
<td>kanja-nka-rla-</td>
<td>nga-rni-rla</td>
<td>ma-rni-rla</td>
<td></td>
</tr>
<tr>
<td><strong>infinitive</strong></td>
<td>kanja-nja-</td>
<td>nga-rni-nya-</td>
<td>ma-nja-</td>
<td></td>
</tr>
<tr>
<td><strong>participle</strong></td>
<td>kanja-nja-</td>
<td>nga-rni-</td>
<td>ma-ni-</td>
<td></td>
</tr>
<tr>
<td><strong>= non-past(a)</strong></td>
<td>kanja-nja-</td>
<td>nga-rni-</td>
<td>ma-ni-</td>
<td></td>
</tr>
<tr>
<td><strong>Members</strong></td>
<td>± 3</td>
<td>± 1</td>
<td>± 3</td>
<td></td>
</tr>
</tbody>
</table>

Table 10b: Warlpiri (Laughren p.c)

Finally, there are cases where the Non-Past functions as the stem for nearly all of the verbal paradigm, as in Yidiny.
It may be observed that for the L and R conjugations, the Non-Past functions as the stem for the rest of the conjugation, with two exceptions – the Imperative, and the Antipassive in the L conjugation.

The use of Past tense forms as stems varies according to the status of the conjugation. In large or open conjugations, the use of Past tense forms as stems is very rare. The only example in available materials is from Guugu Yimidhirr (Table 1), where the Past forms the stem for the Subordinate 1 tense. The Subordinate 1 tense, in turn, serves as the stem for the Precautionary. The Subordinate 1 functions as a Past Perfective when it appears independently. In co-ordinate clauses, it signals causal or simultaneous relations. The Precautionary appears only in co-ordinate clauses, where it indicates the undesirable result of a potential causal chain, which may be avoided by performing an action.

The relevant point is that these three tenses are semantically connected. The Past and Subordinate 1 by their common marking of past time, and Subordinate 1 and Precautionary by their common concern with causal chains. As such, the use of the Past and Subordinate 1 as stems has a semantic motivation.

### Table 11a: Yidiny (Dixon 1977:212)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imperative</td>
<td>wonta 'fall'</td>
<td>paja 'bite'</td>
<td>paja 'leave'</td>
</tr>
<tr>
<td>Non-Past</td>
<td>wonta-n</td>
<td>paja-l</td>
<td>paja-r</td>
</tr>
<tr>
<td>Past</td>
<td>wonta-ng</td>
<td>paja-al</td>
<td>paja-ar</td>
</tr>
<tr>
<td>Purposive</td>
<td>wonta-ana</td>
<td>paja-al-na</td>
<td>paja-ar-na</td>
</tr>
<tr>
<td>Dative Sub</td>
<td>wonta-nyunta</td>
<td>paja-l-nyunta</td>
<td>paja-r-nyunta</td>
</tr>
<tr>
<td>Causal Sub</td>
<td>wonta-anyum</td>
<td>paja-al-nyum</td>
<td>paja-ar-nyum</td>
</tr>
<tr>
<td>Lest + Abs</td>
<td>wonta-anyji</td>
<td>paja-al-ji</td>
<td>paja-ar-ji</td>
</tr>
</tbody>
</table>

### Table 11b: Yidiny Comitative (Dixon ibid: 216)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>'eat'</td>
<td>puka-aji-n</td>
<td>wungapa-aji-n</td>
<td>paja-aji-n</td>
</tr>
<tr>
<td>‘walk up’</td>
<td>majinta-nga-l</td>
<td>maki-l-manga-l</td>
<td>payka-r-manga-l</td>
</tr>
</tbody>
</table>

### Table 11c: Yidiny Antipassive (Dixon ibid: 218)

Yidiny has a very complex pattern of distribution of long vowels, correlating partly with the appearance and non-appearance of word-final vowels (Dixon 1977:42-56). The variation between paja-l and paja-al and between paja-r and paja-ar in (Tables 11a-c) follows from the phonological factors conditioning the distribution of long vowels. It does not follow from any morphological or semantic factors.
Similarly, the use of the Imperative and other forms with irrealis reference, as stems, is restricted to categories which are also irrealis. In Warlpiri, the Imperative is the stem for the Irrealis (see also the Nyangumarta Imperative and Future in Table 12). By contrast with the use of the Past and Imperative as stems, no equivalent semantic connection can be proposed for a number of the uses of the Non-Past as a stem for at least Past Realis Imperfective and Non-Finite forms.

I propose that the use of the Non-Past as a stem does not depend on semantic connections to the categories where it functions as a stem. Rather, it follows from the fact that forms with present time reference function as citation forms for verbs. Thus in Warlpiri, the Non-Past is the citation form of the verb (M. Laughren p.c.), and it functions as the stem for non-finite forms. The central role of present time reference is brought out by a variation in Dyirbal, discussed by Dixon (2002:210-211).

In Dyirbal, there are two tense inflections, one which always refers to future and one which always refers to past time. In southern dialects the past form covers present time, giving a future/non-future contrast (note that the non-future is used as citation form). In northern dialects the future form also covers present time, so that we have here a past/non-past contrast (in these dialects the non-past form is used in citing the verb).

The effects of variation in the coding of present time reference for stem formation are particularly clearly illustrated in Nyangumarta.

<table>
<thead>
<tr>
<th></th>
<th>NY</th>
<th>RN</th>
<th>N</th>
<th>NG</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of members</td>
<td>42</td>
<td>151</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Imperative</td>
<td>-a/-i</td>
<td>-IV</td>
<td>-rra</td>
<td>-wa</td>
</tr>
<tr>
<td>Potential</td>
<td>-u</td>
<td>-lka</td>
<td>-nku</td>
<td>-ngku</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>-nyV</th>
<th>-rnV</th>
<th>-na</th>
<th>-nya</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>-yi-nyV</td>
<td>-rni-nyV</td>
<td>-ni-nyV</td>
<td>-nga-nyV</td>
</tr>
<tr>
<td>Purp/Advisory</td>
<td>-nya-ku</td>
<td>-na-ku</td>
<td>-ni-nyaku</td>
<td>-nga-nyaku</td>
</tr>
<tr>
<td>Present CF</td>
<td>-nyi-ka</td>
<td>-rni-ka</td>
<td>-na-nyaka</td>
<td>-nga-nyaka</td>
</tr>
<tr>
<td>Imperfective</td>
<td>-nyV-lingV</td>
<td>-rnV-lingV</td>
<td>-ni-(nyV)lingV</td>
<td>-nga-nyVlingV</td>
</tr>
<tr>
<td>Remote Future</td>
<td>-nya-ngkuliny</td>
<td>-rni-ngkuliny</td>
<td>-na-ngkuliny</td>
<td>-nga-ngkuliny</td>
</tr>
<tr>
<td>Past CF</td>
<td>-ma</td>
<td>-rni-ma</td>
<td>-na-ma</td>
<td>-nga-ma</td>
</tr>
<tr>
<td>Remote CF</td>
<td>-(nyV)-mal</td>
<td>-rni-mal</td>
<td>-na-mal</td>
<td>-nga-mal</td>
</tr>
<tr>
<td>Remote Past</td>
<td>-nyV(pV)</td>
<td>-rni(pV)</td>
<td>-na(pV)</td>
<td>-nga(pV) - nyal(pV)</td>
</tr>
</tbody>
</table>

Table 12: Nyangumarta (Sharp 2004:161-165)

It may be observed that the Non-Future functions extensively as a stem, including for the Remote Future category. This usage is to some degree obscured by phonological changes. Also, the NG conjugation shows a -nga stem formative, when its Non-Future is -nya. However, the two verbs in this conjugation ka
take’ and yi/yu ‘give’ may be reconstructed with *ng as their predominant marker (1). The most plausible explanation for this discrepancy in the NG conjugation pattern is that *-nga was historically the Non-Future suffix, and served as the stem for the derived tenses. However, it was subsequently displaced by as the Non-Future marker by the allomorph from the major NY conjugation.

The Nyangumarta Non-Future suffixes have cognates over a wide area, but always as exclusively Past suffixes. For example, the Watjarri (Table 5) and Warlpiri (Tables 10a-b) Past suffixes are cognate. The non-future meaning range in Nyangumarta is a development from this Proto-Past meaning. The relevant point is that in no language, where these suffixes maintain their exclusively past time reference, do inflected verb forms involving these suffixes function as stems. I propose that their use as stems in Nyangumarta was subsequent to and dependent on the expansion of their range to include present time reference.

It may be noted that Nyangumarta has a Present tense, distinct from the Non-Future tense. However, the Present tense is defective.

In the minor conjugations there is a formal distinction (for all persons) for non-future and present tense … however, in the major conjugational classes, non-future forms are distinguished from present tense forms only in 3SG forms. Non-future is the inflection which surfaces in forms in which no distinction is made between non-future and present tense. (Sharp ibid:168)

I propose that defective tenses cannot function as stems, at least in large and/or open conjugations.

The situation may be different in small, closed conjugations. Certainly, small closed conjugations can show patterns of stem formation which are not found in large and/or open conjugations. In Guugu Yimidhirr (Table 1), the Past functions as a stem for a wide range of categories with the irregular verbs. This includes categories such as the Purposive, which have no semantic connection with the Past.

There does not appear to be a consistent pattern to the development of stems in small, closed conjugations. In Australia, most of the reconstructed verb roots in these conjugations are monosyllabic. In many areas of Australia, there has been a change with monosyllabic roots becoming proscribed. To accommodate this change, disyllabic inflected forms are adopted as the new roots.

Dixon (2002:229-230) examines this change with the ‘see’, ‘hit’ and ‘go’ verbs, and shows that almost any inflected form may be adopted as a root. This is a classic ‘token’ effect pattern. Given that the most frequent token can vary widely from language to language, the stem forms may vary.

I have already proposed that token effects may be important in the diachrony of small closed paradigms, whereas they will not play a role in the history of large open paradigms. The diachrony of both types of paradigms may
involve type effects. I propose that the principal type effect in the diachrony of ‘conjugation markers’ is that set out in (6).

(6) Forms with present time reference may be used as stems in any category

I propose that there is another type effect which is also relevant to the history of conjugation markers.

(7) Forms which do not have present time reference may be used as stems only where there is a connection in the definitional semantics of the categories – e.g. common past time reference, common irrealis status, common marking of causal chains

Given these type effect principles, inflected forms with present time reference will be the preferred stems in large and/or open conjugations. They will also be used as stems in small, closed conjugations, but these may also use other forms as stems as a result of token effects. Given further that the status of reconstructed verbal paradigms varied along the closed – open continuum, it may be predicted that there was variation in the forms used as stems.

Stems from open conjugations should be reconstructed with present time reference, usually Non-Past as the Future vs Non-Future distinction is uncommon. As discussed in (2), there is evidence for the L class being an open class. The root + L suffix form can be reconstructed with a Non-Past meaning.


The Proto-Gunwinyguan form *ma-r ‘do/say-NONPAST’ belongs with this set as it is a reflex of the *ma-l ‘do, make, tell, say’ verb. The important point to note is that there is no language in which the reflex of *–l functions as a Past tense inflection. Similarly, the root + RR suffix appears to be reconstructable as a Non-Past.


It does not appear possible to reconstruct root + N suffix, or root + Y suffix, or root + Ø (vowel conjugation) forms. It may be noted that Gugu Yalanji has –y as the inflection for the Non-Past and Imperative in the Y conjugation (Patz 2002:92). However, there are no other languages where –y appears by itself as a suffix in the Y conjugation.

The reflexes of stems from the closed M and NG conjugations show a quite different pattern to those from the L and RR conjugations. There are reflexes of

---

4 The languages listed in (9) are those which have distinct reflexes of the RR conjugation from the L conjugation.
root + M and root + NG forms in Proto-Gunwinyguan and its daughter languages (Table 9). However, these are Past Perfective forms, and not Non-Past forms.

This difference in the distribution of the conjugation markers, with the L, RR and Y markers appearing in Non-Past tenses, whereas the M and NG markers, at least, appear in Past tenses, accords with the predictions of the hypothesis that the conjugation markers originated as inner, stem-forming suffixes. We predict that open conjugations, such as the L and RR conjugations, will take ‘Present’ forms as stems, whereas closed conjugations, such as the M and NG conjugations, may take any form as a stem.

The other hypothesis proposes that the conjugation markers originated as the final segment of roots. If they all originated as root-final segments, then there is no explanation for the difference in distribution between the L and RR conjugation markers on the one hand, and the M and NG conjugation markers on the other hand.

Therefore, the evidence is against the hypothesis that the conjugation markers originated root-final segments, and in favour of the hypothesis that they originated as inner, stem-forming suffixes.

References
Alpher, Barry, Nicholas Evans & Mark Harvey. 2003. “Proto Gunwinyguan verbal suffixes”. The Non-Pama-Nyungan Languages of Northern Australia: Comparative studies of the continent’s most linguistically complex region ed. by Nicholas Evans, 305-352. Canberra: Pacific Linguistics
1. Introduction

There is one linguistic matter in which Harold Koch has a failing: he finds it difficult to say ‘no’. Even if a request is on the verge of being outrageous, and even if saying ‘no’ would save him a lot of time and trouble, he has always found it very hard to utter that word. Because of this many of us who are contributors to this volume owe him a debt of gratitude. We take this opportunity of offering him many different ways of saying ‘no’.

A striking feature of negatives in South-Eastern Australian languages is the diversity of words for ‘no’ or ‘not’. This diversity is probably at its maximum in the languages of Victoria, particularly those spoken along the Murray River, whatever their genetic affiliation. Many of the languages of Victoria incorporate the negative word as the language name, either reduplicated, as in Wembawemba, or in combination with a word meaning ‘language’, a feature found in the Eastern Kulin languages.

This diversity of negative words is also present, but much less prominent, further west, as in the Thura-Yura languages of South Australia which comprise Parnkalla, Nukunu, Narangga, Kuyani, Ngadjuri, Adnyamathanha and Kaurna1 with Nauo and Wirangu as outliers, in an area stretching from the Adelaide plains to the head of the Great Australian Bight. There is also less diversity further north in those of the Karnic languages we mention here, namely Arabana-Wangkangurru to the west and north of Lake Eyre, and Yaluyandi and Diyari to the north-east and east of Lake Eyre.

There could be many kinds of avenues to explore: the use of negative words for different types of negation; privatives and negative imperatives. In this paper we have confined ourselves to exploring the diversity of negative forms, and examining some possible origins for them.

---

1 Examples in Thura-Yura languages are from Schürmann 1840, (Parnkalla), Hercus 1999 (Wirangu), Hercus MS (Kuyani), Hercus 1992a (Nukunu), Schebeck 2000 (Adnyamathanha).
2. The Languages of Victoria

The use of a reduplicated negative as a language name is found along most of the length of the Murray River. Many of the languages that show this feature are from the Western Kulin group, but this feature is also found in languages more distantly related: Yotayota and Yabulayabula, once spoken on the Murray above Echuca; Yithayitha and Tarritarti, once spoken on the lower Murrumbidgee and downstream of the meeting of the Murrumbidgee and Murray; and Yuyu and Yakko-Yakko once spoken further west.

Much of the linguistic diversity of these areas was lost in the early days of settlement. Our comments on the negatives in these languages are therefore tentative at best. We do not have the full picture. Nevertheless, several patterns emerge. In some languages, like Wembawemba, the form used in the language name was fully productive as a negative in sentences and texts; but in others the form in the language name is listed as a negative, but never recorded in texts, as in Werkaya. With so many negative words in close genetic or areal proximity, we would expect some crossover of forms, and this we find.

Based on work first done by Schmidt (1919), the Kulin languages were grouped into ‘Eastern Kulin’ and ‘Western Kulin’ (Hercus 1986).

2.1 Eastern Kulin

Blake and Reid (1998) divided Eastern Kulin into Central Victoria (comprising Woiwurrung and Boonwurrung from the Melbourne area, and Thagungwurrung from the Seymour area) and Wathawurrung from the Geelong area. Mathews (1903:244) reported that the languages

are named after their equivalents for the English word “no”, to which is affixed the native word “wurru”, signifying “lip” and hence speech. The names of the dialects therefore mean, “Wo-i speech” and “Boon speech”, respectively.”

Of the large number of different words recorded for these languages for ‘no’ and ‘not’, some are found only in word lists, the earliest of which were produced in the 1840s, ranging through to the work of R.H. Mathews at the turn of the 20th century. Some negative words are found only in text examples recorded by these early researchers.

In word lists for Woiwurrung, Boonwurrung and Thagungwurrung, the word for ‘no’ is often given as something like <Tâ-goong> (Green in Smyth 1878:104), but this word is never found in sentence examples. In the texts themselves, the form of the negative is /ngabun/~/abun/, /ngia~/aia/ or /ngutha~//(y)utha/, all words with initial velar nasals in variation with vowel initials. The sentence <Ngabenan wanin wongin pawn> ‘I will not give you some water’, recorded by Green in Smyth (1878:115) has as its first element a word that we analyse as /ngabun-an/ ‘NEG-1SG’. It is an example of the negative, when it is the ‘headword’ of a sentence being suffixed with the bound subject person
NEGATIVES IN SOUTHEASTERN AUSTRALIA

marker (Hercus 1986:56) as we can also see in examples (1) and (5). This feature of negative words appears to have applied widely across the Kulin languages and is therefore likely to have been a feature of Proto-Kulin.

2.2 Western Kulin

Western Kulin, a much larger grouping, included most of the languages of Western Victoria except for the languages of the Warrnambool/Mount Gambier area.

Within Western Kulin, the negative word is found in the language name only in the northern varieties, not in the two southernmost languages of the group, Tjapwurrung and Djadjawurrung. The form of Western Kulin described in most detail is Wembawemba (Hercus 1986). There are many examples\(^2\) in the texts recorded of the use of the word \textit{wemba} as a negative, as in (1):

\begin{verbatim}
(1) wembanda yanginyin
    not-1SG go-FUT
    “I won’t go” (Hercus 1986:24)
\end{verbatim}

Not only person suffixes could combine with the negative in Wembawemba: there is a related form \textit{wembality} ‘if not’. This word appears to be a combination of the negative and a potential suffix \textit{-ity}, with the \textit{-l-} probably being the frequentative suffix \textit{-ila} (Hercus 1986:48). Stone (1911) listed two other related forms, \textit{wembaken} ‘never’, and \textit{wembungga} ‘to lose’. The former word may be the an example of the negative word with the past participle \textit{-an} (Hercus 1986:43). The latter word is a derivative of the negative with the intensive verbal affix \textit{-ungga}. In Wembawemba the negative \textit{wemba} was clearly a vital part of the vocabulary.

The nearest related language to Wembawemba was Perapaperapa, and in the 1960s Wembawemba speakers reported that the two languages were practically the same, and that some people living at the main reserve at Moonacullah used \textit{perapa} for ‘no’.\(^3\) In fact, \textit{perapa} is recorded with a negative meaning in the Wembawemba texts, in the song \textit{Sentai – the lazy dog}, composed early in the 20\(^{th}\) century by Bob Taylor (Nyaui), as (2). It is the one example of a negative other than \textit{wemba} being used in a Wembawemba text:

\begin{verbatim}
(2) perep-urung werkitiya-l-iny
    not-? work-FREQ-FUT
    “If you don’t work” (Hercus 1986:67)
\end{verbatim}

\(^2\) The Wembawemba examples in this paper are spelled following the conventions established in Hercus (1992).

\(^3\) The spelling of this language name with a schwa follows Hercus (1986).
There is ample evidence that in Perapaperapa as in Wembawemba, the
negative word used in the language name was also the normal negative in use.
In his notebooks R. H. Mathews recorded it in a prohibitive construction as
<Booreba Tyilbak> ‘hit not!’, where <tyilba> is the verb ‘beat, hit’ and <-k> is an
imperative marker (Mathews Notebook I:46).¹ That this word is parallel to
wemba is indicated by the alternative <Wamba Thikkak> ‘hit not’, written
immediately below, and which we presume is an example of the Wembawemba
prohibitive.

There is some tragic material in the Thomas papers from interviews with
the two men who were hanged in 1847 for the murder of Andrew Beveridge.
Much of the material was taken down from another man – Warrigle Jemmy alias
Kittermin, who was Perapaperapa from Gunbower. He had been sent down
earlier to prison and was acting as interpreter. Here he is translating sentences
promising the two condemned men better things in the next world:

(3) No more cry = barbun mummelin
No more frightened = barbun barmbin
No more ill = barbun gillegin
and again, all gone cry = barraban lumelun (Thomas MS 214.21)⁵

The final -n that shows up on all of the negative words in (3), as in barbun, is
perhaps the past participle (see Hercus 1986:43), unless it is a transcriptional
error for the future -iny.

For Werkaya, (Tyatyalla) the form werkaya is confirmed by several early
sources (such as Mathews (1902a:83) who gives <wrekkea> as the word for ‘no’
but this word is not found in any texts.

There is evidence of a negative form <ngooat> given in a Werkaya sentence
by Hagenauer (Smyth 1878:42), but the best attested form of the negative in this
language is <bawa>, bewa> This is used for instance in a Werkaya text recorded
by Rev. Hartmann in Smyth (1878:53-54) and edited by Hercus (1986).

(4) ba tyaŵrak bewa woartin
   “and as he did not come”

Hartmann elsewhere too (Smyth 1878:52) recorded <bawa> as a negative, in
the sentence <Nyo bawa mamek> ‘He is not my father’. Furthermore Mathews
(1902:82) has a similar form in his example <Bowen takak> ‘beat not!’, where the
verb is <taka> ‘beat’, followed by the imperative suffix <-k>. The function of the
-n on bowan in not clear in this case.

¹ Original spellings are shown in <> brackets except in tables, where they are omitted.
Transcribed data from modern recordings are given in italics.
⁵ Transcription by Edward Ryan, whose indefatigable research into all matters related to the
indigenous history of Victoria continues to turn up gems of this kind.
A Tyatyalla sentence, somewhat curious in meaning, was recorded by Mathews from ‘Ned McClenann, Pine Plain Language’ (Mathews MS p.22) <bawan gimba> ‘I am not here’. This is analysable as

(5)  bawa-n  gimba  
not-1SG here

All this evidence shows that the negative word in Werkaya is indeed <bewa>, <bawa>, and that werkaya was perhaps only used as a negative interjection.

In the neighbouring Watiwati language spoken at Piangil, <bawa> meant ‘by and by’ and is listed as such by both T. Macredie in Curr (1886:448), and by E.M. Curr himself, (1886:451). This link between ‘by and by’ and the negatives is not restricted to this example, but will be seen again with <darti> in (6).

For Watiwati as spoken near Swan Hill, Beveridge (in Curr 1886:441-443) records <wotti> with negative meaning in several sentences. There is a short text included in his presentation, which contains the following example, presented here as it appears in the original, with the English written above the Watiwati.6

(6)  By and by I will get others  
Darti  yanda  ngurmin  yooia  
and I not throw away addled eggs  
nga  yetti  wotti  winia  wertawil 7 mikko  
I will give (them) to you  
yetti  wongna  nginma

The example clearly shows that the negative word which marks the ethnonym was also used as an ordinary negative, just as in Wembawemba. The example has another feature of interest: the use for ‘by and by’ of <darti>, which is also the negative word and language name of the neighbouring but not closely related language, Tartitarti.

Blake and Reid (1998) drew a distinction between Watiwati (Piangil) and Watiwati (Swan Hill). There is further early evidence on Watiwati from George Augustus Robinson in Clark 2000:227. This source gives both <but.rat.ba> and <wat.tee> for ‘no’, as will be discussed along with Letyiletyi in 3.1.1.

The three languages Watiwati, Letyiletyi and Mathimathi were very similar to each other and formed a distinct northwestern set of Kulin languages. In

---

6 A practically identical form, tharti ‘by and by’ is found in Yalarnnga in far western Queensland (Breen & Blake 2007:97). It remains uncertain whether this is a Pama-Nyungan cognate.
7 <wertawil> is analysable as wirtu-wil ‘young-having’.
Mathimathi, for which there is a considerable amount of text, the negative mathi appears in four different forms:\(^8\)

- matha – negative adverb ‘not’
- mathawa – prohibitive adverb ‘don’t’
- mathi – negative particle ‘no’
- mathim – ‘no longer’

The negative adverb is exemplified in Hercus (1986:127) and also in her texts (1986:137,138). It seems that in these three languages the link between the language names and the negatives was very much alive.

2.3 Non Kulin languages

In Yotayota sources (Bowe & Morey 1999), the word yota (there spelled <yorta>\(^9\)) is frequently found in negative sentences, both in everyday language, as in (7), and in traditional songs\(^10\), as in (8)

(7) \text{yorta nha-n}  
\text{NEG see-NON,FUT}  
“Not seen” \hspace{1cm} \text{(Bowe & Morey 1999:95)}

(8) \text{yorta bara walitja}  
\text{NEG red kangaroo fat}  
“Not red kangaroo fat” \hspace{1cm} \text{(Bowe & Morey 1999:112)}

Yotayota differed from Wembawemba and Perapapapa in that with a negative imperative, a different form gathagana was employed (Bowe & Morey 1999:95-96). The language next to Yotayota was called Yabulayabula, for which Curr (1886:587) recorded the negative word as <yabbala> and exemplified it with sentences (1886:582).

One of the varieties of Yotayota recorded by Curr was the Pikkolatpan where the negative was given as <yalliba>, presumably a metathesis of <yabula>. Unfortunately there are no sentences recorded for this variety, but the group was referred to by Rev. Shadrack James as Yalliba (Bowe & Morey 1999:8).

\(^8\) There is an error in Hercus (1986) where some of these words are spelled without an underline on the stop, the method of showing the dental laminal in that source. All the words should have the form math-.

\(^9\) This spelling reflects the modern usage by members of the Yortayorta community. The spelling with <r> reflects the fact that the word is now usually pronounced with a low back vowel.

\(^10\) This song was recorded by Curr (1887:579). It is the only traditional Yortayorta song for which any analysis can be proposed. Fortunately it contains two negatives, the latter of which, given here, is a negative plus a noun phrase.
The Yithayitha language to the southwest of Booligal, and along the Murray in the Kulkyne area, Tartitarti and Yarri-Yarri or Yerriyerri language all had names based on the negative. For Yithayitha, the word for ‘no’ is <yitha>, recorded in various spellings (see Horgen 2004:320). It is also found once in a sentence, noted by Macdonald in Curr (1886:287) given as in (5) where the morpheme by morpheme gloss is added by us,

(9) moothoort yitta kaangil nunna mowa lewin
    said  NEG  good  man  one  sit-PRES
    “God said not good man alone to dwell.”

The Mathimathi speaker Jack Long (Hercus 1986:232) recorded the form tarti as meaning ‘no’ in Tartitarti, but there are no text examples. A vocabulary in the Thomas papers of the closely related ‘Yerree Yerree or Lyart Tribe’ from ‘O’Brien of Ki’ (below Euston) has a sentence containing binter for ‘no’ and a similar form is found in McFarlane’s contribution to Curr (1886:282) ‘From Mallee Cliffs to Wentworth’: <pinta ngaia yoorun> ‘I don’t know’. The Thomas papers have a dramatic example of the same negative in an account where Kulkyne Bobby, who belonged to the same group as O’Brien, stands trial for murder:

(10) Kulkyne Bobby, did you put old lubra in ground and lubra not dead?
    Bint.
    No.  (Thomas MS 214.21, transcribed by Edward Ryan).

So it seems that in these languages the word for ‘no’ incorporated in the language name was different from the one in common use.

2.4 Conclusion: negatives in Kulin and neighbouring languages

Our findings relating to the use of negatives in the Kulin languages and some immediately adjoining languages are summarised in Table 1.

The fact emerges that in some of the languages whose name is derived from a negative, that particular ‘no’ word was not the one actually being used at the time that the language was recorded. This leads one to the conclusion that the language names, where they are different from the negatives actually used, probably represent an archaism: the ‘no’ word in the language name had been replaced in current speech by a newer form. Negatives are frequently uttered with emphasis, and so speakers tend to seek novel and arresting ways of expressing the notion of ‘absolutely not’, ‘not at all’, ‘no way’ and so forth. This naturally leads to innovations in expressing negativity.

11 He, along with ‘Young Man Tommy’, was acquitted.
Table 1. Summary of Negative Words in Kulin and Murray River languages

<table>
<thead>
<tr>
<th>Language</th>
<th>Negative word in lang. name</th>
<th>Negative word attested as negative ‘no’?</th>
<th>Same negative word attested in examples?</th>
<th>Other negatives</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EASTERN KULIN</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Woiwurrung</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
<td>&lt;ngabun&gt;,,&lt;n’uther&gt;/-yootha&gt;, also &lt;ngia&gt;</td>
</tr>
<tr>
<td>Boonwurrung</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
<td>&lt;eota&gt;</td>
</tr>
<tr>
<td>Thagunwurrung</td>
<td>YES</td>
<td>YES as thagung, thago</td>
<td>NO</td>
<td>/ngabun/-/abun/, /ngia/-/aia/ or /ngutha/-/yutha/</td>
</tr>
<tr>
<td><strong>WESTERN KULIN</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wembawemba</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>perapa</td>
</tr>
<tr>
<td>Perapaperapa</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>Mathimathi</td>
<td>YES</td>
<td>YES,</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>Werkaya</td>
<td>YES</td>
<td>YES,</td>
<td>NO</td>
<td>&lt;bawa&gt;, &lt;bewa&gt;, &lt;ngooat&gt;</td>
</tr>
<tr>
<td>Watiwati</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>bu-rat-pa</td>
</tr>
<tr>
<td>Letyiletyi</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>Poraipar (Letyiletyi)</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
<td>&lt;latto&gt;</td>
</tr>
<tr>
<td><strong>OTHER</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yotayota</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>&lt;yalliba&gt; in Pikkolatpan</td>
</tr>
<tr>
<td>Yabulayabula</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>pinta, pinter</td>
</tr>
<tr>
<td>Yithayitha</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>Tartitarti</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
<td>pinta, pinter</td>
</tr>
<tr>
<td>Yerriyerri</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
<td>binter, bint</td>
</tr>
</tbody>
</table>

One of the sources of new negatives appears to have been words meaning ‘by and by’. The course of this semantic shift would seem to be in the direction ‘by and by’ > ‘not’.

A person being asked to do something will say they will do it ‘by and by’, meaning ‘not now’, and may end up not doing it all: in other words ‘by and by’ becomes a euphemism for a negative and ultimately an outright negative.

Apart from this likely source of new negatives the origin of the negatives in Victorian languages is obscure. There is no single negative that is shared between Western and Eastern Kulin and which could be traced back to Proto-Kulin. One can however trace back to Proto-Kulin the syntactic feature of a negative as headword being marked for person.

There is one negative base that can be traced back to Proto Western Kulin. Wembawemba has a word *lathuk* ‘naked’. It consists of a stem *lath* followed by the 3rd person singular possessive marker, which also functions as default possessive marker. The obvious meaning is ‘nothing-his/hers’: so there was a stem *lath* meaning ‘nothing’. This must go back to an earlier phase of Western Kulin, as it is also the basis of the word *letyi*, see 3.1.1. below.
3. In Thura-Yura

Speakers wanting to be innovative might borrow the word for ‘no’ from another language and this leads to such words becoming an areal feature, and spreading across languages. There are minor examples of cross-language features as mentioned in Kulin (<thagung>, <bawa> and <darti>) but negatives as areal features are prominent in the languages further to the west in the Thura-Yura languages of South Australia.

3.1 Negatives as areal features

There are at least two forms in the Thura-Yura languages that are of particular interest as areal features, as they have spread across language boundaries.

3.1.1 yakko. This word for ‘no’ in Kaurna, the Adelaide language, goes back to Pama-Nyungan and has cognates far afield:

<table>
<thead>
<tr>
<th>Language</th>
<th>Word/Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Western Bundjalung</td>
<td>yagam(be) ‘no; not; none’ (Sharpe 1992:123)</td>
</tr>
<tr>
<td>Gupapuyngu</td>
<td>yaka ‘not; no’ (Lowe n.d.)</td>
</tr>
<tr>
<td>Pitta-Pitta</td>
<td>-yaku privative suffix ‘without’ (Blake 1979:199)</td>
</tr>
</tbody>
</table>

Kaurna <yakko> ‘no’ has as far as can be ascertained, only one possible cognate in the other Thura-Yura languages and that is Parnkalla <yakkudla> ‘I don’t know, cannot tell’ (Schürmann). There is however evidence of <yakko> in languages further to the north-east of Adelaide, outside the Thura-Yura language group.

Eyre (1845:331) gives a list of words in a language he calls ‘Yak Kumban’:

‘Yak kumban’ the dialect spoken by the Paritke tribe, or natives inhabiting the scrub to the west and north-west of the Murray, which extends along the range of hills from Mount Bryant to the Darling near Laidley’s Ponds, and forms in its variations the language of the Darling itself.’(p.398-399).

The language name ‘Yak Kumban’ is based on the word for ‘no’ and the people were also referred to as ‘Jakojako’, Yakko-yakko, ‘Yokka-yokka’ (Tindale 1974:211). Berndt and Berndt (1993:305) refer to them as ‘Yakamuldak’ and have them ‘located in country around Mildura’. The alternative name used by Eyre, the ‘Paritke tribe’, comes from the Paakantyi word <parri> ‘scrub’; they were literally ‘the Scrub-dwellers’. The ‘Yak Kumban’ language as exemplified in the 25 words listed by Eyre is practically identical to the Marawara dialect of Paakantyi, the Darling River language. The Eyre list shows similarities with Marawara not shared by the other Paakantyi dialects, such as the lenition of ty and k in the sequence itya, ika and uka, as is illustrated in the Table 2.
As in some of the Victorian languages, the word for 'no' given in the list (for Yak-kumban) is not yakko, on which the language name is based, but par-u-wer, which appears to be a borrowing from the neighbouring language, 'Poraipar'. Eyre goes on further to explain the situation where people belonging to three different language groups live in proximity to one another:

...[Aiawong, Boraipar, Yak-kumban]; these tribes meet upon the Murray at Moorunde, and can only communicate to each other by the intervention of the Aiawong dialect, which the north-western or south-eastern tribes are compelled to learn, before they can either communicate with each other, or with the natives of the Murray, at their common point of rendezvous.

'‘Boraipar’, as is evident from the list given by Eyre, is a western Kulin language, although despite some similarity in sound, it is not the same as Perapaperapa. Tindale (1974:206) identifies it as ‘Latjilatji’, and indeed the word for 'no' given in Eyre’s list is latto (Eyre 1945:397), which is obviously closely related to letyi, and to layong found in the Letyiletyi sentences from McLeod, quoted in Curr (1886:438), and it is also the negative found in the form of Watiwati quoted by Robinson (2.2. above). As shown above it is a negative word that strongly suggests that it goes back at least to Proto Western Kulin. The name Boraipar is also based on a word for 'no'. Robinson in Clark (2000:227) gives both <but.rat.ba;> and <wat.tee> for ‘no’ in Watiwati, which is closely related to Letyiletyi. The word <but.rat.ba> is no doubt the same as that the negative used in Bura-Bura described in R.H.Mathews’ manuscript grammar entitled ‘The Burabura language’, which can be identified fairly well with the Swan Hill dialect of Wati-Wati: this is also the conclusion arrived at by Blake (n.d.).

There was probably more communication between the groups than Eyre indicates, for instance the Boraipar material in Eyre (pp 356-7) contains information about the mythological being called 'Nooreele': this is Nhuurali, the Paakantyi term for ‘Ancestor’. As far as can be ascertained, it was not known to Kulin people in general.

It seems therefore highly likely that <paruwer> ‘no’ was borrowed into Yak-kumban from Boraipar and superseded <yakko>, which was retained only in the language name.

There was a further exchange of negative words in this area. In the Ngayawang vocabulary by E.B. Scott (n.d.) there is an entry <loitwun> which ‘appears to be a word used when anything has left, or departed from a place, as Tawpool loitwun the Teal has flown away’ (transcription by Edward Ryan).

<table>
<thead>
<tr>
<th>English</th>
<th>Yakkumban</th>
<th>Marawara</th>
<th>Southern Paakantyi</th>
</tr>
</thead>
<tbody>
<tr>
<td>father</td>
<td>kanbea</td>
<td>kampea</td>
<td>kampitya</td>
</tr>
<tr>
<td>fire</td>
<td>koon-nea</td>
<td>kuneia</td>
<td>kunika</td>
</tr>
<tr>
<td>moon</td>
<td>paity-o-way</td>
<td>pitoa(Bulmer)</td>
<td>paatyuka</td>
</tr>
</tbody>
</table>

Table 2: Some examples of Yak-kumban vocabulary
(Marawara data – Tindale MS and Bulmer 1876, Southern Paakantyi data – Hercus 1993).
Many Australian languages have an intransitive verb derived from the negative adverb ‘not’ or from ‘no’ and this verb usually means ‘to come to an end’, ‘to come to nothing’, ‘to disappear’. The Ngaiawang word <loitwun> ‘departed from’ would have to be an example of the past participle of just such a verb. The best-known example of a parallel form is Western Desert wiyaringu ‘it has come to nothing, it is finished’, derived from wiya ‘no’, ‘not’.

The following development appears to have taken place in this area of South Australia: the negative <paruwer> was borrowed into Yak Kumban from ‘Boraipar’. A form of the Boraipar negative <latto> which appears to go back to Proto Western Kulin, was borrowed into Ngaiawang where, it is indirectly attested in the vocabulary by E.B. Scott (n.d.)

The earlier history of <yakko> however remains obscure. Was <yakko> an old Thura-Yura word that spread from Kaurna to Yak Kumban, or was it the other way round? The language immediately adjacent to Yakkoyakko to the west and closest geographically to Kaurna is Ngaiawang, Eyre’s Aiawang. His list for that language does not give us a word for ‘no’ and Moorhouse’s work of 1846 has a somewhat different form of this word, ‘yeamko’. We could perhaps consider this to be a local variant that arose in Ngaiawang after the word had been transmitted between Kaurna and Yak kumban. The way in which these negatives are interchanged is indeed an areal feature.

<table>
<thead>
<tr>
<th>Language Name</th>
<th>Word used for ‘no’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boraipar</td>
<td>latto</td>
</tr>
<tr>
<td>Yak kumban</td>
<td>paruwer</td>
</tr>
<tr>
<td>Ngaiawang</td>
<td>yeamko, loit</td>
</tr>
<tr>
<td>Kaurna</td>
<td>yakko</td>
</tr>
</tbody>
</table>

Diagram 1: Cross-borrowing of negatives

3.1.2 maga, maka. Another negative that has crossed language boundaries, but in a much more conspicuous way, is maga, maka. It is widely used in the southwest of South Australia. It has been attested from the earliest European records for all the related Thura-Yura languages along the coast: it is attested in Wirangu as muk-ka in the list entitled ‘Head of Great Australian Bight’ (Eyre 1848:397), and for Parnkalla as makkka (Schürmann 1844:25). The only sentence that Schürmann has noted for Nauo in his manuscripts contains the word twice:

Ngai Ngai malpu makkka ai makkka. ‘I am not the murderer, not me’ (Schürmann 1987:136-7).

There is an equally early record for maka across a language boundary in the non Thura-Yura language Mirniny to the west of Wirangu, namely an entry for ‘no’ in Eyre’s list as ger-ga and as muk-ka (Eyre 1845:397).
It is difficult to know in which language subgroup or subgroups the word *maga* originated, but one thing is certain: it spread further. As Kukata people, speakers of a Western desert language, gradually moved into the area they too adopted this word, which is not found in any of the other Western Desert languages. It also spread further into the southern part of Kuyani country. Evidence for this comes from a shrewd observer, the pastoralist R. Bruce. He was a correspondent for R.H. Mathews and wrote reminiscences and epic poems based on Aboriginal stories and containing some local words. He had spent a lot of time at Wonoka, near Parachilna on the Western side of the Flinders Ranges, close to the Arkaba (red ochre) mine. He stated (1902:171):

The word *mucca*, so generally used nowadays, was unknown to the Arkaba blacks forty years ago; at least I never heard one use it.

The introduction of *maka* into southern Kuyani must have occurred by the mid-1880s, as *mukka* ‘no’ occurs in the wordlist from the ‘Arkaba Tura’ at Wonoka by W.M. Green, in Curr (1886:124).

The word also came to be widely used in pidgin and R. Bruce (op cit.) gave as example someone talking about not liking pork and saying:

Me no can’t tellum you, only mucca me like him

A.G. Bolam, who was station master at Ooldea until 1925, wrote a book called *The Trans-Australian Wonderland* in 1923 and that wonderland was the Ooldea area. The book deservedly became highly popular. It contains many dialogues between Europeans and Aborigines speaking a form of pidgin and *mucka* occurs over and over again, as in the following passage (p.115):

White: “How old track?”
Black: “One day mucka (no) more!”

The widespread use of Aboriginal negatives in pidgin is well known from New South Wales data, particularly *baa, bael, bail* in the sense of ‘no, not’, and so the spread of *makka* is by no means unique.

3.2 The possible origin of some of the negatives

It is possible at least in some instances to postulate a probable etymology for words of negation. They come to be negatives from a variety of different semantic concepts, notably ‘bad’, ‘inadequate’, and as already mentioned ‘by and by’ i.e. ‘not now’.
Wi-ring-ung Parnkalla  Kuyani  Adnya-mathanha  Nukunu  Kaurna

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>maga</td>
<td>makka</td>
<td>yakko</td>
<td>not,</td>
<td>no</td>
</tr>
<tr>
<td>guda</td>
<td>kutta</td>
<td>kuta</td>
<td>uta</td>
<td>not,</td>
</tr>
<tr>
<td></td>
<td>not,</td>
<td>not,</td>
<td>not,</td>
<td>don't</td>
</tr>
<tr>
<td></td>
<td>no,</td>
<td>no</td>
<td>marla</td>
<td>a little bit</td>
</tr>
<tr>
<td></td>
<td>madla</td>
<td>marla</td>
<td>marlaka</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>-wakka</td>
<td>-wakanha</td>
<td>-wakanha</td>
<td>-wakanha</td>
</tr>
<tr>
<td></td>
<td>without</td>
<td>without</td>
<td>without</td>
<td>without</td>
</tr>
</tbody>
</table>

Table 3: Some negative words in the Thura-Yura languages (only the most relevant forms are listed here).

3.2.1 Kaurna <madlanna>. A form madla ‘no’ and its derivatives is found throughout the Thura-Yura language sub-group, except for Wirangu. The adjacent languages of the Karna sub-group all have an identical word, but it means ‘bad’:

   Arabana, Yandruwandha, Thirarri madlanthi
   Wangkangurru madla
   Diyari malhantyi

   It would seem likely that the Thura-Yura negatives and the word for ‘bad’ represent one and the same inherited word whose original meaning was ‘bad’ and that the negative meaning developed from there, probably via the various nuances of ‘inadequate’ > ‘lacking’ > being without > no, not. Parnkalla <madla> ‘no’ <mai madla madla > ‘having no food’.

3.2.2 Parnkalla <wakko>. A similar semantic connection can be seen within Thura-Yura, where we get the following situation:

<table>
<thead>
<tr>
<th>Meaning</th>
<th>Parnkalla</th>
<th>Kuyani</th>
<th>Adnya-mathanha</th>
<th>Nukunu</th>
<th>Kaurna</th>
</tr>
</thead>
<tbody>
<tr>
<td>bad</td>
<td>wardlapu</td>
<td>wadli</td>
<td>wadli</td>
<td>wakkinna</td>
<td></td>
</tr>
<tr>
<td>broken</td>
<td>waka</td>
<td>waka-ri-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-to break (v. intr)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>negative</td>
<td>-wakka</td>
<td>-wakanha</td>
<td>-wakanha</td>
<td>-wakanha</td>
<td></td>
</tr>
<tr>
<td>without</td>
<td>without</td>
<td>without</td>
<td>without</td>
<td>without</td>
<td></td>
</tr>
<tr>
<td>wakko ‘no’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Some words for ‘bad,’ ‘broken’ and ‘without’ in Thura-Yura (only immediately relevant forms are shown).

   From this evidence it seems logical to deduce that the base waka- originally meant ‘broken,’ ‘bad, no good’ and that the negative meaning arose from there.
3.2.3 Proto Central Thura Yura *wadLV 'bad'. The above table also shows evidence of an earlier word *wadLV 'bad' and there is further evidence in Simpson and Hercus (2004). It is tempting to think that this word represents the origin of warla 'not' in the Malyangapa language. Malyangapa, along with Wadigali and Yardlyawara belonged to the Yarli language sub-group, and adjoined Thura-Yura to the north-east. There has been borrowing between Kaurna and Malyangapa, possibly via the little-known Thura-Yura language, Ngadyuri. A striking example is the purposive/allative suffix -itya, (Hercus & Austin, 2004:220), and so a connection between Kaurna <wadli> 'bad' and Malyangapa warla 'not' is a distinct possibility.

3.2.4 And there should be more. There are a number of trails that tempt us to go further with questions such as:

What is the relationship between Kuyani purlu, 'can’t', and similar forms in the Karnic languages to the north, Arabana -Wangkangurru pudlu found only in yarri-pudlu 'deaf' lit. ‘ear disabled’, Yaluyandi pulu ‘not’ and Yandruwandha pudlu ‘can’t’?

How is it that in Wangkangurru walya means ‘by and by’, but in Yaluyandi it means ‘don’t’, and further to the east, in Yandruwandha and Wangkumara it means ‘not’?

So far our work has led us to just two conclusions, one, that some of the ‘no’-languages have changed their word for ‘no’, and two, that negative words in general are subject to fascinating changes.

References
Blake, Barry J. n.d. The Mathi Group of Languages. Ms, La Trobe University, Melbourne.


MOORHOUSE, M. 1846. *A Vocabulary and Outline of the Grammatical Structure of the Murray River Language spoken by the Natives of South Australia from Wellington on the Murray, as far as the Rufus*. Adelaide: (Printed by) Andrew Murray.


SCHEBECK, Bernhard. 2000. *An Atnyamathanha-English “research dictionary” Version 0.02*. Port Augusta: The author

SCHürMANN, Clamor Wilhelm. 1844. *A vocabulary of the Parnkalla language. Spoken by the natives inhabiting the western shores of Spencer’s Gulf*. To which is prefixed a collection of grammatical rules, hitherto ascertained. Adelaide: George Dehane.


TEICHelmANN, Christian Gottlob, & Clamor Wilhelm Schürmann. 1840. *Outlines of a grammar, vocabulary, and phraseology, of the aboriginal language of South Australia,*

Thomas, W. MS 214/21, Ms, Mitchell Library, Sydney.
§1. A morphological revolution has occurred in Indo-European studies in the years since our honorand and I were fledgling Indo-Europeanists at Harvard in the 1960’s. Almost nothing in PIE morphology looks the same as it did forty years ago; the discovery of new ablaut patterns, new derivational processes, and new grammatical categories has rendered obsolete such formerly canonical works as Kuryłowicz’s L’apophonie en indo-européen (1956), the structuralist classic that in our student days seemed to represent the last word in sophisticated IE scholarship. Perhaps the single most important development of the past decades, underlying all the others, has been the rediscovery of philology — the renewed realization, partly lost during the theory-driven disputes that accompanied the mid-century “laryngeal wars” — that the evidence of actual forms in actual texts can sometimes tell us much more than the cleverest theorizing.

The achievements of late twentieth-century Indogermanistik, as well as some of its shortcomings, are clearly seen in the impressive Lexikon der indogermanischen Verben (LIV), published in 1997 under the general editorship of Helmut Rix (second edition 2001; see bibliography). LIV sets out to do for PIE what Whitney’s Roots of the Sanskrit Language (1885) does for Sanskrit: to give for each verbal root in the language an account of its primary averbo — present(s), aorist, perfect, causative, and so on. This is a remarkable goal, not only because such a task could never have been attempted forty years ago (the factual knowledge did not exist), but also because no one working within the structuralist tradition of the period would even have thought to attempt it. The fact that LIV is a “first” makes its many successes — hundreds of lucid and sensible articles — all the more admirable. It is incontrovertibly a major resource, comparable in many ways to Pokorny’s Indogermanisches Etymologisches Wörterbuch (1959), which it partly replaces. Yet the quality of the work, taken as a whole, is uneven. Many articles fall short of the standard set by the stronger parts, underscoring the fact that for many PIE verbs the business of sorting out the inherited repertoire of stem-forms is still an ongoing concern. A case in point is the well-attested root meaning ‘extinguish’ (trans.) or ‘go out’ (intrans.), long familiar to Indo-Europeanists as *(z)g‘es- (Pokorny, etc.), but now listed in LIV, in an article signed by Reiner Lipp (541-3), as *(s)g‘esh₂-.

---

1 I am grateful to Alan Nussbaum, Martin Peters, and Jeremy Rau for discussion of the ideas in this paper. All errors are of course my own.
§2. Let us begin with the form of the root itself. The *-h/g1350- that LIV adds to the standard reconstruction is unnecessary. None of the forms cited to justify the laryngeal — the Greek aorist /g1098/g814/g798/g803 (supposedly < *e-sg/g951eha < *e-sg/g951esh/g1350-t) and present /g802/g801/g1185/g809/g797/g808/g801/g809: /g814/g798/g793/g809/g809/g816/g808/g801/g809 (Hesych.), the Vedic aorist d/g227s/g269t ‘despaired’ (for *j/g227s,g269t), and the Tocharian A preterite 3 pl. mid. kaks/g227nt ‘quenched’ (supposedly < *-g/g951sh/g1350- — are decisive. In Greek, as shown by García Ramón (1982: 112 ff.), the presents *g/g951es-na- (i.e., /g802/g801/g1185/g809/g797-) and *zg/g951es-nu- (oβěvνυ-) are parallel nasal stems built to the transitive aorist *(z)g/g951es-s-; neither has any claim to IE antiquity. eβη, as will emerge in §§10-11, is not an anomalous contracted root aorist, but a normal intransitive aorist in -/g803-. Ved. dásit, if it belongs here etymologically, is an “improved” version of earlier *jás(-s)-t, showing the same morphological renewal as in ásit ‘was’ for earlier *á-s-t. Toch. A kaksá- < *kakásá- is a regular class II (causative) preterite, with the stem-final -á- common to all such forms, whether historically set or not.2

The real problem with the traditional formula *(z)g/g951es- (or *(s)g/g951es-),3 of course, is the status of the initial sibilant. Here LIV does nothing to resolve the puzzlement of the standard etymological dictionaries (cf., e.g., Chantraine, DELG 992; Frisk, GEW 686, Pokorny, IEW 479 f.), which either implicitly treat the *z- as a case of “s-mobile” or ponder taking it from a reduced preverb akin to Gk. εξ. Neither is a viable option. The phenomenon of s-mobile is basically confined to roots beginning with a voiceless stop (cf. *spek- beside *pek- ‘look’, *steg- beside *teg- ‘cover’, etc.); a few cases have been claimed before a voiced aspirate (LIV gives only *(s)dherbh- ‘grow stiff’),4 but not a single instance, other than *(z)g/g951es- itself, can be cited before a plain voiced consonant. A development of *(e)ks-g/g951es- (vel sim.) to *zg/g951es- would be unparalleled and, if anything, even more improbable than s-mobile. Whatever the source of the *z-, however, it is confined to Greek. The Indo-Iranian, Balto-Slavic, Tocharian, and Germanic cognates of oβěvνυμι have simple *g/-: cf. Ved. jása- ‘be exhausted’ (preses. jásá, jasya-), Lith. gėsti ‘grow dim’ (OLith. pres. 3 p. ėgėsa), OCS u-gasniti/u-gasiti ‘die out/pout out’, Toch. käs- ‘extinguish/be extinguished’ (B pres. mid. keštar), Go. fra-qistnan (denom.) ‘perish’. Indeed, if García Ramón (op. cit., 106 ff.) is correct in taking Hesychian oβěvνυμ as an Arcadian form (recte *ζάνμυν) with ζ- (d2-) < *g/-, the cluster *zg/- is not even pan-Greek. The replacement of *g/g951es- by *zg/g951es- was a post-IE, and probably a post-Common Greek event.

---

2 -ά- [-a-] may fairly be called the “general” preterital stem vowel in Tocharian; while of laryngeal origin, it has been so widely propagated that it is of no diagnostic value whatever. The lack of palatalization in kaksát marks it as a recent, or recently remade form. Toch. A kūkso, cited as a past participle by LIV, is ill-formed; if genuine, it would imply a root *kās-.

3 Since the PIE phoneme /s/ was unmarked for voice, the notations *(z)g/- and *(s)g/- are completely equivalent.

4 Roots given as “?*(s)bheng-” ‘shine’ and “*(s)g’hh1el-” ‘stumble’ are also listed, but here the onsets are simply the LIV notational substitute for “classical” (and likewise unsatisfactory) *(s)ph- and *(s)k’h-.
§3. The problem of how PIE *g̣h₁es- became Gk. *g̣h₁es- is best approached indirectly. The clearest fact about the behavior of the root *g̣h₁es- in the parent language — a fact missed, ironically, by LIV — is that it made an s-aorist.° The aorist stem *g̣h₁es-s- is directly attested in Greek (ἐσθαναντός) and perhaps Ved. dāṣīt. As we shall see in §6, traces of such a stem are also preserved in Balto-Slavic. The decisive evidence, however, comes from Tocharian.

The conservative treatment of the PIE s-aorist in Tocharian is well known. In the active, the Tocharian “s-preterite” is conspicuously non-sigmatic outside the 3 sg. — a morphological peculiarity that also characterizes the formal counterpart of the s-aorist in Hittite.° The unexpected agreement between Tocharian and Anatolian in this detail lies at the heart of the “h₂e-conjugation” theory of the s-aorist presented in Jasanoff (2003: 174-214). According to this theory, the PIE s-aorist was originally a specially inflected type of root aorist in which the 3 sg. active form, for reasons now lost within the prehistory of PIE, was built from a suppletive sigmatic stem with “Narten” (*ē : *e) ablaut.° Such sigmatic 3 sg. forms first became established in diathetically bivalent verbs of the type *nēk- ‘destroy/ perish’ and *nēH- ‘lead/turn’, where the *s- of the 3 sg. active provided a means of enhancing the saliency of the active/transitive : middle/intransitive distinction. The locus of the s-aorist was in 3 sg. pairs like *nēk-s-t (for theoretically expected non-suppletive, non-sigmatic **nōk-e) ‘destroyed’ vs. *nōk-o ‘perished’ (cf. Toch. A ȵakās : nakät (*nōk-[t]o) ‘id.’), and *nēH-s-t (for theoretically expected non-suppletive, non-sigmatic **nōH-e) ‘led’ vs. *nōH-o ‘turned (intrans.)’ (cf. Hitt. naiš : nēa[t]’id.’).

§4. The Toch. A pattern ȵakās (*nēk-s-t) ‘destroyed’ : nakāt (*nōk-[t]o) ‘perished’, with -s- wholly absent from the middle, is a synchronically irregular archaism. In “normal” verbs the middle forms of the s-preterite are sigmatic in both Tocharian languages; cf., e.g., A act. prakās : mid. pārkāt, B act. preksā : mid.

° LIV’s decision to set up a root aorist rather than an s-aorist is driven by the editors’ apparent commitment to taking ἔσθανεν from *g̣h₁esht-t (cf. §12 below). Of the other forms cited in support of a root aorist, Ved. jāśāmāna-, dāsāmāna-, and dasāt are better referred to a thematic present (cf. §7); OCS u-gas is simply the productively formed aorist to u-gasnten; and Toch. B ksetär points positively to an s-aorist (cf. §§4-5).

° The Hittite category corresponding to the s-aorist is the preterite of the h₁i-conjugation. Compare the Toch. B s-preterite of prek- ‘ask’ and the Hittite preterite of the h₁i-verb dā- ‘take’:

<table>
<thead>
<tr>
<th>Toch.</th>
<th>Hittite</th>
</tr>
</thead>
<tbody>
<tr>
<td>prek-wa ‘I asked’</td>
<td>prek-anam dā-ḫḫun ‘I took’ dā-ween</td>
</tr>
<tr>
<td>prek-asta</td>
<td>prek-as dā-tta dā-tten</td>
</tr>
<tr>
<td>prek-s-a&lt;*-s-(a)t</td>
<td>prek-ar dā-ṯ ḫ &gt;*-s-t</td>
</tr>
</tbody>
</table>

The parallel was first pointed out by Watkins (1962: 61 ff.).

° The central claim of the h₂e-conjugation theory is that PIE had grammatically active presents and aorists which took endings similar to those of the perfect and middle (1 sg. *h₂e, 2 sg. *th₂e, 3 sg. *e-, etc. — hence the term “h₂e-conjugation”). The ancestor of the s-aorist, the “presigmatic” aorist, was precisely such a formation. The introduction of forms with *s-, probably of desiderative origin, into the presigmatic aorist was a gradual process, never completed in Anatolian or Tocharian.
The verbs näk-, päk-, tsäk-, and täm- have other features in common as well. All make intransitive subjunctives of class III (3 sg. A näkatär, pkatär, tskatär, cmatär, B nketär, pketär, tsketär, cmetär < *-otor), with sparsely attested transitive counterparts of class I (B inf. nakti, tskaksi). In addition, näk-, päk-, and tsäk-make transitive s-presents (3 sg. A nkä/g931, pkä/g965, tskä/g965; B nakä/g965, pakä/g965, tsakä/g965), which historically continue PIE s-aorist subjunctives (*näk-se/o-, *päk-se/o-, *tsäk-se/o-). The roots corresponding to the oldest s-aorists in Tocharian thus present a distinctive morphological “profile.” Accidental gaps aside, such roots typically have 1) a transitive active class III (s-) preterite (e.g., A näkäs); 2) an irregularly s-less middle preterite (A näkät); 3) a transitive active class I subjunctive (B nakti); 4) an intransitive middle class III subjunctive (A nkät); and 5) a transitive active class VIII (s-) present (A nkäs). Using this information, we can add two further roots to the Tocharian s-aorist “core.” One is näm- ‘bend (tr) /bend, bow (intr)’, with an s-present, a class I/III subjunctive, and an s-preterite (though without an attested middle) in both languages. The other is käs-.

§5. Despite some confusion in the handbooks, the descriptive facts in the case of käs- are clear. The key s-preterite forms, both active/transitive and

8 Omitted from this list is lyuk- ‘shine’ (= PIE *leuk-), which likewise has a preterite of the näkäs : näkät type: näkät ‘make ripe, cook/grow ripe’ (= PIE *pek- ‘id.’), tsäk- ‘burn (tr.)/burn (intr.)’ (= PIE *dhég/h- ‘id.’), täm- ‘engender/be born’ (no IE etymology), and näk- itself (= PIE *nek-). All are transitive in the active and intransitive in the middle. Of the three with good IE etymologies, two, päk- and tsäk-, have well-established s-aorists elsewhere in the family (cf. Ved. subj. pákṣa-; Gk. ἐπηγα, Lat. coxi < *pēk*-s-; Ved. ádhāk, OCS žax < *dhēg/h-s-). The one exception is näk-, where there is no s-aorist outside Tocharian — no doubt because the transitive functions of the primary verb *nek- were transferred to the iterative-causative *no/g3369e/o- in the “inner” IE languages (cf. Ved. n/g227/g314áya-, Lat. noce/g302).

9 Liv surprisingly denies an s-aorist to *pek- while setting one up for the parallel *dhég/h-.

10 I use the terms “inner IE” and “inner IE languages” to refer to the IE branches that remained after the departure of Anatolian and Tocharian from the rest of the family. Inner IE, in my view, was a proper subgroup of IE, characterized, inter alia, by common innovations in the form of the s-aorist and the inventory of simple thematic presents. Cf. note 19.

11 The origin of the class I/III subjunctive complex, which is intimately linked to the origin of the s-aorist, is discussed in Jasanoff (2003: 199-203). The transitive forms are mistakenly assigned to class III by Krause and Thomas.


13 It is precisely because of the lack of an intransitive s-less middle in Toch. A (*namät < *nöm[t]o) that näm- was not included with näk-, päk-, tsäk-, and täm- in our original inventory. The gap is surely accidental. A cognate s-aorist appears in Vedic Sanskrit (3 pl. subj. namsante, post-Rigvedic 3 sg. indic. anān).
THE PIE ROOT FOR 'EXTINGUISH/GO OUT' 159

middle/intransitive, are found in Toch. B (e.g., 2 sg. act. kesasta ‘you quenched’, 14 3 pl. mid. kessante ‘were extinguished’). There are no quotable finite forms of the preterite in Toch. A, and hence no attested s-less intransitive middle *kasät < *g/g951ós-[t]o parallel to nakät, tamät, etc. But just as the Toch. B counterpart of A nakät is neksate, with the regular sigmatic element -sa- mechanically inserted into the inherited o-grade form *nekte (= A nakät), so B kessa[nte] shows the mechanical insertion of -sa- into *kes[än]te (= A *kasä[n]t).15 The preterite of käs- was thus clearly of the same archaic type as the preterites of nák-, pák-, tsük-, and täm-.

The remaining forms confirm the s-aorist profile. The regular intransitive class III subjunctive of käs- is attested in both languages (cf. A abstract ksälune, B 3 pl. ksentár); its transitive counterpart appears in B inf. kasti ‘to quench’ (class I). The present, though wrongly assigned to the simple thematic type (class II) by Krause and Thomas (199), is in fact the structurally expected class VIII present in *-se/o-, disguised only slightly by the tendency of the -s- of the suffix to merge with the -s- of the root.16 The original stem was thus the inherited s-aorist subjunctive *g/g951ês-se/o-, parallel to *né/g374-se/o-, *pék/g951-se/o-, etc. In Toch. A this gave käs-/käs- followed by the thematic endings, with simplification of the geminate -ss/-ss- and analogical (but quasi-regular) de-palatalization of the initial consonant. In Toch. B the vocalism of the s-preterite, where *-e- and *-o- merged as -e- (cf. kesasta (act.) < *g'ës-, kessante (mid.) < *g'os-), was imported into the present, yielding the attested forms 3 sg. kesäm, mid. késtår, ptcp. keššënica, etc. A near-exact parallel can be seen in A prak-, B prek- ‘ask’, where the vocalism of the present (A 1 sg. praksam, B 3 sg. preksäm) copies that of the preterite (3 sg. A prakäs, B preksa < *-ë-) and class I subjunctive (B 3 sg. preksam < *-o-).

§6. Taken together, the robust presence of the s-aorist ēqēc(ō)ξ in Greek and the “embeddedness” of the s-preterite *kes-(s-) in Tocharian establish the PIE antiquity of the aorist *g/g951-s-s. Further evidence for the s-aorist, though indirect, comes from Balto-Slavic. In Slavic, the most remarkable fact about the root *g'ës- is that all the attested forms have a- (< */g302- ) vocalism: cf. pres. u-gasn, inf. -gast, aor. -gas ‘die out’ (perfective); u-gaš, inf. -gasati (+ aor. -axš) ‘id.’ (imperfective); causative u-gašo, inf. -gasiti (+ aor. -išš) ‘put out’. LIV (crediting Vaillant 1916: 252 f. and Rasmussen 1990: 189) sees the locus of this vowel in -gasiti, presupposing a PIE lengthened-grade iterative-causative of the type *sëp-e/o (or *sëp-ëje/o-) ‘put to sleep’ (cf. Lat. sōpiō, -īre ‘id.’, ON sæfa ‘kill’).17

14 Wrongly glossed as intransitive (‘erloschst’) by LIV. The passage is given by Adams (1999, s.v.).
15 This is the regular pattern; cf. further B teṃtsate, (1 sg.) teksamai, lauksate (cf. note 8) beside A tamät, tsakt, lykät.
16 The mistaken assignment to class II is repeated in LIV, where the forms are tentatively traced to an underlying perfect.
17 It is immaterial for our present purposes whether the original form of the suffix in the “sōpiō-type” was *-ëje/o- (i.e., the “normal” iterative-causative suffix), or *-e/o- (so LIV, following
This is probably correct. But why should the root *g'ës- have formed a lengthened-grade iterative-causative? The substitution of *-ë- for “regular” *-o- in an iterative-causative is normally linked to the presence of a long vowel elsewhere in the extended paradigm — either because the root was of the “Narten” type (cf., e.g., OIr. sáidid ‘sets, fixes’, OCS saditi ‘plant’ < *sôd-ë-o-; Arm. utem ‘eat’ < *hôd-ë-o-), or for some phonological reason, such as the presence of a root-internal laryngeal (cf. OCS laziti ‘climb’ (iter.) < *lôhôg-ë-o-; so LIV 400), non-iter. pres. lêçô < *lêhôg-). The root *g'ës- belonged to neither category; its iterative-causative should therefore have been *g'ës-gë-ë-o-.

What *g'ës- did do, of course, was form an s-aorist. The active s-aorist of *g'ës-, in its classical (i.e., “Inner IE”) shape, would have had ê-vocalism and *-s- throughout the paradigm: *g'ës-s-m, *g'ës-s-s, *g'ës-s-t, ... 3 pl. *g'ës-s-s. Such forms, had they received the normal s-aorist treatment in Slavic, would have surfaced as *žasê, *žëse, *žëse, ... 3 pl. *žasê, with the distinctive inner-Slavic replacement of the 2, 3 sg. of the historical s-aorist by non-lengthened-grade imperfect-based forms. There were obvious reasons, however, for the s-aorist of *g'ës- not to show the normal treatment. Following the early pre-Slavic simplification of *-ss- to *-s-, the inherited s-aorist paradigm would have given *g'ës-sm, *g'ës(s), *g'ës-t, ... 3 pl. *g'ësnt (vel sim.) — no longer a synchronous s-aorist at all to the ear of a native speaker, but a long-vowel “root aorist” of the pre-Slavic type *sedm, *sê(d)s, *sê(d)st, ... 3 pl. *sêñt ‘sat’ or *lêzm, *lé(z)s, *le̬st, ... 3 pl. *lêñt ‘climbed’ (> OCS sêd, sêdê, sêdê, ...; lêzô, lêze, lêze, ...). The latter forms, as we have seen, were associated with the long-vowel iterative-causatives *sôditi (> saditi) and *lôžiti (> laziti). At some point following the separation of Slavic from Baltic, the pattern must have been generalized: pairs of the type *sêd- : *sôditi, *léz- : *lôžiti, etc. induced the replacement of the inherited iterative-causative *gôsiti (< *g'ës/gës-ë-o-) by *gôsiti (> *g'ës/gës-ë-o-). The aorist *gës- itself was subsequently lost in Slavic; a possible Baltic reflex may underlie Latv. pret. dzèsu (beside dzèsu) ‘I extinguished’.18

§7. As with many roots, it is easier to reconstruct the aorist of *g'ës- than its present. The Greek nasal present òbavum (< *òbèva-v-) is an innovation on the basis of the s-aorist ōbèva(ò)a (cf. §2); the relationship between the two simply imitates that of pres. èvum ‘put on’ (< *fèva-v-; cf. Arm. z-genum ‘id.’) to aor. è̄va(ò)at. The productively formed intransitive nasal presents of Baltic and Slavic — Olith. gěsa (mod. Lith. gęsta) and OCS u-gasñòti — are likewise (pace LIV) clearly...
unoriginal. The best window into the PIE situation is afforded by Vedic Sanskrit. Here the root jas- (das-) makes two weakly attested present stems — jása- (pres. ptcp. mid. jásamāna- ‘despairing, exhausted’; also dásamāna-) and jásya- (2 pl. impv. nī jásayata ‘despair’; also dāsayati). The presence side by side of jása-/*g951és-e/o- and jásya-/*g951és-/e/o- recalls the two presents of PIE *pekʷ- ‘cook’ — *pékʷ-e/o- in Ved. pácati (transitive), Lat. coquō, OCS pekō, etc.; and *pékʷ-je/o- in Ved. pácyate (‘ripen’, intransitive) and Gk. nέω. The roots *g951es- and *pekʷ- have been seen together before; they belong to a morphological “family” that also includes *dhegʰ- ‘burn’, *nek- ‘destroy/ perish’ and *nem- ‘bend, bow’ (cf. §§4-5). Of the latter three items, *dhegʰ- is represented in Vedic by a transitive thematic present dāhati (= YAv. dažaiti, Lith. degi, Alb. djeg); *nek- is represented by an intransitive je/o-present násyati (= YAv. nasieiti); and *nem- is represented by a thematic present námati, -te (= YAv. namaitē) which is transitive in the active and intransitive in the middle. The obvious inference is that the simple thematic stems were originally transitive, and that the stems in *-je/o- were intransitive. In some cases the intransitive sense was taken over by the simple thematic middle, as in jásamāna- or (presumably) námate; in others the je/o-present was extended into the transitive sphere, as in Gk. nέω. But for the late protolanguage the simplest assumption is that *g951és-e/o- meant ‘extinguish’ and *g951és-/e/o- meant ‘go out’.19

Reflexes of *g951és-e/o- and *g951és-je/o- are found in other IE branches as well, though not always in their original value. OCS u-gaš ‘die out’ continues the sense of, and presumably goes back to, PIE *g951és-je/o-, with the regular Slavic generalization of -a-. Latv. dzešu, by contrast, maintains the e-grade of *g951és-je/o- but has taken on the transitive meaning of *g951és-e/o- (cf. nέω). Transitive *g951és-e/o- itself may survive in Doric 2 pl. impv. ἐβιτε ‘put out!’ (Sophron, 5th c.), if this is a dialectal contraction of *ἐβι[θ]ετε < *(z)g951és-e-te (cf. Schwzyer 1939: 743, note 1).20

§8. With the core components of the *g951es- averbo thus resolved,21 we can return to the question set aside in §3 — the origin of the specifically Greek root shape *zg951es-. Neither of the Greek forms discussed thus far — the present ἐβέννυμι (+ ἐβιτε?) nor the s-aorist ἐβέσα(σ)α — sheds any light on the problem. ἐβέννυμι, together with its intransitive partner ἐβέννυμα, was formed within Greek by adding a nasal suffix to the full-grade root-form *g951es-; later, for reasons yet to be discovered, but unconnected with the form of the suffix itself, *g951es-nu- was altered to *zg951es-nu- (*g951es-na- remained unchanged; cf. §2). So

19 It is important to say “late” protolanguage in this context, because under the view of the thematic conjugation adopted in Jasanoff (2003: 224 ff.), *g951és-e/o- would have been a “type II” thematic present, only created in the Inner IE period.
20 The idea is also endorsed by Schmidt (1968: 86, note 52).
21 No special discussion is needed of the perfect, which was formed in the regular way (stem *g951gʰós-/ *g951gʰ isValid origin value in ni jajása ‘is over’ (AV, quoted in LIV). 2du. impv. jajastám (RV), despite having taken on transitive meaning, is probably also a perfect.
too in the s-aorist, which was purely transitive in Greek: *g′es-s- was remade to *zg′es-s-, but clearly not for any reason related to the structure of the s-aorist as such. The cluster *zg′- must therefore have originated somewhere other than in these two formations. Inevitably, our attention is drawn to the main tense + voice combination that the present in -νυ- and the s-aorist leave uncovered — the intransitive aorist, represented in Greek by the paradigm ἔσοβην, ἔσοβης, ἔσοβη ... ἔσοβεν.

As we saw in §5, the oldest intransitive 3 sg. aorist of *g′es- — contrasting with the transitive sigmatic 3 sg. *g′ês-t — was *g′ês-o, an archaic root formation still partly preserved in Toch. B kes[sa]nte. Such forms did not generally survive into the “classical” IE languages, being either radically remade (e.g., as thematic or normal root aorists), or replaced altogether. In Greek the prevailing tendency was for all middle aorists with an oppositional intransitive sense to be replaced by aorists in -νυ- (whence later, in part, -η-). Much has been written about the origin of this formation, which is generally agreed to contain the “stative” suffix *-έν (≡ *-eh1- of Lat. maneō, -έννε ‘remain’, Lith. minėti ‘remember’, OCS pri-lopeti ‘stick to’, Hitt. dammatte(s)i ‘is/becomes empty’, and similar forms. This is alone sufficient reason to be skeptical of attempts to explain the final vowel of /γ1098/γ814/γ798/γ803/γ1098/γ814/γ798/γ803/γ686/γ809/γ797/γ805 as something other than stative *-έν- — attempts like the root aorist-based theory of LIV (*έσ-γ′εσ-ε < *e-σg′es-eha < *e-sg′es-s-s-t; cf. §2), or Risch’s proposed derivation of 2 sg. ἐσβης from the s-aorist (* *e-zg′es-s-s), with subsequent reanalysis and generalization of the stem (έ)σβη (1937: 209). But it is unclear how the only intuitively attractive point of departure — a combination of the root *(z)g′es- with the suffix *-έν- could have led to σβη.

§9. Like all -έ-stative formations, the h-aorist was characterized by zero grade of the root. Let us first consider, then, how the potential input sequences *g′s-έ- and *zg′s-έ- would have been treated in Greek. In the case of *g′s-έ- the voiced stop would have been devoiced before the -έ-, giving *ks′s-έ- > *ψη-. Initial *zg′s-έ- would have yielded a voiceless cluster as well, though here, given the involvement of three consonants, the final result is harder to predict with certainty (*sk′s- > *ψ?- *στ?-). In neither case would an initial voiced ψ- or *zg′ had been the phonologically regular output of the zero-grade root. Yet it is notable that just such a development has repeatedly been claimed — first by Mahlow (1926: 433 f.) and, following him, Schwyzzer (1939: 743, note 1); then by Schmidt (1968: 86, note 51, and 1976). According to Schmidt’s fuller 1976 account, the pre-Greek intransitive aorist *g′sê- first gave *bsê-, which then

22 It is irrelevant for our present purposes that the original behavior of this morpheme ranks among the most controversial topics in IE comparative grammar.
23 Frisk (GEW 685; seconded by Chantraine, DELG 992) favors a vague analogy: “Dazu trat als Neuerung ἔσβην, ὑπῆν (nach ἔσβην, ἐκῆν, ἐγκýν usw...”) It is hard to see how this would have worked in detail.
24 This is only one of the objections to starting from full-grade *σβη[ν]ή-, as briefly proposed by Wilhelm Schulze in 1909 (see Schulze 1966: 547, note 2).
underwent metathesis to sbē; from i the new cluster *sb- spread dialectally to other forms of the root *ges-; replacing *g- (or the reflex of *g- — *b-, *d- etc.) wherever it occurred.25 The present stem *ges-nu- (or *des-nu-) was in this way remodeled to *bes-nu- in pre-Attic-Ionic (cf. ζέναμεν). In the Arcadian dialect, where these developments were resisted, the parallel present *ges-na- (*dēses-na-) was retained unchanged (cf. ζέναμεν).

The cardinal virtue of this scenario — and it is a very attractive feature indeed — is that it explains the structure of the intransitive aorist ζόβη and the apparent replacement of *g- by *zg- in a single stroke. But it achieves this at the unacceptable price of assuming a pre-metathesis cluster *g's-/*bs- that could never have existed within the proper history of Greek. The Greek voicing assimilation rule — the rule governing routine alternations like /ag- 'lead' vs. fut. ἀγω 'lead', or οἶδα 'I know' vs. 2 sg. οἶδα (⇔ oid-tha) — was an inheritance from PIE. The zero grade of the root *ges- was "always," so to speak, *k's-; a surface *g's- or *bs- could only have arisen through analogical restoration of the voiced stop under the influence of the full-grade forms. Instances of such restoration are known, a familiar example being associated with Lachmann's Law in Latin (cf. Jasanoff 2004).26 In the case of ζόβη, however, it is beyond belief that pre-Greek speakers would have overridden their voicing assimilation rule to create a nonce cluster *g's- or *bs-, only to eliminate it almost immediately through the operation of an ad hoc metathesis rule. It is no wonder that the Mahlow-Schmidt account of ζόβη and the cluster ζβ-/*zg- has never been widely accepted, and that LIV ignores it entirely.

§10. It would be premature, however, to close the book permanently on the metathesis approach. It is true enough that the zero grade of the root *ges- would have been realized phonetically as [k's-], not [g's-], and that such a sequence could never have been metathesized to [zg-] by Neogrammarian sound change. But not all metathesis is Neogrammarian sound change. The annals of linguistic history are full of cases where what looks like a phonetic change — including a change in the order of elements — is sensitive to or triggered by morphological factors. In a recent paper on Latin (Jasanoff to appear), I pointed out that the verb pandō 'spread out', which goes back to *padn (from still earlier *ptn), owes its -nd- not to a regular metathesis rule, but to a morphological change that aligned pandō with tangō 'touch', scindō 'split', tendō 'stretch', and other presents containing a nasal + stop cluster. Many instances of this kind of "morphological metathesis" are attested in Germanic. The Old Saxon

25 Schmidt also cites the late-attested thematic aorist ζέβεο (*ges-ē/ē-; 44) and the privative adjective ζέβεος (*ges-ē-to-; 45-7) as further zero-grade forms.

26 Note, however, that in the case of Lachmann's Law the restored voiced + voiceless clusters (as, e.g., in *ag-tos 'driven') had been rendered phonotactically admissible through the operation of a prior syncope rule, and that the sound law that operated on the restored cluster (*ag-t- → *ak-t-, etc.) was precisely the kind of change that might have been phonetically expected in that environment.
strong verb *gi-fregnan ‘learn by asking’ had a regular preterite *gi-fragn, but also a metathesized preterite *gi-frang — the latter altered to agree with the common preterites of the type *sang ‘sang’, *swang ‘swung’, *band ‘bound’, etc. Early Germanic cases of the type *brukanaz ‘broken’, with *-ru- substituted for “correct” *-ur- (-*-.r-) under the influence of pres. *brekanaN, pret. *brak, etc., are legion.

It is worth reflecting on how changes like these happen. The locus of innovation, as in almost all primary change, is the first language learner, who must construct a mental grammar and lexicon on the basis of the imperfect, incomplete, and often acoustically degraded information that constitutes his/her primary linguistic data. Errors are frequent, and those that go uncorrected — as some inevitably do — become innovations that may in principle spread to other speakers.27 As the child’s grammatical and lexical knowledge grow, so does his/her ability to make predictions based on the emergent but still imperfectly acquired system. A learner of English will predict *mouses as the plural of mouse and *goed as the past of go; such cases are traditionally classified as “proportional analogy” because the basis for the incorrect prediction can be expressed as a quasi-algebraic proportion (house : houses :: mouse : X, etc.). Not all wrong predictions, however, are translatable into proportional terms. A more advanced learner of English, hearing an unfamiliar Latinate adjective in [-/g536r] preceded by a voiceless velar and a lateral, may guess that the word fleetingly heard was *nucular rather than nuclear. Here we speak of “contamination” (with words like particular, secular, etc.) rather than analogy, yet the basic mechanism — making a wrong prediction on the basis of a perceived synchronic pattern — is the same.

In the cases of morphological metathesis discussed above, *padnō was remade to pandō, and -fragn to -frang, because speakers had principled expectations for what a Latin present and an Old Saxon preterite, respectively, ought to look like, and were prepared, at least for a time, to stand by their predictions in the face of evidence to the contrary. The change of *burkanaz to *brukanaz was comparable, except that here the error that produced the metathesized form was partly prompted by the acoustic phonetic similarity, and hence confusability, of -rV- and -Vr- sequences.

§11. Let us now return to the intransitive aorist of *g951es-. The stem [k/g951s-/g245-], representing underlying /g951s-ē/-, would have presented obvious difficulties for the language learner. First of all, it contained an opaque segment: while all the “normal” forms of the root began with voiced *g951-, the initial stop in [k/g951s-/g245-] was voiceless. Second, it was too short: the root component, which in ordinary inflected forms took up one or more syllables, was reflected in the stem [k/g951s-ē-] by a hard-to-parse, non-syllabic stop + sibilant cluster. A third factor was the crosslinguistic susceptibility of sibilant clusters to metathesis — a well-docu-

27 On the crucial distinction between primary change, which is rooted in the language acquisition process, and sociolinguistic diffusion, which is governed by non-linguistic factors, see the admirably clear statement by Hale (2003: 344 f.)
mented tendency that Blevins and Garrett (2004: 9) attribute to the perceptual effect known as “auditory-stream decoupling.” The position of [k’s-ê-] in early Greek would thus have been highly unstable. Young speakers had trouble processing the form they heard; they were unsure, at some level, of the order of elements in the cluster, and they “expected” a [g-] that was phonotactically debarred from occurring before a sibilant but licensed after it. The result was the sporadic one-step voicing-cum-metathesis of [k’s-ê-] to [zg-ê-] — a “morpho-logical” speech error that, being more transparent than the form it replaced, lent itself to imitation by other speakers. Eventually, initial *zg- began to encroach on *g- in the historical full-grade forms as well. For the phonological change of *zg- to *g- before front as well as back vowels, cf. García Ramón (1982: 102-4).

§12. We can now take stock. LIV presents our verb as “*(s)g’esh,” a root of the same morphological type as *leik- ‘leave’, *jeug- ‘join’, and *k’leu- ‘hear’, with an active root aorist (allegedly seen in εοβη < *e-sg’esh-t) and a nasal present (in οθενμη and Arcadian ζειμμεν < *sg’es-nh). This account, as we have seen, is fundamentally flawed. In fact, *g’es- (sic recte) was a bivalent root of the type *pek- ‘cook/ripen’, *dheg’h- ‘burn (tr.)/burn (intr.)’, and *nek- ‘destroy/perish’, with a transitive s-aorist, a transitive thematic present (in the “inner” languages), and associated intransitive forms. The fit with the data is better under this interpretation than under the LIV account — so much better, in fact, that one wonders whether the editors of LIV would have pressed their “*(s)g’esh” theory at all if they had seen a way to avoid positing a contraction, and hence an active root aorist, for εοβη.

εοβη is admittedly a difficult form, and time alone will tell whether the metathesis-based explanation proposed here is correct. But there is also a methodological point to be made. In Anglo-Saxon jurisprudence there is a maxim, “Hard cases make bad law.” The meaning is plain: the labored solutions that we contrive to difficult problems cannot be allowed to set the parameters for the solutions we find to simpler ones. In dealing with the facts and forms that have engaged us here, εοβη should be the place to finish, not the place to start.

---

28 “While there is still much work to be done on the acoustics and perception of sibilant noise, a number of studies suggest that, in consonant clusters containing sibilants, the sibilant noise somehow distracts the listener, leading to high confusion rates with respect to the linear order of segments… Specifically, there is a tendency to decouple sibilant noise from the rest of the speech stream, and this decoupling can result in dramatic misperceptions” (Blevins and Garrett, ibid.). The authors elsewhere provide (31) a useful table of stop + sibilant and sibilant + stop metatheses.

29 A potential counterexample to García Ramón’s rule, as Alan Nussbaum (p.c.) points out to me, is οθενος ‘strength’, if this goes back to *zghyenos (so Seebold 1983: 32). On the other hand, the clusters *zg- and *zghy- are not quite parallel, and the treatment of *zghy- could principle have been influenced by the maintenance of the sequence *seghu- or *zghu- (with syllabic *-u-) in related forms. Cf. Nussbaum (1998: 525 ff.).
References


1. Introduction

The language of Central Australian Aboriginal songs differs markedly from everyday speech. Early research on Aboriginal song texts suggested that the differences between speech and song were insurmountable. In 1886 F. J. Gillen, the postmaster at Alice Springs who was familiar with the Arrernte language, stated:

‘What about the words’ etc, I have never yet been able to find out the meaning of any of their Arunta [iltarte ?ceremony?].’ (Mulvaney et al. 1997:130).

Difficulties in translating songs has sometimes been put down to their archaic nature (Strehlow 1947:xx). More recent research shows that interpretations of songs do relate to the identifiable words in song texts. We suggest that the multi-dialectal nature of song, the presence of metrical requirements which force phonological alteration of words, as well as the methods of song transmission and interpretation, mean that caution should be taken before assigning the category of ‘archaic form’ to unfamiliar words in song. One way to ascertain whether a form in song is archaic, is to see how it accords with Harold Koch’s meticulous reconstruction of proto Arandic.

In this chapter we draw primarily upon examples of women’s songs from Kaytetye (K) of the Artuya subgroup, and Arrernte (Arr), which along with Anmatyerr (An), Alyawarr (Aly) make up the Urtwa subgroup.

1.1 Extra textual meaning

Words of Aboriginal songs sometimes only account for a fraction of the total meaning conveyed by a song (Berndt 1965:254). This is partly because meaning
may be conveyed through extra-textual means, such as the painted designs and
dance patterns. As well, words carry overtones of meaning, which may only be
known by elders familiar with the narratives associated with songs and who
maintain their social authority through control of such knowledge (Strehlow
1971:195-7; Sutton 1987:89). Such cultural knowledge, which goes far beyond the
text itself, plays a critical role in a singer’s ability to interpret songs.

1.2 Definitions
We regard the song text as a component of song, along with melody and
rhythm. A song text is divided into text lines. In Arandic songs there are usually
two text lines per song. We use the term 'song item' (Moyle 1974:2) to refer to a
single unbroken stretch of singing of a discrete text set to a specific melody.
Following Ellis and Barwick (1987:43) we call a group of consecutive song items
set to the same text a 'small song' and the total number of small songs in a single
performance a 'song series'. Song series are associated with tracts of land known
in Aboriginal English as ‘countries’, which are networks of Dreamings and
mythological stories which loosely correspond to a geographic track of land
(Moyle 1983:67-69). Songs within a song series all have the same melodic
contour. A knowledgeable person can thus identify the country of a song even
when the text is unfamiliar. Small songs can refer to specific sites and
performance of a song series assists people in remembering significant features
of sites, such as where particular plants and waterholes are, and activities that
occurred there in the Dreamtime. The structure of a typical Central Australian
song series performance is shown in Figure 1. The first small song is a song text
containing the word awere (for convenience identified as song text 12), which is
sung twice (song items numbered 1 and 2). The second small song is the song
text containing the word elpere (song text 13), which is sung three times, etc.

2. The role of songs in Central Australian society
In Central Australia all songs are perceived to be from the Dreamtime,4 which
refers to both the time when ancestral beings created the physical world and
the Aboriginal law and belief systems. Dreaming songs are usually seen as ‘pre-
existent’, that is, created by the Dreaming.

4 This contrasts with other parts of Australia where there are also consciously composed songs.
These are most often individually owned, even when jointly composed, and may be of either
traditional or introduced genres. See examples in Grau (1983:55) and Marett (2005).
One of the most important reasons Aboriginal people sing is to assert their relationship to the ancestral beings who shaped the earth, created the social order and cultural practices. The very power and presence of these beings is recreated through performance of songs and can affect the natural environment. For example, songs may be performed to bring or stop rain, to increase the availability of certain animals or plants, or to force a whirly-wind away (see example (9)). Many of these songs may also be performed just for fun at a social gathering; in such situations much of the ritual paraphernalia is left out. Songs also figure prominently in initiation rites, revealing appropriate layers of meaning as the initiate gains various levels of knowledge. While painting a canvas, women will often sing the songs associated with a particular Dreaming, (Barwick 2000) suggesting that song, ceremonial design and Dreaming are all aspects of the same thing.

Performance of songs also reaffirms social organisation and people’s obligations and relationships to one another and can exert societal control. For example, when a local Aboriginal man from northern New South Wales ran a crooked two-up game, a local singer made up a song describing his actions and sang it to the local community in order to shame him. Other changes in society can be brought about by love songs, sung by both men and women to attract the opposite sex, or in performances of songs to harm or to heal people.

### 2.1 Longevity of songs

Pre-existent songs are either handed down from the Dreaming through generations or are given to people through direct communication with particular spirits, often when people are asleep or in an altered state of consciousness. The language of these recently ‘found’ songs is nevertheless perceived to be the language of the Dreaming ancestors and not of contemporary speech. Certain people have a predisposition to receive these songs,

---

5 The Akwelye song series was recorded by G. Koch in 1975 and by M. Turpin in 1999. In analysing the language of the songs the authors draw upon Harold Koch’s extensive work on Kaytetye.
6 Marett and Barwick (2007:4) discuss examples of songs as responses to social change.
7 Wild notes that people can also receive songs when they are sick or away from the rest of the community (1987:2).
which are then passed down through generations. In any given song series, some songs are well known, having been handed down from generation to generation, whereas others appear briefly then disappear forever. Midway between the two extremes, songs may be put away until a suitable period of time has passed following the death of a person associated with the song. A variation on this practice is where songs are often deliberately forgotten when the composer dies and rediscovered at a later date, and their original meaning and origin not mentioned or forgotten (Wild 1987:109).

2.2 Subject matter of songs

The interpretations of Central Australian songs given by singers often refer to events in the Dreamtime, uses of plants and animals, social practices and associated emotions and places, yet are rarely about contemporary events, as often happens elsewhere in Australia. In Central Australia, women’s songs reflect different topics to men’s songs, and children’s songs have different topics from adult songs. A Kaytetye girl’s song performed at sunrise translates as ‘(Sun) give me blonde hair, (sun) give me black hair’, while a Kaytetye boy’s song, performed on the full moon, translates as ‘(Moon make) my brother shorter, me the taller’ (Turpin 2003:61-65). As in many Aboriginal cultures, in Kaytetye the sun is associated with women and the moon associated with men. Both the children’s songs refer to growing up and can be seen to be individualistic, in contrast to Kaytetye women’s awelye songs, which focus more on social practices.

3. Difficulties posed by the language of song

There are many reasons why it is difficult to identify words in song and their meanings. Even when words are apparent in song texts, it is not always clear how their meaning relates to the explanations given by singers (Clunies Ross 1987:2). The practice of learning song texts separately from their meanings may be one factor (Donaldson 1995:149); another is the nature of the performance itself, which presents multiple inter-related elements of a story, or aspects of the Dreaming, rather than a sequential narrative as is common in European cultures (Ellis et al. 1990:115).

Song texts use words from many languages and words may be obscured by the substitution of individual phonemes. This can be done deliberately in order to disguise words when they are sung as part of a ceremony with age or gender restrictions, or it may be the result of adhering to metrical requirements. Furthermore, word boundaries in song texts may be obscured because the

---

8 In contrast, the Tiwi value the process of song creation and discard their Kulama songs after their initial performance (Grau 1983:70).
syllables that are prominent in song may be different from those in ordinary speech.

Song texts often contain segments with no known semantic meaning (Dixon & Koch 1996:12). Sometimes this is to achieve a given number of syllables per line (see § 3.1), or in the case of some Arandic songs, consonants may be inserted before a vowel-initial word to ensure that the text line begins with a consonant (Turpin 2007b:107). Merlan compares the verbal elements in song that have no semantic content with 'non-representational graphic forms' in art (1987:146). The analogy is a good one as it takes into account that in song the voice is a musical instrument as well as a source of language. Furthermore, in some contexts these purely musical segments are interpreted by singers as meaningful. This is an important feature of traditional songs: they can be interpreted differently in different contexts. When considering a word in song as a possible archaism it is important to know whether there are metrical requirements that force phonological alterations, including addition or deletion of syllables in the line.

Before considering the processes involved in creating a song text in more detail, it is necessary to consider the differences between the spoken and sung versions of a song text.

3.1 The transformation of ordinary language into song language

G. Koch (1987)* proposes a useful framework for song text analysis which involves making a distinction between two levels of description. The first is the text as it exists in the abstract, as some kind of verse distributed in definite lines. This is the version spoken in response to a researcher who asks for the text of a song. Kaytetye singers use this form to suggest which song to sing next and to tell others how the text of a song goes. The second level of description refers to the text as it is actually sung. If we recognise both a spoken and a sung version of a song text, we can also distinguish between two different processes that actually produce these texts:

1. The underlying prose, or ordinary language sentences are changed into the spoken version of the song text by what we might call the 'versification process.'

2. The spoken version of a song text becomes the sung version: through what we might call the 'performance process.'

The difference between the spoken song text and the sung text is the result of performance processes that require a text to fit with a musical structure. Leslie Russell, the maker of a Diyari song, stated the sentence in (1) to explain the theme of his song before performing it for Peter Austin (Donaldson 1979:78).

---

* Harold Koch came up with the general framework in discussion with G. Koch.
(1) *Nganha pinya-li rdiya-yi* (speech)
   “It is me the revenge party strikes”

The versification process produced the ‘spoken’ song text in (2), as recited by Leslie Russell after another recorded performance.\(^\text{10}\)

(2) *Pinya-li lha* (spoken song text)
   *Nganha rdiya*

In creating the spoken song text the word order has been changed (Diyari has free word order) and the sentence has been spread over two verse lines, each consisting of four syllables. In the process, the verb *rdiya- ‘strike’* has lost the final syllable that must always be present in prose to indicate the tense, in this case the non-past inflection *-yi*; and the first line has been filled out to four syllables by the addition of a form, *-lha*, which occurs in ordinary language only as a clitic particle signalling new information. Austin’s statement that ‘words of three syllables must have a syllable added to make them fit a full line’ (Austin 1978:531) is a metrical requirement applied to the level of versification.

When the text is sung, a further transformation takes place. The resulting text is presented in (3) with syllable boundaries represented by a space and morpheme boundaries (-).

(3) *Pi nya -nga li-lha-nga* (sung text)
   *Nga nha -nga rdi ya -nga*

Austin documents a ‘nonsense syllable’, *-nga* being inserted after each two syllables in order to make text lines of six syllables when the song is sung. This insertion is part of what we have called ‘the performance process.’ When we consider songs as possible repositories of archaic forms we need to consider whether additional syllables might not be the result of such performance processes.

4. **Phonological processes operating in song**

It can be difficult to know whether a phonological variant in song reflects an earlier form or whether it is the result of deliberate alteration, either to disguise words and so control knowledge or to fit the phonological or aesthetic requirements of song. Phonological variation can be either (i) applied systematically to all words in the song series; (ii) determined by the position of the syllable in the text line; or (iii) word specific. When considering a form in song as a possible archaism it is necessary to rule out (i) and (ii) as motivations

\(^{10}\) This example can be found on Archive Tape 3899 track B at 4 minutes 14 seconds into the tape. The tape is held at the Australian Institute of Aboriginal and Torres Strait Islander Studies Audiovisual Archive.
for the phonological variation. We now consider some systematic phonological alterations in the Kaytetye Akwelye songs.

Prestopping is contrastive in Arandic languages, such as Kaytetye and Alyawarr,\textsuperscript{11} yet in Akwelye prestopped nasals are always sung as plain nasals. For example \textit{atnerte} ‘stomach’ (An, Aly, Arr) is sung as a plain nasal, as shown in the text line in (4).

\begin{align*}
(4) & \quad [n \quad a \quad t \quad o \quad \text{wan} \quad \text{t'ja}] \quad \text{(song text)} \\
& \quad [a'\text{n}a\text{t}\text{a} \quad a\text{n}t'\text{ja}] \quad \text{(speech equivalents)} \\
& \quad \text{stomach} \quad \text{jump-(TNS)}
\end{align*}

H. Koch (1997a) proposes that initial consonant dropping and the appearance of the velar glide that occurred in Arandic languages came after the languages developed prestopping. If the plain nasals in song were due to an archaism we would expect to hear vowel-initial words with an initial consonant, and we would not expect /h/ but its ancestor, which Koch (1997a) proposes as /k/. The Akwelye songs, however, do have velar glides and the internal words of text lines show no additional initial consonant. This suggests that the plain nasals in song do not reflect an archaism.\textsuperscript{12} The modification of prestopped nasals to plain nasals in song has correspondences in some Arandic languages where prestopping is being replaced by plain nasals.\textsuperscript{13} Parallel phonological processes thus operate in both song and spoken languages.

Another neutralisation in song with a parallel in speech can be seen in relation to rounded consonants. Arandic languages show an unusual rounding realisation, whereby rounding is phonologically associated with consonants rather than vowels, although phonetically rounding is manifested on the surrounding vowels.\textsuperscript{14} In song, all rounded consonants other than velars are sung as unrounded, that is, there is no rounding on the surrounding vowels. H. Koch (1997a) proposes a sound change in Arandic languages in which rounded vowels induced rounding on the following consonant. Again, there are reasons why the unrounded forms in song are unlikely to be archaisms. An account of the discrepancy between rounded forms in terms of being an archaism would involve postulating that the initial vowels of words were dropped as they were incorporated into song texts. In a reduplicated form this means postulating initial vowel deletion for both elements in the reduplication, for example *utyert-utyerte ‘grass species’ would be versified as tyerte-tyerte and then we

\textsuperscript{11} For example, Kaytetye \textit{anteyane} ‘standing’ and \textit{anteyane} ‘sitting’.

\textsuperscript{12} A possible motivation for the change to plain nasals is the desire for increased sonority in song.

\textsuperscript{13} Breen notes that prestopping has been lost in the Arandic languages Western Anmatyerr and Ayerrereng, both of which now have only nasals, although Strehlow in 1944 recorded a distinction between long and short nasals (2001:56, 63). Green and Turpin show examples of free variation between nasals and prestopped nasals in Eastern Anmatyerr (2001:103).

\textsuperscript{14} Consider Kaytetye \textit{ike} ‘shell, crust’ with \textit{ikwe} ‘skin name’.
would have to propose that an initial [a] be inserted to the base and the
reduplicated part, as in atyerte-atyerte, which is a cumbersome account. Loss of
rounding has also occurred in some spoken Arandic languages.\(^{15}\)

In Akwelye there are other neutralisations limited to particular syllabic
positions in the song text, and again these have parallels in the spoken
languages. In Kaytetye and Alyawarr there are three sets of contrastive apical
consonants: post-alveolar, alveolar and prepalatal, while in other dialects there
are only two sets. In Akwelye when these consonants occur on an unstressed
metrical position this distinction is neutralised.\(^{16}\) In terms of identifying speech
equivalents this is not just a problem for the researcher, but also for singers of
the songs who are trying to identify the speech equivalents, as can be seen in
(5).

(5) (after singing song text [la]para[para])

\[
\begin{align*}
\text{L1} & \quad \text{S1: Arntwe-ne nte! fish- ... aherrk-akake yaye?} \\
& \quad \text{tell-IMP 2sgERG sun-PROP isn't it} \\
& \quad \text{“Tell us (what its about)”} \\
& \quad \text{S2: aylperre/ arlperre/ alperre} \\
& \quad \text{“fish/ whitewood/ leaf”} \\
& \quad \text{Aye?} \\
& \quad \text{“What?”}
\end{align*}
\]

\[
\begin{align*}
\text{L2} & \quad \text{S1: ‘rwarlperre-arlperre’ wante-rtame?} \\
& \quad \text{(song text line) what-CNTR} \\
& \quad \text{“What’s text line 11b (mean)?”} \\
& \quad \text{S2: arlperre-inge rarte arteye!} \\
& \quad \text{whitewood-EMPH 3sg-DEF tree} \\
& \quad \text{“Whitewood, that tree!”}
\end{align*}
\]

Speaker one (S1) is well versed in the songs and asks speaker two (S2), for the
interpretation of the song just sung. S1 hears the quality of her aunt’s lateral
apical as a prepalatal, and translates the word accordingly: aylperre ‘fish’,
although with some uncertainty.\(^{17}\) On this occasion S2, who did not appear to
hear speaker one’s mishearing as ‘fish’, interprets the word in the song as
arlperre ‘whitewood’. The example shows that while singers draw upon
phonological correspondences to interpret songs this is not always enough to
identify conclusively the word in a song text, as phonological reductions in song
can create numerous possible speech equivalents. Thus learners of songs are

\(^{15}\) Loss of rounding on bilabials has occurred in Western Arrernte (Breen 2001:65) as well as loss
of historical rounding of some Pama-Nyungan reflexes in all the Arandic languages. Consider the
reflex of proto-Pama-Nyungan *mulha which is elhe or alhe in Arandic languages.

\(^{16}\) Briefly, an unstressed metrical position in Akwelye is a syllable that falls on a short note and
between beats. For further discussion see Turpin (2005).

\(^{17}\) It may be that other speakers present heard speaker two’s apical differently, and so translated
the word as alperre ‘leaf’ or arlperre ‘whitewood’.
forced to rely on knowledgeable interpreters of the songs, as Merlan (1987:146) and Strehlow (1971) have noted.

Alterations to specific words in song may also be the result of various types of sound patterning. Reduplication is a widely attested poetic device in songs. In Akwelye, a common sound pattern for five syllable textlines is to have a pattern of reduplication where all but the initial consonant is reduplicated, and the first syllable must contain a coda. Such textlines, as the one under discussion in (5), are shown in (6). (Textline final vowels are also subject to a pattern of alteration occurring as a performance process, see Turpin 2007b:5).

(6)  Textline  speech equivalent
i.  [ol paael pə qa]  [oɭ pa əɭ]  “quick”
ii.  [la[pa ra]pə ra]  [a[pa qa]  “whitewood”
iii.  [ŋor pa[rer pə[a]  [kopa[apa][ə]  “bellbird”

The speech equivalent in (6iii) has no coda in the initial syllable, yet [r] is inserted as a coda in song to create the favoured VC.CV.CV pattern of reduplication.

It is worth considering whether such reduplication has a semantic effect. For example, in many Arandic languages a reduplicated plant term means ‘country dominated by this plant species’. However, from interpretations of the meanings of the songs there appears to be no semantic change and thus the reduplication is merely stylistic.18 The same pattern of reduplication is a common method of word formation in Arandic languages (cf. kwepalepale ‘bellbird’ above). For example, while the form pwelyerre does not exist in speech, the reduplicated pwelyerrelyerre (K) means ‘fine dust’.

In some cases the metrical account of the phonological alteration coincides with a postulated archaic form. For example some Akwelye songs have the form welyele ‘in the shade’. This has the speech equivalents elye-nge (K) and ulye-le (Arr, An, Aly) ‘shade-LOC’. The song form may be based on the (K) form elye, with a consonant inserted to adhere to the song’s required CV syllable structure, thus ELYE. This makes a stem of CVCV structure, and so it would be necessary to use the (K) -le suffix in place of -nge (which only goes on words of (V)CV structure). The form wilje ‘shade’ is also a possible Arandic reflex; Harold Koch argues that in Arandic wiC (where C is a palatal consonant) the vowel assimilated to wuC in (Arr, An, Aly) languages, whereas it remained wiC in (K) (2004:135).

In Akwelye there are very few phonological alterations to speech words that can not be explained on metrical or stylistic grounds. Sometimes these coincide with a postulated archaic form, and nearly always the processes creating these forms have parallels in spoken Arandic varieties.

18 The same pattern of reduplication is also attested in Nurlu songs of the West Kimberleys (Keogh 1990:65, 98). In Tiwi, similar patterns of reduplication occur to meet metrical requirements (Osborne 1989:206).
5. Syntax and morphology

5.1 Syntax

Generally songs have a much simpler syntax than that of everyday speech. Akwelye textlines often consist of a single noun followed by a verb, or just one or two nominals. Subjects are frequently omitted in Central Australian song texts (Keogh 1990; Strehlow 1971; Austin 1978). When subjects are present they are usually a pronoun, whose referent is not stated, as in the Akwelye text line in (7).

(7) [ləŋəŋə̱tər baŋə̱ra la wi ja] (song text)
    aylernanthe-arpe  ikngwe-rrr-lewe-yewe  (speech equivalent)
    1du.ex.OM.NOM-only  invite-RECIP+go&do.quickly-PURP

“We two ask each other to go off together alone.”

While this is also a feature of speech, the context of the utterance enables one to recover the subject in everyday speech, whereas the only way to recover subjects of Australian Aboriginal song texts is through knowledge of what the song is about. In the case of (7) singers translated the referents as two female cross-cousins in some contexts, while in less public contexts they were identified as a man and woman of the opposite patrimoie—potential spouses. Such vagueness means that interpretation relies on the authority of the recognised people.

Verb-final word ordering is widespread in Central Australian songs.19 This is in contrast to spoken Aboriginal languages which have pragmatically determined word order, although this may result in preferred orders.20 Given that verbs are often non-finite in song (see 5.2 below), their tendency to occur in final position may assist speakers in identifying them.

5.2 Morphology

A number of researchers have stated that Australian songs often have non-finite verb forms (Strehlow 1971, Austin 1978, Marett 2005). This is certainly borne out in the Akwelye songs. A relationship can be drawn between verbs without tense and the belief that songs contain the ever-present power of Dreaming ancestors rather than describe actions that happened in the past.

Another explanation for the absence of tense morphemes on verbs in song texts may be related to a process identified by Hale as ‘consonant transfer’

---

19 Strehlow notes this is the basic structure of Arrernte and Luritja songs (1971:213). It is also a strong tendency in Alyawarr awelye (Moyle 1986) and Nurlu songs from the Kimberleys (Keogh 1990:74), in Pitjantjatjara songs (Ellis 1998:435) and Diyari songs (Austin 1978:531–2).

20 Wilkins (1989) notes that in Arrernte there is a tendency for grammatical roles A, S and O to precede the verb. Breen (2008, pers.com.) notes this tendency in the Arandic language Antkerreopenh.
(1984) or transformations (Henderson pers. com.). Consonant transfer is when the final consonant of a text line is carried over to the beginning of the next text line. Consider (8) where the initial consonant /w/ may come from the final consonant of the suffix -yewe ‘tread-PURP’ (Aly):

(8) [wa lə mə tə nə wəŋ pi ja]

\[w=aleme \quad itenye\text{-}we \quad arnpe\text{-}(yewe)\]?

stomach love-DAT tread-(PURP)

? “(We) walk for love”

While tense marking morphemes are absent in songs from many areas of Australia, inflections for other grammatical categories such as associated motion, aspect, number, patrimoieties and generation are common in Arrernte and Kaytetye songs.

Some verbs in songs have a following element which has no known speech equivalent. In the Akwelye songs there is evidence that the element -rne which occurs after the frequently encountered verb arre- ‘put, make, create’ (K) corresponds to an archaic tense marker. Harold Koch (2002) argues that verb classes existed in a language ancestral to Proto-Arandic, and that -rne could have been a tense morpheme of one of the verb classes, which in (Arr, Anm, Aly) languages became reanalysed as part of the verb stem. Such verb classes would account for numerous cognate forms between (K) and (Arr, Anm, Aly) languages, identified by Koch (2002).

Further evidence for this comes from Warlpiri, where -rni--ni is the present tense morpheme of one of the four Warlpiri verb classes, in which class the possibly related yirrarni ‘put’ is also a member. If the -rne on these verb forms once signalled a tense, this is now no longer the case. This ‘poeticisation of grammatical forms’ is common in verbal art (Sherzer, 1990:18). Alternatively, it may have been that the verb stems of the non-Kaytetye (A) forms (arrerne-) rather than the Kaytetye form (arre-) were versified.

A clear case where the form of a word in song is archaic can be seen in an Arrernte charm sung to send away an approaching whirly-wind. The text of this song is shown in (9).22

---

21 It is difficult to come up with a free translation for this text line. Speakers explain it as a boy and girl wanting to marry. Aleme ‘stomach’ is the seat of emotion in Kaytetye. Under the entry for arnpe- ‘step’ in the Arrernte Dictionary an example sentence translates as “an older boy might follow a girl’s track stepping in the footprints, so that they’ll be married.” (Henderson & Dobson 1994:226).

For part A of this song, the spoken version is slightly different to the sung version. The spoken version is shown in (10):

(10) *mwere-angkwe-ktwe*  
*mother.in.law-2sgKin.POSS-AVERS*  

H. Koch proposes that the Arandic aversive morpheme, the ubiquitous ‘lest’ case in Aboriginal languages, derives historically from the dative case plus *+tye*. He states that kinship section marking morphemes (*-anthe*, *-ake~akerre*) may well have intervened (pers.com.), as they do in other case marking derived historically from a dative plus another element (Koch 1997a). In the above song we find a the 2sg possessive kin suffix *-angkwe* can intervene, thus providing further evidence for the origins of the Arandic aversive as the dative case plus *+tye*.

6. The register and language of song vocabulary

In the *Akwelye* series 98% of confirmed words are from everyday speech, however, in other genres this is much less.\(^{23}\) The remaining percentage are either archaic words or words of a special register known as ‘song language’ (discounting fillers to meet the metrical requirements as discussed in §3.1).

Strehlow (1947:xx-xxi) states that Arrernte songs abound in ‘archaic and obsolete words, no longer used in current diction, but traditionally preserved in these instances’, but in some cases these words are speech words from other Arandic varieties. Strehlow states *erare* [irrare] as one such form, but this word occurs in Western Arrernte meaning ‘lonely, homesick’ (Breen 2000:18). It is possible that such forms were borrowed for the purposes of the song and never existed in the spoken language and hence are not archaisms.

Determining the language and register of the words in songs is problematic when there are cognates with different forms, as song texts often contain words from neighbouring languages. Consider the text line of the Kaytetye song in (11).

\(^{23}\) In the Nyigina Nurlu song series 80% are from a spoken language (Keogh 1990).
There are two possible cognate forms for the segment of the text line meaning ‘alone’. Translation (a) shows the text based on the Kaytetye word angarrpe ‘alone’ with the Kaytetye pronoun aylewe, while (b) shows the text based on the related (Arr, An, Aly) suffix -arrpe ‘alone’ with Arandic arlenge ‘far’. In the case of this song text, the speech words corresponding with the first part of the song text do not reveal whether angarrpe or -arrpe is more likely. Both meanings, arlenge ‘far’ and aylewe ‘1du.SMSG’, are consistent with interpretations of this song by the singers (the final syllable of aylewe is elided in song because it begins with a glide, see Turpin 2007a). To complicate matters, the speech equivalents -arrpe and angarrpe also have a corresponding form in the avoidance register — ngketharrpe (K, Anm) ‘on your own’ (Jenny Green, pers.com). Assigning a single language and register to the form -arrpe in this song text obfuscates its relationship to these derived forms and cognates.

Merlan (1987:145) notes that singers themselves are reluctant to assign words in song texts to a particular language, and that when identification of a particular language is made by singers this often refers to territorial information rather than information about the text. The reluctance of singers to see a song text as being of a particular contemporary language may reflect the perception that a song text is the language of Dreaming ancestors (who cross linguistic boundaries) and not of a spoken language, as well as the difficulties inherent in assigning words from ancient texts to modern languages. There may also be a practical element to this, as ceremonies bring people together and this can be done iconically by combining languages in song texts.

7. Semantic extension in songs

Indigenous songs can have many, often divergent, interpretations in the vernacular language. Ellis suggests that multiple interpretations are the norm in Pitjantjatjara songs (1985:63), where the interpretations are the result of highlighting various meanings of the one word. Multiple meanings may also be the result of different people’s perceptions of what the words in the song are (Keogh 1990:85). For reasons of space we consider only meanings where a word

---

24 The first element of this word occurs in certain other respect vocabulary items; compare respect register ilkawngkethe ‘company’ with the neutral register equivalent ilkawtherre.

25 This was noted as early as 1897 by Roth (1897:118-120). Sutton (1987:85), Ellis (1997:33) and Marett (2005).
recurs in a number of songs with a broader meaning than in speech. When such a word recurs in the same position to the same rhythm in various text lines, this is called a ‘formula’ (Rumsey 2001:194). Formulas are common in Central Australian songs (Austin 1978:530), particularly at the end of text lines (Barwick & Marett 2003:26-27).26

In this section we consider the most common formula at the end of Akwelye textlines. In §5.2 we considered the verb form in song arrerne which has the speech equivalents arrerne- (Aly, Anm, Arr) are- (K) ‘put’. While in some songs this is interpreted as ‘put’, in other songs it is interpreted as meaning ‘call (a person “cousin”)’ and in others ‘create, make’, where the textline as a whole means ‘create a women’s ceremony’, ‘create river red gum manna’, and ‘make a girl healthy’. These meanings are not attested in spoken Arandic languages, but there is evidence to suggest that this polysemy may have existed in Kaytetye. The Kaytetye verb are- includes the meanings ‘put up (a construction)’ and ‘make (an object) from something else’. Furthermore, in Kaytetye we find many transitive verb stems that end in -are-. Some of these verbs can be analysed as consisting of a bound cran-type morph (as in cran-berry) plus -are-, which has a transitiviser function, as in (12).27 This suggests that -are- was a formative with a transitivising function, possibly related to the modern verb are- (K) arrerne- (Arr, Aly, An) ‘put’.

(12)  
\begin{align*}
\text{akem-ane- (v.i.)} & \quad \text{‘sit on your own’ (lit. ?-sit)} \\
\text{akem-are- (v.t.)} & \quad \text{‘make someone sit up’} \\
\text{akake.we- (v.t.), akak.arre- (v.t.)} & \quad \text{‘knock something in to stop something from coming loose or falling out’}
\end{align*}

The semantic association between ‘call’, ‘create’ and ‘put’ may reflect the belief that in naming the country the ancestors created it, and that in putting or placing their objects in the country the ancestors created it.29 The semantic extension from ‘put’ to ‘create’ may be motivated by a metaphor whereby the actions of the Dreamtime women, such as naming and putting on ceremonial stripes, are perceived as having the effect of creating, in this case creating the ceremonies which continue to be practised today.

27 There are also transitive verbs that have the segment -are- where the meaning of the other morphemes is known in everyday speech, such as rlweth-arl-are- (lit. half-ALL-put) ‘go through the middle of something’ (lit. put (oneself) into half (of a place)).
28 The verb we- ‘throw’ is a formative in many Arandic verb stems.
29 Strehlow argues that ‘create’ is the appropriate English word for the process involved in many Dreaming narratives where ancestral being ‘create’ the country (1947:xiii).
8. Conclusion

In this paper we have considered Arandic song texts in terms of how they can contribute to our understanding of language change and reconstruction. The multi-dialectal nature of song, the presence of metrical requirements and sound patterning which force phonological alteration of words, as well as the methods of song transmission and interpretation, mean that caution should be taken before assigning the category of archaism to unfamiliar words in song. Taking such caution into account we have looked at how words in Arandic songs accord or otherwise with some of the changes to Arandic languages as proposed by Koch. While an analysis of the song texts reveals some archaic morphological forms, the phonology of songs is on the whole not archaic, yet the types of phonological and semantic processes that are employed in the creation of song texts do resemble diachronic and synchronic processes in spoken Arandic registers.

With further cross-disciplinary work on songs and increased knowledge of linguistic change in Arandic languages we may find further archaisms in song, yet without this knowledge we are merely speculating on the origins of this unique and highly significant oral literature of Aboriginal Australia.

References


1. Introduction

According to one recent authority, noun class systems in Australian languages have been subject to widespread diffusion over northern Australia:

The variation in noun classes — their number, form and places of marking — is consistent with the hypothesis presented here, that noun classes have developed recently, as an areal phenomenon, within the prefixing region. It is basically the category of noun classes that has diffused, with each language developing the actual marking for itself, out of its own internal resources. (Dixon 2002:471)

While Dixon allows that noun class systems can sometimes be traced back to a proto-language, he considers these cases to be the exception rather than the rule. In particular, he implies that the areal diffusion scenario holds for the noun class systems of the Worrorran languages, which he considers to represent a “small areal group” (Dixon 2001:102, 2002:xli, 672-674). He allows that the languages might form a genetic group at a considerable time-depth, but avers that

the evidence is greatly in favour of the alternative scenario — that they are simply three languages which have been in contact for a long time, so that they have grown similar in their typological profile and have borrowed between each other a fair number of lexemes, together with just a few grammatical forms. (Dixon 2002:674)

These proposals are not backed up by any empirical evidence; it is simply claimed that there are few grammatical items shared among the languages, and that numerous differences exist in the grammatical categories. This is false: numerous cognates can be identified among the grammatical items of the languages, sufficient to permit reconstruction of some of the grammar of the proto-language, as demonstrated in McGregor and Rumsey (forthcoming). Furthermore, we have identified a number of lexical correspondences that are more likely to represent retentions than borrowings; these include some thirty-

---

1 A preliminary version of this paper was presented at the Lingvistkredsens Workshop, Genetic vs. Contact-based structural change, Copenhagen University, December 2004. I am grateful to the audience for useful comments, and to Alan Rumsey for detailed comments on an earlier draft, as well as for many discussions on comparative Worrorran. Thanks also to two anonymous referees and to the editors Bethwyn Evans and Claire Bowern for helpful comments. The usual disclaimers apply. It is a privilege and pleasure to contribute to this volume in recognition of Harold Koch’s important contribution to Australian Aboriginal historical linguistics.
six items from the hundred basic lexemes that have been shown to be stable cross-linguistically, including in Australian languages (e.g. Breen 1990; Black 1997). Some of the corresponding items are from non-contiguous languages, rendering a borrowing account problematic. In addition, we have identified a few regular sound changes.

As Dixon says, the systems of noun classes and the forms and loci of the markers, vary across the languages. But this neither argues against retention of noun classes from a proto-language nor supports the areal proposal. To motivate either requires the construction of a plausible diachronic scenario. For retention from a proto-language, reconstruction of a proto-system and indication of how the modern systems came about are requirements. For the areal proposal a plausible mechanism is required for borrowing of grammatical systems independent of the morphemes that realise them. Dixon’s assumption that borrowing represents the default situation, and requires neither argument nor evidence, is untenable.

The present paper elaborates on the story of development of noun classes in Worrorran sketched in McGregor and Rumsey (forthcoming), providing details beyond the scope of that study. Specifically, it provides additional evidence in support of the scenario outlined in that paper, not so much by developing the formal aspects of the reconstruction, but more through an elaboration of the restructurings of the systems and the accompanying semantic changes, particularly in Northern Worrorran.

2. Some background on the Worrorran languages

The Worrorran languages constitute a group of a score or so of named varieties spoken in the Kimberley bloc (McGregor 2004:42-44, 50). Many of these are, to greater or lesser extents, mutually intelligible. Nevertheless, it is important to maintain the distinctions among the varieties, which I will usually refer to (with reservations) as languages. To collapse them into just three languages, as does Dixon (2002), obscures major divergences among them. For instance, Unggumi and Worrorra are grouped together as dialects in Dixon (2002) even though they share only about 50% of similar lexemes from a basic 105 item vocabulary list (McGregor & Rumsey forthcoming). They also show significant grammatical differences that call to question their status as dialects.

The score or so of Worrorran varieties fall, McGregor and Rumsey (forthcoming) argue, into three genetic groups: Western (Worrrioric), Eastern (Ngarinyinic), and Northern (Wunambalic). These three groups are as shown in Figure 1.
3. **Overview of noun class systems in the Worrorran family**

In all Worrorran languages noun classes are agreement classes defined by prefixes and/or suffixes to the dependents of a nominal (adjectives and determiners) and verbal prefixes cross-referencing intransitive subject and transitive object NPs. They are thus systems of nominal gender in the terminology of Corbett (1991). The following (partly constructed) examples illustrate some of these agreement patterns for demonstratives and verbs in Gunin/Kwini (Northern Worrorran).\(^2\)

(1) \textit{benyjin bi-nya bi-yangga} \\
Gunin/Kwini  \\
man HUM-this HUM-goes  \\
“This man is walking.”

(2) \textit{leewa a-nya gadi a-yangga} \\
Gunin/Kwini  \\
dog AN-this run AN-goes  \\
“This dog is running.”

\(^2\) For illustrative purposes I have added in to some of the examples demonstrative forms not in the original utterances, but which are appropriate for NPs of that gender, as tested in elicitation sessions.

\(^3\) The following abbreviations are used: AN: animate, FEM: feminine, HUM: human, MAS: masculine, NEUT: neuter, O: object, PA: past, PL: plural, S: subject, and SG: singular. Morphemic analyses of inflecting verbs are incomplete.
Free third person singular pronouns also come in different forms according to the noun class of the referent: thus five different third person singular forms could be used in place of the NPs in the above examples: bini, aani, wini, mini, and 'nani. In most Worrorran languages some interrogative determiners – especially the 'where' interrogative – also occur in different forms depending on class. What-cha-ma-call-it items sometimes also get class marking.

Class marking occurs in the same morphological loci as person and number marking in all Worrorran languages. Adjectives that take class prefixes also take person and number prefixes: the adjective -yaba 'good, nice' in Wunambal for example takes class prefixes bi-, a-, etc., as well as person prefixes nga- ‘I’, gV- ‘you (singular)’, nyarr(a)- ‘we (exclusive)’, birr(a)- ‘they’, etc. In some Western Worrorran languages, adjectives and determiners occasionally take class-marking suffixes, usually in addition to prefixes. For instance, in Unggumi -neya ‘good’ always takes an agreement prefix, as in i-neya ‘MAS-good’; it occasionally also takes a suffix in addition, as in nyi-neya-ngarrri-nya ‘FEM-good-COM-FEM’. A few adjectives take just class-marking suffixes, e.g. jowingarri-ye ‘big-MAS’.

Inflecting verbs (IVs) also take pronominal prefixes in the same position as class prefixes, as in e.g. Wunambal ngu-wane ‘I fell’, ngarr-wane ‘we (inclusive) fell’, gu-wane ‘you fell’. However, pronominal prefixes indicate gender of cross-referenced intransitive subject and transitive object NPs only. Pronominals cross-referencing transitive subjects and oblique do not distinguish gender – they use a single general 3rd person singular form. A partial exception occurs in Northern Worrorran languages, where transitive subject prefixes make a binary contrast between human and non-human genders. And according to Capell and Coate (1984:120), in Western Worrorran languages a minimal gender distinction is made in the oblique, where FEM is specified by the augment -nya.

Classes are not normally marked on the nouns themselves, except to a limited extent in Western Worrorran languages. For instance, Unggumi has four gender-marking suffixes, the same ones as taken by some adjectives (see above):

\[\text{gunu} \text{rru mi-nya bugur}r \text{ mi-rranggu} \text{Gunin/Kwini}
\]

wind NEUT M-this blow NEUT M-be

“This wind is blowing.”

\[\text{duurr } \text{ nin-da } \text{ warna ni-nya} \text{Gunin/Kwini}
\]

put NEUT$_n$-put honey NEUT$_n$-this

“Put down this honey.”
-ya MAS, -nya FEM, -ngga NEUT_w, and -ma NEUT_m. For some nouns these suffixes seem to be invariably attached to the lexical noun, as in (6a-c), except when it hosts a case-marking clitic; for other nominals, they are optional, as in (6d). For many nominals, e.g. *gilinygiliny* ‘galah’ and *walarri* ‘river gum’, however, the suffixes are not attested at all. Motivations for these differences are not known.

(6)  

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>inya</td>
<td>iri-ya</td>
<td>(MAS)</td>
<td>Unggumi</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>this man</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“this man”</td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>nyinye</td>
<td>jili-nya</td>
<td>(FEM)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>this woman</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“this woman”</td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>wun.ga</td>
<td>winhdhalgi-ngga</td>
<td>(NEUT_w)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>this fire</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“this fire”</td>
<td></td>
</tr>
<tr>
<td>d.</td>
<td>manma</td>
<td>dambi-(ma)</td>
<td>(NEUT_w)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>this place</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>“this place”</td>
<td></td>
</tr>
</tbody>
</table>

In Worrorra the suffix-status of the apparently cognate final segments is less clear-cut: they show more phonological variation, they appear to belong to the stem, and associations with classes are statistical, as shown in Table 1. About 11% of Worrorra nouns have initial segments that indicate gender. According to Clendon (2001:111), gender is signalled by the final segment for only 62% of nouns, by initial and final segments in 9% of nouns, and by initial segment only in 2% of nouns. It is the FEM -nya that has the best status as a class-marking affix, being detachable from the nominal, indicating that it is not (yet) a part of the stem. The other formatives appear not to be detachable.

The situation is less certain in the other Western languages, where information is piecemeal. Examples cited in Capell and Coate (1984) tend to suggest that the situation in Yawijibaya, Umiida, Winyjarrumi, and Unggarrangu is as in Worrorra: statistical correlations between final segments and noun class.

In Eastern and Northern Worrorran languages such gender-indicating endings (or initials) are not discernible on nouns; nor are there strong phonological correlates of the noun classes. In Ngarinyin there are just a handful of regularities. Thus about 90% of words ending in /n/ are in the NEUT_w class (Rumsey 1982:37).
In all Worrorran languages there exists a relatively small class of bound prefix-taking nouns, mainly designating parts of the body. Prefixes to these nouns cross-reference the inalienable possessor in terms of person, number, and class. In Eastern and Western Worrorran languages prefix-taking body part nouns have inherent gender, as indicated by the form of agreement markers, as in Ngarinyin (Rumsey 1982:55-56). In Unggumi bound nouns often take class-marking suffixes as well, as in (7), where the prefix *nga-* to the bound noun indicates that the possessor is the speaker, while the suffix *-ma* indicates the class of the bound noun *milarr* ‘forehead’, also indicated by the form of the demonstrative. The situation in Northern Worrorran languages is uncertain: Vasse (1991:30) avers that prefixing nouns do not have inherent genders, and that gender agreement is always with the possessor, as in (8). Non-prefixing body part nouns in all Worrorran languages have agreement genders, indicated by the form of co-occurring demonstratives.

(7) manma nga-milarr-ma Unggumi
    this:NEUT$_m$ 1SG-forehead-NEUT$_m$
    “my forehead”
    Forest River (Yilji?)

(8) nga-ninga ngi-ninga ngi-yaba nga-lali garij ngambinabang
    1SG-name 1SG-this 1SG-good 1SG-new call they:did:to:me
    “They gave me a good new name.” (Capell & Coate 1984:71)

Table 1: Associations between final segments and classes in Worrorran
(after Clendon 2001:110, slightly modified)

<table>
<thead>
<tr>
<th>Final Segment</th>
<th>MAS</th>
<th>FEM</th>
<th>NEUT$_m$</th>
<th>NEUT$_w$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-Vja</td>
<td>11</td>
<td>10</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>-nya</td>
<td>1</td>
<td>93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-(r)ljia</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-j</td>
<td></td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>-ya</td>
<td>21</td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>-(r)lyja</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-(r)nja</td>
<td>18</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>-i</td>
<td>45</td>
<td>1</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>-(b(a) = -m(a)</td>
<td>7</td>
<td>119</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>-u</td>
<td>8</td>
<td>3</td>
<td>70</td>
<td></td>
</tr>
</tbody>
</table>

% of Ns in the class marked by these segments 57 100 79 57

There are grounds for caution in accepting this observation (cf. also Carr 2000: §3.1.1, note 10). It is typically the case in languages of the region that external possession constructions are normally employed for inalienable possessions; in these the person rather than the part is cross-referenced in the verb (McGregor 1999), as in (8). Further inquiries should be made to ascertain what happens in circumstances where the body part is detached (e.g. ‘His [the dead person’s] arm was twitching’, ‘I saw his leg/tail (but not the entire individual)’, etc.). (For the same reason I question the validity of the claim that in most class-marking Australian languages body part nouns lack their own gender (Dixon 2002:487).)
Northern Worrorran languages typically distinguish five noun classes – four in southern Wunambal according to Capell and Coate (1984:63, 65) – while the Western and Eastern languages distinguish four. In addition, the third person plural category might reasonably be included among the noun classes, as it is by Capell and Coate (1984:81) (where it is dubbed Class III) and Rumsey (1982:32-41). This is partly because it is marked by a morpheme that occurs in the same order-class positions as the class markers, and partly for reasons that will become apparent later. For now we exclude the plural category from our discussion.

The five classes typical of the Northern Worrorran languages are:

- **HUM** a human class, made up exclusively of human nominals;
- **AN** an animate class;
- **NEUT\_w** a neuter class, typified by an initial /w/ in agreement affixes;
- **NEUT\_m** a neuter class, typified by an initial /m/ in agreement affixes;
- **NEUT\_n** a neuter class, typified by an initial /n/ in agreement affixes; this gender is absent from Southern Wunambal.

In the Western Worrorran and Eastern Worrorran languages the four noun classes are as follows:

- **MAS** male humans and some higher animate males; various other nominals;
- **FEM** female humans and some higher animate females; various other nominals;
- **NEUT\_w** a neuter classes, typified by an initial /w/ in agreement affixes;
- **NEUT\_m** a neuter classes, typified by an initial /m/ in agreement affixes.

Categorisation of human nouns is almost completely predictable. In Northern Worrorran languages human nouns are universally assigned to the HUM class, while in the Western Worrorran and Eastern Worrorran languages they are assigned to MAS and FEM classes according to sex. For non-human nominals things are less consistent and clear-cut, and they are generally spread across the classes, with the exception of HUM in Northern Worrorran. Although class membership for non-human nouns is not predictable according to semantic characteristics (e.g. Capell & Coate 1984), it is not random either. We return to semantics below.

The final aspect of noun classes that needs to be mentioned is their formal realisations. We begin with the overview provided in Capell and Coate (1984), as shown in Table 2, since this provides information on almost all of the languages, including many for which data is scarce. The forms should be taken with a grain
of salt because in many cases the authors recorded very limited data on the languages, and, for the better documented languages, did not push their analyses very far. Their classes, indicated by Roman numerals I-VI, correspond in a quite messy way to the above class labels, and serve more to obfuscate correspondences than reveal them. We return to the forms shortly, and present an alternative tabulation.

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yawijibaya</td>
<td>i-</td>
<td>ny(a)-</td>
<td>ya-</td>
<td>yw-</td>
<td>m-~m_b</td>
<td></td>
</tr>
<tr>
<td>Winyjarrumi</td>
<td>ya-</td>
<td>ny(a)-</td>
<td></td>
<td>w-</td>
<td>m-~ma</td>
<td></td>
</tr>
<tr>
<td>Worrorra</td>
<td>i-</td>
<td>ny(a)-</td>
<td>a-~ya-</td>
<td>w-</td>
<td>m-~m-b, ba</td>
<td></td>
</tr>
<tr>
<td>Unggumi</td>
<td>yi-</td>
<td>ny(a)-</td>
<td>arr-</td>
<td>ew-~ngga</td>
<td>m-~ma</td>
<td></td>
</tr>
<tr>
<td>Unggarrangu</td>
<td>i-</td>
<td>ny-</td>
<td>ya-</td>
<td>w-</td>
<td>m-</td>
<td></td>
</tr>
<tr>
<td>Umiida</td>
<td>i-</td>
<td>ny-</td>
<td>i-</td>
<td>w-</td>
<td>m-</td>
<td></td>
</tr>
<tr>
<td>Ngarinyin</td>
<td>a-</td>
<td>ny-</td>
<td>birri-</td>
<td>w-~g-</td>
<td>m-</td>
<td></td>
</tr>
<tr>
<td>Guwij</td>
<td>a-</td>
<td>ny-</td>
<td>birri-</td>
<td>w-~g-</td>
<td>m-</td>
<td></td>
</tr>
<tr>
<td>Wilawila</td>
<td>a-</td>
<td>ny-</td>
<td>birri-</td>
<td>w-~g-</td>
<td>m-</td>
<td></td>
</tr>
<tr>
<td>Munumburru</td>
<td>a-</td>
<td>ny-</td>
<td>birri-</td>
<td>w-~g-</td>
<td>m-</td>
<td></td>
</tr>
<tr>
<td>Wolyamidi</td>
<td>a-</td>
<td>ny-</td>
<td>birri-</td>
<td>w-~g-</td>
<td>m-</td>
<td></td>
</tr>
<tr>
<td>Waladjangari</td>
<td>a-</td>
<td>ny-</td>
<td>birri-</td>
<td>w-~g-</td>
<td>m-</td>
<td></td>
</tr>
<tr>
<td>Forest River</td>
<td>b-</td>
<td>a-</td>
<td>birra-</td>
<td>w-~g-</td>
<td>m-</td>
<td>n-</td>
</tr>
<tr>
<td>Gunin/Kwini</td>
<td>b-</td>
<td>a-</td>
<td>birra-</td>
<td>w-~g-</td>
<td>m-</td>
<td>n-</td>
</tr>
<tr>
<td>Gambera</td>
<td>b-</td>
<td>a-</td>
<td>birr-</td>
<td>w-~g-</td>
<td>m-</td>
<td>n-</td>
</tr>
<tr>
<td>Wunambal N.</td>
<td>b-</td>
<td>a-</td>
<td>birr-</td>
<td>w-~g-</td>
<td>m-</td>
<td>n-</td>
</tr>
<tr>
<td>Wunambal S.</td>
<td>b-</td>
<td>a-</td>
<td>birr-</td>
<td>w-~g-</td>
<td>m-</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 2: Agreement class prefixes in the Worrorran languages  
(after Capell & Coate 1984:63; spelling adjusted)

In conclusion, it will be observed that the claim that there is little commonality in the forms, loci, systems, and semantics of the Worrorran gender systems is at variance with the facts presented in this section.

4. Reconstruction of noun class system of proto-Worrorran

I submit that proto-Worrorran had a system of noun classes marked by agreeing determiners and bound pronominals to dependents in the NP, as well as to IVs. Furthermore, it distinguished four classes: a masculine, a feminine, and two neuters, a W neuter and an M neuter. This system remained largely intact in Eastern and Western groups (though local changes occurred, including reclassifications), but was restructured in the Northern group, as follows:

- A new HUM category emerged via backformation from the original plural category, and human nominals were consistently assigned to this class.
- The original MAS class became the AN class.
The original FEM class became the NEUT\textsubscript{w} class; this class eventually disappeared (or almost disappeared) in Southern Wunambal.

I now present the evidence for this scenario, acknowledging that I do not yet have a complete and watertight account of the changes.

Let’s begin with the proto-Worrorran system. It is suggested that it distinguished four noun classes by prefixes to dependents in the NP, and bound pronominals to IVs cross-referencing intransitive subjects and transitive objects. The prefixes are tentatively reconstructed as follows:

\[
\begin{align*}
*a^- & \quad \text{MAS} \\
*N(a)y & \quad \text{FEM} \\
*g(V) & \quad \text{NEUT}\textsubscript{w} \\
*m(a) & \quad \text{NEUT}\textsubscript{m}
\end{align*}
\]

Evidence in favour of the two neuter categories is formal and semantic. The formal correspondences across the languages is obvious from Table 2, and from the additional data from better attested languages presented in Table 3. (The allomorph sets are incomplete, a number of irregular forms have been ignored, and only ordinary indicative mood forms are given for the pronoun prefixes to IVs.)

Prefixes: Adjectives Determiners Intransitive subjects

| Worrorra | N-\textsuperscript{*}wuN- | ma- | wu-\textsuperscript{*}ku-\textsuperscript{ka-} | ma- | guN- | ma- |
| Unggumi | \textsuperscript{*}wu- | ma- | wu- | ma- | gu- | ma- |
| Ngarinyin | wu- | ma\textsuperscript{a}- | gV-\textsuperscript{*}wV- | mV- | w,\textsuperscript{u}- | ma\textsuperscript{a}- |
| Wunambal | w-\textsuperscript{a} | mi-\textsuperscript{*}mungg- | w-\textsuperscript{a} | mi-\textsuperscript{*}mungg- | wV- | mV- |
| Gunin/Kwini | w-\textsuperscript{w}V- | m-\textsuperscript{*}mV- | w-\textsuperscript{w}V- | m-\textsuperscript{*}mV- | w-\textsuperscript{w}V- | m-\textsuperscript{m}V- |

Table 3: Forms of bound class-marking affixes for two neuter classes in five languages

Prefix allomorphs mostly show initial \textsuperscript{w} and \textsuperscript{m}, followed by a vowel, in most cases \textsuperscript{u} and \textsuperscript{a} respectively. We also find a scattering of allomorphs of the former prefixes with an initial \textsuperscript{g}, and in Worrorra a prefix initial or final assimilating nasal (the former doubtless arises through loss of the preceding segments). It seems likely that the proto-form had the initial velar stop, and that the glide results from lenition. This lenition process is attested elsewhere in Kimberley languages, and is also consistent with the fact that there is evidence of a \textsuperscript{w} \sim \textsuperscript{b} alternation, where the stop occurs after nasals, which McGregor & Rumsey (forthcoming) trace back to \textsuperscript{*}w.

That the corresponding classes are more than merely formally related emerges from their semantics. Table 4 summarises the semantic content of the two classes in four languages for which sufficient reliable data is available.
The NEUT\textsubscript{W} and NEUT\textsubscript{M} classes contrast on a number of dimensions. Nominals denoting abstract entities (such as languages) are typically assigned to NEUT\textsubscript{W}, whereas those denoting places and surfaces go to NEUT\textsubscript{M}. The contrasts can be roughly encapsulated as an opposition between things or individuated entities (NEUT\textsubscript{W}) and places or mass, non-individuated, substances (NEUT\textsubscript{M}). The correlations are striking, and more consistent than expected if the classes arose through areal borrowing of the notion of noun classes, as per Dixon (2002:471).

<table>
<thead>
<tr>
<th></th>
<th>NEUT\textsubscript{W}</th>
<th>NEUT\textsubscript{M}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Worrorra</td>
<td>times, bright, translucent, abstract entities, high, raised things (trees, hills, islands)</td>
<td>places, locations, dark, opaque, low, flat places, some edible roots, grasses</td>
</tr>
<tr>
<td>Unggumi</td>
<td>things, a few temporals, trees, wood, fire, inedible vegetable matter, abstract entities</td>
<td>places, surfaces, times, majority of artefacts</td>
</tr>
<tr>
<td>Ngarinyin</td>
<td>time, trees, wood, and their products, rocks and minerals, many body parts, abstract entities</td>
<td>places, topographic phenomena, plants, especially edible, many body parts</td>
</tr>
<tr>
<td>Wunambal</td>
<td>objects, things, individuated entities, abstract entities, shiny objects, inedible fruit bearing trees; a few edible poisonous roots requiring preparation</td>
<td>places, topographic feature, surfaces, dark objects, food-bearing plants and trees, root foods</td>
</tr>
</tbody>
</table>

Table 4: Summary of semantic contents of the two neuter classes NEUT\textsubscript{W} and NEUT\textsubscript{M}

According to Dixon (2002:493), around half of the gender-marking non-Pama-Nyungan languages have an M class consisting primarily of nouns designating vegetable foods. Consistent with this, in Ngarinyin and Wunambal edible vegetable matter typically goes to the NEUT\textsubscript{M} class (cf. discussion of example (9) below), while inedible vegetable matter, including the generic term ‘tree’ and many trees and tree products are assigned to the NEUT\textsubscript{W} class. The latter assignment is consistent across Worrorran languages, and may well form the basis for the association of the W class with things (in many Australian languages the word for ‘tree, wood’ is also used generically for ‘thing’). The connection between NEUT\textsubscript{W} and edible vegetable food does not, however, hold in Western Worrorran languages: in Unggumi vegetable matter denoted by NEUT\textsubscript{M} nominals is consistently inedible – edible foods go primarily into the MAS class; and in Worrorra foods of all types are primarily assigned to a
collective class, a small and marginal fifth class defined by the third person plural as the agreement marker (Clendon 2001:143-144).

This is consistent with the genetic relatedness of the two neuter classes: it is plausible that the proto-Worrorran \textit{NEUT}_m contained most edible vegetable foods, and that the system was restructured in Western Worrorran, where vegetable foods were recategorised. How might this have come about? In all Worrorran languages \textit{NEUT}_m is strongly associated with places, an association that may have initially arisen through metonymic transfer from vegetable matter to associated locations. This association may have strengthened to the point of pushing vegetable foods out of the class in the Western Worrorran languages; alternatively, vegetable foods may have been reclassified to distinguish them from the locations they characterised.

A number of non-Pama-Nyungan languages also show a \textit{gu(n) ~ wu} class (Sands 1995:313-314; Dixon 2002:496-497), although it is less semantically consistent than the M class cross-linguistically. In the absence of any serious attempt at higher level reconstruction, it is impossible to determine whether or not Worrorran \textit{NEUT}_w and \textit{NEUT}_m classes are historically related to these classes in other non-Pama-Nyungan languages. It is equally plausible that the similarities are due to accidental similarity (the forms are short and not highly marked) or to borrowing, in which case the formal correspondences may be reflections of use of similar lexical material, such as the widespread \textit{mayi} ‘vegetable food’. A lexical source for the \textit{gu(n) ~ wu} class marker is less evident, but could possibly lie in a demonstrative \textit{ginya} ‘this, here’ found in a scattering of Australian languages.

The two other non-human classes in Northern Worrorran languages, AN and \textit{NEUT}_n, can be plausibly equated respectively with the MAS and FEM classes in the other two groups. Again, the reasons are both formal and semantic. Table 5 shows a selection of allomorphs of the relevant class marking prefixes.

<table>
<thead>
<tr>
<th>Class</th>
<th>Adjectives</th>
<th>Determiners</th>
<th>Intransitive subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Worrorra</td>
<td>i-~a-</td>
<td>nyi(N)-</td>
<td>a-~i-~ga- nyV- ga- nyi-</td>
</tr>
<tr>
<td>Unggumi</td>
<td>i-</td>
<td>nyi-</td>
<td>i- nyV- ga- nyi-</td>
</tr>
<tr>
<td>Ngarinyin</td>
<td>a2-</td>
<td>nya-</td>
<td>a-ji- nya2- a1- nya2-</td>
</tr>
<tr>
<td>Wunambal</td>
<td>a-</td>
<td>na-</td>
<td>a- ni(ng)- a~da- nV-</td>
</tr>
<tr>
<td>Gunin/Kwini</td>
<td>a-</td>
<td>n(V)-</td>
<td>a- n(V)- a- n(V)-</td>
</tr>
</tbody>
</table>

Table 5: Some allomorphs of two class-marking prefixes in a selection of languages

The phonological correspondences are not as straightforward as in the case of the two inanimate genders, and I cannot provide a full account of the development of the modern forms. Nonetheless, each language has at least one allomorph of the MAS class marker with the low vowel. Perhaps this was the proto-vowel, the high vowel deriving via subsequent morphophonemic or analogical processes.
The correspondence between nyV- in the Eastern and Western groups and nV- in the Northern group is not implausible. Laminal-apical correspondences are well attested in Australian languages (e.g. Evans 1988), and can be identified for stops and glides in Worrorran (McGregor & Rumsey forthcoming). There is evidence that proto-Worrorran recognised three phonemic series: lamino-dentals, lamino-palatals, and apico-alveolars. It is possible that the FEM marker in proto-Worrorran was *NHa or *NHi. This could have developed into na in Northern Worrorran languages with the loss of the lamino-dental series (a process that seems to have been ongoing in historical times with the lamino-dental stop — McGregor & Rumsey forthcoming). It could have changed into nyi- in the other groups, under influence of the high front vowel, which may explain why Unggumi has an initial lamio-palatal rather than the expected lamino-dental nasal. The different shapes of the initial nasal, that is, may result from of an earlier change in vowel quality in certain groups.

Independent support comes from semantics. Table 6 shows some of the main semantic principles underlying the AN and NEUTn classes in Wunambal.

The semantic content of these two classes resembles that of the MAS and FEM classes of Eastern and Western Worrorran languages, minus human males and females. First, notice the categorisation of artefacts metonymically related to males and females tend to be accorded to the AN and NEUTn classes, respectively, as is common in gender languages, including Ngarinyin (Rumsey 1982:37-39) and Worrorra (Clendon 2001:334).

<table>
<thead>
<tr>
<th>AN</th>
<th>NEUTn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human-like: 'wanjina', 'devil', 'dead person', 'rain', 'lightning', 'thunder'</td>
<td>Flighted animals: birds (generic), many bird species, insects?</td>
</tr>
<tr>
<td>Land animals: 'horse', 'goanna', 'dog', 'emu'; 'fruit bat'</td>
<td>Things associated with female sphere: 'coolamon', 'honey', 'eggs', 'sun' (mythologically female)</td>
</tr>
<tr>
<td>Things associated with animals or males: 'spear', 'axe', 'knife', 'headband', 'sinew', 'moon'</td>
<td>Concave shapes: 'canoe', 'coolamon', 'woomera'</td>
</tr>
<tr>
<td>Long straight shapes: 'spear'</td>
<td>Plants: 'scrub nettle', 'flower'</td>
</tr>
<tr>
<td>Plants: 'vegetable food', various vegetable food types; 'ashes', 'ochre'</td>
<td>Plants: 'vegetable food', various vegetable food types</td>
</tr>
</tbody>
</table>

Table 6: Semantic composition of two noun classes in Wunambal (after Carr 2000, slightly revised)

Second, perhaps the majority of animates in Eastern and Western Worrorran languages are assigned to the MAS and FEM categories. There is a tendency in Worrorra for squat or rounded animals to be assigned to the FEM gender; this plausibly accounts for the association of NEUTn with birds.6

6 Another possible explanation is provided by Mark Harvey’s observation (Harvey 1997) that in many languages of northern Australia the MAS to FEM opposition for human beings is paralleled by an ground dwelling to air/tree/water dwelling opposition for other animates.
Third, associations between elongated objects and the AN class and concave objects and the NEUT\textsubscript{N} class are paralleled by the ‘extended’ vs. ‘rounded’ shape correlation in Worrorra (Clendon 2001:334).

Fourth, the sun is categorised as NEUT\textsubscript{N} in Northern Worrorran languages, but as feminine in Eastern and Western languages, as in many other gender-marking Australian languages. This has a mythological basis — the sun was a woman in the Dreamtime. Likewise, the moon is classified as MAS in Eastern and Western languages, and as AN in (northern) Wunambal (Capell 1941:306): the moon was a man in the Dreamtime. These two correlations can hardly be fortuitous.

In sum, there is too much in common between the AN and MAS classes and the NEUT\textsubscript{N} and FEM classes to be accidental or the result of areal diffusion of the idea of noun classes. Granted this, the HUM class of the Northern group evidently emerged simultaneously with the emptying of the AN and NEUT\textsubscript{N} classes of human nouns. This class is marked in Wunambal by the adjectival agreement prefixes \textit{b}\textsuperscript{i}- and the intransitive subject pronominal prefix \textit{b}\textsuperscript{V}-; the corresponding free pronoun is \textit{bini}. Where does this come from? Most likely it derives via backformation from the third person plural prefix, \textit{b}\textsuperscript{Vrra}-, which is a probable reflex of proto-Worrorran \textit{*birra} (McGregor & Rumsey forthcoming).

As observed already, there is a strong correlation in Worrorran languages (as there is cross-linguistically) between plural and human nouns: if any nouns show distinct treatment of singulars and plurals, this will be maintained for human nouns. Typically, in Worrorran languages the plural category is marked in agreement morphology for nouns with human referents. Given this strong correlation, and the recurrent -\textit{rrV}- of non-singular pronominal forms, it seems natural that \textit{rr} would be analysed as a plural marker. By backformation, the preceding segment \textit{b}\textsuperscript{V}- might be analysed as a human class marker. Additional support comes from the observation that in Northern Worrorran languages a similar process seems to have occurred in the second person, where the singular pronominal prefix to prefixing nouns appears as \textit{g}(V)-, a form quite unlike the forms in other Worrorran languages — \textit{ngun}- (Worrorra, Yawijibaya, Unggarrangu, Umiida) ~ \textit{nyun}- (Unggumi) ~ \textit{nyunga}- (Ngarinyin). Compare, however, the Northern Worrorran second person plural prefix \textit{grr}- (Carr 2000; McGregor & Rumsey forthcoming).

The pronominal prefixes to transitive verbs in Northern Worrorran languages are unlike those in the other two groups in having different forms according to whether a third person singular agent belongs to the human class or one of the other classes (Carr 2000; Capell & Coate 1984:233). In the Eastern and Western Worrorran groups the prefixes are identical for all classes of agent. It is possible that the distinct set of HUM agent prefixes in the Northern Worrorran languages also arose via reanalysis of the corresponding plural forms. Although there is some suggestive evidence in favour of this, many of the
forms are not amenable to a simple explanation by reanalysis; further work is
required to motivate this scenario.

One further observation is pertinent. The HUM classes in Northern
Worrorran languages is constituted primarily of nouns with human referents.
Unsurprisingly, animate beings are sometimes personified in myths, and
temporarily reassigned to the HUM class. More interestingly, the HUM class also
includes collective or mass nouns. Nouns normally assigned to other classes are
sometimes reassigned to the HUM class to indicate collectivity. In (9), mee
‘vegetable food’ normally goes in the AN class, while the other two nouns
denoting vegetables are NEUT\textsubscript{w}. But the demonstrative occurring with mee is the
HUM one, while the cross-referencing pronoun in the verb is the form for HUM
objects. Examples in Carr (2000) also illustrate reclassification of AN and NEUT\textsubscript{w}
class nouns to signal masses.

(9) mee binya: yarn\textsubscript{gu}: gumbia:, Wunambal
vegetable this kapok potato
\textit{garmmarn\textsubscript{gu}:, ngarrambu} \textit{an-birr-mra-ngi}
long: Yam root: food: type HUM.O-3PLS: grab- PA
“They got these foods: kapok roots/ shoots, potatoes, long yam, and
ngarrambu.” (Carr 2000:§3.1.1)

Significantly, in Ngarinyin a similar situation obtains: the plural category is
used as a type of collective or mass category for non-human nouns (Rumsey
1982:39):

(10) ngala nyindi ngala birri Ngarinyin
meat FEM: this meat they
“this meat” “a mass of meat, some meat”

Recall also the marginal fifth class in Worrorra signalled by the third person
plural prefixes, that consists of nouns specifying collectives. These nouns
usually have no corresponding singular class.

It may thus be hypothesised that the Wunambal plural category was
previously also a neuter collective category, but with the emergence of the
singular HUM category, neuter collectives previously in the plural class were
reassigned to HUM.

The final things to consider are the class-marking noun suffixes in Western
Worrorran languages, and/or the phonological correlates of the classes. The two
most obvious scenarios are: (a) that just Western Worrorran languages acquired
the affixes; and (b) that proto-Worrorran had class-marking on its nouns, but
this was subsequently lost in Eastern and Northern Worrorran daughter
languages.
My preference is for scenario (a). Although there are nouns in Eastern and Northern languages that show phonological indications of class membership, they are few in number. The suffixes in Unggumi are formally almost identical with the prefix forms; the most divergent form, the NEUT<sub>w</sub> suffix -ngga, may well be the result of a hardening process preventing the disappearance of the class-marker at the morpheme boundary, which would normally put the initial consonant at risk in intervocalic position. (Such one-off morphophonemic processes are attested in other languages of the region, e.g. Gooniyandi — McGregor 1990:103-104.) Perhaps previously determiners followed nouns, as in some modern languages (e.g. Ngarinyin — Rumsey 1982:136-138). If the determiners had an initial class-marking prefix as in the modern languages, it may have glommed onto the nominal, the remainder of the determiner ultimately being lost. Some support for this suggestion comes from Ngarinyin, where gender marking anaphoric pronominals are regularly cliticised to nouns with little apparent anaphoric function (I am grateful to Alan Rumsey for drawing my attention to this fact). Ultimately these elements could have become class-markers. The class-marking suffixes of some adjectives may have arisen by the same process. It is possible that this system of suffixes underwent levelling and attrition in Worrorra. Similarly, the few phonological correlates of classes on Ngarinyin nouns may be relics of an earlier cycle of grammaticalisation from enclitic determiner to suffixed noun class marker.

5. Conclusions

Noun class and classifier systems have been popular topics in Australianist linguistics, and a fair amount has been written about their origins and development. Noun class systems are widely believed to have grammaticalised from noun classifier systems in which a generic precedes a specific nominal. It is believed that generics grammaticalised first to classifiers, then to gender markers. This has sometimes been presumed to have happened in neighbouring languages through areal influence (Dixon 2002). While it is possible that (some) class-marking prefixes in Worrorran languages derive ultimately from generic nouns, I have not been able to reconstruct lexical sources for all of them. Indeed, some appear to have sources in pronominal elements.

Application of the comparative method permits reconstruction of a plausible gender system for proto-Worrorran, including a set of markers and a sequence of formal and structural changes that would result in the modern systems. Some of the Worrorran class-markers, as we have seen, bear formal similarities to class-makers in some other non-Pama-Nyungan languages. We cannot presume cognacy (indeed, anything) from formal similarities in an unprincipled sample of languages. To demonstrate historical relatedness of the class-markers requires reconstruction of plausible systems in proto-languages (which themselves remain highly speculative); whether these might be class-systems or classifier systems, or just words remains to be seen. What we need to
do at this point in time, I would argue, is to focus on low-level historical-comparative reconstruction. After this has been done is the time to address questions of higher-level relatedness.

References


HITTITE DUWĀN (PARĀ)

H. CRAIG MELCHERT
University of California, Los Angeles

I am pleased to offer the following lines in tribute to Harold Koch, with fond memories of the times shared in Room B of Widener Library at Harvard when we began graduate study of historical linguistics together nearly four decades ago.¹

The Hittite adverb duwān in those rare instances where it occurs alone is generally interpreted as ‘hither’ (Beckman 1996:157, Tischler 2001:182, Rieken 1999:69), and the equally rare combination duwān pari as ‘hitherto, heretofore’ (Friedrich 1952:231, Güterbock & Hoffner 1995:128-129 and Tischler loc. cit.). The far more frequently attested correlative duwān…duwān is typically glossed as ‘hither and thither’ (both Friedrich and Tischler, loc. cit.).

Nevertheless, the most popular etymology for duwān is that which derives it from PIE *dweh₂m (*[dwām] by ‘Stang’s Law’ as discussed below) and sees it as forming a direct word equation with Grk. ὅψις ‘for a long time, long ago’: see Tischler 1994:491-492 with references to Pisani, Schindler and others. The adverb duwān is further analyzed as a fossilised accusative singular of a root noun also seen in Hittite tūwa ‘(a)far’ and tūwaz ‘from afar’ (see e.g. Melchert 1984:30, following Schindler, and Rieken 1999:70). For correlated duwān…duwān there is a competing etymology based on PIE *two-…two- ‘the one…the other’ (Benveniste 1962:84-85), but see the serious reservations of Cardona (1987:5).²

Neither etymology explains the assigned synchronic meaning ‘hither’ for duwān alone. Furthermore, correlated duwān…duwān means simply ‘in one direction…in the other’ (Hoffner 1997:133). There is no evidence that in the correlated instances the first direction is near-deictic, indicating motion towards the speaker. It is thus far from clear that single duwān and correlated duwān…duwān are related. Finally, duwān (parā) should mean ‘for a long time; long ago’ by the first etymology. In what follows I will address these discrepancies as well as briefly review the formal problems of deriving Hittite tūwa and tūwaz from the same prehistoric paradigm as duwān, which have never been fully acknowledged.

I begin with the combination duwān parā. The overall usage of Hittite parā is compatible with either ‘heretofore’ or ‘long ago, for a long time’ for duwān parā.

¹ I presented a preliminary oral version of this analysis at the Twenty-Fifth East Coast Indo-European Conference at The Ohio State University in Columbus, Ohio, June 21, 2006. I am grateful to members of that audience, in particular to Jay Jasanoff, for helpful critical remarks. I, of course, remain solely responsible for the views expressed here.

² Tischler (1994:492) is wrong in implying that Cardona uncritically endorsed Benveniste’s comparison.
In both cases \textit{parā} would mean 'before (in time), previously', for which one may compare \textit{parā} \textit{hand(afetyara} ‘(divine) providence’ (actually ‘(divine) pre-ordination, preparation’, as per Puhvel 1991:105 contra Güterbock & Hoffner 1995:130) and \textit{pērān parā} ‘previously, beforehand’, with renewal by addition of the synchronic adverb for ‘before’. Compare also \textit{pēran parā} UD.KAM-an ‘the day before’ (thus contra Güterbock and Hoffner 1997:303).

The contexts of two of the attested instances of \textit{duwān parā} also permit either 'heretofore' or 'long ago, for a long time'. The first is found in an oracular inquiry (KUB 5.1 iii 53-54, NH/NS):

\begin{quote}
BAD-an=ma=mu=za duwān parā GIM-an SAG.KI-za ātina=ya QATAMMA kēdani=za kan Līl-ri kwutan imma kwutan neyānāri nu=mu pean hūyāši 'If you, Storm-god of Nerik, are my personal deity, as you (have) protected (me) for a long time/heretofore, will it now also be likewise? Wherever I turn on this campaign, will you run ahead of me?'
\end{quote}

The second appears in a prayer of Muršili II regarding the plague (KUB 14.14 Ro 37-39, NH/NS). I give in parentheses a translation of several preceding clauses in order to make clear the overall context: ('The land of Hatti prospered...You gods have now proceeded subsequently to take vengeance on my father for that affair of Tuthaliya the Younger. My father [-ed] on account of the blood of Tuthaliya. Whatever princes, lords, commanders, officers went over to the side of [my father] died from [that] affair. That same affair also reached the land of Hatti, and the land [of Hatti] began to die because of [that] affair.') nu KUR URUGIDRU-ti duwān parā duwān parā namma da[ššēš]ta KUR URUGIDRU-ti hinganaz [mekki?] dammešhāttat 'The land of Hatti [has] -ed for a long time/heretofore, but now the plague has become even worse. The land of Hatti has been [much] oppressed by the plague.'

The decisive third example of \textit{duwān parā} again comes from an oracular inquiry (KBo 2.2 iii 19-27, NH/NS): ('If you alone, Sun-goddess of Arinna, are angry because of the vows (made to obtain) offspring, and in addition no other deity is joined with you, let the signs be favorable: ...unfavorable. § Seeing that this was the outcome,') [DINGIR-LIM-ši duwān parā duwān parā \textit{arha} zaluki ta]

\begin{quote}
3 My textual citations follow the conventions of the Chicago Hittite Dictionary. The sigla OH, MH, and NH (Old Hittite, Middle Hittite, and New Hittite respectively), refer to the date of the composition of the texts, while OS, MS, and NS (Old Script, Middle Script, and New Script) indicate the relative age of the manuscripts. For the various text editions see Laroche (1976) or visit the online Konkordanz der hethitischen Texte created by Silvin Košak: http://www.hethport.uni-wuerzburg.de/hetkonk.

4 Likewise indeterminate is KUB 6.9+ i 12-13 (restored with Van den Hout 1998:16): \textit{[AN}A }"\textit{UTU-ši} kan kuit LUGAL-ua[nni aššūd duwān] parā arha zalukiša 'As to the fact that [for his Majesty] the a\textit{cession} to kingship has been heretofore/for a long time postponed.' We may safely assume that the important enthronement ritual usually took place soon after the de facto assumption of the kingship, so that a delay in the ceremony from one year to the next (see van den Hout 1981:279-283 and 1998:85-88) could well have been regarded by the king and his advisors as 'for a long time'.
Because I had offended the deity through hubris long ago, are you the deity angry on that account? Let the signs be unfavorable. If you the deity are angry only on that account, because I had offended you through hubris long ago…'.

As translated by Güterbock and Hoffner (1995:129), the clauses with *duwan parā* as ‘Because I [o]ffended the [dei]ty heretofore…because I have offended you heretofore.’ In so doing, they overlooked the past tense of the auxiliary verb *hark*-. Hittite is quite strict in its use of tense in the analytic perfect construction with the verbs *hark-* ‘to have, hold’ and *ēš- ‘to be’ plus the past participle. It without exception employs present tense of the auxiliary for the present perfect and the preterite for the past perfect (see Hoffner & Melchert forthcoming §§22.22-23). Since the use of ‘heretofore’ implies that a past action is described with reference to the present time, it is only compatible with a present perfect, for which we would expect *šallakartan ḫarmī ‘I have offended* with present tense of the auxiliary. The use of the preterite *harkun* shows that the action of the clause is being described with reference to another past action, and the preceding context makes it clear that that other action was: namely, the vows made to the Sun-goddess to obtain offspring. Having learned that the latter action is one cause of divine anger, the inquiry then asks whether it is the only one. Informed that it is not, the inquiry then proceeds to investigate other possible earlier causes. That one of them might be characterised as having happened ‘long ago’ is entirely reasonable. The interpretation ‘for a long time, long ago’ thus seems justified for the combination *duwān parā*.

Evidence for *duwān* alone is limited to four assured examples. The first occurs in a text of Muršili II regarding arbitration of disputes in Syria (KBo 3.3 iii 29-33, NH/NS): (‘But if there is some legal dispute, let the Priest intervene in disputes among you, and let him investigate the disputes for you.’) *mān DINU=ma kuitki šallešzī n=at arḫa ēppūwanzi UL tarāłtēni n=at kan duwān MAḪAR 4 UTU-ŠI parā naišten n=at 4 UTU-ŠI arḫa ēpzi ‘But if some legal dispute becomes too big, and you cannot dispose of it, send it duwān before His Majesty, and His Majesty will dispose of it.’ Beckman (1996:157) translates *duwān* as ‘here’, following the conventional interpretation. The context obviously is consistent with such a sense, but by no means imposes it. One wonders why such a specification of place or direction is needed, since the location of the Hittite king is hardly in doubt. We might also expect the normal word for ‘here, hither’ in Hittite, which is *kā*, though obviously the existence of another term for this notion is possible.

The other three previously overlooked examples of *duwān* alone provide us with the crucial evidence to determine its true meaning. All of these appear in a...
single text, one of the Middle Hittite Maṣat letters. The first two occur in the same passage and must be treated together, with the preceding and intervening context (HKM 66: 23-25 & 31-32, MH/MS): ('Also the servant of Saparta whom I sent to the Gasga country (with the words): 'Let him proceed to find the son of Saparta!' n=a[\ldots]m\[\ldots]n k[a]r\[\ldots]\ pānza n-an=m[u]=kan duwān parā nai 'If he has/is already gone, send him to me duwān.' ('If he hasn't gone yet, write to Lullu and Zuwanna. Let one take three Gasga men in his place. But let him go and find the son of Saparta.)') namma=as mah\[\ldots]EGIR-pa paizzi n-an=mu=kan duwān parā nai 'Then when he returns, send him to me duwān.'

The third instance of duwān comes at the very end of the letter and deals with a completely different topic (HKM 66: LeftEdge 1-5):

nu ammuk duwān h\[\ldots\]a\[\ldots\]k\[\ldots\]treškatten "\[\ldots\]Tahazzilim=n a kuit walh[e]r\[\ldots\] nu=\[\ldots\]\ kāša L[U] TEMI awan a\[\ldots\] et SIG5-anza=wa-e[a n]u=\[\ldots\]=\[\ldots\]kan lē kuwataqa lahlahjiyaši 'You (pl.) kept writing to me duwān that [they?] had also beaten/struck Tahazzili. A messenger has just come from him: 'I am well.' Do not worry at all about him.'

The use with an imperative excludes a meaning 'long ago, for a long time' for duwān alone, while the fact that the addressee is ordered to send the servant duwān in the event that he has already gone likewise eliminates 'hither, here'. The addressee cannot reasonably be ordered to send the servant anywhere if he has already departed on his previously assigned mission. The only plausible sense I can see for duwān in this context is 'later, subsequently'. The addressee is to send the servant only later, after he has returned from his mission. If the man has already left, then there is nothing more to be done. If he has not yet left, the addressee is to take the opportunity before he leaves to secure three men to carry out his regular duties in his absence.

A sense 'later, subsequently' also fits the first example cited above from KBo 3.3 iii 29-33: 'But if some legal dispute becomes too big, and you cannot dispose of it, send it subsequently before His Majesty, and His Majesty will dispose of it.' The meaning 'later' will not work for the last example cited with a preterite verb, but here we may assume rather 'late(ly), (of) late': 'You (pl.) kept writing to me lately that [they?] had also beaten/struck Tahazzili. A messenger has just come from him: 'I am well.' Do not worry at all about him.' Thus Hittite duwān

---

6 It is crucial to note that in this passage the adverb parā is functioning as a preverb with the imperative verb nai, the combination of which means 'to send' (nai alone would mean 'turn!'), and thus parā does not form a syntagm with duwān. Confusion on this point long prevented me from understanding the passage correctly and may have caused others to overlook the evidence here for duwān alone.


8 In purely formal terms the verb h[a]treškatten of the first clause could also be interpreted as imperative second plural: 'Keep writing to me...', but the following context argues strongly for a preterite second plural.

9 The adverb 'late' alone is attested well into the 19th century in English in the sense 'lately, of late'. Since Hittite has no morphological marker of the comparative, we would not expect any difference between 'of late' and 'later'.
'late' covers the range of English 'later' and 'lately', having the former value with a present-future or imperative verb, and the latter with a preterite.

Reconciling the difference between 'late' for duwān alone and 'long ago, for a long time' for duwān parā is not difficult. One may compare Latin sērō 'late' which forms a direct word equation with Old Irish sīr 'long(-lasting)' and Welsh hir 'long' (spatial and temporal) < PIE *seh₁-ro-. For the Hittite I assume preservation of an older meaning 'long' (of time) in the fixed phrase duwān parā (note again the Greek cognate διήν) and a shift to 'late(r)' in simple duwān.

The attested meaning of the well documented correlated pair duwān...duwān clearly is 'in one direction...in the other', and in some cases it appears to be equivalent to kēt...kētt—the 'to/on the one side ...to/on the other side'. The latter sense is typical in ritual contexts, as in the following (KBo 26 Ro 23-24, OH/OS):

```
pēdešmi-pat tuwān 1-ŠU waḥanzi 10.MES ALAM.ZU[...tuwa]ann=a 1-ŠU waḥanz[i] 'The [ ] turn once in place in one direction. The 'performers' [ ] turn once in the other direction.' The cited example with its specification of 'in place' makes it clear that the synchronic sense of correlated duwān ...duwān contains no inherent element of distance.
```

However, some examples do incorporate the notion of movement across a distance. I cite first a famous example from the Hittite Laws (KBo 6.26 i 35-38, OH/NS): ('If someone sows seed upon (another’s) seed, his neck shall ascend the distance. I cite first a famous example from the Hittite Laws (KBo 6.26 i 35-38, OH/NS): 10 [ ]

```
SIMDI GU₃.H.1 A turiyanzi kēl mene-šīt duwān kēll=aa mene-šīt duwān nēyanzi LŪ-eš aki GU₃.H=ya akkanzi 'They shall hitch up two teams of oxen. They shall turn the face of one in one direction and of the other in the other direction. The man shall be killed, and the oxen shall be killed.' While the text does not spell out the gory details, it is clear that the oxen are to be driven far enough in opposite directions to pull apart and kill the perpetrator.
```

In two other instances the distances involved are specified. 11 The first is from a late version of the Hittite Laws (KBo 6.4 i 11-13, NH/NS): takku ÜL=ma A.GAR dammel pēdan duwan 3 DANNA duwan=a 3 DANNA nu=kan kuš kuš URU-aš anda SīxSĀ-ri... 'But if it is not a cultivated land, but a virgin place, (they measure) three DANNA’s in each direction, and whatever city is determined (to be) within (that radius)...'. The second appears in the Myth of Appu (KBo 19.108: 10, OH/NS): [...duwan IKU-an SUD-at duwānm=a IKU-an SUD-a[t]] '...drew [one IKU in one direction] and drew one IKU in the other direction.'

I see no way to exclude entirely derivation of correlated duwān... duwān from a virtual *tweh₃₃,m...tweh₃₃₃₃ to the 'oppositional' stem *two- à la Benveniste, thus assuming a completely different historical source for this expression from that

---

10 Compare also instances with papparā- 'to sprinkle (liquid)' in KUB 12.40 ii 7-8 (NS), with īšuwa- 'to scatter, sprinkle (dry material)' in KUB 10.72 ii 23-24 (NS), and with lājuwa- 'to pour' in KUB 41.33 Ro 14 (MH/NS).

11 For an estimated value of fifteen hundred meters for the Hittite DANNA and fifteen meters for the IKU see Melchert (1980). Compare also van den Hout (1990:517-521).

12 See also the very similar example with terippiya- 'to plow' in KUB 13.1 iv 21 & 23 (MH/MS).
for duwān (parā). However, as already pointed out by Cardona (1987:5), there is not only no other evidence for a stem *two- in Anatolian, but also none for an accented pronominal stem *to- on which the stem *t-wo- was probably built (for this analysis see the arguments of Cardona 1987). It thus seems far more economical to suppose that the correlated use of duwān...duwān began in examples where the original meaning was ‘for a distance... for a distance’. Reanalysis to ‘in one direction...in the other direction’ led to a ‘semantic bleaching’ whereby the expression could be used merely for ‘on/to the one side...on/to the other side’.

According to this scenario, the accusative singular *dweh₂m of a PIE root noun ‘distance’ was used as an ‘accusative of extent’ both spatially and temporally. The meaning *‘for a distance’ is attested indirectly in Hittite correlated duwān... duwān, while the sense of ‘for a long time, long ago’ is reflected directly in Greek ὑπὸ and Hittite duwān parā. Hittite duwān ‘late(r), (of) late’ represents a secondary development of the temporal use.

The attested usage of Hittite duwān therefore presents no obstacles in semantic terms to its derivation from PIE *dweh₂m and its equation with Greek ὑπὸ, as widely assumed. The derivation of the Hittite also faces no formal problems, if one accepts the minor PIE synchronic rule known as ‘Stang’s Law’, by which word-final *-eh₂m was realised as phonetic *[−m] (see Meier-Brügger 2003:97 with references). Indeed, the extraparadigmatic status and hence likely early morphological isolation of Hittite duwān makes it one of the strongest pieces of evidence for this particular application of ‘Stang’s Law’.

The derivation of Hittite tūwa ‘far’ and tūwaz ‘from a distance’ from the same root noun paradigm as that of duwān is far less straightforward, despite the claims in Melchert (1984:30) and Rieken (1999:69-70). Any analysis of these forms must begin with the attested orthographic facts. The adverb duwān is spelled overwhelmingly with initial tu-wa- or du-wa-: tu-wa-a-an (12x), tu-wa-an (2x), du-wa-a-an (19x), du-wa-an (20x) vs. only three instances of tu/du-u-wa-. On the other hand, tūwa appears at least a dozen times always as tu-u-wa, and tūwaz likewise shows overwhelmingly tu-u-(wa)-: tu-u-wa-az (at least 33x), tu-u-wa-za (9x), tu-u-az (4x). In view of these figures we may safely regard the hapax tu-az (KUB 25.36 v 4, MS) as an error.

There can be no doubt that the pattern just described reflects a real linguistic contrast between duwān on the one hand and tūwa/ tūwaz on the

---

13 I reconstruct a virtual *tweh₂m ...tweh₂m because *twom would surely have led to Hittite *tān (see Melchert 1994:128 with references).
14 I am much indebted to Professor Theo van den Hout for sharing with me the files of the Chicago Hittite Dictionary on duwān, which considerably amplified my own data. The three exceptions are du-u-wa-an in KBo 4.4 i 38 (NH/NS), and two instances of tu-u-wa-an (KBo 25.3 RICol 6’, OH/OS’ and KUB 23.34.5, MH/NS). The spelling with initial tu- is clearly the older practice, gradually replaced by du-, which alone is attested in New Hittite compositions. Since voiced stops are devoiced word-initially in Hittite (see Melchert 1994:18-20), the spelling of the initial stop as tu- or du- has no functional significance.
other. The claim of Rieken (1999:70) that both spellings represent [twa-] is not remotely credible. If this were the case, we would expect the distribution of tu/du-wa- and tu/du-u-(wa)- to be roughly the same in all three words. We must conclude rather that duwān spelled with tu/du-wa- does have the shape [twa:n], as we would expect from a preform *dwām, while the spelling tu-u-(wa)- shows that the other two must be disyllabic [tu:wa] and [tu:wats].

How we should account for this discrepancy remains unclear. One approach is to attribute it to the effects of the PIE synchronic rule known as ‘Lindeman’s Law’. By this rule all monosyllabic forms of the original paradigm of the putative root noun for ‘distance’ would have had beside them disyllabic variants: nom. sg. *d(u)weh₂s, acc. sg. *d(u)weh₂m (phonetically *[d(u)wām] as per above), and arguably an endingless locative *d(u)weh₂. If both variants persisted into pre-Hittite, we may assume that the *dw- of the shorter variants blocked the regular dissimilation *duw- > *dum- in the longer (see Melchert 1984:23-27). If we further assume that by this stage there was no longer any unitary paradigm, but merely two isolated adverbs *d(u)wām and *d(u)wā, nothing precludes that the former generalised the monosyllabic variant and the latter the disyllabic. The latter then developed regularly into tüwa, and the ablative tüwaz was formed secondarily on the basis of the presumed endingless locative.

If *dweh₂- ‘distance’ continued to form a unitary paradigm into the prehistory of Hittite, it seems to me ad hoc to suppose that the accusative singular generalised the monosyllabic variant *dwām, while the presumed endingless locative generalised disyllabic *duwā. There is an alternative. If we consider again the pre-Hittite paradigm of *dweh₂-, we may reasonably assume that the nominative singular *dweh₂s became pre-Hittite *twās, either by regular phonological change or by reshaping from *twaḥs after the accusative singular *twaḥm from PIE *dwām. The regular outcome of the weak stem *duh₂- in Hittite would have been *tuh₂- (cf. tuḥḥi- ‘cough, gasp’ < *dhuh₂-). Thus genitive singular *tuḥḥaš, dative-locative *tuḥḥi, allative *tuḥḥa, ablative *tuḥḥaz. That such an irregular paradigm would have been reshaped seems clear. The question is: just how? A complete levelling after the strong cases would certainly have been possible: hence gen. sg. *twaḥ, dat.-loc. *twi, and so forth. However, monosyllabic forms in the non-direct cases of a substantive

---

15 For further examples of initial Cu-u-(wa)- as a spelling for [CuwV-] see Melchert (1984:28).
16 For the PIE phenomenon known as ‘Lindeman’s Law’ see among others Meier-Brügger (2003:91) with references. For this basic approach to the problem of tüwa and tüwaz see already Melchert (1984:30), following Jochem Schindler, but without full recognition of the complications involved.
17 Since tüwa ‘far’ is attested in the sense of both an allative and a locative, it may plausibly be derived in functional terms from either a locative or an allative. For other Hittite adverbs in -a showing synchronically both functions compare kā ‘here, hither’ and apiya ‘there, thither’.
18 The apparent retraction of the accent in tüwa < *duwā is reminiscent of that in CLuvian pīya- ‘give’ or tüwa- ‘put, place’ (see Melchert 1994:89), however these are to be explained.
would have been highly unusual in Hittite. I submit then that a direct replacement of *tuh- by *tu- as the stem in these cases is equally plausible. The insertion of a hiatus-filling -w- would then have been entirely regular (see Melchert 1984:28-29).

A choice between the alternatives just presented depends crucially on whether or not a unitary paradigm of *dweh- ‘distance’ persisted into the prehistory of Hittite. I see at present no basis for deciding this question and thus for a reasoned choice between them.19

References


---

19 Rieken (1999:70 with note 332) objects on phonological grounds to the alternative derivation by Eichner (1978:160) of *tuwa* and *twaz* from an adjective *duh₂-yo- (endorsed in Melchert 1994:128). Her objection may or may not be valid. For the loss of prehistoric *y* in Hittite in a sequence *-Vh₂yV- one may compare present third plurals in *-nzi < *-eh₂yenti (Melchert 1994:130). On the other hand, the shape *duh₂-yo- is quite close to that of *teh₂-yé/ó ‘to steal’, which does appear as Hittite *týew- with preserved *y-, as stressed by Rieken. In any case, reconstruction of an adjective stem *duh₂-yo- ‘far, distant’ just to provide the basis for the two adverbs *tuwa* and *tuwaz* seems at best uneconomical, if not ad hoc.


1. Introduction

Australianist literature has often presented morphological ‘reconstructions’ posited through the comparison of paradigms across sets of languages whose forms are assumed to have a shared phonetic basis. The resulting ‘proto-paradigms’ have then been used to argue for the genetic relationship of the languages involved (see e.g. Dixon 1980, Blake 1988, Harvey 2003, Green 2003). This practice will be referred to in this paper as ‘inspectional’ morphological reconstruction. The use of the word ‘inspectional’ is not to be taken literally, hence the inverted commas. In doing ‘inspectional’ reconstruction practitioners do not carry out a simple inspection, but consider phonetic factors, plausible changes etc., that make their assumption of shared phonetic basis more or less plausible. Nevertheless, this is not reconstruction in the traditional sense. The Comparative Method requires that morphemes contributing to a reconstruction (grammatical or not) follow regular patterns of sound correspondence, since it is their adherence to this pattern that establishes their cognacy. This point is spelled out very clearly in Koch’s general procedures for morphological reconstruction (Koch 1996:220-221), the first of which is reproduced below:

Match tentative morphs, that is, formal bits that are potentially cognate according to established phoneme correspondences and changes. These matches may be found in the same language (so we have internal reconstruction), in different but related languages (so we have comparative reconstruction), or in different but not necessarily related languages (so we have an analysis of borrowing). Note that this assumes the prior establishment of phonological changes on the basis of lexical cognates. (emphasis mine)

Works such as Koch (1996) have greatly increased awareness of methodological issues among Australianists, but the type of non-standard reconstruction described above continues to be present in the literature. The starting point of this paper is that there has been insufficient critical examination of the validity of these methods and the resulting ‘reconstructions’, and that this is necessary for a proper evaluation. This paper suggests that there are two key reasons for the development and persistence of these practices. Firstly, the nature of the patterns of similarity among Australian languages is...

1 My interest in methodological issues in historical linguistics was first sparked by Harold Koch’s teaching. I am therefore delighted to have the opportunity of presenting a paper on such a topic in this book in his honour. I wish to thank Alan Dench, Mark Ellison, John Henderson, Kim Schulte and my co-editors for reading and commenting on this paper. All faults remain my own.
unusual (section 2). Secondly, the privileged status enjoyed by non-standard morphological reconstruction compared to other non-standard practices is due, at least in part, to its presumed association with the views of Indo-Europeanist and theorist Antoine Meillet (section 3).

2. **Australian ‘unorthodoxy’ as a consequence of the data**

Australian comparative data displays an ‘unusual’ pattern of similarity. Firstly, morphemes that are potentially cognate are very similar in phonological form. See, for example, the set of forms below, taken from languages ranging from Cape York Peninsula in the north-east, to central Australia, to the Pilbara region of north-western Australia, all associated with the meaning ‘vegetable food’ (data from Alpher 2004):

<table>
<thead>
<tr>
<th>Language</th>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uradhi (Atampaya)</td>
<td>mayi</td>
<td>“vegetable food”</td>
</tr>
<tr>
<td>Linngithigh</td>
<td>ayi</td>
<td>“vegetable food”</td>
</tr>
<tr>
<td>Wik-Mungknh</td>
<td>may</td>
<td>“vegetable food”</td>
</tr>
<tr>
<td>Umpila</td>
<td>mayi</td>
<td>“vegetable food”</td>
</tr>
<tr>
<td>Aghu-Tharrnggala</td>
<td>yi</td>
<td>“vegetable food”</td>
</tr>
<tr>
<td>Djabugay</td>
<td>ma:</td>
<td>“vegetable food, food other than meat”</td>
</tr>
<tr>
<td>Yidiny</td>
<td>mayi</td>
<td>“vegetable food, food other than meat”</td>
</tr>
<tr>
<td>Bidyara</td>
<td>mayi</td>
<td>“tucker, fruit”</td>
</tr>
<tr>
<td>Atnyamathanha</td>
<td>mayi</td>
<td>“food (vegetable)”</td>
</tr>
<tr>
<td>Karlamayi</td>
<td>mayi</td>
<td>“vegetable food”</td>
</tr>
<tr>
<td>Pitjantjatjara-Yankunytjatjara</td>
<td>mayi</td>
<td>“food from plants (as opposed to meat)”</td>
</tr>
<tr>
<td>Warlpiri</td>
<td>miyi</td>
<td>“vegetable food”</td>
</tr>
<tr>
<td>Walmajarri</td>
<td>miyi</td>
<td>“vegetable food”</td>
</tr>
<tr>
<td>Nyangumarta</td>
<td>mayi</td>
<td>“vegetable food”</td>
</tr>
</tbody>
</table>

Although this kind of similarity is sometimes encountered in well established language families – e.g. Austronesian *mata ‘eye’ and *lima ‘five’ are retained relatively unchanged across the Austronesian world because they involve stable segments – in the Australian case, it is the vast majority of potential cognates that are identical or near-identical. Despite this high degree of formal similarity, there are, however, fewer potential cognates than might be expected at both the wide-scale comparative level exemplified above, as well as at lower levels. O’Grady, Voegelin and Voegelin (1966) write, for example, that the cognate density that is found between dialects of the same language in some Australian cases is the same as what would be found between different languages within a language family in other parts of the world:

> Something special is going on in Australia. To draw attention to this we use the term FAMILY-LIKE LANGUAGE for Australian languages having neighbor intelligibility but otherwise having as low a cognate density as exists between languages in the usual language families of the world ... it would be misleading to use the simple term ‘language’ for a speech community whose cognate density lies in the middle range of what characterizes a ‘language family’ for the rest of the world (O’Grady, Voegelin & Voegelin 1966:11)
The high degree of similarity in phonological form means that regular sound correspondences are very frequently between identical morphemes, so that these correspondences do not provide the evidence of systematic phonological change that would, by the logic of the comparative method, distinguish inheritance from borrowing. Capell (1956:83) states, for example:

> Usually in the Australian field words are either fairly obviously cognate as between languages, or equally obviously non-cognate. There has not appeared the same necessity of establishing sound laws to prove connections ... ²

The low density of cognates then further contributes to ‘demoting’ the importance of regular sound correspondences for two reasons. Firstly, because the low number of potential cognates makes recognising any ‘patterns’ of correspondence difficult and, secondly, because rival correspondences often all occur in one or two items, making it almost impossible to decide which of the ‘patterns’ reflects inheritance and which borrowing – which are the regularities and which are the exceptions – as Blake (1988:4) describes in the quote below:

> In Australia ... it [is] hard to find sufficient examples to establish a correspondence with a proper account of the conditioning involved and then to be able to form a hypothesis about exceptions being borrowings.

It is not difficult to understand how these factors have resulted in a tradition that favours ‘inspectional’ reconstruction. In some cases, we find a potential lexical cognacy rate of well below 10%, but grammatical morphemes are so similar that the ‘inspectional’ reconstruction of paradigms is possible, as in Harvey (2003) where he presents a reconstruction of pronominal prefixes for the northern Australian languages. The forms on which he bases the reconstruction of the 1minimal pronoun *nga- (Harvey 2003:490) are given in Table 1 overleaf.

We can contrast this comparative data to a data set from the region of New England in northern New South Wales (Table 2 overleaf), where the potential cognate sets do provide evidence of systematic phonological change and therefore constitute reconstructions in the traditional sense. The data is from Crowley (1976:35).

² This quote also shows that, from very early on, there has been a clear bias towards inheritance as a default explanation of similarity between Australian languages (see Evans 2005 for a more recent example). This bias probably stems from a lack of non-linguistic clues to suggest otherwise – there is no clear evidence of multiple migrations to the Australian continent and a certain degree of cultural uniformity is observable. Evans (2005) argues that in his recent work Dixon (e.g. Dixon 2002) has taken the opposite position and treats diffusion as the default. This interpretation is arguable, but, nevertheless, both positions are wrong: there is no default.
Table 1: Minimal comparative data – northern Australian languages

<table>
<thead>
<tr>
<th>Gloss</th>
<th>Nganjaywana</th>
<th>Djangadi</th>
<th>Yugambal</th>
<th>Other Languages</th>
<th>Reconstruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘two’</td>
<td>dwala</td>
<td>buta</td>
<td>buda</td>
<td>bula</td>
<td>*buLal</td>
</tr>
<tr>
<td>‘magpie’</td>
<td>mbuda</td>
<td>nambul</td>
<td>–</td>
<td>ya:mbul</td>
<td>*pambul.L</td>
</tr>
<tr>
<td>‘hill’</td>
<td>guuda</td>
<td>bugul</td>
<td>–</td>
<td>–</td>
<td>*bugul</td>
</tr>
<tr>
<td>‘fool’</td>
<td>ngada</td>
<td>mangal</td>
<td>–</td>
<td>–</td>
<td>*mangal</td>
</tr>
<tr>
<td>‘father’</td>
<td>bida</td>
<td>–</td>
<td>babil</td>
<td>–</td>
<td>*babil</td>
</tr>
<tr>
<td>‘food’</td>
<td>gyada</td>
<td>–</td>
<td>–</td>
<td>wigal</td>
<td>*wigal</td>
</tr>
<tr>
<td>‘cry’</td>
<td>dwa</td>
<td>wuta</td>
<td>–</td>
<td>–</td>
<td>*wuLa</td>
</tr>
<tr>
<td>‘dog’</td>
<td>iقرأyayga</td>
<td>–</td>
<td>–</td>
<td>migila</td>
<td>*migiLanj</td>
</tr>
<tr>
<td>‘cod’</td>
<td>ruda</td>
<td>–</td>
<td>gurul</td>
<td>–</td>
<td>*gurul.L</td>
</tr>
<tr>
<td>‘urinate’</td>
<td>–</td>
<td>gita</td>
<td>–</td>
<td>gila</td>
<td>*giliA</td>
</tr>
<tr>
<td>‘lip’</td>
<td>–</td>
<td>witiŋ</td>
<td>–</td>
<td>witiŋ</td>
<td>*wiliŋ</td>
</tr>
<tr>
<td>‘we’</td>
<td>di-</td>
<td>nati</td>
<td>nadi</td>
<td>nali</td>
<td>*nali</td>
</tr>
</tbody>
</table>

Table 2: New England comparative data

Though many Australianists have responded to the unusual pattern of similarity by adopting a non-standard strategy of reconstruction, Dixon has recently claimed that the Comparative Method is just not applicable to Australian languages. In the preface to his 2002 book, for example, he states the following:

... I had assumed that the methodology which applies so well for the languages of Europe and North America and Oceania would also be appropriate for the linguistic situation in Australia. It is not, but it took me a long time to realise this. (Dixon 2002:xvii)

The problem, however, is not one of applicability. As Hoenigswald (1990:382) writes, “... there are degrees, not of amenability to what we call technically the “comparative” method ... but of yield”. That is, the Comparative Method is always applicable, but it does not give equally conclusive results every time it is applied - the ‘yield’ being dependent on the nature of the data:

The comparative method and its adjuncts are in principle general rather than language-specific ... [but] [i]f the descendant languages are too “similar”, reconstruction is uninteresting ... [i]f the languages have come to differ typologically
from one another or from the proto-language, the reconstruction gains in interest. (Hoenigswald 1990:382)

Unfortunately, in the Australian case, due to the pattern of similarity described, the yield is quite low. The Comparative Method is only able to conclusively establish the genetic relationship of languages and aspects of their subsequent history when the languages display enough non-identical systematic sound correspondences. Identical sound correspondences, even when systematic, fail to give a conclusive answer on the question of whether or not the independent systems compared derive from the same ancestor system – in the usual sense of genetic relationship. Nevertheless, an inconclusive answer is an answer. The Comparative Method therefore always reveals something, even if not necessarily what was hoped for.

3. The status of ‘inspectional’ morphological reconstruction

As mentioned in the introduction, ‘inspectional’ morphological reconstruction is not reconstruction via the Comparative Method, but its practice enjoys a privileged status in comparison to other non-standard methods due to its association with Antoine Meillet and his supposed preference for irregular morphology as evidence in genetic arguments. So much is made of irregular morphology that it is almost forgotten that the practice of reconstruction, without established sound correspondences, is actually non-standard – be it applied to lexical or grammatical morphemes. For example, Evans’ evaluation (2005:257) of a proposed relationship between Murrinh-Patha and Ngan.gi-tyemerri (two neighbouring languages of Northern Australia) illustrates both the importance given to aberrant/irregular morphology and the fact that it is considered to be clear evidence of genetic relationship:

Green (2003) has recently made a convincing case for their relatedness based on far-reaching similarities across dozens of irregular prefixal paradigms, but this is not matched by significant similarities elsewhere in their grammar or lexicon ...

(emphasis mine)

In an unpublished manuscript, Harrison convincingly demonstrates that Meillet’s views on genetic relationship have been misinterpreted, and that at the basis of the misinterpretation is the treatment of Meillet’s notion of fait particulier – literally ‘particular fact’ – as an equivalent to the modern notion of ‘shared aberrancy’, as is seen in the following description from Campbell (2003b:69):

---

3 See Campbell (2003a) for a description and discussion of other non-standard methods and practices.
Meillet favoured ‘particular processes’, ‘singular facts’, ‘local morphological peculiarities’, ‘anomalous forms’, and ‘arbitrary’ associations (i.e. “shared aberrancy”). (emphasis mine)

Harrison argues that Meillet’s ‘particular facts’ must have encompassed a much broader range of linguistic evidence than is covered by the notion of ‘shared aberrancy’. That this must be the case is obvious from a closer reading of Meillet’s statements on the importance of ‘singular facts’. His examples have been cited repeatedly to illustrate his presumed preference for ‘shared aberrancy’, but a different picture emerges when they are seen in the actual context of his overall discussion. Harrison (section 4.4) writes the following about Meillet’s most often cited example of a ‘singular fact’, the aberrant 3rd person of the verb ‘to be’ in Indo-European languages, reproduced below (e-grade stem in the singular, ø-grade in the plural):

<table>
<thead>
<tr>
<th></th>
<th>3s</th>
<th>3p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sanskrit</td>
<td>ásti</td>
<td>sánti</td>
</tr>
<tr>
<td>Latin</td>
<td>est</td>
<td>sunt</td>
</tr>
<tr>
<td>Gothic</td>
<td>ist</td>
<td>sind</td>
</tr>
</tbody>
</table>

From the context in which Meillet introduces his example in Sur la Methode de la Grammaire Comparée ... we are led to conclude that the singularity or particularity of his faits particuliers is not their irregularity but their arbitrariness. Meillet was not opposing regular and irregular, but arbitrary (in Peircean terms, symbolic) and general (schematic, typological, natural, or iconic) ... It is not shared anomalies that demonstrate genetic relatedness, but the shared (arbitrary) symbols that are the common currency of the standard comparative method.

That is, in advocating the importance of ‘singular facts’ Meillet was warning against the use of mere resemblances (lexical look-alikes) and typological similarities, and his preference for ‘singular facts’ within the realm of morphology, was exactly because it was ‘morphology’, whether or not it then also displayed aberrancy.

Harrison points out that, because examples such as the one above involve both ‘morphological similarity’ and ‘shared aberrancy’, the distinctness of the two notions is often overshadowed, as is the fact that they contribute differently to genetic arguments. ‘Shared aberrancy’ is valued as against the possibility of chance similarity, the idea being that an unusual pattern is unlikely to have arisen independently more than once; ‘Morphological similarity’ is instead valued against the possibility of borrowing, grammatical morphemes typically being thought of as more resistant to borrowing than lexical morphemes. Harrison (section 3) writes:

... what Meillet found significant about morphology was that it was resistant to borrowing. Meillet believed morphological evidence to be a privileged sort of evidence for genetic relatedness, but only insofar as it meets the other requirements of the comparative method, including the demonstration of regular sound correspondences.
Being well aware of the fact that regular sound correspondences can arise as a result of heavy borrowing as well as from inheritance, Meillet believed that if the forms relatable through sound correspondences included the grammatical morphemes, there was no doubt on the question of genetic relationship. Although the idea that morphology is resistant to borrowing is still considered to be true in relative terms, our understanding of language contact and borrowing has greatly expanded since Meillet’s time. The borrowing of morphology is no longer thought to be as rare as once believed (see, for example, works such as Thomason and Kaufman 1988) and the reliability of assumptions regarding morphology and borrowing needs to be carefully reassessed.

The following quote, where Meillet himself describes his overall position on the issue of genetic relationship, should leave no doubt as to whether he would have considered ‘inspectional’ morphological reconstruction and aberrant morphology adequate ‘proof’ of genetic relationship:

Everywhere where the phonetic system and the grammatical system show precise correspondences, where regular correspondences permit the recognition of the single origin of the words and of the phonetic system and where the system of grammatical forms is explained from a point of common origin, [genetic] relatedness is obvious. (Meillet 1965:88, translation by Harrison)4

Any association of ‘inspectional’ morphological reconstruction with the views of Antoine Meillet is therefore not justifiable.

4 “Partout où le système phonétique et le système grammatical présentent des concordances précises, où des correspondences régulières permettent de reconnaître l’unité d’origine des mots et du système phonétique et où le système de formes grammaticales s’explique en partant d’un original commune, la parenté est evidente.”

4. Concluding remarks

This paper has attempted to clarify the difference between ‘inspectional’ reconstruction and reconstruction via the Comparative Method. In reconstruction via the Comparative Method, a proto-morpheme is reconstructed only when regular sound correspondences (the majority of which are usually non-identical) exist across the forms compared, or where plausible departures (via analogy) can be identified. This contrasts with the practice of ‘inspectional’ reconstruction, where the forms are simply assumed to be relatable because of their similarity and general knowledge of phonetic factors and plausible sound changes. As already mentioned, the difference between the two types of reconstruction has become blurred in part due to a misunderstanding of what constitutes a ‘singular fact’ in the historical
linguistics of Antoine Meillet. The lack of a clear sense of two distinct practices, has resulted in ‘inspectional’ reconstructions rarely being explicitly presented for what they are, and this, though not necessarily intentional, is potentially misleading.

The paper has also argued that Australian comparative data displays a pattern of similarity that is in part responsible for the development and persistence of non-standard methods. However, the fact that ‘inspectional’ reconstruction lacks the same built-in evidential strength of the Comparative Method cannot be forgotten or ignored. It would be desirable to develop a more principled way of assessing ‘inspectional’ reconstructions using quantitative probabilistic analysis for example, rather than relying on general assumptions of likelihood. This would strengthen their value, but they may never reach the level of validity of reconstructions that meet the logic of the Comparative Method.

References


Harvey, Mark 2003. “Reconstruction of pronominals among the non-Pama-Nyungan Languages”. The non-Pama-Nyungan languages of northern Australia: Comparative studies of the continent’s most linguistically complex region ed. by Nicholas Evans, 475-513. Canberra: Pacific linguistics.


WARLPIRI VERB ROOTS IN COMPARATIVE PERSPECTIVE

DAVID NASH

Australian National University and
Australian Institute of Aboriginal and Torres Strait Islander Studies

1. Introduction

When it comes to verb roots, Warlpiri is the odd language of the Ngumpin-Yapa subgroup (McConvell and Laughren 2002) – compared to the other half dozen, it has twice as many simple verb roots. Warlpiri has over a hundred monomorphemic verb roots whereas Western Ngumpin (Walmajarri and Jaru) have around forty, and Eastern Ngumpin (such as Gurindji and Mudburra) 30-35, and only 44 have been recorded in Warlmanpa (Warlpiri’s congener Yapa language).

The Ngumpin-Yapa languages each express much the same range of meanings by lexical verbs (simple and compound), so languages with fewer monomorphemic verbs employ a greater number of coverb-verb combinations. This paper focuses on the monomorphemic verb roots; the morphosyntactic status of coverbs also varies somewhat across the Ngumpin-Yapa subgroup, but is not investigated here. The conclusion is that just about all of Warlpiri’s verb roots are inherited, and thus there is little support for Dixon’s (2001, 2002) speculation “that Warlpiri increased the number of simple verbs under areal pressure”.

2. Profile of Warlpiri verb root inventory

Warlpiri has up to 130 recorded monomorphemic verb roots (counting separately the arguably homophonous roots with divergent senses). There is synchronic and diachronic texture to the inventory.

---

1 This paper has grown out of Nash (1982). I am grateful for feedback from the participants in the March 2007 Australianist workshop (University of Sydney Crommelin field station, Pearl Beach). I tentatively offer this to Harold Koch, who has long encouraged solid historical linguistic work on Australian languages.

2 Pace McConvell and Laughren’s (2004:172) summary that “Warlpiri and Western Ngumpin have a relatively high number of monomorphemic verbs (around 100)”.

verb root | gloss and case frame | NgY?
---|---|---
> ka-nyi | ERG transport ABS | NgY, CA
> nya-nyi | ERG perceive ABS | NgY, CA
> pi-nyi | ERG impact on ABS | NgY
> yi-nyi | ERG give ABS to DAT | NgY, CA
> nga-rni | ERG ingest ABS | NgY, CA
> ji-ni | ERG scold ABS | NgY
> ma-ni | ERG get ABS | NgY, CA
> -ma-ni | ERG cause ABS become... | NgY
> ya-ni | ABS go | NgY, CA
> janka-ni | ABS burn | 
> -jarri-ni | ABS become ... | NgY
> karr-ni | ABS stand | NgY
> ngana-ni | ABS recline | 
> nyina-ni | ABS sit, be | CA
> -pardi-ni | ABS arise | Y
> purka-ni | ABS run | 
> purra-ni | ERG (person) burn ABS | 
> wangka-ni | ABS speak, | CA
> wanti-ni | ABS drop | NgY
> wapa-ni | ABS move about | 
> -jirri-rni | ERG act forcefully on ABS | 
> kiji-rni | ERG throw ABS | 
> luwa-rni | ERG hit ABS with missile | NgY
> marda-rni | ERG hold ABS | Y
> ngaarr-ni | ERG scold ABS | Y
> pugi-ni | ERG cut ABS | NgY
> paka-ni | ERG strike ABS | Y
> panti-ni | ERG pierce ABS | 
> -parri-ni | ERG act on ABS | 
> yirra-rni | ERG place ABS at LOC,ALL | NgY

Table 1: 30 Warlpiri verb roots with many lexical coverb-derived stems; NgY=Ngumpin-Yapa (see Table 3), CA='Common Australian'; verbs are cited with NonPast suffix; root meaning often diluted in meaning in combination with a coverb.

2.1 Synchronic profile

Each Warlpiri verb root conjugates by tense-aspect suffix in one of five conjugations. Two of the conjugations account for over 100 of the roots, and for these there is a fair correlation with case frame (transitivity). All roots in the two large conjugations have at least two syllables, and the inflected verb usually has three or more syllables (the exception being the Non-Past in the intransitive conjugation where -ni alternates with zero); the roots in the small conjugations are typically monosyllabic and the inflected verb has two or more syllables.

---

About thirty Warlpiri verb roots occur in various lexicalised combinations with a coverb. A coverb (also called preverb) is an optional uninfl ecting accompaniment to an inflected verb root. Some coverbs can combine with any verb and the combination’s meaning is predictable (e.g. pina ‘return’); these are the productive coverbs. Other coverbs are limited to combining with particular verb roots, in idiosyncratic ways, so that each combination requires its own lexical entry. Nash (1982) studied the combinatory subclassification of the coverb class, and identified that only a minority of Warlpiri verb roots, those listed in Table 1, occur in combination with many lexical verbs. These include all the monosyllabic verb roots, that is, all the members of the three minority conjugations; these are listed first in Table 1.

A further 19 verb roots account for the remaining verb roots which are found in lexical coverb-coverb combinations. These are listed in Table 2.

<table>
<thead>
<tr>
<th>verb root</th>
<th>gloss and case frame</th>
<th>NgY?</th>
</tr>
</thead>
<tbody>
<tr>
<td>jiti-mi</td>
<td>ABS descend</td>
<td></td>
</tr>
<tr>
<td>karrka-mi</td>
<td>ABS proceed</td>
<td></td>
</tr>
<tr>
<td>-para-mi</td>
<td>ERG follow ABS</td>
<td></td>
</tr>
<tr>
<td>pura-mi</td>
<td>ERG follow ABS</td>
<td></td>
</tr>
<tr>
<td>purla-mi</td>
<td>ABS shout to DAT</td>
<td></td>
</tr>
<tr>
<td>yarnka-mi</td>
<td>ABS start on journey</td>
<td></td>
</tr>
<tr>
<td>kati-rni</td>
<td>ERG weigh down on ABS</td>
<td></td>
</tr>
<tr>
<td>pangi-rni</td>
<td>ERG scratch ABS</td>
<td></td>
</tr>
<tr>
<td>purta-rni</td>
<td>ERG take ABS away from DAT</td>
<td></td>
</tr>
<tr>
<td>ingeki-rni</td>
<td>ERG set fire to ABS</td>
<td></td>
</tr>
<tr>
<td>yurrpa-rni</td>
<td>ERG grind ABS</td>
<td></td>
</tr>
<tr>
<td>kampa-mi</td>
<td>ABS burn</td>
<td></td>
</tr>
<tr>
<td>kulpa-mi</td>
<td>ABS return home</td>
<td>NgY, CA</td>
</tr>
<tr>
<td>yula-mi</td>
<td>ABS cry</td>
<td></td>
</tr>
<tr>
<td>parpa-rni</td>
<td>ABS wait for DAT</td>
<td></td>
</tr>
<tr>
<td>V-yalpi-mi</td>
<td>...return (from hunting) and V</td>
<td></td>
</tr>
<tr>
<td>-pirri-rni</td>
<td>ERG act on ABS...</td>
<td></td>
</tr>
<tr>
<td>-nga-rni</td>
<td>ABS move</td>
<td></td>
</tr>
<tr>
<td>-ma-ni</td>
<td>ABS make noise</td>
<td>NgY</td>
</tr>
</tbody>
</table>

Table 2: Warlpiri verb roots with a few lexical coverb-derived stems

Thus almost half of Warlpiri verb roots occur in some lexical combination with a coverb. The remaining verb roots, about sixty in number, do not; they combine only with productive coverbs, in combinations which do not call for individual lexical entries.

2.2 Diachronic profile

2.2.1 Shared across Ngumpin-Yapa. Inspection of the verbal inventories of representative NgY languages reveals that about half the verb roots of each NgY language have cognates in other NgY languages, as set out in Table 3.
<table>
<thead>
<tr>
<th>gloss</th>
<th>Western</th>
<th>Ngumpin</th>
<th>Eastern</th>
<th>Yapa</th>
</tr>
</thead>
<tbody>
<tr>
<td>burn</td>
<td>kampa-</td>
<td>kampa-</td>
<td>kamb-</td>
<td>kampa-</td>
</tr>
<tr>
<td></td>
<td>lwara-</td>
<td>rr</td>
<td>u</td>
<td>mi</td>
</tr>
<tr>
<td>hit</td>
<td>la-</td>
<td>la-</td>
<td>lang-</td>
<td>la-</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>n</td>
<td>ku</td>
<td>ka</td>
</tr>
<tr>
<td>with</td>
<td></td>
<td></td>
<td>bi-</td>
<td>ka</td>
</tr>
<tr>
<td>missile</td>
<td></td>
<td></td>
<td>'bite',</td>
<td>ka</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>piya-ka</td>
<td>rni</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>paja-</td>
<td>rni</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>warlmanpa</td>
<td>rni</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>warlpiri</td>
<td>rni</td>
</tr>
<tr>
<td>spear</td>
<td>paja-</td>
<td>paja-</td>
<td>paja-</td>
<td>paja-</td>
</tr>
<tr>
<td></td>
<td>rra</td>
<td>n</td>
<td>rni</td>
<td>rni</td>
</tr>
<tr>
<td>bite</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sit, stand</td>
<td>karri-</td>
<td>karri-</td>
<td>karri-</td>
<td>karri-</td>
</tr>
<tr>
<td></td>
<td>wanti-</td>
<td>wanti-</td>
<td>wandi-</td>
<td>wanti-</td>
</tr>
<tr>
<td>fall</td>
<td>ma-</td>
<td>ma-</td>
<td>ma-</td>
<td>ma-</td>
</tr>
<tr>
<td></td>
<td>nyja</td>
<td>nyja</td>
<td>nyja</td>
<td>nyja</td>
</tr>
<tr>
<td>say</td>
<td>nga-nyja</td>
<td>nga-nyja</td>
<td>nga-nyja</td>
<td>nga-nyja</td>
</tr>
<tr>
<td>eat</td>
<td>yu-ngka</td>
<td>yu-ngka</td>
<td>yu-ngka</td>
<td>yu-ngka</td>
</tr>
<tr>
<td>give</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>carry</td>
<td>ka-ngka</td>
<td>ka-ngka</td>
<td>ka-ngka</td>
<td>ka-ngka</td>
</tr>
<tr>
<td>see</td>
<td>nya-ngka</td>
<td>nya-ngka</td>
<td>nya-ngka</td>
<td>nya-ngka</td>
</tr>
<tr>
<td>hit</td>
<td>pu-ngka</td>
<td>pu-ngka</td>
<td>pu-ngka</td>
<td>pi-nyi</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>get</td>
<td>-ma-nta</td>
<td>ma-n</td>
<td>ma-n</td>
<td>ma-n</td>
</tr>
<tr>
<td>go</td>
<td>ya-nta</td>
<td>ya-nta</td>
<td>ya-n</td>
<td>ya-n</td>
</tr>
<tr>
<td>scold</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>put</td>
<td>ngaja-rra</td>
<td>yuwa-rr</td>
<td>yuwa-rr</td>
<td>—</td>
</tr>
<tr>
<td>void</td>
<td>yinpa</td>
<td>yinpa</td>
<td>yinpa</td>
<td>—</td>
</tr>
<tr>
<td>sing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>leave</td>
<td>wanyja-rra</td>
<td>kuma-rrn</td>
<td>wanyja-rra</td>
<td>—</td>
</tr>
<tr>
<td>cut</td>
<td>warnta-n</td>
<td>kuma-rrn</td>
<td>kuma-rrn</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cry</td>
<td>lu-ngka</td>
<td>lu-ngka</td>
<td>lu-ngka</td>
<td>lu-ngka</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>scratch</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>wet</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>tie</td>
<td>karrpi</td>
<td>kurnta-n</td>
<td>kurnya-</td>
<td>way-</td>
</tr>
<tr>
<td>count</td>
<td>17</td>
<td>17</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>total V</td>
<td>39</td>
<td>41</td>
<td>60%</td>
<td>43%</td>
</tr>
<tr>
<td>%</td>
<td>44%</td>
<td>44%</td>
<td>60%</td>
<td>43%</td>
</tr>
</tbody>
</table>

Table 3: Corresponding verb roots across Ngumpin-Yapa (NgY), with fraction of all verb roots in each language

All the Warlpiri roots in Table 3 happen also to be in Table 1: that is, the Warlpiri roots with corresponding forms across Ngumpin-Yapa are ones (but not the only ones) which are found in a lot of lexical coverb-verb combinations in modern Warlpiri. Thus the degree to which a verb root enters into lexicalised

---

4 That is, 55% of 30 in Table 1.
<table>
<thead>
<tr>
<th>pPNy</th>
<th>gloss</th>
<th>ps</th>
<th>Warlpiri</th>
<th>NgY?</th>
</tr>
</thead>
<tbody>
<tr>
<td>*ka:mpa-</td>
<td>'cook in earth oven'</td>
<td>vt</td>
<td>kampa-mi</td>
<td></td>
</tr>
<tr>
<td>*nga:ci-</td>
<td>'lay (egg), give birth to (young)'</td>
<td>vt</td>
<td>ngaja-mi</td>
<td></td>
</tr>
<tr>
<td>*nga:ci-</td>
<td>'lay (egg), give birth to (young)'</td>
<td>vt</td>
<td>ngaja-mi</td>
<td></td>
</tr>
<tr>
<td>*nga:ci-</td>
<td>'lay (egg), give birth to (young)'</td>
<td>vt</td>
<td>ngaja-mi</td>
<td></td>
</tr>
<tr>
<td>*paj(a)- / *pa:ca-</td>
<td>'bite' / 'taste it'</td>
<td>vt</td>
<td>paja-rni</td>
<td></td>
</tr>
<tr>
<td>*parnti-</td>
<td>'smell' vt</td>
<td>Warlpiri</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*parnti-</td>
<td>'smell' vt</td>
<td>Warlpiri</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*paka-</td>
<td>'dig'</td>
<td>vt</td>
<td>paka-rni 'strike'</td>
<td>Yapa</td>
</tr>
<tr>
<td>*paka-</td>
<td>'dig'</td>
<td>vt</td>
<td>paka-rni 'strike'</td>
<td>Yapa</td>
</tr>
<tr>
<td>*wanga-</td>
<td>'peck'</td>
<td>vt</td>
<td>wangka-</td>
<td>Yapa</td>
</tr>
<tr>
<td>*karrli/a-</td>
<td>'pierce'</td>
<td>vt</td>
<td>karla-mi</td>
<td>Warlpiri</td>
</tr>
<tr>
<td>*nita-</td>
<td>'sit'</td>
<td>vi</td>
<td>nyina-mi</td>
<td>Warlpiri</td>
</tr>
<tr>
<td>*paca-</td>
<td>'hit'</td>
<td>vt</td>
<td>paji-rni 'cut'</td>
<td>Warlpiri</td>
</tr>
<tr>
<td>*wapa-</td>
<td>'go'</td>
<td>vi</td>
<td>wapa-mi</td>
<td>Warlpiri</td>
</tr>
<tr>
<td>*karri-</td>
<td>'stand'</td>
<td>vi</td>
<td>karri-mi</td>
<td></td>
</tr>
<tr>
<td>*nga-</td>
<td>'eat'</td>
<td>vt</td>
<td>nga-rni</td>
<td></td>
</tr>
<tr>
<td>*panga-</td>
<td>'dig'</td>
<td>vt</td>
<td>pansi-rni</td>
<td></td>
</tr>
<tr>
<td>*ngiwa-</td>
<td>'lie'</td>
<td>vi</td>
<td>nguna-mi</td>
<td>Warlpiri</td>
</tr>
<tr>
<td>*parnta-</td>
<td>'smell'</td>
<td>vt, vi</td>
<td>parnta-mi</td>
<td>Warlpiri</td>
</tr>
<tr>
<td>*parnta-</td>
<td>'smell'</td>
<td>vt, vi</td>
<td>parnta-mi</td>
<td>Warlpiri</td>
</tr>
</tbody>
</table>

Table 4: Pama-Nyungan verb roots (after Alpher 2004) attested in the Ngumpin-Yapa subgroup; reconstructions to pPNy above double line. Abbreviations: ps= part of speech, v=attested combinations with coverbs is an indication of its age within Warlpiri, and while this appears to be generally so, there are two exceptional pan-NgY roots whose Warlpiri reflexes lack lexical combinations: not in Table 1 but in Table 3 are *pangi-rni and *kampa-mi (the latter root further being Capell’s (1956:85-94) ‘Common Australian’ widespread beyond NgY).9

---

1 I list this bound verb root (found in just two verbs in the Eastern dialect only) as a reflex of *ra-somewhat speculatively. Apart from not understanding the historical morphology and semantics, it has not undergone the NgY lateralisation *r→rl (McConvell & Laugren 2004:155), as it has in Warlmanspa (Yapa) la- (McConvell & Laugren (2004:156) see Warlmanspa la- as a reflex of *ruwa- not *ra- but this does not affect the point about lateralisation.) It may be that lateralisation does not apply in the third syllable: however relevant known pPNy roots are attested in Yapa with *r→rr (*kakara ‘moon’, *micara ‘cold’ > Warlpiri kakarra-, Warlmanspa mijarra, Alpher 2004).

6 For this meaning Warlpiri uses yula-mi, which may well be a loan from WD, see below. McConvell and Smith discuss the sound change -u deletion after ng, giving Guridji particle lungkarra not *lungwarra


8 Also Jaru has the verb root parntiny ‘ABS smell’, in northern dialect. Warlpiri parnti is also a coverb, in parnt-nya-nya ‘ERG smell ABS’.

9 Also there is at least one root with lexicalised combinations which appears to be a recent loan, namely -alpi-mi from Arrernte (see below); if so, it is unusual in having been borrowed not as a stem but a derivational affix.
2.2.2 Inheritance from Pama-Nyungan. Of the 45 proto-Pama-Nyungan verb roots reconstructed by Alpher (2004), 16 have a Warlpiri reflex.\(^\text{10}\) These are set out in Table 4 supplemented by two ‘Common Australian’ roots (*ya- and *yu-). At the bottom of Table 4 are a further five roots attested in PNy (Alpher 2004) outside NgY but not to pPNy, with NgY reflex. Taking these verb roots together, it is striking that 15 attested across Ngumbin-Yapa are attested in particular in Warlpiri, a further two roots (*paka- and *wangka-) are retained only in the Yapa subgroup, and a further six roots are retained in Warlpiri uniquely within Ngumpin-Yapa. Only one of these “established” verbs attested in NgY is not attested in Warlpiri: *rungka- with Yapa (Warlmanpa) reflex lung-. A further 14 Warlpiri verb roots have correspondences elsewhere in Ngumpin (Table 5) or just in the Yapa subgroup (Table 6).\(^\text{11}\)

<table>
<thead>
<tr>
<th>pNgY</th>
<th>gloss</th>
<th>ps</th>
<th>Warlpiri</th>
<th>other languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>*jarri-</td>
<td>'become'</td>
<td>vi</td>
<td>-jarri-mi</td>
<td>Wmj -jarri-, Wpa -ja-nya</td>
</tr>
<tr>
<td>*ji-(^\text{12})</td>
<td>'scold'? 'burn'?</td>
<td>vt</td>
<td>-ji-ni</td>
<td>Gur, Jaru ju-n 'scold', Mud jundurru 'scold', Wpa ji-nya 'burn'</td>
</tr>
<tr>
<td>*ma-(^\text{13})</td>
<td>'say'</td>
<td>vi</td>
<td>-ma-ni</td>
<td>pan-Ngumpin</td>
</tr>
<tr>
<td>*mapa-(^\text{14})</td>
<td>'spread on, rub'</td>
<td>vt</td>
<td>mapa-rni</td>
<td>Wmj mapa-</td>
</tr>
<tr>
<td>*pirrki-rni</td>
<td>'make by trimming'</td>
<td>vt</td>
<td>pirrki-rni</td>
<td>coverb pirrka (E Ngumpin)</td>
</tr>
<tr>
<td>*yunpa-rn</td>
<td>'sing'</td>
<td>vt</td>
<td>yunpa-rni</td>
<td>pan-Ngumpin</td>
</tr>
</tbody>
</table>

Table 5: Other Warlpiri verb roots shared with some Ngumpin languages

These tables account for some 37 Warlpiri verb roots. It is clear that Warlpiri has retained more verb roots from pNgY than the other NgY languages have. In this light we can look at the balance of the Warlpiri simple verb inventory.

\(^{10}\) There are 1130 sets, of which 357 with a pPNy reconstruction. Of the 1130 sets, 211 are verbs, of which 45 verbs have a pPNy reconstruction, and 24 have a Warlpiri reflex (whether or not reconstructed to pPNy).

\(^{11}\) Note that the proto forms proposed in the first columns have not yet been properly demonstrated. We can note however that the dropping in Warlmanpa of the second syllable of *jarri* and *ngarri*- appears to be regular, compare also *karrri* (Warlmanpa ka-nya), but does not apply to *warrri* (which perhaps is *wayi-.*

\(^{12}\) Pintupi/Luritja tju-nu 'put, place', also used as 'Cause' (Hansen & Hansen 1992


\(^{14}\) Also attested outside NgY, in WD including Pintupi/Luritja maparnpa ‘doctor man; doctor man’s spirit’, Arrernte aperne-me, and Warumungu apa-nta.

\(^{15}\) See McConvell and Laughren (2004:154,172-3).
2.2.3 Verbs with etymological clues (more recent than PNy). The Warlpiri verb roots remaining to be discussed are each unique to Warlpiri within Ngumpin-Yapa; and they number are about eighty. There has been little relevant reconstruction, and so comparative remarks here are speculative. I offer them as a preliminary organisation of relevant information, which also can be brought to bear on previous speculation. Another gross shortcoming of my historical comparative discussion of verbs here is the lack of attention to inflected forms and the role of paradigms (conjugations).

The verb roots unique to Warlpiri within NgY can be grouped according to several cross-cutting criteria:

(A) verbs with a clearly corresponding form in a neighbouring language
   from recent borrowing
   from shared inheritance

(B) verbs with a discernible internal derivation, resulting from fusion of a coverb-verb combination

<table>
<thead>
<tr>
<th>pY</th>
<th>gloss</th>
<th>ps</th>
<th>Warlpiri</th>
<th>Warlmanpa</th>
</tr>
</thead>
<tbody>
<tr>
<td>*marta-</td>
<td>'hold, have'</td>
<td>vt</td>
<td>marda-rni</td>
<td>marta-nya</td>
</tr>
<tr>
<td>*murla-</td>
<td>'copulate'</td>
<td>vt</td>
<td>murla-rni</td>
<td>murla-nya</td>
</tr>
<tr>
<td>*ngarri-</td>
<td>'tell; call s.o. s.th; scold'</td>
<td>vt</td>
<td>ngarri-rni</td>
<td>nga-nya</td>
</tr>
<tr>
<td>*nguka-</td>
<td>'drink'</td>
<td>vt</td>
<td>nguka-rni</td>
<td>nguka-nya</td>
</tr>
<tr>
<td>*pali-</td>
<td>'die'</td>
<td>vi</td>
<td>pali-mi</td>
<td>palu-ngunya</td>
</tr>
<tr>
<td>*parti-</td>
<td>'rise, set off'</td>
<td>vi</td>
<td>parti-mi</td>
<td>partu-ngunya</td>
</tr>
<tr>
<td>*warrka-</td>
<td>'climb'</td>
<td>vi</td>
<td>warrka-rni</td>
<td>waka-nya</td>
</tr>
<tr>
<td>*warri-</td>
<td>'search for'</td>
<td>vt</td>
<td>warri-rni</td>
<td>wayi-nya</td>
</tr>
</tbody>
</table>

Table 6: Warlpiri verbs shared only with Warlmanpa in NgY

(A) I list 36 Warlpiri verb roots in Table 7 (6), Table 8 (25) and Table 9 (5) which appear to be shared with a (non-NgY) neighbouring language. The form or meaning correspondence is not always straightforward, and in some may prove unsustainable. Also, I generally cannot tell whether Warlpiri shares the verb root owing to borrowing or shared inheritance. That each root does not otherwise occur in NgY might be taken to point to borrowing, but in some instances the kind of form or meaning change points more to inheritance than borrowing, and as we have seen above Warlpiri does retain a number of verb roots uniquely in NgY.

16 This verb exhibits medial lateralisation when compared with Pintupi/Luritja mura-rnu 'copulate' (Hansen & Hansen 1992).
17 There is also a coverb pali 'dead' in Warlpiri and Warlmanpa; and the verb root pali- 'die' is also in Ngarti.
18 Compare also Pintupi/Luritja warrka-rnu 'generic verb; motion; moved; CRS' (Hansen & Hansen 1992).
19 Fitzgerald (1997:235) set 5.230 adds Kaurna warre-ndi 'to look for', warri.appe-ndi 'to seek, pick up, find', Diyari warra.y.ma-l 'find, meet', Woiwurrung wooru.nderoneit 'pick up'.
### Table 7: Warlpiri verb roots which could be loans from Arrernte (including Anmatyerr)

<table>
<thead>
<tr>
<th>Warlpiri</th>
<th>gloss</th>
<th>East &amp; Central Arrernte</th>
</tr>
</thead>
<tbody>
<tr>
<td>V-alpi-mi</td>
<td>... return and V</td>
<td>alpeme 'return, go back' (also derivational)</td>
</tr>
<tr>
<td>yampi-mi</td>
<td>ERG leave alone ABS</td>
<td>impeme 'leave alone'</td>
</tr>
<tr>
<td>yarlti-rni</td>
<td>ERG bite ABS</td>
<td>arlkweme 'eat, feed on, bite on'</td>
</tr>
<tr>
<td>yilya-mi</td>
<td>ERG send ABS to DAT</td>
<td>ilantheme 'send'</td>
</tr>
<tr>
<td>yulka-mi</td>
<td>ABS cherish DAT</td>
<td>ilkelheme 'like'</td>
</tr>
<tr>
<td>wirnti-mi</td>
<td>ABS (women) dance</td>
<td>urnteme 'dance, (men) dance'</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Warlpiri</th>
<th>gloss</th>
<th>Pintupi/Luritja</th>
</tr>
</thead>
<tbody>
<tr>
<td>japi-rni</td>
<td>ERG ask ABS about DAT</td>
<td>tjapi-ruu 'asked; requested'</td>
</tr>
<tr>
<td>jartti-rni</td>
<td>ERG scrape ABS</td>
<td>tjartti-ruu 'wiped; cleaned; cleared; to clean up an area for camping or sitting'; tjartti-riri 'open place; cleared place; clearing'</td>
</tr>
<tr>
<td>jija-mi</td>
<td>ABS succumb to DAT</td>
<td>tjija-ngu 'was hurt; accidently hurt; suddenly hurt'</td>
</tr>
<tr>
<td>jiti-mi</td>
<td>ABS descend</td>
<td>tjiti-ngu 'dribbled; dripped; ran out; poured out'</td>
</tr>
<tr>
<td>kardi-rni</td>
<td>ERG fetch ABS (water)</td>
<td>kutu-ruu 'lifted; raised the level; filled up; lit. 'made above'; to raise the level in a container by filling it up'</td>
</tr>
<tr>
<td>kipi-rni</td>
<td>ERG winnow ABS</td>
<td>kipi-ruu 'winnowed'</td>
</tr>
<tr>
<td>kulpa-mi</td>
<td>ABS return</td>
<td>kurlpa-ngu 'returned; went back; came home'</td>
</tr>
<tr>
<td>mila-rni</td>
<td>ERG select (best of) ABS</td>
<td>mili-ruu 'took, took out'</td>
</tr>
<tr>
<td>mirri-rni</td>
<td>ERG erase ABS</td>
<td>mirri-ruu 'killed; murdered'</td>
</tr>
<tr>
<td>nyunjji-rni</td>
<td>ERG kiss ABS</td>
<td>nyunjju-ngu 'kissed; quietened'</td>
</tr>
<tr>
<td>parpa-rni</td>
<td>ABS wait for DAT</td>
<td>parta-ruu 'waited'</td>
</tr>
<tr>
<td>parlji-rni</td>
<td>ERG wash ABS</td>
<td>palji-ruu 'washed; cleaned'</td>
</tr>
<tr>
<td>parnti-rni</td>
<td>ERG touch ABS</td>
<td>pampu-ruu 'touched'</td>
</tr>
<tr>
<td>karda-rni</td>
<td>ABS accompany</td>
<td>parta-ruu 'attached; fixed on; joined; to fix or attach items to larger objects'</td>
</tr>
<tr>
<td>tirli-rni</td>
<td>ERG flake ABS</td>
<td>tirli-ruu 'broke off; to split off a stone sliver to be used as a knife by hitting two stones together'</td>
</tr>
<tr>
<td>wardi-rni</td>
<td>ERG straighten ABS</td>
<td>wartu-ruu 'stretched; after sitting in one position for a long time'</td>
</tr>
<tr>
<td>wirli-rni</td>
<td>ERG poke ABS</td>
<td>willi-ruu 'poked with finger'; willi-pungu 'pointed with the hand'</td>
</tr>
<tr>
<td>yaja-rni</td>
<td>ERG enlist ABS</td>
<td>watja-ruu 'said; told; spoke; commanded; promised; accused'</td>
</tr>
</tbody>
</table>

---

20 Note also Luritja yulkanyi 'affectionate; to have affection for others especially children; from Aranda JH’ (Hansen & Hansen 1992).

21 ratwem seems to apply more generally (men and women) in western Anmatyerr (Jennifer Green p.c.).


23 Possibly relatable to pPNy *mi:l 'eye' (Alpher 2004:470); cf. Yolngu Matha mila- 'take, take out'.

WARLPIS VERB ROOTS 229

Table 8: Warlpiri verb roots in common with Pintupi/Luritja

<table>
<thead>
<tr>
<th>Warlpiri</th>
<th>gloss</th>
<th>Pintupi/Luritja</th>
</tr>
</thead>
<tbody>
<tr>
<td>yarli-rni</td>
<td>ERG soak ABS</td>
<td>yarlu-rnu 'uncover'</td>
</tr>
<tr>
<td>yarri-rni</td>
<td>ERG build fire ABS</td>
<td>yarrpu-ngu 'scolded; rebuked'</td>
</tr>
<tr>
<td>yirriti-rni</td>
<td>ERG knock over ABS</td>
<td>yinti-rnu 'dripped; dribbled; poured'</td>
</tr>
<tr>
<td>yirnti-rni</td>
<td>ERG knock over ABS</td>
<td>yinti-rnu 'dripped; dribbled; poured'</td>
</tr>
<tr>
<td>yula-mi</td>
<td>ABS cry, wail</td>
<td>yula-ngu 'cried; wailed'</td>
</tr>
</tbody>
</table>

Table 9: Warlpiri verb roots with possible non-verb cognate in Pintupi/Luritja

<table>
<thead>
<tr>
<th>Warlpiri</th>
<th>gloss</th>
<th>Pintupi/Luritja</th>
</tr>
</thead>
<tbody>
<tr>
<td>mati-rni</td>
<td>ABS go in procession</td>
<td>martu yanu 'went across'</td>
</tr>
<tr>
<td>ngarlki-rni</td>
<td>ERG block ABS</td>
<td>ngarlkin-ma-nu 'protected; argued against; shielded; to protect another in a fight'</td>
</tr>
<tr>
<td>rdirrji-rni</td>
<td>ABS start a fight</td>
<td>turrtja-nura 'satisfied'</td>
</tr>
<tr>
<td>wipi-mi</td>
<td>ABS radiate</td>
<td>wipu 'tail'</td>
</tr>
<tr>
<td>yipi-rni</td>
<td>ERG pick out pimple ABS</td>
<td>yipi 'milk, mother, breast'</td>
</tr>
</tbody>
</table>

(B) Warlpiri verb roots with a possible compound origin

<table>
<thead>
<tr>
<th>Warlpiri</th>
<th>gloss</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVC+{pi-} (cf. root pu- ~ pi- Table 4) &amp; CVC+{ji-} (cf. root -ji- Table 5)</td>
<td>cf. tirpi-mi and others in Table 8 &amp; cf. rdjrji-rni, nyunjji-rni, parli-rni (Table 8, Table 9)</td>
<td></td>
</tr>
<tr>
<td>karlp-mi</td>
<td>ERG (temperature) cause suffer ABS</td>
<td>cf. kmal-mi, kmalpi-mi, kmalnji-mi, kmalnji-rni, kmalnji-rni</td>
</tr>
<tr>
<td>wrrpi-rni</td>
<td>ERG eat ABS [avavoidance]</td>
<td>Warlmanpa wrrpiny (coverb) 'completely' (esp. with 'eat', 'drink')</td>
</tr>
<tr>
<td>jampi-rni</td>
<td>ERG lick ABS</td>
<td>with 'eat', 'drink'</td>
</tr>
<tr>
<td>marnpi-rni</td>
<td>ERG touch ABS</td>
<td>with 'eat', 'drink'</td>
</tr>
<tr>
<td>larrji-rni</td>
<td>ERG scratch ABS</td>
<td>with 'eat', 'drink'</td>
</tr>
<tr>
<td>walji-rni</td>
<td>ERG pluck ABS</td>
<td>with 'eat', 'drink'</td>
</tr>
<tr>
<td>winji-rni</td>
<td>ERG pour ABS</td>
<td>with 'eat', 'drink'</td>
</tr>
<tr>
<td>others</td>
<td></td>
<td>with 'eat', 'drink'</td>
</tr>
<tr>
<td>waraparnpi-mi</td>
<td>ABS sing out</td>
<td>wara 'tall; long'</td>
</tr>
<tr>
<td>uuurlparra-rni</td>
<td>ERG dehair ABS</td>
<td>uuurl-panpi-rni 'dig up earth', uuurl-parntarrinjarra-karrini 'be stooped over'</td>
</tr>
</tbody>
</table>

25 Cf. Northern Nyangumarta yajana- (V2) 'follow'.
26 Also Parnkalla windu-tu 'to drip, trickle down', Ngarluma yinti-ku 'to drip, descend, land', Payngu yirrtkala-L- 'pour' (Fitzgerald 1997:249).
27 This root is presumably related to pTangkic *yula-ca (p.c. B. Alpher from N. Evans).
28 Fitzgerald (1997:226) set 5.159 links this stem with Umpila walki- 'stop, prevent'.
29 Also Nyangumarta wipiri 'long and thin', Payngu (Fitzgerald 1997)
30 Alpher (2004) reconstructs pPNy *yipi "woman"; hence to sustain the link with the Warlpiri verb I need to propose a changes along the lines 'woman' > 'breast' > 'nipple' > 'pimple' as well as zero conversion of removal. An alternative is the putative correspondence set (Alpher p.c.) *cimpi 'squeeze it' (YOront thew 'defecate' with kun 'faeces', Ngarluma jimpi-lku 'to squeeze-of blister, pimple').
Table 10: Warlpiri polysyllabic verb stems with possible compound etymology

Of those verb roots in Table 10 which prove to have a demonstrable compound origin, their age may predate pNgY given that there are similar verb roots (in Table 8 and Table 9) with apparent cognates outside NgY. On the other hand, some (particularly those last listed) may be coverb-verb combinations fused into a single verb stem since the differentiation of NgY.

Nash (1982) proposed that some Warlpiri verb roots could have originated recently from conjugation of an erstwhile coverb. The two instances (pali-mi and parnti-mi) I now see as erstwhile verbs because they have cognate verbs in languages besides Warlpiri.

2.2.4 Warlpiri verbs with no local provenance. There remain some 34 Warlpiri verb roots not assignable to any of the categories above: they are unique within NgY, no corresponding form found in any neighbouring language, and I am not aware of a possible compound origin. Of course, it is possible that some were borrowed into a predecessor of Warlpiri from an erstwhile neighbouring language, with the particular evidence for the source no longer in modern languages. For a few I note possible cognates further afield in P Ny, and future research may also reveal possible cognates further afield in Australia for these and a number of the others. Generally, we can expect the Warlpiri verb roots in Table 11 are mostly retentions from pNgY or earlier.
<table>
<thead>
<tr>
<th>Warlpiri</th>
<th>gloss</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>maja-rni</td>
<td>ERG straighten ABS</td>
<td></td>
</tr>
<tr>
<td>ngarti-rni</td>
<td>ERG erect ABS</td>
<td></td>
</tr>
<tr>
<td>nganya-rni</td>
<td>ERG warm ABS</td>
<td>Baagandji nganya 'flame, light' Fitzgerald (1997:183) set 4.82</td>
</tr>
<tr>
<td>-nga-rni</td>
<td>ABS move</td>
<td></td>
</tr>
<tr>
<td>ngarrmi-rni</td>
<td>ERG ritually increase ABS</td>
<td></td>
</tr>
<tr>
<td>nyurla-mi</td>
<td>ERG knead ABS</td>
<td></td>
</tr>
<tr>
<td>paja-rni</td>
<td>ERG taste ABS</td>
<td></td>
</tr>
<tr>
<td>papi-mi</td>
<td>ABS ignite</td>
<td></td>
</tr>
<tr>
<td>-para-mi</td>
<td>ERG follow ABS</td>
<td>cf. pura-mi</td>
</tr>
<tr>
<td>parnta-rni</td>
<td>ERG withdraw-from-fire ABS</td>
<td></td>
</tr>
<tr>
<td>-pirri-rni</td>
<td>ERG act on ABS...</td>
<td></td>
</tr>
<tr>
<td>punta-rni</td>
<td>ERG take ABS away from DAT</td>
<td></td>
</tr>
<tr>
<td>pura-mi</td>
<td>ERG follow ABS</td>
<td>cf. -para-mi</td>
</tr>
<tr>
<td>purla-mi</td>
<td>ABS shout to DAT</td>
<td></td>
</tr>
<tr>
<td>rdipi-mi</td>
<td>ABS gather</td>
<td></td>
</tr>
<tr>
<td>turlka-mi</td>
<td>ERG pinch ABS</td>
<td></td>
</tr>
<tr>
<td>wanja-rni</td>
<td>ERG take ABS (sip) off DAT</td>
<td></td>
</tr>
<tr>
<td>wari-rni</td>
<td>ABS tie ABS</td>
<td></td>
</tr>
<tr>
<td>wayi-rni</td>
<td>ABS tie ABS</td>
<td></td>
</tr>
<tr>
<td>wirnpiri-mi</td>
<td>ABS whistle</td>
<td></td>
</tr>
<tr>
<td>yarrka-mi</td>
<td>ABS start on journey</td>
<td></td>
</tr>
<tr>
<td>yingki-rni</td>
<td>ERG set fire to ABS</td>
<td></td>
</tr>
<tr>
<td>yirdi-mi</td>
<td>ABS be frightened</td>
<td></td>
</tr>
<tr>
<td>yuko-mi</td>
<td>ABS enter</td>
<td>Pintupi/Luritja yuKarli-ngu 'climbed down; descended; alighted'</td>
</tr>
<tr>
<td>yurrrpa-rni</td>
<td>ERG grind ABS</td>
<td>?: Pintupi/Luritja yurlpu 'powdery substance', yurlpu-rni 'made powdery'; Warlpiri yurlpu - yurlpa 'red ochre'</td>
</tr>
</tbody>
</table>

Table 11: Warlpiri verb roots unique in NgY and absent from neighbouring languages; with comment on possible wider correspondence

2.3 Summary

<table>
<thead>
<tr>
<th>Table</th>
<th>category of Warlpiri verb roots</th>
<th>count</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>from Pama-Nyungan and attested in the Ngumpin-Yapa subgroup</td>
<td>23</td>
</tr>
<tr>
<td>5</td>
<td>others shared with some Ngumpin languages</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>shared only with Warlmanpa in NgY</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>could be loans from Arrernte</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>in common with Pintupi/Luritja</td>
<td>25</td>
</tr>
</tbody>
</table>

When the noun munga 'night' was proscribed, the word yuka substituted for it, also yula:yuurtu 'morning' for mungalyurru; cf. wanta ka:yuurtu 'the sun is setting (literally, entering)'. 
3. **Previous study**

This study has to some extent duplicated the investigation of McConvell and Laughren (2002:172-3).\(^{32}\)

Some verb roots in Warlpiri can be shown to be Arandic loans, but other than those it is likely that a large number of monomorphemic roots shared by Western Ngumpin and Yapa might be reconstructible to proto-NGY, some having been lost in Eastern Ngumpin. However we have decided to be more conservative in this paper and only consider as candidates for proto-NGY those verbs which have witnesses in all three branches, Yapa, Western Ngumpin and Eastern Ngumpin. This reduces the pool to less than forty.

Of these 30-40 verbs investigated, all but one turned out to be inherited from some higher level grouping like proto-Nyungic, or in a number of cases from Proto-Pama-Nyungan, which excludes them from consideration as shared innovations in NGY.

McGregor (2004) is another who has noted the geographic cline in the coverb-verb construction (his CVC):

Most Pama-Nyungan languages of Australia’s north-west and west show CVCs of the PV-IV type, including all of the languages along (or very close to) the border with non-Pama-Nyungan languages: Karajarri, Mangala, Walmajarri, Jaru, Gurindji, Mudburra, Warlmanpa, Warumungu, and Bilinara. Most of these languages appear to have a small number of IVs — around 40 — covering similar basic semantic domains.[...]

Superclassifying CVCs can be found a considerable distance into the Pama-Nyungan region. As one moves further from the non-Pama-Nyungan border, however, the construction appears to reduce in significance, and IVs tend to increase in number while PVs become fewer. For instance, there is some evidence that CVCs exist, to different degrees, throughout much of the Western Desert sprachbundt, the chain of linguistic varieties spoken throughout a large part of the central desert region of Australia. In the northernmost varieties Wangkajunga, Yulparrja, Kukatja, and Pintupi the system is most prominent — though considerably less so than in their northern neighbours.

Dixon (2001:74-75, 2002) discusses Warlpiri and Warlmanpa for which the only reference given is Nash (1982), which is presented as countexemplifying his

---

\(^{32}\) The proposal that NgY has about 50 simple verbs, and that daughter languages other than Warlpiri have lost simple verbs, and more so in Eastern than Western Gurindji, was made in the unpublished presentation by McConvell and Schultze-Berndt (2001), and adverted to by McConvell (2003:257).
proposed cycle (5) of change (whereby a language reduces its simple verb inventory):

Proto-WJb was of type (b), like Warlmanpa

Warlpiri has increased the inventory of simple verbs, moving into type (c).

Nash (1982) suggests that this may have happened through Warlpiri having (i) reanalysed some coverb-plus-simple-verb combinations as new simple verbs; and (ii) accorded simple verb status to what were coverbs, so that they now take the suffixes associated with simple verbs.

The suggestion (that “Warlpiri has increased the inventory of simple verbs”) has not been made by me. The attributed suggestions are embedded in a bigger more complex picture sketched by Nash (1982:185–200), and elaborated above. So, at most 17 modern Warlpiri roots exhibit the change which Dixon (2001:73-4) labels “(i) A coverb-plus-simple-verb combination coalesces into a single verb” (Table 10 above), and at most two roots exhibit the change he labels (ii) “accord […] simple verb status to what were coverbs” (p.74) (pali-mi and parti-mi, which I now see as erstwhile verbs). The two processes thus categorise at most 19 out of 130 Warlpiri verb roots.

Dixon does not mention the Warlpiri verb roots which could be recent loans from Arandic or WD (Table 7 and Table 8), and ignores the problem his view faces in the numerous (34) Warlpiri simple verbs for which a source in reanalysis or borrowing is highly unlikely (Table 11).

Dixon (2001:74) concludes that:

it is likely that Warlpiri increased the number of simple verbs under areal pressure from its south-westerly neighbour WD, the Western Desert language, which is of type (c), with a couple of hundred simple verbs.

It is more likely, on the evidence to this point (and compatible with the sparse speculation at the end of §4, Nash 1982:200), that within Ngumpin-Yapa a lot of simple verbs dropped from the ancestral languages, especially if they did not enter into lexical coverb-verb combinations, and many are retained only in Warlpiri. In this alternative view, Ngumpin-Yapa verbs are not a counterexample to Dixon’s pattern (5) (2001:73), and we do not need to countenance the kind of “areal pressure” which would cause a language to add simple verbs, many of undetectable origin.

Note no page reference is given, and no quotation. The point of my 1982 paper was to show that the difference between Warlpiri and other Ngumbin-Yapa languages in verbal lexicalisation — in D’s terms, that Warlpiri is (c) and the rest are (b) — is reduced when one sees the texture of Warlpiri coverb-verb organisation. I did not propose a direction of change. (I am to blame for Dixon’s misunderstanding, in that (a) I talked of “the covert class of Warlpiri verb roots which are presumably of more ancient vintage in the language, viz. the ’core roots’ of 3.1” (1982:200), and (b) in my conclusion (§5, 1982:200) I focussed on reanalysis of Dixon’s type (i), because of its resolution of structural tension morphology vs. semantics, and did not repeat my other points in the conclusion. At a seminar (’Ergativity and accusativity in Australian languages’, 4pm Monday 3 March 1997, Coombs Seminar Room A, ANU), when I first heard how Dixon had misunderstood my paper, I told him so, but the misunderstanding persists in Dixon (2001).
References


OUJIANG WU TONES AND ACOUSTIC RECONSTRUCTION

PHIL ROSE

Australian National University

1. Introduction

We do not expect tones, as phonological phenomena, to vary without limit, but to show similarities, or so-called tonological universals, across varieties (Maddieson 1978). Nevertheless, when one samples tones from different Asian tone languages, or even different dialects, they often show a great variety of pitch shapes. Zhu and Rose (1998) for example demonstrated no less than 25 different tones in four varieties from different sub-groups within the Wu dialects of east central China. (I use the term tone in this paper in a sense analogous to phone, that is, as a constellation of audibly different properties, with pitch predominating but also including length and phonation type, that constitutes observation data for phonological – usually tonemic – analysis.)

In exploring the nature of phonological variation in Language, we naturally concentrate on differences. This paper, on the other hand, makes use of how similar tones can be. It looks at variation in tones at seven localities distributed over a fairly small geographical area where the varieties, which belong to the same dialectal sub-group, have been described as homogeneous. Although this study reveals some interesting synchronic linguistic-tonetic variation, the main reason for undertaking it is not typological, but, as befits a festschrift for Harold Koch, historical. As mentioned, the Wu dialects of Chinese are well-known for their tonal complexity. Not only do they exhibit probably the most complex tone sandhi in the world (Rose 1990), but they also show a bewildering variety in their citation tones. Most Wu varieties have eight citation tones, but these can differ considerably in their shape. It is a major challenge to find out how it got that way. In order to apply the comparative method, it is normal to start from the smallest well-defined groupings. That is what this paper does for a first order sub-group of Wu called Oujiang. It brings together acoustic descriptions of tones of speakers from several different Oujiang sites, in order to reconstruct two of the tones of proto-Oujiang.

Attempts to reconstruct actual tonal values from modern source material are rare. Possibly because of the bewildering synchronic variety of tonal shapes, and few clear indications of common ways tones can change, it is normal to reconstruct abstract tonal categories with no phonological content (e.g. Brown 1985). One notable, and relevant, exception is Ballard (1969: 70, 92-98) who reconstructed pitch values for proto-Wu tones from cognate sets, and tone sandhi behaviour, in 16 Wu sources. This paper is novel in its attempt to go one further and reconstruct acoustic values for proto tones. It has the following
structure. I first briefly characterise the Wu dialects and their Oujiang sub-group. Next, acoustic data of Oujiang citation tones are described from seven sites. Then, as an important precursor to acoustic reconstruction, a method of quantifying homogeneity in tonal acoustics is presented. Reconstructed acoustics and measures of proto variation are then derived by means of normalisation of two demonstrably homogeneous tones.

2. Wu and Oujiang

The Wu dialects, of which Shanghai is a well-known variety, are spoken in the two eastern Chinese coastal provinces of Zhejiang and (southern) Jiangsu (Rose 2001). They are considered a first-order sub-group of Middle Chinese, along with five other major dialect groups of Chinese: Mandarin, Hakka, Yue, Xiang and Gan. The sixth major dialect group of Sinitic, Min, is assumed to have diverged before Middle Chinese (Norman 1988).

Oujiang is one of the sub-groups currently recognised for Wu (Zhengzhang 1987). Sources differ a little on the number of sub-groups. Zhengzhang (1987) lists six: Taihu, Taizhou, Oujiang Wuzhou Chuqu, Xuanzhou. The Oujiang sub-group, eponymous with the river bisecting it in a N.W. to S.E. direction (jiāng = river), is located in the S.E. corner of Zhejiang province, and is approximately congruent with the Wenzhou administrative area. About five million people speak Oujiang varieties, of which Wenzhou is probably the best known.

As Harold is fond of pointing out, it is an important methodological principle in historical linguistics that sub-groups can only be established on the basis of shared, unusual innovations. The sub-group status of the Wu dialects, resting as it does on a retention from Middle Chinese, is therefore suspect from an orthodox point of view. However, from a Bayesian point of view, the traditional position articulated by Chao is perfectly reasonable. Since the Wu dialects are very nearly the only varieties of Sinitic to have retained such a structural property (the vast majority of the other dialects has devoiced the Middle Chinese voiced series of obstruents leaving only a two-way contrast), the
existence of a systematic three-way contrast in an idiolect is strongly diagnostic of Wu-hood.

Most of the Wu sub-groups, too, have not been established on innovations, but are typically characterised as showing various constellations of phonological, lexical and morpho-syntactic features. Oujiang, however, is different. Oujiang can be considered a *bona-fide* sub-group because of its unusual development of two proto-Wu tones. Proto-Wu is reconstructed with eight tones, two of which - tones *IVa and *IVb - occurred on syllables with short Rhymes and final stops. Elsewhere in the Wu area, reflexes of these tones are usually short, with a word-final glottal-stop, e.g. Zhenhai [pʰ̚ ʂ] (← proto-Wu *pak 45) 百 hundred; & [bʰ̚ ʂ] (← proto-Wu *bak 23) 白 white (reconstructions from Ballard 1969: 70). Oujiang dialects, however, show an unusual compensatory lengthening, whereby the proto-Wu final stop in *tones IVa/b has been lost, and the original short tone has developed a long, complex pitch. The cognates of the Zhenhai IVa & b tone morphemes hundred & white in the Oujiang dialect of Wenzhou, for example, are [pʰ̚ ʂ 312] & [bʰ̚ ʂ 212].

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>high level</td>
<td>high rising</td>
<td>high falling</td>
<td>mid dipping</td>
</tr>
<tr>
<td>b</td>
<td>mid falling</td>
<td>low rising</td>
<td>low level</td>
<td>low dipping</td>
</tr>
</tbody>
</table>

Table 1: Proto Oujiang tones reconstructed from auditory data.

The published auditory descriptions of Oujiang tones show them to be generally homogeneous (Fu et al. 1985: 20). With the exception of Wencheng, Oujiang dialects have very similar pitch shapes for reflexes of Middle Chinese (and by extension proto-Wu) tones. This makes the reconstruction of most proto Oujiang tones from the auditory descriptions non-controversial. Such a reconstruction is shown in Table 1. The tones show a neat pattern with three pitch contours – level, rising and falling – each with a relatively higher and lower pitch resulting from the depressor effect. Depression is a phonological effect associated with word-initial position in many Wu dialects, whereby the pitch onset of a tone is lowered, so for example a depressed falling tone will have a convex pitch shape (Rose 2002). One interesting fact is the asymmetry in the level and falling tones; the upper and lower paired contours are not reflexes; the level tones come from *Ia & *IIIb; the rising tones from *IIa & *IIib; and the falling tones are from *IIIa & *IIIb. In each case, the a reflex is higher than the b. The two remaining tones – the elongated reflexes of *IVa & *b – also typically form a higher and lower pair. Although the acoustics will uncover some interesting variation in this pattern, it helps to keep it in mind as a useful synchronic template when the individual speakers’ tones are discussed.
3. Acoustic data

Acoustic descriptions of tonal fundamental frequency (F0) and duration are available for seven Oujiang sites: Wenzhou, Xiāngyang, Xiāngyang, Xīngjiā, Xīngjiā, Yongjia, Yuèqíng, and Wéngjiā. Figure 1 shows their location in the Oujiang dialect area as delineated by Wurm et al. (1987). This selection of sites gives a fairly evenly distributed sampling of all the Oujiang area except for the N.W, where the Oujiang dialects border on the Chuqu sub-group of Wu. The data are from my measurements and analysis of some very fine recordings made by W.L. Ballard in the 1980’s and by Zhu Xiaonong in the 1990’s. The procedure used to extract the acoustic tonal features can be found in Rose (1990, 2002).

Figure 2 shows the mean citation tonal acoustics for the seven Oujiang sites. Each panel has the tones of a single speaker (means of between five and ten tokens per tone). F0 is plotted as a function of absolute duration: this is because there are some characteristically big differences in the duration of different tones in Oujiang which would otherwise be obscured. The top three rows show Wenzhou, Xiangyang, and Yongjia, each row with two speakers shown side by side. The remaining four sites are represented by one speaker each, and are shown in the bottom two rows. The tones are identified by their Middle Chinese tonal categories: Ia/b, IIa/b, IIIa/b, and IVa/b. This makes it easier to see which
tones are cognates, an important consideration because there are some potentially confusing differences between speakers and varieties. The tones are also labelled with simple pitch descriptors, like 'mid level'; ‘+D’ indicates depression.

3.1 Wenzhou citation tone acoustics

The two Wenzhou speakers’ tones are shown in the first row of Figure 2: a middle-aged male recorded by Ballard, and a young female I recorded in 2006. The corpus and elicitation is described in Rose (2002), and was exactly the same for both speakers. Both show the three level, rising and falling F0 contours already described: tones Ia & IIb have level pitch; tones IIa & IIb rising; and tones IIIa & Ib falling pitch.

Although both speakers conform largely to the typical Oujiang (OJ) configuration, they differ from the auditory descriptions, and each other, in several ways. (1) The female does not distinguish an upper and lower rising tone: both her IIa and IIb tones are realised with the same low rising pitch/F0 (F0 differences at onset and peak are not statistically significant). Compare this with the male’s separate reflexes of IIa and IIb, which present a typical Oujiang configuration of higher and lower rising tones. Thus the female can be said to only have seven citation tones. (2) Whereas the male has typical long complex falling-rising realizations for his IVa and IVb tones (the falling part being the more salient), the female lacks the rising component for both these tones. They consequently have higher and lower mid falling pitch. This realisation for IVa/b, without a rise, does not occur in any published OJ transcription. (3) The female’s mid falling tonal F0 (Ib), seems to fall from a relatively higher value than the male’s, and this is presumably related to the fact that she already has a mid falling tone: IVb. It can be seen that the F0 of her IVb tone has a very similar level-falling shape to that of the male’s mid falling Ib tone. The female’s Ib tone also has a clearer rising onset perturbation from the depressor effect than the male. In her speech, the depressor was conditioned by the [+/− sonorant] feature of the onset consonant: it was audible on syllables with an obstruent onset, but not on those with sonorants. (4) The speakers differ in the pitch height of their Ia tone. The female’s Ia tone has a upper mid level pitch and corresponds to the auditory transcription in Table 1 of [44]. The pitch of the male’s Ia tone, on the other hand, sounds lower, in the middle of his pitch range, and would be best transcribed as [33]. The difference in F0 height corresponding to this pitch difference is easy to see in figure 2, with the male’s Ia F0 lying considerably lower in his F0 range than the female’s. It is of interest to note that the historically paired Ia and Ib tones, although they have different contours, appear to be linked in the sense that they both involve similar initial targets. This relationship can also be seen in the other speakers.
Figure 2: Citation tone acoustics of speakers from seven Oujiang varieties.
3.2 Xiangyang citation tone acoustics

The two Xiangyang speakers, a brother and a sister, were recorded by Ballard with the same corpus and elicitation as the Wenzhou speakers above. No published auditory descriptions exist, therefore I have given mine for the two speakers in Table 2. It can be seen that both have eight tones, configured in the typical OJ way, with the paired level (or quasi level), rising, falling and complex contours seen in the Wenzhou male. There is not a lot of difference between the speakers in the tones’ pitch features. Their acoustics are shown in the second row of Figure 2. As with Wenzhou, tone IIIa and tone Ib constitute a high and mid-falling pair, the latter also showing a clear depressed onset for both speakers. It can be seen in figure 2 that both speakers have very similar F0 shapes for their IIIa tone, with an F0 which falls through most of their F0 range. They also have similar rising-falling F0 contours for their Ib tone, (the rising portion being a reflex of the depressor) but in slightly differing locations: the F0 peak of the female’s Ib lies slightly higher than that of the male. This is also reflected in their pitch transcription ([231 vs 341]). The Xiangyang speakers are also similar to the majority of the Wenzhou data in showing a pair of rising-pitched tones for tone IIa & tone IIb. There is little between-speaker difference in the pitch of these tones, with the female having just a slightly lower pitch onset to her upper rising tone IIa. Figure 2 shows the F0 of these tones rising into the upper half of the speakers’ F0 range from low (IIb) and either a mid or low-mid location.

Tones Ia and IIIb constitute a further contour pairing, as in Wenzhou, but the details are a little more complicated. Both speakers have a mid or low level pitch with depressed onset for IIIb; their Ia tone has very similar pitch level as their IIIb, but for the female it falls slightly [43], whereas for the male it is level [22]. These relationships can be seen in the F0 data. The male’s Ia and IIIb F0 shapes lie close together in the lower third of his F0 range, with the IIIb shape lying slightly lower and showing a depressed onset. The female’s Ia & IIIb lie slightly higher in her F0 range, with the Ia shape falling steadily to an offset which is similar to her IIIb tone. Her IIIb shape also shows a depressed onset. The fourth pairing obtains, as with Wenzhou, between tone IVa and tone IVb. In both speakers, IVa has a falling-rising pitch, with the falling component more salient, whereas IVb has a low level-rising pitch (thus this – low level rising – is another pitch shape for IVb). The speakers differ in the location of these complex contours, with the female’s tones located slightly higher in her pitch range. The tonal F0 for tones IVa & IVb reflects their pitch details nicely. Figure 2 shows that both speakers show remarkably similar F0 contours for tones IVa & IVb, but with the male’s tones located lower.
3.3 Yueqing citation tone acoustics

The Yueqing speaker’s tones are plotted in the fourth row of Figure 2 and are from a recording made by Zhu Xiaonong of a female speaker. She has eight tones with the usual OJ configuration of paired level, rising and falling contours. Her tonal pitch characteristics correspond largely with the published descriptions, which show the common pairing of level, rising, falling and dipping contours already documented for other sites. Figure 2 shows the two falling F0 shapes for her tones Ib and IIIa; the two rising F0 shapes for tones IIA and Hb, and a high level F0 shape for tone Ia. Her tones agree too, by and large, with the previous speakers. There are two main differences. (1) Zhu’s speaker has a new shape for IVa: a slightly falling pitch [32] in her mid pitch range. The corresponding gradually falling F0 shape for her IVa tone can be easily picked out in figure 2. The published descriptions indicate an expected mid-dipping [323] for this tone, so perhaps this female lacks the final rise. (2) She also has two allotones for the reflex of IIIb. One is the low level pitch version common to most other speakers; the other, plotted with dots in Figure 2, is a depressed low falling tone [121]. There is no obvious segmental conditioning to the split which would predict which morphemes are said with level pitch and which with falling, and it may be that her depressed low falling IIIb morphemes represent a merger in progress with Ib which also has a depressed falling pitch/F0. Against this hypothesis is the fact that, as can be seen from Figure 2, the falling IIIb morphemes have much lower F0, and sound much lower in pitch than the Ib morphemes.

3.4 Qingjiang citation tone acoustics

The Qingjiang speaker’s citation tone acoustics are on the right of the fourth row in Figure 2. They are from recordings made by Zhu of a young male speaker. There are no published auditory descriptions. From Table 3, which contains my pitch transcription of Zhu’s speaker, it can be seen that these data are very similar to those already presented for the other speakers, comprising four contours, each with an upper and lower version. Three of the contours are the same as Wenzhou, Xiangyang and Yueqing, i.e. level, rising and falling. The F0 corresponding to these can be seen in Figure 2. Of note is that the lower level tone IIIb lacks the depressor found in the other speakers, and the high falling tone IIIa has a slight onset shoulder. Of greatest interest is the contour for the fourth pair of tones – tone IVa & tone IVb – which is delayed low rising, and therefore differs from that found in the other speakers who have falling-rising, or falling pitch contours. Comparison with the other speakers’ tone IV F0 shapes

<table>
<thead>
<tr>
<th></th>
<th>Ia</th>
<th>Ib</th>
<th>IIa</th>
<th>IIb</th>
<th>IIIa</th>
<th>IIIb</th>
<th>IVa</th>
<th>IVb</th>
</tr>
</thead>
<tbody>
<tr>
<td>female</td>
<td>[43]</td>
<td>[341]</td>
<td>[24]</td>
<td>[14]</td>
<td>[51]</td>
<td>[233]</td>
<td>[423]</td>
<td>[223]</td>
</tr>
<tr>
<td>male</td>
<td>[22]</td>
<td>[231]</td>
<td>[34]</td>
<td>[14]</td>
<td>[51]</td>
<td>[122]</td>
<td>[312]</td>
<td>[112]</td>
</tr>
</tbody>
</table>

Table 2: Pitch of Xiangyang citation tones.
in Figure 2 shows that the Qingjiang shapes are clearly different. The concentric F0 for this pair of tones starts to rise about a third of the way into the Rhyme (it is likely that their initially falling F0 is a consonantly perturbed onset perturbation and not perceived as a fall in pitch). Although this contour is not found in the other speakers, something obviously like it – [23] and [12] – is noted for IVa and IVb in two of the Wenzhou sources (ZH 1962, CH 1964).

<table>
<thead>
<tr>
<th></th>
<th>Ia</th>
<th>Ib</th>
<th>IIa</th>
<th>IIb</th>
<th>IIIa</th>
<th>IIIb</th>
<th>IVa</th>
<th>IVb</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[33]</td>
<td>[33]</td>
<td>[34]</td>
<td>[14]</td>
<td>[441]</td>
<td>[22]</td>
<td>[223]</td>
<td>[113]</td>
</tr>
</tbody>
</table>

Table 3: Pitch of Qingjiang citation tones.

3.5 Yongjia citation tonal acoustics

The Yongjia tonal acoustics in the third row of Figure 2 are from two sources: a male speaker recorded by Ballard, and the female speaker described by Chao way back in 1928. In his pioneering description of Wu dialect phonetics, Chao transcribed tonal pitch musically, using a sliding pitch pipe. Chao’s musical description of Wu tones remains unsurpassed, and his notation permits easy transformation into F0 values on the simplifying assumption that tonal pitch is a primary function of F0 – see Rose (1993: 215, 216) for similarly reconstructed Shanghai tonal F0. Unfortunately, the duration of the tones cannot be reconstructed, and so the tonal F0 is plotted as a function of equalized duration in Figure 2.

Yongjia tones Ia and Ib show the same values as in the previous sites: upper-mid, or mid, level, and (depressed) mid-falling respectively. The acoustics of these tones in Figure 2 show nothing new. There is a lot more variation between the different sources’ descriptions for the other Yongjia tones, however, and they show somewhat less agreement between the four Yongjia sources than is found for Wenzhou and Xiangyang.

Firstly, some variation is evident for tone IIIa. Two sites (Fu et al. 1995, Cao 2002) show the same [42] high-falling tone as for the other sites’ tone IIIa. However, Chao’s and Ballard’s speakers are a little different. The male speaker Ballard recorded had a fall with a much lower onset: [31], and interestingly Chao also noted a mid fall (with an initial level component) for his speaker in 1928. This lower onset is shown clearly in Figure 2, where the F0 shape for tone IIIa can be seen to fall from a lower position in the F0 range for both speakers.

There is considerable between-source variation in the Yongjia reflexes of tones IIa and IIb. Ballard’s Yongjia speaker, as well as one of the sources (Fu et al. 1985), has the same typical OJ pattern for IIa/b as already observed in all except the female Wenzhou speaker, namely realised as a pair of upper and lower rising pitch tones. The concentric F0 shapes of these two rising tones are easily identified in Figure 2. However, as also can be seen, Chao’s 1928 Yongjia speaker had a very different pattern. His tone IIa had high falling pitch (musically transcribed as 53⁰), and his IIb tone had a mid convex pitch (24¹) which looks...
interpretable as its depressed counterpart (IIb is shown in Figure 2 with a dotted line for clarity). Finally, in the last source (Cao 2005) both IIa and IIb are shown as convex pitch tones. Thus Cao’s notation seems to partake of features of the other three. The difference between rising and falling pitched tones, as demonstrated for Yongjia IIa and IIb, is very big. An extrinsic falling pitch/F0 for these two tones would of course present a problem for reconstruction, given that the other speakers and other sites have rising pitches for IIa and IIb.

Zhengzhang (1995: 359), who is a native Wenzhou speaker, has commented on this variation. He notes the falling pitch for these tones in the earliest description of Wenzhou tones, by Montgomery in 1893, as well in Chao’s 1928 Yongjia description. He says that a falling pitch for these tones occurs if the speaker tries to prolong their intrinsically short duration, as is the case when Wenzhou scholars declaim citation tones. (The falling part would then be construed nowadays as an intonational L# boundary tone.) Thus, he says, Chao’s 1928 transcription was in error in interpretation of the pitch fall in IIa/b. It is difficult to dismiss the inside information from a very experienced native-speaker linguist like Zhengzhang, even if it impugns the field-work of a titan like Chao. And it is worth noting that the Pingyang speaker below shows free variation of a very similar kind, at least for IIa, between a stopped high rising pitch and an unstopped rise-fall pitch.

There are also between-source differences in the description of tone IIIb. For two sources (Cao 2002, Fu et al. 1995), IIIb is the same as for the other sites: a low level [22] tone, pairing tonetically with Ia. In contrast, Chao’s 1928 speaker is shown as having a IIIb tone with a long low falling pitch component, followed by a little rise (2'13") (its reconstructed F0 can be seen clearly in Figure 2). A very similar F0 contour for Ballard’s speaker’s IIIb tone, although it lies a little higher in the F0 range, is unlikely to be a coincidence, and so it seems that both Chao’s and Ballard’s speaker share a different reflex for IIIb than the low level reflex, found in the other descriptions.

Tones IVa and IVb are a little less complicated. Firstly, although three descriptions show the usual pairing between high and low reflexes, Chao says his speaker lacks a reflex for tone IVb, and that it has merged with IIIb. This is why there are only seven tonal shapes shown for his speaker in Figure 2. Apart from this, there is little discrepancy between the descriptions of Yongjia IVa pitch.

3.6 Pingyang citation tone acoustics

The male Pingyang speaker’s tones in the last row of Figure 2 are from a recording made by Zhu. The speaker shows the same basic configuration as the others, and the published sources, with two exceptions. (1) His IIa morphemes showed free variation between a short stopped high rising pitch [34], and an unstopped high rise-fall pitch [342]. The F0 corresponding to these two variants can be easily seen in Figure 2, where the Pingyang stopped IIa tone F0 is plotted
with dots. It is of interest to note, too, that one of the modern sources gives a slightly falling [54] pitch for this tone, while the other cites the typical OJ high rising [45] pitch. We may see here, then, an actual example of the problem mentioned above for Yongjia, whereby the high rising tone is intrinsically short, but in citation form may optionally acquire a falling component. This could be modeled by having a LH(L) melodic component with the first two tones associated with a single mora, and the last, optional, L tone associated with a second mora. One can note further that the source that describes a falling pitch for IIa also describes a falling pitch component [243] for IIb, whereas Zhu's Pingyang speaker does not have this falling component for his IIb. This is perhaps an additional indicator of the optionality of the falling tail for these tones. In any case, the Pingyang example is important historically because it shows a possible source of a falling from a high rising tonal pitch. (2) The Pingyang speaker's IIIb tone appears to have an overall rising F0 contour, but this is due to two factors: a long depressed onset, and a clear final glottal stop, which raised its F0, and pitch, a little at the end. (3) The published sources show mostly non-complex rises for tones IVa/b. Zhu's Pingyang speaker, however, had rather long IV tones, with a substantial level pitch before the final rise. The F0 corresponding to this long level then rising pitch contour can be easily seen in Figure 2.

3.7 Wencheng citation tone acoustics

Wencheng is dramatically different. The acoustic data on the last row of Figure 2, from a recording of a male speaker by Ballard, show far more differences from the other dialects than similarities. There are, to be sure, paired higher and lower level tones, and higher and lower rising tones, but they are not the same reflexes as for the other OJ varieties. The lower rising tone, for example, is a reflex of IVa, not IIb, and the lower level tone is a reflex of Ib, not IIIb. There are not even any falling tones. The only thing that Wencheng clearly shares with the other varieties is a mid level tone for Ia, and a high rising tone for IIa (its final falling F0 is an intrinsic effect from a [?]). The other tones do not appear elsewhere in the OJ group examined. For example, there is no depressed upper-mid level tone (IIb), neither is there the marvelous super-complex reflex for IIIb, which seems to consist of a lower-mid level target, most likely a depressor, followed by a dipping contour. Even the low dipping contour for IVb rises much higher than in the other varieties, and the apparent merger of Ib and IIIa is a very rare phenomenon in Chinese.

<table>
<thead>
<tr>
<th>Ia</th>
<th>Ib</th>
<th>IIa</th>
<th>IIb</th>
<th>IIIa</th>
<th>IIIb</th>
<th>IVa</th>
<th>IVb</th>
</tr>
</thead>
<tbody>
<tr>
<td>33</td>
<td>22</td>
<td>34</td>
<td>23</td>
<td>22</td>
<td>3323</td>
<td>23</td>
<td>213</td>
</tr>
</tbody>
</table>

Table 4: Pitch of Wencheng citation tones.

There are also some discrepancies between the published descriptions of Wencheng tones and Ballard's Wencheng speaker's acoustics. My pitch
transcriptions are therefore given in Table 4. There is, by and large, agreement for tones Ia, Ib, IIa IIIb IVa and IVb. For example, corresponding to Ballard’s speaker’s [33] and [22] for tones Ia and Ib, both published sources describe a bigger contrast in pitch height, between high and low level pitches for Ia and Ib, with one source having a slight final rise: [44(5)] and [11(3)]. For tone IIIb, which has a [3323] pitch, no depressor is transcribed, but otherwise the low dipping contour [312, 313] is the same. The biggest difference is for tone IIIa, which has merged with tone Ib in Ballard’s speaker (the clear difference in duration between IIIa and IIb in figure 2 is due to an intrinsic truncation effect from the initial aspirated consonant in the IIIa morphemes), but which has a separate high rising [334] or dipping [434] pitch in the sources. There is also clearly a difference in the IIb realisations. Ballard’s speaker has a depressed mid-level pitch [233], whereas one of the sources shows a mid level pitch without the depressor, and the other shows a mid dipping [324] pitch.

Despite these differences, Wencheng still shows elongated reflexes of proto-Wu *IVa/b tones, and thus must count as a synchronic Oujiang variety. However, whether it is diachronically an Oujiang variety that has undergone considerable subsequent changes independently of the other Oujiang dialects; or whether it is originally not Oujiang but has changed under the latter varieties’ influence, remains to be determined.

4. Quantifying tonal homogeneity in Oujiang

The sections above have described the tonal acoustics of speakers from seven different sites in the Oujiang sub-group. It is clear that some tones show very little variation from site to site. The reflex of Ia has mid or upper-mid level pitch in most sites, for example, and the reflex of IIb has a low rising pitch. How can one quantify this variation, however? One possibility is to measure the degree of homogeneity of the different speakers’ tones by how well they cluster after normalisation. Normalisation is a mathematical procedure for getting rid of the individual characteristics in the speech acoustics to leave a linguistic-phonetic residue - what is common to the variety (Rose 1987, 2000). As well as speech-dependent information, a particular tonal F0 value contains speaker-dependent information. It is determined not only by the particular tonal target a speaker is aiming for, but also by the length and mass of their vocal cords.

A very simple normalisation method is to first find a speaker’s mean and standard deviation F0 over all their tones, and then transform their individual F0 values by subtracting them from the mean and dividing by the standard deviation. This is called a z-score normalisation, after the equivalent transform in statistics. This method maximally eliminates variation in F0 due to between-speaker differences in cord mass and length. Speakers’ normalised F0 values can then be averaged to derive a representation of the variety, as well as a quantifi-
cation of the amount of variation in the variety, which can be modelled as a function of the standard deviation around the mean normalised curve.

Now, if it can be shown that the amount of variance around the mean normalised F0 shape for a given tone across the Oujiang varieties is the same as that found for a comparable tone in a single dialect, it can be claimed that the tones from the different Oujiang sites are as homogenous as in a single dialect. This idea can be demonstrated using two tones from the Wu dialect of Shanghai, a dialect for which there is data on 18 speakers’ citation tones (Rose 1993, Zhu 1999), and for which there is consequently a reasonable estimate of the within-dialect between-speaker variance in normalised tonal F0. Shanghai (SH) has five tones, one of which has a high falling pitch (Tone 1) and another a low rising pitch (Tone 3). These sound very similar to the high falling IIIa and low rising IIb tones of Oujiang, and it was these two tones - high falling in SH and OJ, and low rising in SH and OJ - that were compared by z-score normalisation.

Figure 3 shows the z-score normalised values for the four tones plotted against equalized duration. The thin dotted lines represent the normalised tones for the 18 Shanghai speakers and the thick interrupted line their mean normalised value. The thin solid lines represent the normalised OJ tones and the thick solid line their mean normalised value. It can be seen that the OJ falling and rising tones lie less than a standard deviation above the corresponding SH tones, but that otherwise the two varieties’ normalised contours are remarkably similar. (The vertical displacement is probably because OJ has more tones distributed in the lower F0 range than SH, and consequently a lower mean F0, leading to higher normalised values.) The standard deviations at each of the sampling points are given in Rose (2004: Table 2), which showed generally very
little difference between OJ and SH in these tones (the differences are of the order of one tenth of a standard deviation, which is exiguous).

Thus, in spite of the fact that the Oujiang descriptions span some twenty years, it is possible to claim considerable homogeneity for these two tones in the Oujiang dialects examined. The same could probably be demonstrated for all the other OJ tones except reflexes of IVa/b, which, as the results of the acoustic analyses above have shown, do vary over the Oujiang area.

5. Acoustic reconstruction of proto-Oujiang tones

An important point, and one that does not seem to have been realised before, emerges from the demonstrated homogeneity of the normalised tonal acoustics in the Oujiang area. This is that they can be interpreted in a completely new light, namely as historical constructs. The normalised tonal F0 contours in Figure 3, together with a quantification of their variation, constitute the best estimate of the proto-Oujiang tones from which the modern forms derive. As an example, Figure 4 shows the acoustic reconstruction of two proto-Oujiang tones. On the left is the high falling proto-Oujiang tone IIIa, based on F0- and duration-normalised values from all varieties except Yongjia (which, it will be recalled, has a mid falling IIIa) and Wencheng (which has a mid level IIIa). The thick line shows the z-score normalised mean F0, which falls from about two standard deviations above the mean to about 1.5 sds below it. This is the best estimate of the F0 of proto-Oujiang IIIa and, together with the normalised duration quantifies what is meant by the reconstructed “high falling” for IIIa of Table 1. The slightly thinner lines mark one standard deviation corridors above and below the reconstructed mean, and give a crude idea of the variation around it: something that is not really possible with auditory based reconstructions.
more sophisticated model of the variation must take into account the fact that values at the different sampling points are not independent, and give the probability of a particular set of values conditional upon a starting value for the tone, and the mean contour. A further sophistication would be to model the reconstructed F0 time-course with discrete cosine transforms or polynomials.

On the right of Figure 4 is the reconstructed mid falling proto-Oujiang tone 1b, based on F0- and duration-normalised values from all varieties except Wencheng, which had a lower mid level reflex for 1b (a standard deviation corridor is not shown). The reconstructed F0 shows that the tone is not straightforwardly mid falling, but has an initial, probably depressed, shoulder. (This is notated only in a few of the published sources, most of which describe the OJ 1b reflex as [31]). It also falls from higher than the middle of the range than implied by the synchronic acoustic descriptors.

Of course, one cannot just mechanically reconstruct across the board from mean acoustic values: this approach would not work with the OJ reflexes of IVa and IVb, which have been shown to differ considerably over the area (see Figure 2). There needs to be demonstrated homogeneity, and an indication that a more abstract reconstruction is not indicated, based on what we know about the way tones change (Brown 1985).

6. Summary and Envoi
This paper has described the citation tones in the small Wu Chinese sub-group of Oujiang in order to show how proto tones can be reconstructed from their modern acoustics. The key to this is a novel interpretation of the results of normalisation – the mean normalised values and the variation around the mean – as historical constructs.

The proto-Oujiang tones thus estimated may then be compared with similarly reconstructed cognates from other Wu dialect sub-groups for which there are modern acoustic data. It can be appreciated that the quantified aspect of this acoustic reconstruction approach, where each level of abstraction is explicit, and one operates with continuous variables, opens up the possibility of a whole new level of precision in historical tonological reconstruction and comparison.

Harold was the one to meet me at Canberra Airport in 1979 when I arrived to start my linguistics career in Australia. Then, the time-consuming process of extracting, measuring and processing tonal acoustics, (bent over the now antiquated analog spectrograph and spectrogram with their ever-present smell) meant the kind of work demonstrated in this paper, where acoustic data from 35 speakers have been used, was not possible. It is now. However, the ease with which we can now process data, and the new insights we win therefrom, should not blind us to the fact that computers will not supplant a prior, orthodox, linguistic analysis of the kind that Harold has always championed.
References


ISSUES IN THE MORPHOLOGICAL RECONSTRUCTION
OF PROTO-MON-KHMER

PAUL SIDWELL
Australian National University, and
Centre for Research in Computational Linguistics (Bangkok)

1. Introduction

The honouree of this volume is famed for his passionate interest in morphology, morphological reconstruction and his championing of etymological methods in reconstruction. In this context, it gives me great satisfaction to honour Harold by discussing some issues in the morphological reconstruction of Proto-Mon-Khmer, and setting forth a reconstruction of various Proto-Mon-Khmer affixes.

The Mon-Khmer family belongs within the Austroasiatic phylum of South and Southeast Asia. The morphological systems of these have been described to greater and lesser extents, and comparative Mon-Khmer studies have been conducted - with stops and starts - for more than century already. It is now clear that Proto-Mon-Khmer bound morphology was characterised (at least) by prefixing and infixing, employed for purposes connected with derivation and verbal semantics. There was also much creativity in the formation of expressive lexicon, with much use of reduplication and assimilation. The latter are not discussed here; instead, it is only affixes that will be examined.

Approaches to Mon-Khmer morphological reconstruction have faced various hurdles; one being that since the foundational studies of Schmidt (especially 1905, 1906) the issues have tended to be seen through the lens of the Austric hypothesis. As formulated by Schmidt, Mon-Khmer and Austronesian are sisters within an Austric macro-phylum, and although many lexical parallels have been adduced, the most important evidence has always been seen to reside in the morphology, specifically in among a small subset prefixes and infixes (see Reid 2004, 1999, 2005 for recent views). Thus, much of the modest relevant research effort has effectively presumed that the Mon-Khmer-Austronesian morphological parallels are archaic and reflect reconstructable morphemes. Such a top-down approach to reconstruction has its own merits, but in the present case we are dealing with very old languages families whose affilitation is not generally accepted, and may well be impossible to demonstrate. In this

1 The author gratefully acknowledges the support of the Mon-Khmer Languages Project by the National Endowment for the Humanities. Any views, findings, conclusions, or recommendations expressed in this publication do not necessarily reflect those of the NEH. Thanks are also due to the Department of Linguistics, Max Plank Institute for Evolutionary Anthropology (Leipzig). Finally the author would like to thank one of two anonymous reviews for their helpful comments and suggestions.
paper 1 advocate a more bottom-up approach, emphasising the comparative analysis of forms which occur within a well-established genetic grouping, and which for formal and distributional reasons are likely to be archaic, and not borrowed.

2. Methodological preliminaries

Mon-Khmer languages have a distinctive morphological typology, characterised by prefixation and infixation of mono- and sesquisyllabic stems. The preferred phonological shape of unaffixed Proto-Mon-Khmer stems (based on Shorto 2006) included:

- Monosyllabic: C(R)V(V)C, where R was an optional medial consonant, one of the set /r, l, j, w, h/, and:
- Sesquisyllabic: CCV(V)C, in which initial clusters of stops (oral and nasal) were permitted, with a weak vowel epenthesis after the first C, creating a light initial syllable or ‘presyllable’.

Initial sequences of two or more consonants could be built up by affixation, with a likely preference for initial sequences of three consonants in morphologically complex words, affording presyllables a coda if the stem was light (CVC) (Anderson (2004) has a similar formulation he calls the “bimoraic constraint”). Thus, some affixes had mono- and bi-consonantal allomorphs, applied according to the weight and/or sonority contour of the stem.

Syntactically the languages are isolating, so that morphological processes are largely restricted to derivational functions, although one finds quite a variety of idiosyncratic developments – for example the case marking of pronouns of (Katuic) Ta’oih (Solntseva 1996) and Pacoh (Alves 2006) are transparently innovative.

Not withstanding the above, speakers of many Mon-Khmer languages have more or less abandoned affixation as a morphological strategy, typically in the context of phonological changes that have restructured the lexicon into monosyllables. In these languages, such as Vietnamese, Nyaheun, and others, compounding has emerged as an important new word-formational mechanism (probably for both internal structural reasons, plus contact with compounding languages such as Chinese, Thai, Lao).

---

2 One of the reviewers of this paper suggested that I should consider the possibility that various items of Mon-Khmer bound morphology may be ancient borrowings from Austronesian. My intuitive response is to dismiss the suggestion as unlikely and difficult to test, although I must admit that it is not impossible. Probably such a hypothesis cannot be tested properly without first conducting far more extensive bottom-up studies of Austroasiatic morphology.
Some reconstructions of affixal forms have already been suggested. Relying especially upon Nicobarese evidence, Reid (1994) offers the following Proto-Austroasiatic affixes:

Causatives: *pa/-ap-, *ka-
Agentives: *-um-, *ma/-am-
Instrumental: *-an-, *-in-
Objective: *-a

The last of these, the Objective suffix, is especially doubtful. The form does not occur elsewhere in Austroasiatic; the Nicobarese “suffixes” are analysed as clitics by Braine (1970), and should be considered innovations. The other affixes reconstructed by Reid do have cognates in other branches, and therefore do represent ancient morphemes. However, placing weight on Nicobarese led Reid to reconstruct these affixes with CV/VC structure, even specifying the quality of the vowels. This may be correct for Nicobarese, but it is quite possible that they were purely consonantal in Proto-Mon-Khmer, and may have been similarly so in Proto-Austroasiatic. Compare with Anderson (2004) who, giving special weight to Munda forms, reconstructs the Proto-Austroasiatic causative as a prefix *'b- with CVC stems and as an infix *-'b- with CCVC stems.³ The uncertain place of Nicobarese in the Austroasiatic family, and the lack of a well-developed comparative reconstruction of Proto-Nicobar, means that the use of such evidence for wider comparative analysis is extremely problematic. Clearly there are issues of segmental and prosodic phonology that are far from solved at the Proto-Austroasiatic level, and for now I prefer to focus on reconstruction at a lower phylogenetic level.

Taking the bottom-up approach with the Mon-Khmer family, it is convenient to begin by identifying at least two branches of the family that maintain evidence of rich morphological systems, and yet are widely separated geographically. Guided by these principles we would turn immediately to the Aslian languages which, located on the Malay Peninsula, have escaped the areal effects of contact with Thai/Lao that affected so much of the mainland (pointed out by Diffloth 1979). Fortuitously, the Aslian languages show a morphological complexity which has been well described, and in this study, I draw upon Jahai (Burenhult 2005) and Semelai (Kruspe 2004).

At the geographical extreme from Aslian are the Northern Mon-Khmer languages: Khasi, Palaungic and Khmuic. The first two are not promising; Palungic languages have little morphology to offer, and “morphological processes in Khasi are almost non-existent” (Rabel-Heymann 1976:981). However, Khmuic is a promising candidate for morphological comparison, since documented varieties show considerable morphological complexity (although only modest morphological productivity). Thanks to apparently very

³ I assume this notation represents *b/-b-, since Anderson describes the sounds as “glottalised”.
conservative phonology, extensive complex forms persist, which aid in the historical analysis. For this study I draw upon the Kammu documented by Svantesson (1983) and the Khmu of Suwilai (2002).  

Svantesson anticipates the selection of Aslian and Khmuic as criterion branches for morphological reconstruction:

It is believed that Proto-Mon-Khmer (and Proto-Austroasiatic) had a rich system of derivational affixes (Pinnow 1960, Diffloth 1976), and some of the affixes found in Kammu (e.g. causative /p-/ are widespread in different Austroasitic subgroups. Aside Munda, the Austroasiatic subgroups having the most developed morphological systems are probably Aslian (see e.g. Benjamin 1976a) and Nicobarese. In Vietnamese there is almost no derivational morphology at all, and in most other groups, the situation is similar to that in Kammu, i.e. affixation is common but not completely productive. (Svantesson 1983:73)

There are also other factors, in addition to morphological complexity, that point to both Khmu and Aslian being generally rather conservative as Mon-Khmer languages go. Many Mon-Khmer languages that have been in contact with Tai have undergone a distinctive type of phonological restructuring, in which voiced stops devoiced, and the vocalism was restructured to double the number of contrasts, either by acquiring tones, breathy phonation, a plethora of diphthongs, or a combination of two or three of these factors. Typically the Proto-Mon-Khmer imploded stops continue as a distinct series, often losing implosion to become the ‘new’ voiced series (as happened in Thai and Lao). This did not occur in Aslian languages, which maintain the old voicing contrast (in fact merging the Proto-Mon-Khmer imploded stops into the voiced series), instead many of them lost the length contrast in the vowels, apparently under the influence of Austronesian. Many of the Khmic languages, although spoken in the zone of prolonged Tai contact, have retained the Proto-Mon-Khmer voiced series (also merging the Proto-Mon-Khmer imploded stops into the voiced series as did Aslian), and have resisted vowel restructuring. Suwilai Premsrirat’s Khmu Chuang is such a conservative variety.

Svantesson (1983:83) lists some 27 distinct forms of “affixators” in Kammu. Several of these are transparently innovative forms or specifically related to expressive formations, and these are removed from consideration for now. The remaining forms are reorganised according to function and listed as follows (prefixes indicated by x-, infixes indicated by -x-, allomorphs by /):

- **Verb forming:** gn-, kn-, km-
- **Resultative:** tl-, sl-, hn-
- **Causative:** p-, pn, -m-
- **Reciprocal:** tr-
- **Nominalising:** rn-, -mn/-r/-/n(d)-, -n-, -mn-, sy-, sr-, tr-

---

4 Henceforth forms noted as ‘Kammu’ are all from Svantesson (1983), and those noted as ‘Khmu’ are from Premsrirat (2002).
Prefixes are straightforwardly added to the left edge of stem, while infixes are inserted to the right of the leftmost consonant. Speakers do not readily tolerate initial sequences of more than three consonants, so there are consequent limits on the amount of affixation that may applied, and in some cases there is deletion of segments to accommodate affixation.

A remarkable phonological concomitant of Kammu affixation is what Svantesson calls “coda assimilation”. Affixation tends to create presyllables, and in certain cases, the rhymes of these presyllables (often underlyingly /n/ or /r/) assimilate to the coda of the mainsyllable, e.g. (from Svantesson 1983:58):

- Nominalizing rn-:
  - hāc ‘to carve’ > rchāc ‘carving’
  - wāj ‘to swim’ > rjwāj ‘fin’

- Causative pn-:
  - kāk ‘bent’ > pkkāk ‘to bend’
  - rōos ‘angry’ > pśōos ‘to make somebody angry’

The situation in Aslian is more complex; in particular, there is a far more elaborated version of coda assimilation, which Aslian specialists variously call “coda copying”, “reduplicative infixation”, “incopyfixation”. The difference is that instead of a well-formed segment assimilating to the mainsyllable coda, it is an underspecified or epenthetic segment (or segments) that assimilates. Consider the imperfective aspect marker in Jehai (from Burenhult 2005:94-5): for monosyllabic stems the prefix consists of a copy of the mainsyllable onset plus a copy of the mainsyllable coda, e.g.:

- sut ‘to sob’ > stsut ‘sobbing’
- we? ‘to exist’ > w?we? ‘existing’

For sequisyllabic stems the allomorph is simply a copy of the mainsyllable coda inserted to create a minorsyllable rhyme, e.g.

- tōoc ‘to ask’ > tcooc ‘asking’
- rwis ‘to cut grass’ > rswis ‘cutting grass’

In disyllabic stems the minorsyllable rhyme gains a coda:

- turh ‘to tap poison’ > tuhrh ‘tapping poison’
- gulém ‘to carry’ > gulem ‘carrying’

---

5 This last term coined by Matisoff (2003:28) with characteristic wit.
6 Disyllabic stems are not a general feature of Mon-Khmer languages, and in Aslian can be assumed to have arisen under Austroasiatic influence. Many in fact are obvious Malay loans.
Note that the allomorphy maintains the (historically?) preferred word-shape of morphologically complex forms. This allomorphy also means that the full set of Jehai affixes is difficult to summarise abstractly. Here is table of verbal derivation from Burenhult (p.94):

<table>
<thead>
<tr>
<th>Derivational morpheme</th>
<th>Monosyllabic</th>
<th>Sesquisyllabic</th>
<th>Disyllabic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspect/Aktionsart</td>
<td>/cip/</td>
<td>/kjen/ ‘to listen’</td>
<td>/gulem/ ‘to carry’</td>
</tr>
<tr>
<td>Imperfective</td>
<td>cp-cip</td>
<td>k-jen</td>
<td>g+cp-lcm</td>
</tr>
<tr>
<td>Progressive</td>
<td>b-cip</td>
<td>b-kjen</td>
<td>b-gulem</td>
</tr>
<tr>
<td>Iterative</td>
<td>lp-cip</td>
<td>l-kjen</td>
<td>l-gulem</td>
</tr>
<tr>
<td>Continuative</td>
<td>cip-cip</td>
<td>kjen-kjen</td>
<td>gulem-gulem</td>
</tr>
<tr>
<td>Distributive</td>
<td>cip-cip</td>
<td>k-ip-jen</td>
<td>g-im-lcm</td>
</tr>
<tr>
<td>Reciprocal</td>
<td>ca-cip</td>
<td>k-ar-jen</td>
<td>-</td>
</tr>
<tr>
<td>Affix /m/</td>
<td>mp-cip</td>
<td>m-p-jen (?)</td>
<td>mm-lcm (?)</td>
</tr>
<tr>
<td>Causative</td>
<td>pp-cip, pi-cip, pr-cip, tr-cip</td>
<td>k-ri-jen</td>
<td>g&lt;ri-lcm</td>
</tr>
<tr>
<td>Nominalisation</td>
<td>np-cip</td>
<td>k-nj-jen</td>
<td>g&lt;n&gt;ulem</td>
</tr>
<tr>
<td>Collective plural</td>
<td>—</td>
<td>k-ra-jen</td>
<td>g&lt;re&gt;ulem</td>
</tr>
<tr>
<td>nominalisation</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Relative</td>
<td>t-cip (?)</td>
<td>t-kjen (?)</td>
<td>t-gulem (?)</td>
</tr>
</tbody>
</table>

Figure 2: Paradigm of verbal derivation in Jahai (Burenhult 2005:94, Table 4.5)

So what is the historical context of Aslian incopyfixation? Matisoff suggests that it is not especially unique:

Such processes are still apparently productive in Nancowry Nicobarese (Radhakrishnan 1970:149ff). A few Proto-Waic reconstructions look like they have fossilized incopyfixes, e.g. *rmhom ‘heart, mind’ [Diffloth 1980, p. 148]. Delcros’ Khmu dictionary (1966) has forms like rtjut ‘horrible’ fnbaap ‘immense’. (fn.150, p.30)

My theory is that the origin is phonological, and was connected to the structure of sesquisyllables and associated phonotactic constraints. I expect that the Kammu pattern is more-or-less indicative of the Proto-Mon-Khmer state of affairs; coda assimilation was not necessarily especially productive, and many of the forms still in use are fossils, the accumulated detritus of centuries of sporadic formations. On the other hand, the Aslian languages, once separated from other Mon-Khmer speakers, went on to regularise and elaborate, with innovations such as the phonologically underspecified morphemes. At the earliest stage (at the Proto-Mon-Khmer level?) speakers inserted epenthetic rhymes into consonant clusters, and with euphonic/expressive motivations began to rhyme these with the mainsyllable codas. This is quite consistent with the sorts of word game that are common in Southeast Asia (e.g. swapping

---

7 I haven’t (yet) found Aslian-type epenthetic coda copying described in respect of Munda languages, although expressive reduplication is common to them.
rhymes within couplets of monosyllables is popular in Vietnamese), and expressive reduplication is described in detail for many Mon-Khmer languages.

In the following section, I will present evidence for the reconstruction of a selection of Proto-Mon-Khmer affixes. The presentation is far from comprehensive, due to space considerations, and should be taken as illustrative rather than authoritative.

3. **Infixed**

3.1 **Mon-Khmer Infixed**

Infixedation is an important word-formational process in Mon-Khmer languages. It has been suggested that generally infixedes can be analysed as a type of prefixation, a position that Kruspe takes in her analysis of Semelai, arguing that the infixedes are effectively prefixed to the prosodic head (2004:65-6). Looking for an historical explanation one may observe that among Mon-Khmer languages the most common infixedal segments are nasals, while nasal onsets in sesquisyllabic stems are rare (or nonexistent). By contrast, nasal codas are common in all types of mainsyllables - the proportion typically approaches 50% (except in those languages in which hardening of finals is a regular development). A similar but weaker tendency applies to liquids. This suggests that during a specific historical stage of Mon-Khmer (pre-Proto-Mon-Khmer ?) nasals and liquids were not permitted as leading onset segments of sesquisyllables. Whatever the particulars of this constraint, it may have been associated with a metathesis of sonorant initials, i.e. NC- > CN-. Perhaps we see an echo of just such a process in the behaviour of Jehai affixes (see examples under 3.3).

Another point to consider is whether there is an historical relation between /n, m, r/ as the typically infixedal segments, and the implosives /?/. There is a remarkable symmetry in the fact that the implosives are only reconstructable for Proto-Mon-Khmer as prevocalic segments, but never preconsonantally (in contrast, say, to contemporary Bahnar, see Banker et al. 1979). A general lenition of preconsonantal implosives, followed by metathesis (**CVC > **NCVC > *CNVC) would account for the existence and phonological form of Austroasiatic infixedes. It this context it is intriguing that Anderson’s (2004) reconstruction of the Munda causative suggests an implosive articulation.

3.2 **Nominalising** *-n-

The nominalising *-n- is distributed widely, although in some languages it is no longer productive, or is restricted in its productivity (e.g. only weakly productive in Khu mic). It appears that essentially any verb could be nominalised with *-n-. It is not clear if its usage was originally restricted (to perhaps
instrumentals or resultatives) and subsequently generalised, or the other way around.

**Kammu (instrumental):**

\[ \begin{align*}
    \text{kóh} & \quad \text{‘to cut’} & \text{knóh} & \quad \text{‘cutting board’} \\
    \text{kèêp} & \quad \text{‘to fasten belt’} & \text{knèêp} & \quad \text{‘quiver’}
\end{align*} \]

**Khmu (resultative):**

\[ \begin{align*}
    \text{peêr} & \quad \text{‘to slice’} & \text{pnêêr} & \quad \text{‘sliced pieces’}
\end{align*} \]

There are vestiges in other Northern Mon-Khmer languages, such as Khasi (data from Henderson 1976):

**Khasi:**

\[ \begin{align*}
    \text{shong} & \quad \text{‘to sit, dwell’} & \text{shnong} & \quad \text{‘place, village, town’} \\
    \text{sait} & \quad \text{‘to wash (vegetables)’} & \text{snait} & \quad \text{‘strainer’}
\end{align*} \]

In Aslian the affix is accompanied by a copy of the mainsyllable coda:

**Jahai:**

<table>
<thead>
<tr>
<th>CVC stem</th>
<th>CVC stem</th>
<th>CCVC stem</th>
<th>CCVC stem</th>
</tr>
</thead>
<tbody>
<tr>
<td>sam</td>
<td>to hunt</td>
<td>nmsam</td>
<td>‘act of hunting’</td>
</tr>
<tr>
<td>tboh</td>
<td>‘to beat’</td>
<td>tnhboh</td>
<td>‘act of beating’</td>
</tr>
</tbody>
</table>

**Semelai:**

<table>
<thead>
<tr>
<th>CVC stem</th>
<th>CVC stem</th>
<th>CCVC stem</th>
<th>CCVC stem</th>
</tr>
</thead>
<tbody>
<tr>
<td>c’or</td>
<td>‘treat with fire’</td>
<td>nrc’or</td>
<td>‘firing’</td>
</tr>
<tr>
<td>smaŋ</td>
<td>‘to request’</td>
<td>snmaŋ</td>
<td>‘request for sth’</td>
</tr>
</tbody>
</table>

The infix shows considerable allomorphy in various other Mon-Khmer languages, essentially relating to restrictions on the occurrence and collocations of /\text{n}/. E.g.:

**Katu:**

\[ \begin{align*}
l > a /_l & \quad \text{pleh} \quad \text{‘to turn on road’} & \text{paleh} \quad \text{‘crossroads’} \\
l > a /_r & \quad \text{pruung} \quad \text{‘to blow fire’} & \text{paruung} \quad \text{‘pipe to blow fire’}
\end{align*} \]

3.3 **Nominalising Agentive *-m-*

Aslian provides evidence of a labial nasal agentive affix, e.g.:

**Jahai:**

\[ \begin{align*}
    \text{wî} & \quad \text{‘left’} & \text{mwî} & \quad \text{‘left-handed person’} \\
    \text{tem} & \quad \text{‘right’} & \text{mmtem} & \quad \text{‘right-handed person’}
\end{align*} \]
Semelai:

- $kba$ 'fish with pole'
- $kmba$ 'fishing pole'
- $sdar$ 'to remember'
- $smrdar$ 'a good rememberer'

In Old Khmer we find a rare 'agentive' $-m-$, e.g.:

Old Khmer:

- $s$ 'to reside, preside'
- $sm$ 'in residence; to officiate'

A similar morpheme is found in Mon, and Diffloth (1984) reconstructs a Proto-Monic *$-m-$ agentive. However, the closest match in Kammu/Khmu is an unproductive complex infix $-mn-$ with an instrumental function:

Kammu:

- $tis$ 'to support'
- $kminis$ 'support'
- $kées$ 'to fence in'
- $kminées$ 'fence'

Khmu:

- $koh$ 'to chop'
- $kmnoh$ 'chopping block'

Perhaps the Khumic $-mn-$ is a remnant of a combination of two infixes, the labial reduced to a remnant of the Proto-Mon-Khmer agentive.

3.4 **Expressive of repetitiveness/numerousness $-*l-$**

Matisoff (2003:27) notes that:

Jah Hut has an $-l-$ infix that occurs in expressives to indicate intensity or great numbers (e.g. $saïbyr$ 'sight of dishevelled hair', $saïbyr$ 'long, abundant dishevelled hair'), and which also occurs fossilised in the form $-l-$ in names of animals characterised by rapid, jerky movements ($klbak$ 'butterfly', $hlid$ 'cockroach', $kljeh$ 'kind of small bird').

One finds similar forms in Khmu:

Khmu:

- $klba$ 'wave'
- $klia?$ 'fluctuating movement'
- $khwa?$ 'echo, reverberating sound'
- $kli$as $kli$as 'cluttered'
- $tloook$ 'clf. for cluster of fruit'
- $tluatl$ 'cluttered (small things)'

And in other Mon-Khmer languages, e.g. compare Sre (South Bahnaric) $klleep$ 'centipede' with Khmer $klaep$ 'centipede'. This affix may also be the source of the Khmer iterative infix $-L-$, which derives both nouns and verbs (discussed by Jenner & Pou 1980-81 p.xlix), e.g.:
Khmer:

\begin{align*}
\text{toom} & \quad \text{‘to bear’} & \text{troom} & \quad \text{‘to support patiently’} \\
\text{kik} & \quad \text{‘to hold against side or under the arm’} & \text{klik} & \quad \text{‘armpit’}
\end{align*}

3.5 \textit{Nominalising instrumental *-p-}

This particular infix is very archaic; it is no longer productive in most Mon-Khmer languages, so discussions are absent from descriptive grammars, and reconstruction is dependent on comparative and philological evidence (hence the reliance on Mon and Khmer), Bauer (1988), comparing Old Mon and Old Khmer reconstructs Proto-Mon-Khmer *-p- ‘nominaliser’.

Reflexes of *-p- are well attested in Khmer, e.g.:

\begin{align*}
\text{Khmer:} & \quad \text{dal} & \quad \text{‘to pound’} & \text{tbal} & \quad \text{‘mortar bowl’} \\
& \quad \text{reco} & \quad \text{‘to plait’} & \text{rbaen} & \quad \text{‘basketry’}
\end{align*}

Cognates of the infixed “mortar” etymon are widespread in Mon-Khmer, e.g. (Shorto (2006) entry 1757): Praok (Wa) po, Rumai (Palaung) mpae, Katu tapal, Nyaheun dwaw, Sre mpal, Semang pel and others showing the effects of regular historical sound changes.

Old Mon (Shorto 1971:xxiv) also possessed the same infix, which by regular sound change took the form -w-, e.g.:

\begin{align*}
\text{Old Mon:} & \quad \text{til} & \quad \text{‘to plant’} & \text{twil} & \quad \text{‘cultivable land’} \\
& \quad \text{pa} & \quad \text{‘to do’} & \text{puwa} & \quad \text{‘deed’}
\end{align*}

4. \textit{Prefixes}

4.1 \textit{Mon-Khmer prefixes}

Comparing prefixes across Mon-Khmer languages is extremely problematic, as the published descriptions typically treat phonologically disparate forms as allomorphs, without further analysis or explanation. In this short study, I will focus on just a few in an attempt to tease out the historically underlying forms.

4.2 \textit{Causative *p-}

This prefix has attracted the attention of various scholars, especially as it is well attested in Munda, where it shows an infixed allomorph that is paralleled in Mon-Khmer (see Anderson 2004 for a recent discussion). Within Mon-Khmer the most unmarked reflexes have the form of a voiceless bilabial stop, hence the reconstruction as *p-. 
In Kammu the affix is typically attached to intransitive and stative verbs. It has several allomorphs with specific phonotactic distributions: p-, pn-, -m-. In the following examples note that examples of pn- are subject to coda assimilation:

Kammu:
- p- haan 'to die' phaan 'to kill'
- p- kaa 'to climb' pkaa 'to cause to go up'
- pn- poool 'to hang (intr.)' pipoool 'to hang (tr.)'
- pn- + coda assim. pag末端 'drunk' pypag末端 'to make drunk'
- -m- skar 'straight' mkar 'to straighten'
- -m- rken 'stretched' rnkKen 'to stretch'

In Jahai this prefix has an obligatory augment, either a presyllable coda copied from the mainsyllable coda or a vowel /i/ or syllabic consonant /r/. While it is applied to both transitive and intransitive verbs, it appears from Burenhult’s data that intransitive bases predominate, and adding the prefix does not especially increase transitivity, regardless of the base, but may simply indicate "rather specific and restricted meanings" (2005:106).

Jahai:
- p- + copy coda gej 'to eat' pjagej 'to feed'
- pi- muc 'to eat' pimuc 'to feed'
- pi- + restricted kap 'to bite' pikap 'to tear apart with teeth'
- pr- hit 'tremble' prhit 'cause smwn. to tremble'

For Semelai, Kruspe describes a causative prefix par- which has:

- various allomorphs due to dissimiation of the coda
- an allomorph -r- where the stem already has a full presyllable
- an infrequent allomorph p- which Kruspe assumes is borrowed, since it is difficult to suggest motivation for its distribution

The prefix is applied to transitive, intransitive and stative verbal bases. Kruspe (her table 5.4) also lists it as occurring with nouns, but this is not exemplified in her text.
Semelai:
par- tut ‘to blow’ partut ‘cause smth. to blow’
pan- ca ‘to eat’  panca ‘to feed’
pa- lcm ‘pleasant’  palem ‘to cheer up smwn.’
-r- hampon ‘lightweight’ harpon ‘to lighten’
p- jįji? ‘dirty’  pįji? ‘to dirty smth.’

For Car Nicobarese, Braine (1970:162) describes a causative prefix ha-, with allomorphs ?a- occurring with laryngeal initial CVC stems, and mi-, -um- with CVCVC stems. Word initial /h/ goes back to *p by regular sound change in Nicobarese, so it is straightforwardly identified with the Proto-Mon-Khmer *p-prefix.

Car:
ha- rác ‘hot’ haráč ‘to heat’
mi- huríj ‘black’ miríj ‘to make black’
-r- sirój̄ká ‘ashamed’ sumriój̄ká ‘to shame’

In Katu the prefix can be affixed to both transitive and intransitive verbs, and to stative verbs or even nouns to create a causative verb with the meaning “activity that results in state/creation of thing”. Examples:

Katu:
verb trans. cha ‘to eat’ pacha ‘cause to eat’
verb intrans. mut ‘to run’ pamut ‘cause to run’
noun atet ‘rust’  patet ‘to make to rust’

On the strength of the evidence presented here, one could suggest that the pr/par- forms attested in Aslian have been influenced by the par- causative prefix of Austronesian (via Malay?), so that reconstruction of the /r/ augment to Proto-Mon-Khmer is not indicated. However, the augmented forms have near parallels in Khmu, consistent with a prosodic preference for presyllables with codas affixed to CVC roots.

4.3 Reciprocal *tr-

In Kammu the prefix tr-, with allomorph tj-, forms reciprocal verbs, e.g.:

Kammu: pók ‘to bite’ tıpók ‘to bite each other’
  ràac ‘to scratch’ tjràac ‘to scratch each other’

Costello (1966:70) documents a Katu ta- reciprocal prefix, e.g.:
Katu:

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>kap</td>
<td>‘to bite’</td>
</tr>
<tr>
<td>takap</td>
<td>‘to bite each other’</td>
</tr>
<tr>
<td>leng</td>
<td>‘to kill’</td>
</tr>
<tr>
<td>taleng</td>
<td>‘to kill each other’</td>
</tr>
</tbody>
</table>

This prefix, with effectively the same meaning, form and distribution is also attested in Bahnaric languages such as Jeh (Gradin 1976) and Bahmar (Banker 1964).

In various Aslian languages this affix appears to have been replaced by an Austronesian borrowing, apparently Malay ber-, e.g.: Semelai b(r)- characterised by Kruspe as marking middle voice (fn.6, p117) which includes reciprocity. For Temiar, Benjamin (1976:179-80) describes an equivalent morpheme bar- with a similar range of meanings. According to Burenhult reciprocal is marked in Jehai by an -a- infixed to the verb.

5. Closing remarks

Presently the comparative-historical reconstruction of Proto-Mon-Khmer is still grappling with important issues of phonology, lexicon, and classification. Nonetheless, it is evident that much can be achieved in terms of comparative reconstruction of affixes and morphological processes, and the time is right for a comprehensive bottom-up tabulation and comparative analysis of Mon-Khmer morphology.

References


PART III
PROCESSES OF CHANGE
1. Introduction*

It has been claimed that the theory of case is “one of the most central concepts in modern generative syntax” (Brandner & Zinsmeister 2003). The goal of a case theory is to predict how arguments of a predicate are morphosyntactically realised. Most theories today employ the concept of thematic role to identify arguments in a level of abstraction known as argument structure, and attempt to map these thematic/theta roles onto grammatical functions. These types of theories are known as mapping theories. However, a case theory should ideally make predictions about the relationship between theta roles, grammatical functions and the morphosyntactic case and its morphological realisation. A theory describing this three-way relation is called a licensing theory (after Kiparsky 1997 and others). Most work in case theory could be best described as a mapping theory, taking for granted the relationship between structural cases and grammatical functions, such that morphological realisation is overlooked. However, in this paper I hope to show that the morphological realisation of a case is an important part of understanding case, and with the right theory, still predictable.

It is unfortunate that the morphological realisation of cases has not been the focus of much recent research, leaving potentially many morphological generalisations unrecognised. The context in which I hope to shed light on such generalisations is case marking in four-place predicates. These are necessarily derived predicates, by adding an argument such as causativisation of a ditransitive verb.

There is a substantial literature on causatives, much of which was inspired by Comrie (1976) who investigated the syntax of causative constructions cross-linguistically. Such works include Baker (1988), Cole (1976, 1983), Davies and Rosen (1988), Gerdts (1984), Gibson (1980), Gibson and Raposo (1986), among others. However, few have included data from causativised ditransitive verbs. This omission is likely due to the relative rarity of this construction. Nedyalkov and Sil’nickij noted that if a language has a causative, “it can be established that causative affixes are more productive in combination with [intransitive] verbs than with [transitive] verbs” (1973:7). There are languages that only allow intransitive verbs to causativise, but no languages where only transitive verbs

* Many thanks to the anonymous reviewers for helpful comments. I remain solely responsible for the data and analysis presented here.
are causativised. Typically, one could generalise that the more arguments you have in the base predicate, the harder it is to add a causer to that predicate. Velázquez-Castillo notes that in Guaraní the transitive-based causatives are found considerably less frequently than the intransitive-based causatives, constituting only 16 percent of all morphological causatives in the text she examines (Velázquez-Castillo 2002). Indeed, studies on causatives have recognised the need to distinguish the following types of verbs in order to determine how susceptible a verb will be to causativisation (Shibatani 2002:6):

(1)  
   a. Inactive transitives  
   b. Middle/ingestive verbs  
   c. Active transitives  
   d. Transitive verbs

Conspicuously absent from this list are ditransitive verbs. This is not surprising considering that many languages do not have ditransitive verbs, and of those that do and have a causative construction, many do not allow them to be causativised. However, Basque (isolate, southwestern Europe) does allow ditransitive verbs to causativise as illustrated by (2) and (3) (Donohue 2004, ex. 206).¹

(2)  
   \[\text{Ni-}k \text{ pobre-}ei \text{ diru-}a \text{ ema-}ten \text{ diet}.\]  
   1SG-ERG poor-PL:DAT money-ABS give-IMPF TNS.have.3PL:DAT.1SG:ERG  
   “I give money to the poor.”

(3)  
   \[\text{Apaika-}k \text{ pobre-}ei \text{ diru-}a \text{ eman-araz-}i \text{ zidan ni-ri}.\]  
   priest-ERG poor-PL:DAT money-ABS give-CAUS-PRF TNS.have.1SG:DAT.PST 1SG-DAT  
   “The priest made me give money to the poor.”

Four-place predicates, such as the one formed by causativizing a ditransitive verb as in (3) are rather unusual because there are no underived verbs which take four arguments (at once).² Causativised ditransitives are thus one way to force an extra argument into the argument structure. How languages cope with this can be quite revealing of case licensing principles.

Unfortunately very few descriptions of languages or studies of causatives include data of causativised ditransitives. The earliest notable exception to this is the work of Comrie (1975, 1976). However, Comrie’s study was primarily

¹ The following abbreviations are used throughout this paper: 1,2,3: First, second, third person, ABS: Absolutive, AUX: Auxiliary verb, CAUS: Causative morpheme, DAT: Dative, ERG, Ergative, IMPF, Imperfective, INSTR: Instrumental, LOC: Locative, PRF: Perfect, PL: Plural, SG: Singular, TNS: Tense marker.
² See Donohue (2004: 129-130) for further discussion of possible explanations for why languages do not have underived four-place predicates, given that there are verbs, like transaction verbs, with four conceptual arguments, any of which can be a structural argument.
concerned with what grammatical relation the causee will be assigned, rather than the morphosyntactic case with which it will be realised. Subsequent work on causatives has explored the syntax of causatives in greater detail (e.g. Baker 1988, Alsina 1997, among others). However, most of these studies are couched within a linking theory, where the association between thematic roles and grammatical functions is the primary concern. Those that do address the morphological case marking aspect of licensing theory seldom include data from four-place predicates. However, there are notable exceptions to this trend. In this paper I first present data from Basque causatives that will be the focus for my discussion of case. I then briefly summarise recent work by Joppen and Wunderlich (1995), Joppen-Hellwig (2000) and Donohue (2004). After comparing these approaches, I will then summarise the key generalisations and show how Donohue’s approach can extend to explain the variation in Basque case marking.

2. Basque causatives

Basque (Euskara Batua) is an ergative isolate spoken in southwestern Europe. The Basque examples work as one might expect: when you add an argument, the case markers are assigned as they would for arguments of an underived verb with the same number of arguments. That is, the causativised intransitive verbs result in a case array that resembles a regular transitive case array, as shown in (4)–(6), and the causativised transitive verb results in a case array the same as for a ditransitive verb, as illustrated in (7) and (8). The causativised ditransitive is repeated in (9).

(4) **Intransitive [1-place predicate: ABS]**

\[
\text{Mikel} \quad \text{joan da.} \\
\text{MikelABS} \quad \text{go.PRF TNS.be} \\
\text{“Mikel went.”} \\
\text{(Donohue 2004, ex. 194a)}
\]

(5) **Transitive [2-place predicate: ERG-ABS]**

\[
\text{Soldadu-ek} \quad \text{haur guzti-ak hil zituzten.} \\
\text{soldier-PL:ERG child whole-PL:ABS kill.PRF TNS.PL:ABS.have.3PL:ERG.PST} \\
\text{“The soldiers killed all the children.”} \\
\text{(Donohue 2004, ex. 198a)}
\]

(6) **Causativised intransitive [=case array of 2-place pred.: ERG-ABS]**

\[
\text{Ama-k} \quad \text{Mikel joan-araz-i du.} \\
\text{mother-ERG mikelABS go-CAUS-PRF TNS.have} \\
\text{“Mother made Mikel go.”} \\
\text{(Donohue 2004, ex. 194b)}
\]

In examples (4) – (6) we see that the addition of an argument to an intransitive verb (4) in Basque results in a case array (6) that resembles a regular two-place predicate (5). Below we see that the pattern is similar for causativisation of a
transitive verb (7) which results in a case array that resembles a regular three-place predicate (8).

(7) Causativised-transitive \(\text{[=case array of 3-place pred.: ERG-DAT-ABS]}\)

\[
\]

“Herod made the soldiers kill all the children.” (Donohue 2004, ex. 198b)

(8) Ditransitive \(\text{[3-place predicate: ERG-DAT-ABS]}\)

\[
\text{Ni-k poor-PL:DAT money-ABS give-IMPF TNS.have.3PL:DAT.1SG:ERG}
\]

“I give money to the poor.” (Donohue 2004, ex. 206a)

The is no prediction one could make for causativised ditransitives as there is no underived four-place case array for it to model. In Basque, we have the following:

(9) Causativised ditransitive \(\text{[=4-place predicate]}\)

\[
\text{Apaiza-priest-ERG poor-PL:DAT money-ABS give-CAUS-PRF TNS.have.1SG:DAT.PST 1SG-DAT}
\]

“The priest made me give money to the poor.” (Donohue 2004, ex 206b)

These effects of causativisation on the case arrays in Basque can be summarised in the Table 1 below (after Joppen & Wunderlich 1995).

<table>
<thead>
<tr>
<th>Verb type</th>
<th>Regular case array</th>
<th>Causativised case array</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intransitive</td>
<td>ABS(_x)</td>
<td>ERG – ABS(_x)</td>
</tr>
<tr>
<td>Transitive</td>
<td>ERG(_x) – ABS(_y)</td>
<td>ERG – DAT(_x) – ABS(_y)</td>
</tr>
<tr>
<td>Ditransitive</td>
<td>ERG(_x) – DAT(_y) – ABS(_z)</td>
<td>ERG – DAT(_x) – SEM.DAT(_y) – ABS(_z)</td>
</tr>
</tbody>
</table>

Table 1: Case arrays in Basque

3. Previous studies

There are three studies that have addressed these Basque data. While all three studies address issues relating to the case marking generalisations and are cast within similar case theoretic approaches, they differ in crucial ways that I highlight in the following subsections (see also Donohue forthcoming).

3.1 Theoretical preliminaries

The case theory used throughout this paper is Lexical Decomposition Grammar (LDG; Kiparsky 1997 and elsewhere; Wunderlich 1997 and elsewhere). Unlike most case theories which focus on the mapping between grammatical functions
and thematic roles, LDG is a theory of case licensing, capturing the ternary relation between thematic roles (arguments), grammatical functions or ‘abstract case’ and morphosyntactic (and morphological) case. LDG has constrained principles for relating levels of abstract case and morphosyntactic case by defining them both with the same two relational features [±H(ighest) R(ole)] and [±L(owest) R(ole)]. The theory captures generalisations and predications both about typologically diverse languages and highly complex phenomena within a specific language.

3.1.1 Semantic form. Following Bierwisch (1986 and elsewhere), LDG assumes a level of structure called semantic form (SF) which represents the grammatically relevant parts of a verb’s conceptual structure. It consists of minimally decomposed expressions formulated in predicate logic and expressed using lambda-categorial expressions. SF representations are thus constrained to two basic types: propositions, or constants and individuals, or variables. Consider the verb ‘show’.

\[
\text{show: } \lambda z \lambda y \lambda x \left[ x \text{ cause } [\text{can } [y \text{ see } z]] \right]
\]

In (10), the constants are the units of meaning into which the predicate is decomposed, and the variables are \( x, y, z \), representing the participants. The variables are lambda-abstracted out of the SF, and the resulting lambdas are equivalent to thematic roles, where the (inside out) depth of embedding represents the thematic hierarchy for a given verb.

3.1.2. Abstract case. Abstract case is defined using the same two given relational features. These are assigned to the ‘thematic roles’ according to their relative position in the semantic form. Once \(+HR\) and \(+LR\) have been assigned, the rest can be assigned implicationally.

\[
\text{show: } \lambda z \lambda y \lambda x \left[ x \text{ cause } [\text{can } [y \text{ see } z]] \right]
\]

With the highest and lowest roles identified, all other roles must be marked as non-highest role and non-lowest role to complete the feature specification. Once the abstract case is defined, the morphosyntactic case is assigned through simple unification. The relational case features cross-classify to define four abstract cases:

\[
\begin{align*}
\text{i. A:} & \quad \left[ +HR \right] \quad \left[ -LR \right] \\
\text{iii. O:} & \quad \left[ -HR \right] \quad \left[ +LR \right] \\
\text{ii. S:} & \quad \left[ +HR \right] \quad \left[ +LR \right] \\
\text{iv. D:} & \quad \left[ -HR \right] \quad \left[ -LR \right]
\end{align*}
\]
3.1.3 Morphosyntactic case. These features ([±HR], [±LR]) are also used to specify the morphosyntactic structural case (note that semantic case is not defined in this way). Typically the unmarked case nominative/absolutive is characterised by not having any specified features. The accusative is usually characterised as [-HR] and the ergative [-LR], while the dative is the most highly specified with a negative instance of both features. In Basque, as suggested above, the structural case inventory is taken to be:

Abs: [    ]
Erg: [-LR]
Dat: [-HR, -LR]

There are two conditions which govern the association of morphosyntactic case with abstract case. These are given in (13).

(13) i. Unification: Associated feature matrices must be non-distinct.
    ii. Specificity: Specific rules and morphemes block general rules and
        morphemes in the same context.

Thus, feature matrices will only unify if they are non-distinct. For example, typically the dative is defined as [-HR, -LR] and will unify with the middle role in a ditransitive verb:

(14) show: \[ \lambda z \lambda y \lambda x [x \text{CAUSE} [y \text{SEE} z]] \]

Dative morphosyntactic case [-HR, -LR] thus unifies with \( \lambda y [-HR, -LR] \). The less specific nominative case ([    ]) and ergative case ([−LR]) will not unify with this abstract case due to specificity: the more highly specified case available in the inventory [-HR, -LR] will block the use of a more general morpheme in the same context.

3.2 Joppen and Wunderlich 1995

Joppen and Wunderlich (1995) examine Basque causatives in great detail within the LDG. They view causatives as flat structures which add an argument, but which have no internal structure. Thus, the argument structure of the following two clauses would be considered the same:

(15) a. make eat: \( \lambda y \lambda x \lambda q [q \text{CAUSE} [x \text{EAT} y]] \)
    b. give: \( \lambda x \lambda y \lambda x [x \text{CAUSE} [y \text{HAVE} z]] \)

The causer argument is identified by a ‘q’, but the label of the variables in the semantic form is irrelevant. The two structures in (15) have the same basic
SF and three arguments. In Joppen and Wunderlich’s view the two structures are equivalent.

These can capture the nominal case facts in Basque as I will show below, recalling the Basque examples from §2.

(16) make go: $\lambda x \lambda q [q \text{ CAUSE } [x \text{ GO}]]$

$\text{MS.case} \rightarrow [ ] [-LR]$

Ama-k Mikel joan-araz-i du.

mother-ERG Mikel.ABS go-CAUS-PRF TNS.have

“Mother has made Mikel go.”

(17) make kill: $\lambda y \lambda x \lambda q [q \text{ CAUSE } [x \text{ CAUSE } [y \text{ DIE}]]]$

$\text{MS.case} \rightarrow [ ] [-HR,-LR] [-LR]$

Herodes-ek soldadu-ei haur guzti-ak hil-araz-i zizkien.


“Herod made the soldiers kill all the children.”

(18) make give: $\lambda z \lambda y \lambda x \lambda q [q \text{ CAUSE } [x \text{ CAUSE } [y \text{ HAVE } z]]]$

$\text{MS.case} \rightarrow [ ] [-HR,-LR] [-HR,-LR] [-LR]$

Apaiza-k pobre-ei diru-a eman-araz-i zidan ni-ri.

priest-ERG poor-PL:DAT money-ABS give-CAUS-PRF TNS.have.1SG:DAT.PST 1SG-DAT

“The priest made me give money to the poor.”

The examples in (16)–(18) show that the same principles of unification between abstract case and morphosyntactic case apply, resulting in the indicated morphological cases.

Joppen and Wunderlich’s approach has many virtues. Thematic roles are derived from a semantic form and there is no need to refer to a thematic hierarchy directly. Moreover, simple unification underlies the entire linking theory. LDG correctly generates the morphological case patterns in Basque. However, there is a problem in that this approach does not distinguish between the two datives in the causativised ditransitive (18) and we have seen that it is $\lambda x$ that receives structural case, while $\lambda y$ is rendered semantic: it can no longer govern verbal agreement and may optionally appear in the destinative case.
This problem is addressed in further work by Joppen-Hellwig (2001) that I will discuss next.

### 3.3 Joppen-Hellwig 2000

Joppen-Hellwig (2000) is a cross-linguistic study which provides a typology of case arrays in four-place predicates (that are causativised ditransitives). Specifically, if a language allows four-place predicates through morphological causativisation and does not allow case doubling, one of the arguments must be realised in a semantic case. The chief observation is that which argument becomes a semantic case is predictable based on whether or not the language is ergative or accusative, as illustrated below.

(19) a. If the language is ergative, the lower middle role ($\lambda y$, Recipient) will be realised as a semantic case.
   b. If the language is accusative, the upper middle role ($\lambda x$, Causee) will be realised as a semantic case.

Languages with split ergativity pattern consistently with either accusative or ergative languages, regardless of the split in the case marking system.

Joppen-Hellwig accounts for this by suggesting that inherent control properties of the argument (animacy, etc.) play a role in the argument linking. For this she posits a feature $[\pm C]$. This is assigned cyclically to the arguments in the semantic form. For example:

(20) `make go`: $\lambda x \ \lambda q \ [q \ CAUSE \ [x \ GO]]$

The assignment of $[\pm C]$ starts in the innermost predicate and then extends outwards, such that the highest role typically has the most control in a clause, indicated by its lack of $[\pm C]$ features. Consider the causativised ditransitive:

(21) `make give`: $\lambda z \ \lambda y \ \lambda x \ \lambda q \ [q \ CAUSE \ [x \ CAUSE \ [y \ HAVE \ z]]]$

The 'best' controller corresponds to the highest role, $\lambda q$, as $q$ has no $[-c]$ features, and the 'worst' controller is the lowest role, which has only $[-c]$ features.

From this, Joppen-Hellwig claims that the “relevant” feature for ergative languages is $[-c]$. In this way, $\lambda x$, having only one $[-c]$ feature, in some sense ‘outranks’ $\lambda y$ and is thus deemed more of a controller. The structural dative is then assigned to the causee.

Conversely, in accusative languages $[-c]$ is said to be the relevant feature, and similar results in the recipient ($\lambda y$) being assigned the structural dative case, while the causee is rendered oblique.
This does capture the facts. However, it is stipulative and the definitions of control are vague. Donohue (2004) presents an alternative way to predict the dative/semantic case assignment for Basque without having to refer to any stipulative features such as [+c].

4. Causee case marking generalisations

Donohue (2004) builds on these previous works by extending the theory. She presents an analysis using Lexical Decomposition Grammar but couched within an Optimality Theoretic (Smolensky and Prince 1993) framework (OT-LDG). The analysis relies on two critical assumptions. The first is relatively noncontroversial: causatives are instances of predicate argument composition (e.g. Alsina 1997) and as such recognise the causee as an embedded subject (or a-subject). The second is that causees, as subjects, should be marked like regular (transitive) subjects. It is this second characterisation of the causee case marking process that Donohue develops within OT-LDG. In addition to the extensive cross-linguistic typology presented in Donohue (2004), this section demonstrates that this approach to case marking also accounts for diachronic and dialectal variation.

To ensure that causees are marked as subjects, Donohue proposes the ‘Causee case marking principle’ (CCP) which states that causees will ideally be marked by a case which resembles that of typical transitive subjects and which is a structural case ([+SC]):

(22) **Causee Case Marking Principle:**

Causes will be maximally faithful to the features of a transitive subject: [+HR, –LR] and [+SC]

The cases in Basque can then be defined as shown in Table 2 (see Donohue 2004: 159 for further explanation).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[+HR, –LR] / [+SC]</td>
<td>yes</td>
<td>yes</td>
<td>ERGATIVE</td>
</tr>
<tr>
<td>[+HR, –LR] / [–SC]</td>
<td>yes</td>
<td>no</td>
<td>INSTRUMENTAL</td>
</tr>
<tr>
<td>[–HR, –LR] / [+SC]</td>
<td>(one mismatch)</td>
<td>yes</td>
<td>DATIVE</td>
</tr>
<tr>
<td>[–HR, –LR] / [–SC]</td>
<td>(one mismatch)</td>
<td>no</td>
<td>SEMANTIC DATIVE</td>
</tr>
<tr>
<td>[–HR, +LR] / [+SC]</td>
<td>(two mismatches)</td>
<td>yes</td>
<td>ACCUSATIVE</td>
</tr>
</tbody>
</table>

**Table 2: Case feature matches**

The CCP is implemented as optimality theoretic constraints to derive a typology of case marking. I refer the interested reader to Donohue (2004, in preparation) for details of the OT account. The results of the OT implementation are that the final column listing the cases in Table 2 is essentially the preferred output for the case of the causee, ranked from highest to lowest.
However, to recall the Basque example of a causativised ditransitive we see that the causee is marked by a structural dative case.

(23) Apaiza-k pobre-ei diru-a eman-araz-i zidan ni-ri.
    priest-ERG poor-PL.DAT money-ABS give-CAUS-PRF TNS-have-1SG.DAT-PST 1SG.DAT
    “The priest made me give money to the poor.”

The reason why the causee is not marked by an ergative case is that Basque does not allow any doubling of structural cases. The next best option would be a true instrumental case. However contemporary Basque (Euskara Batua) does not have a true instrumental (or a passive and thus a need for one). It is thus the morphological inventory that actively determines the available options for a given language.

4.1 Old Basque

An interesting difference between old and contemporary Basque is that old Basque had a passive and an instrumental with which to express the demoted agent. With this different morphological case inventory we would predict that the instrumental case is used to express the causee in four-place predicates in old Basque, and this is indeed the case. The causee in a causativised ditransitive is expressed in the instrumental case in 19th century Labourdin Basque shown in (24) (Ortiz de Urbina 2003: 440, taken from Elissamburu Piarres Adame, pp80).

(24) Mutil-ez zain-araz-ten zitian bere arthalde handi-ak.
    boys-INST tend-CAUS-IMPF TNS.3PL:ABS.3ERG his flock large-PL:ABS
    “He had his large flocks tended by boys.”

4.2 Western Basque

It is also important to note that the CCP is only observable when the case assignment cannot operate as usual due to there being too many arguments. This is another parameter in which languages vary. And indeed we find this variation in dialects of Western Basque. In these dialects, with similar case inventories to Euskara Batua, the causee is always specially case marked. Instead of using the CCP as a ‘repair strategy’, it is the norm for causees. All causees are marked using the structural dative, regardless of the transitivity of the base verb. We thus observe an identical paradigm to Euskara Batua, except that with causativised intransitive verbs (25) causees also bear the dative case (Ortiz de Urbina 2003: 435).

(25) Asarre bixi-bixitt-an jarri erazo dauste ni-ri.
    fury alive-alive-LOC get CAUS AUX 1SG-DAT
    “They have made me get very furious.”
5. Concluding remarks

There are several points made in this paper that I would like to reiterate as concluding remarks. First, it is important to take morphology seriously. In the data we have examined, it is clearly critical to carefully consider the actual morphological form of the case when trying to understand generalisations that govern the distribution of case markers. It follows, thus, that it is also important to consider the whole system of cases – the case inventory – for a given language under study. Second, in addition to cross-linguistic data, patterns of variation within a language are also an important source of data for testing theories of case. In particular diachronic and dialectal variation are a particularly good source for patterns of microvariation, which a good theory should be able to easily accommodate.

References


THIRD PERSON PLURAL AS A MORPHOLOGICAL ZERO
OBJECT MARKING IN MAROVO

BETHWYN EVANS
University of Manchester

1. *Introduction*

Asymmetries of grammatical coding, particularly within morphological paradigms, are often considered to be motivated by the principles of iconicity and economy, such that semantically unmarked values and/or frequently occurring values of a grammatical category tend to be coded linguistically in a less complex way, often lacking overt coding altogether (Greenberg 1966, Bybee 1985, Givón 1995, Croft 2003, Haspelmath 2008). Such cross-linguistic tendencies are both synchronic and diachronic. Thus the synchronic tendency for particular values of grammatical categories to lack overt coding cross-linguistically reflects processes of language change that result in such systems (Bybee 1985, Koch 1995).

In Marovo, an Oceanic language of the Solomon Islands, it is the combination of 3rd person and plural number which lacks overt coding within the object marking paradigm. This 3PL zero morpheme in Marovo presents something of a diachronic puzzle, raising the question of what motivated its development. It is argued here that economy is a relevant factor, but that the exact motivation for the 3PL zero morpheme can only be determined when the system of object marking is considered within the broader context of discourse patterns of transitive constructions.

2. *Object marking in Oceanic languages*

Oceanic languages commonly have a set of post-verbal markers which index the person and number of the object argument. For example, in (1) the verbal enclitic =eu denotes that the object argument is 1SG, and in (2) the verbal suffix -di indexes the 3PL object argument, which is also expressed by the clause-initial noun phrase.

---

1 It is with pleasure that I dedicate this paper to Harold, my first teacher of historical linguistics. The research for this paper was supported by a Simon Fellowship at the University of Manchester. This fellowship and also a British Academy Small Research Grant (SG-40401) provided financial support for the fieldwork carried out on Marovo, both of which are gratefully acknowledged. I would like to thank all those in Buini Tusu, a community on a small island in Marovo lagoon, Solomon Islands, who helped me during my fieldwork. I am also very grateful to Claire Bowern, Darja Hoenigman, Luisa Miceli, Louise Mycock, Mary Raymond and two anonymous referees for comments on this paper.
(1) Tama-ku mo ware-eu. (North-East Ambae, Vanuatu)
father-1SGP RL call=1SGO
“My father is calling me.” (Hyslop 2001:337)

(2) Tamóata ú-te-di. (Manam, Papua New Guinea)
man 1SGS.RL-see-3PL
“I saw the men.” (Lichtenberk 1983:134)

Within the paradigm of similar object markers in Marovo there is variation, including a zero allomorph, in the form of the 3PL object marker (Table 1).3

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>SINGULAR</td>
<td>-o</td>
<td>-ho</td>
<td>-a</td>
</tr>
<tr>
<td>PLURAL</td>
<td>-da</td>
<td>-mi</td>
<td>-di, -i, -Ø</td>
</tr>
<tr>
<td>EXCL</td>
<td>-ami</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Object markers in Marovo

In (3) the object argument, expressed by the clause-final noun phrase, ria ihana ‘the fish’, is indexed within the verb complex by the object marker -di.4

(3) Beto ma-[ni la humahuma vae-di\textsubscript{VC} [ria ihana]_{O} (Marovo)
finish then-3SGS go spear.fishing take-3PL ART:PL fish
“Afterwards you can go spear fishing and get the fish.”5

Example (4) has a similar structure, but here the 3PL object argument, denoted by the noun phrase chore tahami kahike ‘our three canoes’, is indexed within the verb complex by the marker -i. In (5) it is the lack of an object marker, or zero, which indicates that the object argument is 3PL, here also denoted by the noun phrase, ria tege ta-gu ra ‘my mat bundles’.


3 The plural object markers in Marovo index all non-singular object arguments. Thus an object argument that is denoted by a noun phrase coded as dual or trial will be indexed by the plural object markers.

4 The syntactic constituent in Marovo that comprises the verb(s), verbal modifiers and markers of tense/aspect/mood, transitivity and participant reference is labelled the verb complex (VC). The object markers occur as the final element of the verb complex and so may be attached to a post-verbal modifier.

5 For ease of interpretation of the examples, the verb complex is enclosed in square brackets and labelled VC. Subject and object noun phrases are also enclosed in square brackets and labelled S and O, respectively.
The use of a zero morpheme to index 3PL objects in Marovo is determined by the behaviour of the verb in transitive constructions with non-3PL object arguments. For example, in (6) the verb vagara ‘to net’ occurs transitively with the transitive suffix -i and the 3SG object marker -a, which alone expresses the object argument, and refers to the participant denoted by ia vasina ‘the area’, the head of the preceding relative clause. In (7), a clause with a 3PL object argument, this verb occurs with the transitive suffix and lacks an overt object marker.

Examples of verbs which occur with a zero ending to indicate a 3PL object argument are provided in Table 2. Comparison with forms of these verbs coded for a 3SG object provides evidence for the zero allomorph of the 3PL object marker.

<table>
<thead>
<tr>
<th>Intransitive</th>
<th>Transitive with 3SG object</th>
<th>Transitive with 3PL object</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>asa</td>
<td>asa-i-a</td>
<td>asa-i-Ø</td>
<td>to grate (sth.)</td>
</tr>
<tr>
<td>om-omi</td>
<td>omi-a</td>
<td>omi-Ø</td>
<td>to see (sth.)</td>
</tr>
<tr>
<td>pocho</td>
<td>pocho-a</td>
<td>pocho-Ø</td>
<td>to squeeze (sth.)</td>
</tr>
<tr>
<td>ruja</td>
<td>ruja-i-a</td>
<td>ruja-i-Ø</td>
<td>to pound (sth.)</td>
</tr>
<tr>
<td>vagara</td>
<td>vagar-i-a</td>
<td>vagar-i-Ø</td>
<td>to net (sth.)</td>
</tr>
<tr>
<td>—</td>
<td>usi ni-a</td>
<td>usi ni-Ø</td>
<td>to use (sth.)</td>
</tr>
</tbody>
</table>

Table 2: Examples of verbs in Marovo which occur with a zero morpheme marking 3PL objects
With other verbs, comparison of their forms in constructions with 3SG and 3PL object arguments indicate that a 3PL object is indexed by -\textit{i}. For example, in (8) \textit{veko} 'to leave' occurs with the 3SG object marker -\textit{a}. When \textit{veko} occurs in constructions with 3PL object arguments, as in (9), it takes the ending -\textit{i}.

(8) $[\text{Veko-a}]_{VC} \ [\text{dekuru} \ ia]_{O}$
\hspace{1cm} $\text{leave-3SGO log} \hspace{1cm} 3SG$
\hspace{1cm} "Leave that log..."$

(9) $[\text{La veko-i}]_{VC} \ [\text{ria labete}]_{O}$
\hspace{1cm} $\text{go leave-3PLO ART:PL timber}$
\hspace{1cm} "Go and leave that timber."

The verbs in Table 3 show the same morphological pattern of object marking as \textit{veko} 'to leave'. That is, they occur with -\textit{i} as the 3PL object marker.

<table>
<thead>
<tr>
<th>Intransitive</th>
<th>Transitive with 3SG object</th>
<th>Transitive with 3PL object</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>holu</td>
<td>holu-a</td>
<td>holu-i</td>
<td>to buy (sth.)</td>
</tr>
<tr>
<td>(heru-)heru</td>
<td>heru-a</td>
<td>heru-i</td>
<td>to carry (sth.)</td>
</tr>
<tr>
<td>ole</td>
<td>ole-a</td>
<td>ole-i</td>
<td>to call (s.o.)</td>
</tr>
<tr>
<td>seke</td>
<td>seke-a</td>
<td>seke-i</td>
<td>to cut (sth.)</td>
</tr>
<tr>
<td>ta-tonu</td>
<td>tonu-a</td>
<td>tonu-i</td>
<td>to do (sth.)</td>
</tr>
</tbody>
</table>

Table 3: Examples of verbs in Marovo with which -\textit{i} indexes a 3PL object

The data in Tables 2 and 3 demonstrate that in Marovo, as in many other Oceanic languages, verbs occur in one of two transitive structures: one, VERB + TRANSITIVE MARKER + OBJECT MARKER, in which the verb occurs with one of the two transitive markers -\textit{i} or -\textit{ni} followed by an object marker, as in (6); the other, VERB + OBJECT MARKER, in which the verb occurs with only a following object marker, as in (8). Verbs which occur in the former structure tend to occur with -Ø to index a 3PL object argument, and verbs which occur in the latter structure tend to occur with -\textit{i} to index a 3PL object. The use of -\textit{di} to index a 3PL object argument appears to be in free variation with the use of -Ø and -\textit{i}. Thus, examples (10) and (11) show the verb \textit{heru} 'to carry' used in very similar contexts, first with the 3PL object marker -\textit{di} and second with -\textit{i}.

(10) $... \ pata \ [la \ heru-d]_{VC} \ [ria \ labete]_{O}$
\hspace{1cm} (Marovo)
\hspace{1cm} in.order go carry-3PLO ART:PL timber
\hspace{1cm} "... in order to carry the timber ..."

\hspace{1cm} 6 There are a few verbs, such as pero 'to scrape' and pocho 'to squeeze', which occur in the structure VERB + OBJECT MARKER and with which a 3PL object argument is indexed by -Ø. It is not clear at this stage exactly why such exceptions occur.
(11) \[ \text{Heru-i} \text{ VC } [\text{hami}]_s [\text{ria labete}]_o, \ldots \]

\text{carry-3PL O} 1\text{PLEXCL ART:PL timber}

“We carried the timber, ...”

While the distribution of the allomorphs of the 3PL object marker can be described in terms of morphosyntactic classes of verbs and free variation, the presence of such allomorphy raises the question of the origins and development of the system. In particular, a zero morpheme to indicate 3PL is not expected cross-linguistically and so what motivates its presence in Marovo represents a diachronic puzzle.

3. **Explanations of morphological zeroes**

Asymmetries within morphological paradigms are not necessarily random since similar kinds of asymmetries are found cross-linguistically. In terms of zero coding within paradigms, there is a tendency for the same values within particular grammatical categories to lack overt coding. For example, within a sample of 50 languages, Bybee (1985:52-53) found that of the 27 languages in which verbs are coded for number, 21 (78%) coded singular with zero. Of the 28 languages in which verbs are coded for person, 15 (54%) coded 3rd person with zero and 4 (14%) lacked overt coding for 1st person. From a diachronic perspective, morphological zeroes can be viewed as reflecting either (i) a ‘gap’ within the paradigm that results from the failure of a marker to develop for a particular value, or (ii) the creation of a zero morpheme through the loss or reanalysis of a previously overt marker.

Ariel (2000) takes the first view as the explanation for the cross-linguistic tendency for person agreement systems to overtly code 1st and 2nd person, but to lack overt coding for 3rd person. She (2000:198) proposes that this pattern is best accounted for in terms of speakers’ choice of referring expression based on an assessment of how accessible the participant is to the addressee. A referring expression which is more informative (i.e. giving a greater amount of lexical material), more rigid (i.e. identifying a participant relatively uniquely), and less attenuated (i.e. lengthier or accented) is likely to be chosen to denote participants of lower accessibility. This is because the addressee will better be able to retrieve from memory and identify a participant of low accessibility if given more formal cues to do so. In contrast, a referring expression which is less informative, less rigid and more attenuated is likely to be chosen to denote participants which are highly accessible (Ariel 2000:204). Ariel (2000:205) characterises the accessibility of participants in terms of entity salience and unity. Entity salience is defined by a number of different criteria where entities on the left under (a)–(e) in Figure 1 are considered more salient than those on
the right. Unity refers to the distance and degree of cohesion between expressions denoting the participant (cf. (f) and (g) in Figure 1).

(a) speaker > addressee > non-participant (3rd person)
(b) high physical salience > low physical salience
(c) topic > non-topic
(d) grammatical subject > non-subject
(e) human > animate > inanimate
(f) repeated references > few previous references > first mention
(g) non intervening/competing references > many intervening/competing references

Figure 1: Antecedent salience (Ariel 2000:206)

The grammaticalisation of independent pronouns which leads to the less informative and less attenuated coding of participants by reduced/cliticised pronouns or agreement markers results from the speakers’ intention to indicate that a referent is highly accessible (Ariel 2000:206). In this way person agreement markers will tend to be created to code 1st (speaker) and 2nd (addressee) person participants, referents that are the topic of discourse and human and animate participants. The lack of overt coding within an agreement paradigm for 3rd person is thus explained by the fact that 3rd person participants are not in themselves highly accessible, and so are unlikely to develop such coding strategies.

As Bybee (1985:55) states, there is also evidence that speakers restructure morphological paradigms in ways that create zero morphemes in positions where the lack of overt coding is expected from cross-linguistic patterns. The creation of zero morphemes through analogical change was first described in detail by Watkins (1962) in the context of Indo-European reconstruction. Watkins (1962) proposes that a typical kind of morphological change is the reanalysis of an inflected verb stem as constituting a zero exponent for a 3SG subject, and the construction of a new paradigm based around the reanalysed stem. For example, Watkins (1962:165-174) demonstrates that the t-preterite verb stems in Old Irish, such as -bert ‘he bears’, can be explained in terms of the reanalysis of a verb root coded for a 3SG subject with the ending -t as comprising a verb root coded for the aorist by -t and lacking overt coding for the 3SG subject. Evidence of the change comes from the Old Irish verb stems -biurt ‘I bear’ and -birt ‘you bear’, which include the original 3SG subject ending. Koch (1995) places Watkins’ Law within a broader framework of morphological change, establishing the general diachronic principle:

A word-form which expresses by means of a non-zero marker a property which is typologically expected to be coded by zero is liable to be reanalysed as containing a zero marker. (Koch 1995:64)

Thus Koch (1995:34-46) demonstrates that similar patterns of reanalysis leading to the creation of zero morphemes also occur for other categories that are
Third Person Plural as a Morphological Zero

Considered to be semantically unmarked and tend to be coded by zero synchronically, such as nominative and absolutive case and present tense. Koch (1995:46-57) also shows that there is a cross-linguistic tendency for the creation of zero morphemes for values of grammatical categories that are considered to be locally unmarked (Tiersma 1982). For example, there is a diachronic tendency for zero morphemes to be created for the locative case with nominals denoting places or times, for 1SG possessors with kin terms and for plural number with nouns denoting objects that are typically referred to in collections.

Koch (1995: 64) concludes that the creation of zero morphemes has at least two different motivations: iconicity and economy. Thus the creation of zero-coding for semantically unmarked values of a grammatical category is iconically motivated in that the semantically least complex value is reanalysed to comprise the least complex linguistic expression, that is, zero. However, the creation of zero morphemes for locally unmarked values of grammatical categories is economically motivated. That is, the most frequent form within the paradigm is reanalysed as having the least amount of linguistic expression. Bybee (1985:57-65) presents similar factors as explaining the basic form around which a paradigm is likely to be restructured, but from a cognitive perspective. She (1985:60) suggests there is a tendency for a paradigm to be organised such that the form that is semantically basic or unmarked and the most frequently occurring tends to become the one from which all other forms are derived. One reason for this is that semantically basic and frequent forms are most likely to be learned and stored independently.

Haspelmath (2006, 2008) argues against the use of a concept of ‘markedness’ in linguistics, proposing that many of the structural asymmetries for which markedness is argued to provide an explanation can actually be directly explained by asymmetries in frequency. Thus Haspelmath suggests that:

All universal morphosyntactic asymmetries can be explained on the basis of frequency asymmetries, i.e. they all show economic motivation: More frequent patterns are coded with less material. (Haspelmath, 2008:185)

Haspelmath (2008) proposes that these economical patterns in language arise primarily through three processes of change, (i) differential phonological reduction, (ii) differential expansion of a new construction, and marginally (iii) morphological analogy. Differential phonological reduction refers to the tendency for frequent expressions to be less carefully articulated and so to undergo phonological change at a faster rate than less frequent, and less predictable, expressions (Haspelmath 2008:206). Most cases of economical coding, Haspelmath (2008:207) claims, result from the differential expansion of a new and more complex construction. New constructions typically arise within a specific context and may be used to highlight a particular meaning or add clarity in situations of potential ambiguity, and such novel constructions often comprise additional linguistic material. While the use of such constructions
tends to be extended to a wider range of contexts over time, they will be ‘inhibited’ from completely replacing the older construction in two ways. First, since frequently occurring combinations of meanings are deeply entrenched in speakers’ mental grammars and tend to be resistant to change, an innovative construction is unlikely to replace the older construction in such contexts. Second, an innovative construction is unlikely to be extended to contexts in which the associated meanings are expected and so explicit linguistic expression is redundant. Thus, in contexts where meanings are predictable and expected, speakers will tend to economise and not use the new, and more explicit, construction (Haspelmath 2008:207-208). This latter factor inhibiting the extension of an innovative construction is also related to frequency since expected and predictable associations of meaning tend to be so as a result of frequent co-occurrence.

4. Zero coding of third person plural

A first step towards explaining the occurrence of a zero 3pl object marker in Marovo is to determine whether it reflects a gap within the paradigm or the creation of a zero morpheme. Comparison of the Marovo paradigm of object markers with those in closely related languages of the New Georgia group shows that a zero morpheme for 3pl is also found in other languages. The object marker paradigms for four New Georgia languages are presented in Table 4: Kubokota in the west, and Roviana, Hoava and Marovo in the east. That Roviana and Hoava also show zero 3pl object markers suggests that this form in Marovo needs to be considered within the broader context of the New Georgia group of languages.

<table>
<thead>
<tr>
<th></th>
<th>1SG</th>
<th>2SG</th>
<th>3SG</th>
<th>1PINC</th>
<th>1PLEX</th>
<th>2PL</th>
<th>3PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kubokota</td>
<td>-ziu</td>
<td>-yo</td>
<td>-a</td>
<td>-Ø</td>
<td>-yita</td>
<td>-yami</td>
<td>-yamu</td>
</tr>
<tr>
<td>Roviana</td>
<td>-u, -au</td>
<td>-yo</td>
<td>-a</td>
<td>-yita</td>
<td>-yami</td>
<td>-yamu</td>
<td>-Ø</td>
</tr>
<tr>
<td>Hoava</td>
<td>-rao</td>
<td>-yo</td>
<td>-a</td>
<td>-yita</td>
<td>-yami</td>
<td>-yamu</td>
<td>-Ø</td>
</tr>
<tr>
<td>Marovo</td>
<td>-o</td>
<td>-ho</td>
<td>-a</td>
<td>-da</td>
<td>-ami</td>
<td>-mi</td>
<td>-di, -i</td>
</tr>
</tbody>
</table>

Data from Kettle 2000, M.Raymond p.c., Corston-Oliver 2003, Davis 2003

Table 4: Object markers in New Georgia languages

The typical origin of person agreement markers is the grammaticalisation of independent pronouns (Givón 1976), and the evident cognacy amongst object markers and independent pronouns in Oceanic languages indicates that this is indeed the origin of object markers within Oceanic. The independent pronouns

---

8 Data from Simbo (Palmer 1996), Nduke (Scales 1997) and Vangunu (Bourchier 2007) were also considered. The data on object markers in these languages are limited, but what is known does not appear to contradict the conclusions which are presented here.

9 The zero morpheme for 3pl objects in Kubokota appears to occur only in applicative constructions which are marked by -ni with singular objects and -di with plural objects. An explanation of these data requires a detailed description of changes within the system of transitivity and object marking in Kubokota and is beyond the scope of this paper.
and object markers for four Oceanic languages are listed in Table 5: Bali-Vitu (Meso-Melanesian, Papua New Guinea), Hoava (Meso-Melanesian, Solomon Islands), Longgu (Southeast Solomonic, Solomon Islands) and North-East Ambae (Southern Oceanic, Vanuatu). In each of these languages, object markers are cognate with the independent pronouns for most person-number values.

<table>
<thead>
<tr>
<th></th>
<th>Bali-Vitu</th>
<th>Hoava</th>
<th>Longgu</th>
<th>North-East Ambae</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PRN OBJ</td>
<td>PRN OBJ</td>
<td>PRN OBJ</td>
<td>PRN OBJ</td>
</tr>
<tr>
<td>1SG</td>
<td>yau -a</td>
<td>rao -rao</td>
<td>na(u) -u</td>
<td>neu -eu</td>
</tr>
<tr>
<td>2SG</td>
<td>eyo -yo</td>
<td>yoe -yo</td>
<td>oe -o</td>
<td>nigo -ko</td>
</tr>
<tr>
<td>3SG</td>
<td>ia -Ø</td>
<td>(i)sa -a</td>
<td>gaia -a</td>
<td>nje -a,-e</td>
</tr>
<tr>
<td>1PLINCL</td>
<td>yita -</td>
<td>yita -yita</td>
<td>gia -</td>
<td>kide -</td>
</tr>
<tr>
<td>1PLEXCL</td>
<td>yami -</td>
<td>yami -yami</td>
<td>ami -</td>
<td>kanai -</td>
</tr>
<tr>
<td>2PL</td>
<td>yamu -</td>
<td>yami -yamu</td>
<td>amu -</td>
<td>kimiu -</td>
</tr>
<tr>
<td>3PL</td>
<td>yiti -nazi</td>
<td>ria -Ø</td>
<td>gira -ra,-i</td>
<td>njire -ra,-re</td>
</tr>
</tbody>
</table>

Data from Ross 2002, Davis 2003, Hill 2002, Hyslop 2001

Table 5: Independent pronouns and object markers in Oceanic languages

Within this context, Corston-Oliver (2003) presents evidence from discourse patterns which suggests that the lack of an overt marker for 3PL object arguments in Roviana can be explained in terms of such a form never having developed. In Roviana, the occurrence of 3PL referents as object arguments is infrequent; only 36 (23.8%) of 151 object arguments have 3PL reference within the data set examined (Corston-Oliver 2003:286). Of these 3PL object arguments only 14 have human reference. In addition, independent pronouns in Roviana rarely occur with non-human referents; 16 (6.8%) of 234 occurrences of independent pronouns have non-human referents. Taken together, these facts indicate that in Roviana, and given a uniformitarian approach also in the language’s history, the 3PL independent pronoun did not occur as an object argument “with sufficient text frequency to be grammaticized as a bound affix” (Corston-Oliver 2003:287). This discourse-based explanation for the lack of an overt 3PL object suffix in Roviana not only fits well with the Roviana data, but is also supported by Ariel’s (2000) explanation of the development of agreement markers. However, the analysis is problematic when object markers in Roviana and other New Georgia languages are considered within the context of object marking and its history in Oceanic more broadly.

A set of object markers with very similar functions and distribution as those found in contemporary Oceanic languages can be reconstructed for Proto Oceanic, the forms of which are given in Table 6. For Proto Oceanic only a partial paradigm of object markers, comprising 1st, 2nd and 3rd person singular and 3rd person plural forms, are reconstructable. It is likely that non-3rd person plural object arguments in Proto Oceanic were denoted solely by independent pronouns (Evans 1995).
The shared origin of the Proto Oceanic object markers and independent pronouns is evident from a comparison of the reconstructions in Table 6, suggesting that the grammaticalisation of pronouns as object markers for 1st, 2nd and 3rd person singular and 3rd person plural must have occurred at some stage prior to Proto Oceanic. Languages such as Longgu and North-East Ambae (Table 5), have conservative systems of object marking, reflecting the Proto Oceanic forms and retaining the same partial paradigm. The grammaticalisation of independent pronouns as object markers has continued throughout the history of Oceanic languages. For example, in Hoava the object markers -ita ‘1PLINCLO’, -ami ‘1PLEXCLO’ and -amu ‘2PL’ reflect independent pronouns of the same forms (Table 5). Thus, in any discussion of the history and development of object markers in contemporary Oceanic languages, the relative chronology of the processes of grammaticalisation that have led to the occurrence of object markers for different person and number values needs to be considered.

The object markers and the morphological structure of transitive verb complexes in Marovo are to a large extent conservative. As described in §2, there are two transitive structures in Marovo, one in which the verb complex occurs with one of the two transitive markers -i or ni and with an object marker, as in (6), and the other in which there is no transitive marker, but only an object marker, as in (8). These same two transitive structures, namely VERB + TRANSITIVE MARKER + OBJECT MARKER and VERB + OBJECT MARKER, are reconstructable for Proto Oceanic. In Proto Oceanic, the distribution of these two structures was apparently determined by the phonological shape of the verb. Consonant-final verb stems and verb stems ending in *-a occurred with the transitive suffix *-i followed by the object markers in transitive constructions, whereas other vowel-final verb stems occurred with only the object markers (Evans 2003:104-117). Thus verb stems like *inum ‘to drink’ and *rubat ‘to be loose’ can be reconstructed as occurring with the transitive suffix *-i in Proto Oceanic, namely *inum-i- ‘to drink sth.’ and *rubat-i- ‘to loosen sth.’ The same is true for verb stems that ended in *-a, so transitive forms with the transitive suffix, such as *soka-i- ‘to pierce sth.’ and *wara-i- ‘to speak to s.o.’, are reconstructable for Proto Oceanic also had two transitive markers, *-i and *akin[i] (see Pawley 1973, Evans 2003), the antecedents of Marovo -i and ni, respectively. For the purpose of the present paper, it is sufficient to discuss the transitive suffix *-i only.
Proto Oceanic. Verb stems ending in other vowels, however, occurred in the second transitive structure and are reconstructable for Proto Oceanic as occurring with only the object markers, as with *piro=a ‘to twist it together’ and *wase=a ‘to divide, distribute it’ (Evans 2003). This distribution of the two transitive structures in Proto Oceanic is schematised in Table 7, along with the corresponding transitive structures in Marovo.

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proto Oceanic</td>
<td>INTR. CVCVC</td>
<td>CVCa</td>
</tr>
<tr>
<td>TR.</td>
<td>CVCVC-i=OBJ</td>
<td>CVCa-i=OBJ</td>
</tr>
<tr>
<td>Marovo</td>
<td>INTR. CVCVCV</td>
<td>CVCa</td>
</tr>
<tr>
<td>TR.</td>
<td>CVCVC-i=OBJ</td>
<td>CVCa-i=OBJ</td>
</tr>
</tbody>
</table>

Table 7: Transitive structures in Proto Oceanic and Marovo

While there are some exceptions, the tendency in Marovo is for the transitive structure of a verb to be determined by its phonological shape. Disyllabic verb stems which end in -a, like picha ‘to crack’, occur with the transitive suffix -i followed by an object marker, picha-i-a ‘to crack it’; this is a direct reflex of the same structure in Proto Oceanic (Table 7, column B). Disyllabic verb stems ending in vowels other than -a, such as golu ‘to clean’, occur with only the object suffixes, golu-a ‘to clean it’; again this directly reflects the Proto Oceanic structure (Table 7, column C). Many polysyllabic verb stems in Marovo have the shape CVCV, for example, vagara ‘to net’. When these verbs are used transitively the transitive suffix -i replaces the final vowel of the verb stem and is followed by the object suffixes, vagar-i-a ‘to net it’. This type of transitive structure in Marovo reflects the Proto Oceanic structure of a consonant-final verb stem occurring with the transitive suffix *-i (Table 7, column A). Proto Oceanic consonant-final stems are reflected in Marovo, and indeed in all other Northwest Solomonic languages (Ross 1988:218), with an additional echo vowel. Thus, Proto Oceanic *onom ‘six’ is reflected in Marovo as onono and Proto Oceanic *lapuat ‘to be big, important’ (Ross 2003:191) as Marovo lavata ‘to be big’. With verb stems this sound change affected the intransitive forms of verbs which were consonant-final, but not the transitive forms which occurred with the transitive suffix *-i and so were not consonant-final. It is this sound change which has resulted in the synchronic morphological pattern.11

---

11 It is difficult to demonstrate these changes with a specific verb stem in Marovo because of the high degree of lexical replacement in New Georgia languages, and I currently lack a clear Marovo reflex of a Proto Oceanic consonant-final verb stem. The change can be shown in Roviana. Roviana yarata ‘to bite’ reflects Proto Oceanic *katat ‘to bite’ with the addition of an echo vowel. The transitive form of this verb in Roviana is yarat-i- ‘to bite sth.’ (Waterhouse 1949:24), reflecting Proto Oceanic *katat-i- ‘to bite sth.’ (see Evans 2003:320-321 for the data supporting these reconstructions).
In terms of the forms of the object markers, the Marovo 1SG, 2SG and 3SG morphemes are direct reflexes of the Proto Oceanic forms. The Marovo 3PL object marker -di is not a direct reflex of Proto Oceanic *=ra, but rather reflects *-dri, an innovative 3PL object marker reconstructable for Proto Western Oceanic, which is regularly reflected in Proto Northwest Solomonic as *-di, and inherited as such into Marovo (Evans 1995:70-71, 85-86).12

While Marovo -di ‘3PLO’ is an inherited form, the other two allomorphs appear to be more recent innovations and represent two stages of development. First, a zero marker for 3PL object arguments was created. Evidence for a zero morpheme in Marovo indexing 3PL object arguments was presented in §2. The second change is that the transitive suffix *-i appears to have been reanalysed as a 3PL object marker. Indirect evidence of this reanalysis comes from the verbs with which there is a paradigmatic relationship between VERB + -a with a 3SG object argument and VERB + -i with a 3PL object argument. This morphosyntactic pattern appears to reflect the extension of *-i, reanalysed as a 3PL object marker, to verbs with which it did not originally occur.

The presence of a zero 3PL object marker in Marovo, and I would argue in other New Georgia languages, represents the creation of a zero morpheme. Before investigating possible motivations for this zero morpheme, it is important to note that it does not reflect the loss of *-di through regular sound change. The clearest evidence of this comes from the Marovo nominal suffix for 3PL possessors -di, as in tina-di ‘their mother (mother-3 PLP)’. Marovo -di ‘3PLP’ is a direct reflex of the Proto Oceanic 3PL possessive suffix *-dri[a] (Lichtenberk 1985, Ross 1988:353-354), demonstrating that *dr > *d > Ø in this environment is not a regular sound change in Marovo.

The loss of -di ‘3PLO’ in Marovo also seems unlikely to reflect differential phonological reduction motivated by economy. Cross-linguistic studies of frequency of different person-number values suggests that it is either 3SG or 1SG which occur most frequently (Bybee 1985, Ariel 2000). Within contemporary Marovo, 3PL is not the most frequently occurring person-number value amongst object arguments. Table 8 gives the number of the different person-number categories of object arguments within a total of 952 clauses. Although these data are from texts of different styles, all were one- or two-person narratives, and this is taken to explain the exceedingly low number of 1st and 2nd person object arguments. However, what is clear from these data is that 3SG object arguments are much more frequent than 3PL object arguments. Thus 69% of transitive clauses occur with 3SG object arguments, while only 30% of transitive clauses occur with 3PL object arguments.13

12 The form of the Proto Western Oceanic 3PL object marker is likely connected to the innovative 3PL independent pronoun *idri[a], which characterises Western Oceanic (Ross 1988:352-357).

13 Similar results were found in Kubokota. Within two glossed texts in Kettle (2000), 54 (78%) of the 69 transitive clauses have 3SG object arguments and only 7 (10%) had 3PL object arguments.
On the basis of these data, it is difficult to explain the loss of the original 3PL object marker in terms of economy. Since 3SG object arguments occur far more frequently than 3PL ones, the creation of a zero morpheme would be expected for 3SG. 14

It is equally difficult to explain the creation of a zero 3PL object marker in terms of iconicity. That is, that 3PL -Ø reflects the tendency for semantically unmarked values of a grammatical category to lack overt coding. Andersen (2001) argues for a conceptual analysis of semantic markedness which is independent of the linguistic characteristics that are often associated with it, such as frequency and formal coding. He proposes that it is the construal of essentially symmetrical oppositions as asymmetrical that underlies semantic markedness. Thus, while exclusive semantic relations are intrinsically symmetrical modes of opposition, it is a characteristic of many pairs of terms that are logically exclusive opposites that one of the terms also functions as the hypernym and is construed as both inclusive and included (Andersen 2001:43).

Following this analysis of semantic markedness, it is the morpheme with the combined values of 3rd person and singular number in the Marovo object marking paradigm which can function as a hypernym and so be considered semantically unmarked. There are instances where the 3SG object marker -a and 3PL object marker -di ~ -Ø ~ -i represent an exclusive opposition in terms of number, as shown by (12) and (13). In (12) the 3SG object marker -a occurring with the verb va-legu ‘to kill’ indexes a single woman, denoted in the first clause by meka ‘one’. In (13) the same verb occurs with the 3PL object marker -di which

<table>
<thead>
<tr>
<th>1ST SG/PL</th>
<th>2ND SG/PL</th>
<th>3SG</th>
<th>3PL</th>
<th>NO. TRANSITIVE CLAUSES</th>
<th>TOTAL NO. OF CLAUSES</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>3</td>
<td>167</td>
<td>73</td>
<td>246</td>
<td>952</td>
</tr>
</tbody>
</table>

Table 8: Frequencies of person-number categories of object arguments in Marovo

14 Even if it is assumed that the creation of a zero morpheme for 3PL object arguments was a common innovation in the history of a number of New Georgia languages (eg. Roviana, Hoava and Marovo), I know of no evidence which could support a claim that at the time when the zero morpheme came into use 3PL object arguments did occur more frequently than 3SG ones, thus explaining 3PL -Ø as economically motivated.
indexes the object noun phrase *ria ihana* ‘the (plural) fish’. The 3SG object marker -a is also used to index object arguments with plural rather than singular reference. In some Oceanic languages, and perhaps even in Proto Oceanic, object arguments with plural reference are indexed by singular object markers if they denote inanimate participants\(^{15}\). In Marovo, objects denoting both animate and inanimate participants can be indexed with the 3PL object marker; see examples (13) and (10). An object argument with plural reference can also be indexed by the 3SG object marker -a; this occurs primarily when the participants denoted are viewed as a group. In (14) the 3SG object marker with seke ‘to cut’ indexes the head of the relative clause, tongania ria tege ‘all the leaves cut for making mats’; in this context the leaves which have been cut are viewed as a single group rather than many individual leaves.

\[
\begin{array}{l}
\text{one} & \text{3PLS} & \text{go} & \text{arrive-3SGO} & \text{3PL} & \text{and.SG} & \text{CAUS-die-3SGO} & \text{3PL} \\
\end{array}
\]
“One, they reached her, and so they killed her.”

(13) *... ma-[ma va-legu-d] [ria ihana]...*
\[
\begin{array}{l}
\text{then-1PLEXCLS} & \text{CAUS-die-3PLO} & \text{ART:PL} & \text{fish} \\
\end{array}
\]
“... then we kill the fish ...”

(14) *Vari-paru hami tongania ria tege*
\[
\begin{array}{l}
\text{RECIP-put.together} & \text{1PLEXCL} & \text{every} & \text{ART:PL} & \text{mat} \\
\end{array}
\]
“... we gathered together all the mat leaves which we three had cut ...”

These data indicate that the -a object marker can be used both as an exclusive term, indexing a 3SG object in opposition to a 3PL one, and as a hypernym, indexing 3rd person objects regardless of number. Thus, in Marovo 3SG can be considered semantically unmarked within the object marking paradigm, which weakens any claim that the creation of the 3PL zero morpheme may have been motivated by iconicity. Rather it seems likely, as Corston-Oliver (2003) suggests for Roviana, that the explanation for the 3PL object marker in Marovo can be found in patterns of discourse. Certain tendencies of discourse in Marovo indicate that the differential loss of 3PL -di, and thus the creation of a zero morpheme, may have been motivated by a reduction of redundancy in transitive constructions.

Object markers in Marovo have two functions; namely, they show either grammatical or anaphoric agreement (Bresnan & Mchombo 1987, Siewierska 2004). That is, they can co-occur with a noun phrase that expresses the object

\(^{15}\) I would like to thank one of the anonymous referees for pointing this out.
argument of the clause (grammatical agreement), as in (13), or they may themselves be the only expression of the object argument within the clause (anaphoric agreement), as in (12). Typically, 1st and 2nd person object markers show anaphoric agreement; it is only in specific discourse contexts that a 1st or 2nd person object marker is used as a grammatical agreement marker. By contrast, 3rd person object markers in Marovo typically show grammatical agreement and co-occur with an inter-clausal object noun phrase. Within 78 clauses with 3PL object arguments, in only 15 (19%) is the object marker the sole indication of the object argument within the clause. For clauses with 3SG object arguments, the proportion in which the object marker is the only indication of the object argument is slightly higher, but still only 45% (49 of 146) of clauses.

In a transitive construction with a 3rd person object argument, the object is usually expressed by a lexical or pronominal noun phrase and is indexed within the verb complex. With 3SG object arguments the object marker provides grammatical coding of singular number, which tends not to be coded within the noun phrase. Thus, of the 3SG object arguments expressed by a noun phrase, only 20 out of 97 (21%) are grammatically coded for number within the noun phrase. Clauses such as (15), with no coding of singular number within the object noun phrase, are thus more frequent than clauses such as (16), in which the object noun phrase is coded for singular number.

(15) [Tavete ni-a]_{VC} [ipacha]_{O} [ra]_{S}, … (Marovo)
work TR-3SG GO bailer 1SG
“I made a bailer, …”

(16) [Kave-a]_{VC} [ia]_{S} [meka mara lavata]_{O}.
pull-3SG 3SG one trevally big
“He pulled in one big trevally.”

With 3PL object arguments, in 76% (48 of 63) of the clauses in which a 3PL object is expressed by an inter-clausal noun phrase, that noun phrase exhibits grammatical coding of plurality. Examples such as (17) and (18), where the object noun phrase is coded for plural number by the plural article ria or a plural quantifier, are more frequent than clauses like (19), where the object noun phrase has plural reference but is not grammatically coded for number.

(17) Ma-[ni lae]_{VC} [ia]_{S} [asa-i-Ø]_{VC} [ria uvikola]_{O} (Marovo)
then-3sgS go 3SG grate-TR-3PL ART-PL tapioca
“Then she went and grated the tapioca roots.”
(18) ...ma-[gu la omi pule-Ø]_{VC} [raka]_{I} [katiga baeni ihana]_{O} ...
then-1SGS go see go.back-3PLO 1SG some school.fish fish
"... then I saw some schools of fish,..."

(19) ... beto asa hami, [ngina pero-Ø]_{VC} [hami]_{I} [ngochara]_{O}
finish grate 1PLEXCL RRR scrape-3PLO 1PLEXCL coconut
"... after we've grated, we will scrape coconuts."

In this way the creation of a zero 3PL object marker in Marovo may have been motivated by the reduction of redundancy. Since both the person and number of a 3PL object argument are typically expressed by an object noun phrase in Marovo, indicating the same information within the verb complex is redundant. This redundancy would have occurred more frequently for 3PL object arguments than 3SG ones, thus explaining the creation of a zero morpheme for 3PL rather than 3SG.16

The 3PL object marker -i in Marovo is only found with verbs which occur transitively in the structure VERB + OBJECT MARKER, such as seke ‘to cut’ which occurs as seke-a ‘to cut it’ and seke-i ‘to cut them’. It is possible that this use of -i reflects a reanalysis of the transitive suffix *-i as a marker of both transitivity and 3PL objects in constructions like that in (17) and its subsequent extension to verbs like seke ‘to cut’. If this is an accurate analysis of the history of -i ‘3PLO’, a possible motivation for the change would be that such verbs are otherwise indistinguishable in their intransitive and 3PL-object transitive forms. For example, in (19) the verb pero occurs in a transitive construction with an object noun phrase, the form of the verb could be either intransitive or transitive. It is likely that avoidance of such constructions in Marovo, a language in which transitivity is typically coded within the verb complex, has led to this use of -i. Perhaps it is simply as a marker of transitivity, but it does occur with 3PL object arguments in contrast to objects of other person-number values, which are coded simply with overt object markers.

The creation of a zero morpheme for 3PL object arguments has been described here for Marovo, but implicit in the analysis is that this innovation began in the common ancestor of Marovo and the closely related languages Hoava and Roviana. While the zero 3PL object marker completely replaced the original marker *-di in Roviana and Hoava, in Marovo -di ‘3PLO’ is still marginally used.

5. Concluding remarks

Previous studies of morphological zeroes from a diachronic perspective demonstrate the need to consider the development of such morphemes within

---

16 It is likely that the frequent coding of plural number within noun phrases in Marovo is a post-Proto Oceanic innovation, but further research is needed to determine the exact chronology of this innovation relative to the changes within the object marking system.
the context of the overall formal and functional organisation of the paradigms of which they are a part (Watkins 1962, Bybee 1985, Koch 1995). This examination of the zero morpheme indexing 3Pl object arguments in Marovo highlights the need to also consider their development within the context of discourse patterns. Thus, the development of -Ø ‘3PlO’ in Marovo (and other eastern New Georgia languages) appears to have been motivated by changes in the structure of noun phrases. The increased grammatical coding of plurality within object noun phrases motivated the loss of overt coding of 3pl objects within the verb complex, thereby reducing redundancy of coding within transitive constructions with 3pl object arguments. The underlying motivation of this change is economy, the primary motivating factor of the development of such grammatical asymmetries (Haspelmath 2008). However, rather than frequency of occurrence and therefore predictability of a construction resulting in a lack of overt coding, I have argued that it is an increase in grammatical coding elsewhere in the system which has resulted in the lack of overt coding for a particular person-number value within the object marking paradigm.

References


Bourchier, Louise. 2007. Vangunu report. Ms, Australian National University, Canberra.


1. Introduction

Paradigm levelling, in which irregularities or other complexities in inflectional morphology are regularised through analogy with other elements in the paradigm, is a common process in morphological change (Kiparsky 1982). Studies of analogical change have shown that change tends to eliminate alternations by replacing less frequent alternants with more frequent ones and marked forms with unmarked ones (Kuryłowicz 1947; Mańczak 1958; Trask 1996; Winter 1971). Moreover, it has often been noted that paradigm levelling affects low frequency forms before high frequency forms and that high frequency irregularities are more likely to be maintained (Bybee 1985; Hooper 1976). However, there are also accounts in which marked forms have replaced unmarked forms: for example the levelling of Old French singular plural oppositions of the type chastel–chasteaux ‘castle’ in which the marked plural has replaced the unmarked singular, giving Modern French château–châteaux. The existence of analogy in such cases accounts for the phenomena, but does not explain why such cases of analogical levelling occur and what processes actually affect analogical change from marked to unmarked change, rather than the alternative. For example, in the case of Old French singular forms, it is true that modern French does not normally have phonetically differentiated singulars and plurals,¹ so the opposition chastel–chasteau is irregular and a possible target for paradigm levelling, but what is not explained is levelling on the basis of the plural rather than the singular, which would have produced a similar regularity.

In the paradigm levelling found in the perfect in varieties of Jersey Norman French (JNF), analogical change often appears to move from more marked and lower frequency forms to influence the whole verb paradigm. The two main areas where analogical change has occurred in the paradigm will be discussed separately — the reduction to essentially a single perfect paradigm across all conjugations; and the syncretism of 2nd and 3rd person plural suffixes for the perfect, with different origins for the forms in different dialects.

The JNF data is notable in that while various varieties have undergone different levelling processes, they almost all involve analogical influences from more marked forms on less marked forms. The paper will attempt to explain

¹ Although cases do continue to exist, as in oppositions of the type cheval–chevaux ‘horse’, or exceptional forms such as os ‘bone’ (/os/-/ø/) or oeil–yeux ‘eye’.
how language internal factors including sound changes have contributed to the levelling processes in JNF and have allowed more highly marked and less frequent forms to influence less marked, more frequent forms.

2. **Verb conjugations in JNF**

Verbs in JNF occur in five basic conjugations (see Table 1), each of which has its own distinctive morphological patterning with phonologically conditioned subvariants resulting from normal processes of sound change (Liddicoat 1994).

<table>
<thead>
<tr>
<th>Conjugation</th>
<th>1</th>
<th>1A</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>4A</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infinitive ending</td>
<td>-e</td>
<td>-i</td>
<td>-i</td>
<td>-i</td>
<td>-r</td>
<td>-o</td>
<td>-e</td>
</tr>
<tr>
<td>Example</td>
<td>dune</td>
<td>kmäfi</td>
<td>fini</td>
<td>dormi</td>
<td>vâdr</td>
<td>bôd</td>
<td>fil</td>
</tr>
<tr>
<td>Gloss</td>
<td>‘give’</td>
<td>‘begin’</td>
<td>‘finish’</td>
<td>‘sleep’</td>
<td>‘sell’</td>
<td>‘drinking’</td>
<td>‘fall’</td>
</tr>
<tr>
<td>Standard French</td>
<td>donner</td>
<td>commencer</td>
<td>finir</td>
<td>dormir</td>
<td>vendre</td>
<td>boire</td>
<td>choir</td>
</tr>
<tr>
<td>Latin conjugation</td>
<td>-are</td>
<td>-are</td>
<td>-ire</td>
<td>-ire</td>
<td>-ère</td>
<td>-ère</td>
<td>-ère</td>
</tr>
</tbody>
</table>

Table 1: Verb conjugations in JNF

Conjugation 1 corresponds to the standard French conjugation in -er, derived from Latin -are verbs. Conjugation 1A differs from the first conjugation in the infinitive and past participle only. In these cases, the final -i is the result of a regular sound change which reflects an Old French ending -ier, which is found after palatalised consonants: for example kmäfi ‘to begin’ (French commencer, Old French commencer). Conjugations 2 and 3 are both derived from Latin -ire verbs and differ by the presence in Conjugation 2 of an increment -is- to the verb root in the present indicative plural, the imperfect and both tenses of the subjunctive. Conjugation 4 and 4A verbs are derived from Latin -ère verbs. Conjugation 4 is characterised by a verb root ending in a consonant, while conjugation 4A has a root ending in a vowel, after which the etymological /r/ has been assimilated in most JNF varieties (Spence 1957; Liddicoat 1994). Conjugation 5 verbs are derived from Latin -ère verbs and typically end in a long vowel.

3. **The perfect in JNF**

Unlike most varieties of French, in all varieties of JNF there is a single set of endings for the perfect which is generalised across all five conjugations. In these forms the thematic vowels of the various conjugations are substituted with /i/ (except for a few verbs discussed below, with a different vowel). The most geographically widespread paradigm of perfect endings is shown in (1). The parenthesised sounds are those found in liaison contexts, and in the context of verb endings appear if the following word begins with a vowel and is closely

---

2 The JNF perfect corresponds historically to the modern standard French passé simple.
3 This same change is also found in the imperfect subjunctive (see Liddicoat 1994).
associated with the verb (for example, a subject clitic pronoun, an indefinite article in an object NP, etc.).

(1) 1sg -i(z) 1pl -i:m(z) 2sg -i(z) 2pl -i:t(z) 3sg -i(t) 3pl -i:t

This paradigm is found in most of the varieties of Norman French spoken on Jersey, with the exception of those in the centre and the extreme north-west of the island (St Ouen parish). It shows a number of features of the system of perfect marking in JNF in addition to the generalisation of the thematic vowel /i/. The first is the syncretism of the singular forms, at least in the non-liaison forms. The second is the syncretism of the 2nd and 3rd person plural forms, again with differential marking in the liaison forms. The paradigms found in the centre and north-west (St Ouen) show similar patterning, but with some variation in the final inflection, particularly in the plural.

(2) 1sg -i(z) 1pl -i:m(z) 2sg -i(z) 2pl -i:t(z) 3sg -i(t) 3pl -i:t

The central dialect shows the same distribution of marking in liaison forms as in (1), although with a different form of the ending for the 2nd and 3rd person plural forms, while the St Ouen forms do not have distinctions for 2nd and 3rd person in liaison forms, as well as having a distinct form of the plural endings.

In addition to the set of inflections with the thematic vowel /i/ described above, there are two other paradigms for the perfect with different thematic vowels: one with the thematic vowel /y/, which is limited to about a dozen verbs of Conjugation 4; and one with a thematic vowel /e/, which is found only for the Conjugation 3 verbs maet ‘to put’ and pràdr ‘to take’, and the Conjugation 4 verbs vnè ‘to come’ and mè ‘to hold’. While these paradigms have a different thematic vowel, the endings otherwise correspond with the /i/ paradigm above.

3.1 Origins and development of the thematic vowel of the JNF perfect

All of the forms attested in JNF are found in Old French and an evolutionary path can be proposed to account for the forms found in the modern varieties based on Old French forms. The perfect in Old French had a number of different paradigms as shown in Table 2.

---

4 This paradigm is also found in Sark Norman French, which is derived primarily from the variety spoken in St Ouen parish (Liddicoat 1994, 2002).

5 In addition, the verb ave ‘to have’ is irregular, although the plural forms have perfect endings related to the other paradigms.
The table shows a number of similarities and differences between the Old French perfect system and the JNF system. Some of the changes which have occurred in the development of the JNF system are shared with standard French and are early levellings of the paradigm: thus weak $i$, weak $i^2$, strong $i$ and strong $s$ perfect endings collapsed together; and in standard French, weak $a$ and weak $a^2$ verbs also collapsed together, resulting in a three-part paradigm with the thematic vowels $a$, $i$ and $u$. The system of distinctions between perfect forms in the different conjugations found in Old French would appear to have continued into Old JNF in much the same way, if one considers that the writings of the Jersey-born twelfth century poet Wace can be taken as indicative of the state of the language. However, it must be acknowledged that Old French texts are not totally reliable as sources for language change in dialects as they involve considerable dialect mixing. With this caveat in mind, it can be seen that in Wace’s Roman de Rou (Holden 1970), Wace’s use of perfect verb inflections essentially makes the same distinctions between conjugations as is seen in Table 2 above.

<table>
<thead>
<tr>
<th>JNF Conj1</th>
<th>Conj1A</th>
<th>Conj2, Conj3</th>
<th>Conj4, Conj5</th>
<th>Conj4</th>
<th>Conj4A</th>
<th>Conj5, ave 'to have'</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>-ai</td>
<td>-ai</td>
<td>-i</td>
<td>-i</td>
<td>-ui</td>
<td>-s</td>
</tr>
<tr>
<td>2sg</td>
<td>-as</td>
<td>-as</td>
<td>-is</td>
<td>-is</td>
<td>-us</td>
<td>-sis</td>
</tr>
<tr>
<td>3sg</td>
<td>-at</td>
<td>-at</td>
<td>-it</td>
<td>-iet</td>
<td>-t</td>
<td>-ut</td>
</tr>
<tr>
<td>1pl</td>
<td>-ames</td>
<td>-ames</td>
<td>-imes</td>
<td>-imes</td>
<td>-umes</td>
<td>-simes</td>
</tr>
<tr>
<td>2pl</td>
<td>-astes</td>
<td>-astes</td>
<td>-istes</td>
<td>-istes</td>
<td>-ustes</td>
<td>-sistes</td>
</tr>
<tr>
<td>3pl</td>
<td>-erent</td>
<td>-erent</td>
<td>-irent</td>
<td>-irent</td>
<td>-i(d)rent</td>
<td>-urent</td>
</tr>
</tbody>
</table>

Table 2: Perfect verb conjugations and forms in Old French (source: Einhorn 1975)

6 The spelling $eü$ the diaeresis indicates two vowel sounds $ay$ whereas the absence of the diaeresis would indicate the vowel $œ$.

7 The variant -$idrent$ is primarily a feature of northern and north-western Old French dialects.

8 Einhorn’s (1975) convention of distinguishing between variant conjugations in Old French using a superscript 2 has been adopted in this paper: weak $a^2$ is therefore a variant of the weak $a$ conjugation.

---

Weak $a$  

\[ \text{Au siege de Roem le \textit{cuiderent} gaber / s'il le peissent prendre ou par force enz entrer. 'At the siege of Rome they \underline{wanted} him to boast / if they could take it or enter it by force.'} \]

\[ \text{Mez quant Henri y vint n'i \textit{ouserent} ester 'But when Henry came there they \underline{dared} not remain.'} \]

\[ \text{Par nuit s'en eschapa, que hons ne l'aperchut, / fors ceuls qui l'\underline{enmenerent}, qu'a conduire l'estut, 'By night they escaped, so that no-one saw them, / except for those who \underline{led} them, for they had to be guided.'} \]

Weak $a^2$  

\[ \text{Lez fiz \textit{mesconseil}lerent por le pere encombrer. 'They \underline{badly advised} the sons to trouble their father.'} \]

---

6 The spelling $eü$ the diaeresis indicates two vowel sounds $ay$ whereas the absence of the diaeresis would indicate the vowel $œ$.

7 The variant -$idrent$ is primarily a feature of northern and north-western Old French dialects.

8 Einhorn’s (1975) convention of distinguishing between variant conjugations in Old French using a superscript 2 has been adopted in this paper: weak $a^2$ is therefore a variant of the weak $a$ conjugation.
Weak i

Qui firent livres et escriz / Des nobles faiz et des bons diz / Que li baron et li seignour / Firent de tens anciannot. ‘Who made the books and writings / Of the noble deeds and fine words / That the barons and the lords / did in ancient times.’

Richart li viex fu fiz Guillaume Longue Espee / que li Flamenz treirent. ‘Richard the old was the son of William long Sword / who the Flemings betrayed.’

Strong i

De Petou vindrent moinegn que lor out amenee, ‘From Poitou came monks that they had brought to them’

Plusors de sez voisinz le tindrent pour seignor; / trente anz tint Normendie en sa bonne vigor. ‘Several of his neighbours held him as lord; / thirty years he held Normandy in good health.’

Weak u

Looyis l’espousa qui out grant mariage; / en Jerusalem furent en lonc pelerinage, ‘Louis married her, who had a big wedding, / they went (lit. were) to Jerusalem on a long pilgrimage.’

Strong u

Mais par les bons clerz qui l’escristrent, / Quiz les gestes es livres mistrent, ‘But because of the good clerics who wrote it, / Who put the acts into books,’

Cil dui vassal qui tant conquistrent, /Tant orent terres et tant pristrent, ‘Those two vassals who conquered so much, / They had so many countries and took so much.’

Tant guerrea Franchoiz et tant lor fist poour / que il s’entracorderent, paiz pristrent et amour. ‘Francis fought so much and made them so afraid / that they made an agreement and took peace and love.’

Strong u

Cil dui vassal qui tant conquistrent, /Tant orent terres et tant pristrent, ‘Those two vassals who conquered so much, / They had so many countries and took so much.’

In these examples a number of things can be noted which have a bearing on the situation in modern JNF. Firstly, there is no evidence for the collapse of the a and u stems with the i stems. Secondly, of the four verbs which have a nasalised thematic /e/ in the perfect in modern JNF, prādr and metr do not have nasalised thematic vowels (pristrent, mistrent), however there is evidence that vnē and tnē have an etymologically consistent nasal form, as is also found in standard French.

Leaving aside changes which have affected most of the dialects of Northern France (such as the loss of the final -ent in 3rd person plural forms), the most notable difference between modern JNF perfects and the forms used by Wace is the loss of the weak a and weak a’ suffix forms and the extension of the weak i forms to replace them. This appears to be an analogical change, and is interesting in that the Conjugation 1 or a verbs are numerically the most common forms in the variety and include most of the regular high frequency verbs. This means that the statistically most frequent verb paradigm has been

---

9 Third person plural forms of weak i verbs – the only form that constrasts with weak i verbs are not found in Wace’s Roman de Rou and have fallen together with weak i verbs with the verb ending -ient: Franchiez par la campaigne ça e la e’espandirent / Heberges, e foillies, e pavillions tendirent. The French spread out here and there in the countryside / They set up houses, leaf huts and pavilions.

10 Approximately 60 percent of the verbs in the dictionary in Liddicoat (1994) are Conjugation 1 verbs.
lost in an analogical change which has levelled these verbs according to a more marked paradigm, that of Conjugations 2 and 3. At the same time the statistically least frequent verb paradigms, weak \( u \) and strong \( u \), have persisted in JNF with perfect forms with a thematic \(/y/\).

The persistence of \( u \) verbs is however only partial, as some verbs have shifted into the \( i \) paradigm for the perfect. In the case of the verb \( liö ‘to read’ \) the historical \(/y/\) perfect (and other forms with \(/y/\), such as the past participle) appear to have converged on the basis of a regular \(/i/\) vowel in all other forms. However, other verbs which have historically shifted from \(/y/\) to \(/i/\) in the perfect, such as \( ge: ‘to fall’, ve: ‘to see’, treö ‘to milk’ \) and \( gre: ‘to prepare, arrange’ \) (and its derivatives), are not so simply explained by analogy with predominating paradigm internal forms. The collapse of the highly marked \( u \) forms with the \( i \) forms in these verbs can in part be explained by sound changes which have occurred in JNF. An unstressed \(/e/\) before \(/j/\) commonly closes to \(/i/\) as in the following forms of \( ge: ‘to fall’:\n
\[
\begin{align*}
\text{present indicative:} & \quad j\hat{g}iö (1pl) \ u: \hat{g}iö (2pl) \\
\text{present subjunctive:} & \quad \hat{g}iöö (1pl) \ u: \hat{g}iöö (2pl) \\
\text{imperative:} & \quad \hat{g}iöö (1pl) \ \hat{g}iöö (2pl) \\
\text{imperfect:} & \quad \hat{g}iöö, \text{etc. (all persons)} \\
\text{present participle:} & \quad \hat{g}iöö
\end{align*}
\]

This leads to a verb such as \( ge: \) having phonologically conditioned variants between \(/i/\) and \(/e/\) as the thematic vowel. This alternation appears to have exerted an influence on the perfect forms, with analogical support from the perfect forms in \(/i/\) in other paradigms. In this case a phonologically conditioned variant has been reanalysed to introduce a morphologically conditioned alternation in the perfect. This explanation does not, however, hold for verbs such as \( kre: ‘to believe’, \) which maintains \(/y/\) in its perfect paradigm. The preservation of such \(/y/\) forms appears to be the result of a conservative influence from past participle forms which have retained \(/y/\), as the verbs which have adopted analogical \(/i/\) perfects also all have past participles in \(/e/\).

The loss of the \( a \) forms in preference for the \( i \) forms is harder to understand than the loss of \( u \) forms, as the expectation would be for the least marked paradigm to influence more marked forms, or at least to be preserved in opposition to more marked forms. However some possible explanations can be traced to regular sound changes which have occurred in the development of JNF.

The diphthong \(/je/\) in the weak \( a' \) 3rd person plural forms (-ierent) would regularly reduce to \(/i/\) \((\text{commencerent} > \text{kumäfìr})\), as also happened in the infinitive \((\text{commencer} > \text{kumäfì})\), past participle \((\text{commencé} > \text{kumäfì})\) and some second person plural forms: present indicative \((\text{commenciez} > \text{kumäfì})\) and subjunctive \((\text{commenciez} > \text{kumäfì})\). Given the presence of \(/i/\) in the past participle, the infinitive and the 3rd person plural perfect, the \( i \) form may have
spread to the other perfect forms by analogy. As was the case with /y/ forms, the weak a² verbs would thus have taken a phonologically conditioned variant and extended it as a morphologically conditioned variant.

It would appear that the falling together of the weak a² verbs and the i verbs weakened the frequency asymmetry between first conjugation and other verbs in JNF, with the /i/ forms becoming more common in the overall system. However, the overall numbers of a² verbs is relatively small and the existence of two classes of verbs in which /i/ forms are found would not seem to indicate per se why the remaining Conjugation 1 verbs should be generalised on the pattern of i verbs in the perfect while they maintain distinct forms for many other tense/aspect paradigms. This seems particularly odd since the least frequent perfect form, /y/, is preserved. However part of the answer to this may lie in the full set of forms of the perfect endings for Conjugation 1 verbs. While forms of such verbs are not attested between Wace and modern JNF, the probable forms following Wace, with relevant sound changes attested in JNF, would be as shown in Table 3.

<table>
<thead>
<tr>
<th></th>
<th>a forms</th>
<th>i forms</th>
<th>u forms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>dane</td>
<td>fini</td>
<td>mule</td>
</tr>
<tr>
<td>'to give'</td>
<td>*dune</td>
<td>fini</td>
<td>muly</td>
</tr>
<tr>
<td>1sg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2sg</td>
<td>*duna</td>
<td>fini</td>
<td>muly</td>
</tr>
<tr>
<td>3sg</td>
<td>*duna</td>
<td>fini</td>
<td>muly</td>
</tr>
<tr>
<td>1pl</td>
<td>*dunam</td>
<td>finim</td>
<td>mulae:m</td>
</tr>
<tr>
<td>2pl</td>
<td>*dunat</td>
<td>finit</td>
<td>muly:t</td>
</tr>
<tr>
<td>3pl</td>
<td>*duner</td>
<td>finir</td>
<td>muly:r</td>
</tr>
</tbody>
</table>

Table 3: Hypothetical perfect endings in JNF (hypothesised unattested forms in JNF marked *)

It can be seen that the vowel in the perfect endings of the a forms of Conjugation 1 verbs is less regular than that of the other conjugations. It is therefore possible that some form of levelling process has occurred here to reflect the regularity of the other forms. The ways in which this change occurred could be explained by a process of abduction: that is the treatment of an occasional feature as if it were a general feature (Andersen 1973, 2005). In this case, the regularities of Conjunction 1A verbs, which pattern identically with other conjugations in the perfect but which are similar to Conjugation 1 verbs in most other forms, could have been perceived as the general feature and extended to the ‘less regular’ verbs of Conjugation 1. This would mean that a regular rule could have been hypothesised based on Conjugation 1A before a variable rule had been developed by learners. The rules for forming the perfect seem to have been reanalysed following a potential vulnerability in the system which results from a phonetically induced change in Conjugation 1 perfect

---

11 The [œ] in the first person plural of u forms is a regular change of /y/ before a nasal consonant and [œ] and [y] are in complementary distribution here (Liddicoat 1994).
forms leading to a consequent internal irregularity – an irregularity which is not mirrored in the perfect in other Conjugations.

Thus while the Conjugation 1 forms of the perfect were more frequent than other forms, and hence less marked in that sense, they were more irregular as a paradigm, and hence reanalysed on the basis of the normally similar Conjugation 1A forms (identical to Conjugation 2 and 3 forms).

3.2 The development of the JNF perfect plural endings

As well as the collapsing together of the paradigms for verbs of different conjugations, the other clearly anomolous feature of the JNF perfect paradigms given in (1) and (2) are the forms of the 2nd and 3rd person plural forms, repeated here for convenience together with the Old French Weak and the standard French forms which have developed from these Old French forms.

(4) General Central St Ouen Old French Standard French

| 2pl  | -i:it(z) | -i:dr | -istes | -istes /-it/ |
| 3pl  | -i:t     | -i:dr | -irent | -irent /-ir/ |

In all three JNF varieties, the non-liaison forms of 2nd and 3rd plural are identical. However while the overall pattern is thus essentially identical, the process of levelling leading to this pattern would appear to be different for each dialect, with the central dialect having expanded the use of the (less marked) 3rd person plural to cover 2nd person plural contexts, the St Ouen variety having expanded an infrequent and highly marked 3rd person plural form to cover all 3rd person plural and 2nd person plural contexts, and the other dialects having expanded a form which was etymologically 2nd person to cover the less marked 3rd person contexts.

It is important to note that syncretism between the 2nd person plural and 3rd person plural perfect verb suffixes has no impact on comprehension, as subjects are obligatorily marked by person clitics in all verb forms except the imperative (Liddicoat 1994:242-243). Moreover, in regular verb paradigms the 2nd and 3rd person singular non-liaison forms have fallen together as the result of regular sound changes in a number of other verb forms as shown in Table 4.

<table>
<thead>
<tr>
<th>dane</th>
<th>fini</th>
<th>dormi</th>
<th>vadr</th>
<th>kre:</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘to give’</td>
<td>‘to finish’</td>
<td>‘to sleep’</td>
<td>‘to sell’</td>
<td>‘to believe’</td>
</tr>
<tr>
<td>C1</td>
<td>C2</td>
<td>C3</td>
<td>C4</td>
<td>C5</td>
</tr>
</tbody>
</table>

| present indicative | ty dun | ty fini | ty dor | ty vä | ty kre |
|                   | i dun  | i fini  | i dar  | i vä  | i kre  |
| imperfect         | ty dun  | ty fini  | ty dormi | ty väd å | ty krig v |
|                   | i dun  | i fini  | i dormi | i väd å | i krig v |
| future            | ty dun na | ty fini da | ty dormi da | ty väd ra | ty krig v da |
|                   | i dun na | i fini da | i dormi da | i väd ra | i krig v da |

Table 4: Syncretism in 2nd and 3rd person singular verbs in JNF
Syncretism of 2nd and 3rd person forms is therefore a common feature of the JNF grammar, although normally in the singular only and not in the plural, where all three forms are usually distinguished (with the 3rd person plural form being indistinguishable from the 3rd person singular in some cases). The existence of these identical forms in the 2nd and 3rd person singular may have exerted a force on the levelling of the plural forms in the perfect; although it does, of course, have no effect on the direction of levelling.

It is also interesting that the levelling for all varieties except St Ouen has occurred only in the non-liaison forms, while in liaison, a liaison consonant (/z/ or /t/) continues to be found, differentiating the 2nd and 3rd persons, both singular and plural. That is, the liaison consonant is not explained by the base form, which would have a /t/ liaison consonant in forms derived from the 3rd person and a /z/ liaison consonant found in forms derived from the second person.

It would appear that the liaison consonants are the result of an analogical reassignment of (or conservation of) liaison consonants regardless of the original liaison form found for each ending. This would in turn pattern with other 2nd person and 3rd person verb forms which regularly have these consonants in other contexts. The use of /z/ and /t/ as liaison consonants for 2nd and 3rd verb forms of various paradigms is common, such as for 2nd person plural present indicative in (5) and the 3rd person plural imperfect in (6).

(5) u: krjiz oz errná?: You, PL believe, PRES, 2PL PREP, DET, ART, PL ghost, PL 'Do you believe in ghosts?'

(6) i pakc a ma fam they speak, IMPF, 3PL to my, F wife 'They were speaking to my wife'

The absence of liaison in the 2nd and 3rd person plural perfect in the St Ouen variety appears to be a simple fact of the phonology of this variety of JNF, in which the liaison consonant for the 2nd person plural is /θ/ rather than /z/; the consonant clusters /drθ/ and /drt/ are not possible clusters in this variety (see Liddicoat 1994:83-84; Spence 1987). Liaison consonants are found in other contexts in this variety and otherwise function identically to the other JNF varieties.

12 Liaison consonants are so strongly markers of person that they may be inserted in contexts where they are not etymological, as in:

i s an vat an vil he 3, REF, PART go, 3, PRES PREP town 'He is going to town' (cf. French il va en ville)
Returning to the non-liaison forms, we need to consider in more detail the source for the three different dialect forms, the direction of any analogical change, whether the analogical change was from a less marked to a more marked context or vice versa, and whether any other factors may have contributed to the change. It is easiest to consider this for each variety separately.

Central varieties of JNF have the non-liaison form -ir for both 2nd and 3rd person plural. This form would be expected (given the collapse of perfect paradigms) for 3rd person plural, having come via regular changes from the Old French -irent. Forms in /r/ are found throughout the set of paradigms of the perfect in Old French, and the presence of /r/ is a typical marker of the 3rd person plural perfect in other French varieties, including standard French.

In these central varieties, then, we need to explain the expansion of a 3rd person plural form to cover the 2nd person plural as well. An argument can be made for the 3rd person being semantically basic and less marked than other persons (cf. Koch 1995). The extension of 3rd person plural to 2nd person plural in central varieties of JNF could thus be argued as the extension of a less marked form to replace a more marked form. The extension of 3rd person plural to 2nd person plural may also have been supported by analogy with the singular in which 2nd person and 3rd person forms are not distinguished in any of the verb paradigms except in liaison: that is in the very context where the distinction is retained in the plural. This process too could be seen as a process of abduction in which the rule set applying to the singular was perceived as applying to the whole paradigm. Such a process of abduction in which person is marked only in liaison form would also explain the retention of the otherwise anomalous liaison consonant /z/ in the 2nd person /r/ forms.

In the most common variety of JNF, the non-liaison 2nd and 3rd person plural forms are both -i. This ending is easy to explain as a 2nd person form, as /i/ is universally found in 2nd person plural forms in Old French throughout the perfect paradigms, and remains the usual marker of 2nd person plural perfect in standard French and other varieties.

An argument could then be made parallel to that for central varieties above, that this etymological 2nd person plural form extended to 3rd person plural contexts. However there is then a markedness problem, in that this would require the extension of a more marked (2nd person) form to cover a less marked (3rd person) context. It would appear that this spread of a more marked form was supported by sound changes which occurred in these varieties.

In some varieties of JNF, historically there was a process of assibilation of intervocalic /r/, while in other varieties /r/ was retained as such. Varieties where /r/ was retained are spoken in central regions; precisely those varieties which have the non-liaison perfect form -ir for 2nd and 3rd person plural. It therefore appears that the forms found in the plural endings of the perfect may have some relationship to this sound change: /r/ is always found in the perfect in varieties where /r/ is conserved and /r/ is never found in varieties where /r/
is assimilated. Assibilation of /r/ is a common phenomenon with a range of variations: /ð/ where the consonant remains intervocalic or becomes word-final after a vowel, or one of a variety of assimilated forms where a previously intervocalic /r/ has come into contact with a preceding consonant (Liddicoat 1994; Spence 1957). Liddicoat (1991) argues, using forms from the related dialect of Sark, that early stages of the assimilation process may have involved affrication, yielding a form /ʃr/ or /dr/, a form which is not found word-finally in either Sark Norman French or JNF. The existence of an intermediate form such as /ʃr/ could explain the syncretism of 2nd and 3rd person plural /t/ forms in the general variety of JNF more effectively than a simple argument in which a marked form has replaced a less marked form: it would be the case that an affricate /ʃr/ was levelled to a dental stop /t/ rather than becoming the expected interdental fricative /ð/, under influence from the 2nd person plural form /t/, supported as was argued for the /r/ forms by the syncretism of the singular forms. That is, the etymological 3rd person plural /ir/ became /it/ through sound change, but rather than continuing on to /ið/, the suffix changed to /it/ because of the association with the 2nd person plural /it/. In this case, similarity of form may have been more influential than markedness.13

The syncretic 2nd/3rd person plural perfect form in St Ouen, -iɛdr, is the most anomalous. In the majority of JNF varieties, the existing forms seem to be the result of a generalisation of one common regular ending to another person category, supported perhaps by analogy from the identity of the singular forms. However, the form -iɛdr in St Ouen cannot be accounted for as easily. While 3rd person forms in -drent are attested in Old French, they are infrequent and are restricted to a small number of verbs in the strong i and strong s perfects (see Table 2). This means that it is unlikely that -drent was a very frequent verb form in JNF in the earlier history of the language. In St Ouen, therefore, the origin of the syncretised forms appears to be a generalisation of a very marked, low frequency form across the entire paradigm.

The most plausible explanation for the form in /dr/ in St Ouen JNF involves a two-stage process: the adoption of a highly marked ending throughout the 3rd person plural, and then the extension of the 3rd person form to the 2nd person. The latter may be considered an example of an unmarked form being extended to replace a marked form, under analogical influence from the singular, as was argued above for the extension of the /r/ form in central dialects. The initial change, the adoption of a form in /dr/ throughout the 3rd person plural, may once again related to an early affrication of intervocalic /r/. It is probably the case that certain /dr/ 3rd person forms did persist in JNF, as vindrent and tindrent are attested in the 18th century diaries of Daniel Messervey (1896), who was from the East of the island, and used these forms in writing a regional form of

---

13 In addition, it may be possible that Old French strong s perfects continued to exist in JNF and that with the loss of prevocalic /s/ in these forms, the persistence of at least some 3rd person plural forms in /t/ may have supported the levelling process. However, there is no independent evidence in any Channel Islands variety that such 3rd person forms did persist.
Standard French. In the St Ouen dialect, it may be the case that an affricated form of /r/ (/'f/' or /dr/) was not reduced to either a stop or a fricative in 3rd person plural perfect forms, but rather levelled on the model of verbs ending in a consonant cluster /dr/. This form was then extended from the 3rd person plural to the 2nd person plural.

4. Conclusions

This paper has examined processes of morphological change in JNF for which marking alone is not the primary influence on analogical change. Analogy, in these cases, interacts with phonological change to produce results which are not consistent with a mechanism for analogical change which argues for a directional influence of unmarked forms on marked forms, of frequent forms on less frequent forms. While such general principles may work to describe grammatical change as a global phenomenon, they are not necessarily adequate to account for patterns of analogical change in particular languages.

In JNF, sound change works in a number of ways to create irregularities in the grammatical system and these irregularities appear to have become sites for analogical change. It appears that it is the presence of the irregularity in the paradigm which becomes the catalyst for analogical change, rather than the frequency of the irregularity in the system. This would appear to be relevant for the case for the levelling of the Conjugation 1 paradigm in particular. The particular ‘strength’ of the /i/ verbs in JNF, in the perfect at least, appears to be their total regularity rather than their frequency. Sound changes may also lead to similarities of form, as was argued for the endings of the 2nd and 3rd person plural, and in contexts where the marking of person by verb endings is redundant, similarities of form may lead to a coalescence of forms through an assimilation of one ending to another. In this way, elements which are of very low frequency in the paradigm may exert an influence on higher frequency forms, because of language internal frequencies of particular phonotactic patterns.

Sound change may also lead to ‘blending’ of paradigms as one of the consequences of internal irregularities of the paradigm. This has been the case in Conjugation 1A, Conjugation 4 and Conjugation 5 verbs in which some sound changes led to similarity in forms between some parts of the verb paradigm and forms in the Conjugation 2 and Conjugation 3 paradigms. These resulting paradigm internal irregularities appear to provide points of contact where regular paradigms exert an influence on the irregular paradigm. As verbs or sub-classes of verbs assimilate to the more regular paradigm, this increases the overall numerical strength of such forms and increases the influence that they may have on other parts of the verb system.

It appears that the location of such irregularities in the verb paradigm also exerts an influence. It is to be noted that for Conjugation 1A verbs in particular, the phonologically conditioned irregularities were found in past contexts — the
past participle and the perfect itself — and this may strength the analogical attraction of another, regular, paradigm. In the case of Conjugation 1A verbs the regular /i/ paradigm contains elements which are identical to the irregularities which have appeared in these verbs with the result that the conservation on non-\textit{i} forms takes on the appeareance of being the irregularity in the system. For Conjugation 4 and 5 verbs, the influence appears to come from the forms of the imperfect, as well as from other non-past verb forms in which an irregularity with /i/ forms is generated. In those verbs where past forms, especially the past participle, act as a counterweight against analogical levelling on the basis of /i/ forms, analogical change to /i/ forms seems to be resisted, at least in some cases, although these forms are also characterised by internal regularity within the perfect.

The development of phonologically conditioned irregularities and the blending of paradigms seen in JNF appear to provide a site in which abductive processes (Andersen 1973, 2005) are at work and in which regularities in other parts of the verb system get incorporated into less regular elements of the system. This would mean that abduction is not an explanatory process in itself, but rather is associated with sites of paradigm internal inconsistency which become regularised through the applications of alternate regular rules.

The evolution of the perfect in JNF shows above all that analogical change within a particular language is a complex phenomenon, and that, while analogical processes may be invoked to explain what has happened in a particular language, such processes are not always predictive of patterns of change. In JNF analogical change cannot be seen as an alternative pathway to regular conditioned sound change, but rather as interacting with sound change in complex ways, within particular parts of the language system. The factors which lead to analogical change functioning in particular ways seem to affect elements of the morphological systems in highly localised ways, with multiple trajectories interacting in the process of analogical levelling.

References


GRAND-DADDY MORPHS
THE IMPORTANCE OF SUFFIXES IN RECONSTRUCTING PAMA-NYUNGAN KINSHIP

PATRICK MCCONVELL
Australian Institute of Aboriginal and Torres Strait Islander Studies

1. Introduction

There are many suffixes on kinship terms in Pama-Nyungan languages in Australia, some readily segmentable in the recent forms of the languages, some where the suffix may only be discerned by internal or comparative reconstruction. Among the most common are one which has been studied in a short article: -rti (Nash 1992) and another which has been mentioned by Harold Koch in writings and presentations: -ji/-ju. The latter has been thought to descend historically from a first person singular possessive suffix; the absorption of possessive suffixes, particularly first person singular, is a particularly common pathway of change, not only in Australia (Koch 1983, 1995, 2003a, b). Another pathway of change which has been identified in Australia and more generally is from a dyadic affix to a kinship-reciprocal affix (Evans 2003, 2006). The focus in the present article is on affixes which may have come from old inflectional and derivational paradigms, but which have become absorbed into the kin-term itself, rather than recently productive suffixes (such as number, gender, case or address/reference) although examples where there are productive paradigms are cited for comparison. Apart from those mentioned above, several other reconstructable suffixes of this type found in quite large areas of the Pama-Nyungan range are examined to see if light can be shed on their origins and distribution. This kind of morphological evidence provides particularly strong support for the notion of a Pama-Nyungan family.

2. Pama-Nyungan and kinship

Harold Koch has been one of the foremost linguists involved in the reconstruction of the Pama-Nyungan family in Australia using the comparative method. As well as his own contributions on the Arandic sub-group, he has provided masterly summaries of the method and its results in Australia (e.g., Koch 1997), and encouraged others to produce work on other sub-groups emulating his own exacting standards (Bowern & Koch eds. 2004). Despite criticism of the ‘Pama-Nyungan idea’ and the comparative method from some quarters (Dixon 2002, Clendon 2006) these efforts have won support and admiration from the great majority of Australianist linguists.
Kinship terminology is a particularly fertile field for Pama-Nyungan comparative studies. Indeed it might well be said that the validity of the Pama-Nyungan family could be justified on the basis of kinship terminology alone, although there is of course a huge amount of other evidence also.

2.1 The case of *kami ‘mother’s mother’

Just look at one proposed PNy reconstructed root *kami (see map, in Peterson, McConvell and McDonald 2006). The reflexes of the form are found throughout the Pama-Nyungan region and just in one area outside, which is accounted for by borrowing.

Semantics is one of the problems of reconstruction. Usually reflexes of a root have different meanings in different languages and the methods of how to attribute which sense or senses to a proto-form are less well established than those of phonological reconstruction. In the case of *kami as we see in the map, while the meaning ‘mother’s mother’ is widespread, there are in some areas additional meanings or extensions such as to ‘mother’s mother’s brother’ in polysemous cases, and some changes to a different meaning (perhaps via an intermediate polysemy) such as ‘father’s mother’ in the Karnic languages. In this case, a different root for ‘mother’s mother’, kanyini, has taken over, as kami changes to FM. There are various chains of meaning change involved with this and other roots. Tracing from current forms back through these chains can be very useful in narrowing down the possibilities for a proto-meaning.

In the case of the domain of kinship there is a significant body of theoretical writing and empirical research on kinship semantics and kinship systems in anthropology as well as linguistics, and to a lesser extent on possible changes in systems. One way to check the plausibility of a single reconstruction is to set it in the context of a reconstructed system as a whole, by reconstructing all the terms in a proto-system. If the reconstructed proto-system is the same as one of the limited number of kinship systems known to exist, there will be fewer doubts about it than if it is an unheard of or bizarre type of system, for instance.

Phonological reconstruction on the other hand, is not so problematic, and we can be fairly confident about the form of *kami in particular. There are divergent reflexes of the term which are accounted for by regular sound change: for instance abmi-ngarr in Oykangand, in Southern Cape York Peninsula. In this case, initial consonant dropping and prestopping of the nasal has occurred (cf. Koch 2003:134-5 for parallel changes in Arandic). In actuality though, as the hyphen in the cited form implies, the terms always occur with a suffix -ngarr in Oykangand.

2.2 Reconstructing kin affixes

This leads us to the theme of this paper: a great many languages have kinship affixes – affixes used solely or primarily with kinship terms – either currently
productive or reconstructable. The reconstruction of morphology constitutes the other major part of the work of kinship term reconstruction alongside that of the roots. Indeed as we shall see demonstrated in this paper it is impossible to carry out adequate reconstruction of kinship terminology in Australia without close attention to the morphological analysis of the kin terms. Affixes which may be buried inside roots and not distinguishable by evidence available in one sub-group may reveal themselves as segmentable suffixes in another sub-group, and reconstructable as affixes rather than part of the root, in the common proto-language.

There is also important research to be done in tracing the pathways of semantic change in suffixes and their effects on the roots they are found with and may merge with. These patterns of change, which may be universal or relate to regional cultures, will assist in reconstruction. I deal with some of these issues in section 3; seminal work on this is to be found also in Dench (1997).

In this paper, we shall look at several examples of this kind within the Pama-Nyungan family which will demonstrate the twin claims made here, that kinship terminology and kinship morphology make a substantial contribution to establishing Pama-Nyungan as a valid family on the basis of the comparative method; and that kinship term reconstruction can only proceed hand-in-hand with morphological analysis and reconstruction.

The first two suffixes dealt with -rti (section 2) and -ji/ju (section 3) are very widespread across the Pama-Nyungan family and are probably to be reconstructed to some high level sub-group within PNy, if not proto-PNy itself. In many languages these suffixes have been ‘buried’ in kinship terms in the sense of being obligatory, having no distinct meaning and no use outside this context. However in other sub-groups they appear as optional, and with a function of adding meaning. These attestations also point towards etymologies that include proto-forms which are more general than kinship alone, particularly in the case of -ji/-ju which is most probably an old first person singular possessive suffix.

The question of whether the specialised reciprocal kinship meaning of a suffix (-jarr and similar forms) descends from the meaning of dyadic and ultimately a more general dual is the topic of section 4. Finally -ny (and variants) discussed in section 5, which is also very widespread in Pama-Nyungan, is traced back and some hypotheses about its original function discussed.

These are only some examples of widespread kinship suffixes in Pama-Nyungan; there are a number of others which we do not have the space to discuss here (eg McConvell 2007b). Other morphological processes apart from suffixation have also played a role in the prehistory of PNy kinship – for instance, partial reduplication of roots – but space limitations did not permit inclusion of a section on this (see McConvell 2007a).

As with roots, the proto-affixes require to have both their phonological form (and phonotactics) reconstructed, and their meaning. The meaning of the affix has to be interpreted along with the meanings of the roots with which it
occurs. Since such affix meanings may be neither very clear, nor monosemous, nor regularly compositional, even in single languages, the task may be complex. In some cases understanding of kinship morphology may be not only useful, but even necessary for full understanding of the meaning of the roots with which affixes combine.

2.3  

Proto-Pama-Nyungan Root+Suffix: *kami-nyjarr

In the case of *kami, the term we began with, as well as reflexes of the plain root, reflexes of the root with suffixes are also widely distributed in the Pama-Nyungan area. One such complex proto-form is *kaminyjarr, reflexes of which are found also in regions very distant from each other geographically and in languages in very different subgroups of Pama-Nyungan. It is found in a number of languages of Cape York Peninsula, for instance in Guugu Yimidhirr, meaning the reciprocal of 'mother's mother' (i.e., 'woman's daughter's child'). In New South Wales *kaminyjarr is found reflected as gamidharr in dialects of Bundjalung with the same meaning. In these two languages reflexes of the basic grandparent root *kami are found as well as this grandchild derivative. In two other, again widely disparate, PNy regions, however, the grandchild complex root is found without a reflex of the *kami grandparent form, which has been replaced by another root. In Ngumpin-Yapa, in the central west of the Northern Territory and into the southern Kimberley, the form is kaminyjarr (e.g., in Gurindji), and in Yolngu languages of North-east Arnhem Land gaminyarr with an expected regular sound change nyj>ny (McConvell 1997).

The existence of such morphologically complex forms in the reconstructed proto-language, with reflexes scattered throughout the PNy zone is again a powerful plank of evidence in support of Pama-Nyungan. It also gives us a window into how morphology has played a role in the kinship terminology throughout the history of Pama-Nyungan, from the start. Morphology, both productive and lexicalised, plays a huge part in the story of kinship terminology and the related systems in Australia and it is only possible to scratch the surface with this brief examination.

3.  

The Australian kinship affix *-rti

This is the title of an article by David Nash (1992). He shows that this suffix is found on a range of kinship terms in quite a number of Pama-Nyungan languages, especially in the west but also in New South Wales and parts of South Australia.1 The status of this as a separate suffix can be demonstrated synchronically for some of the languages because the bare stem appears, for

---

1 Reflexes of the suffix are apparently absent in languages of western Queensland, and the Murray-Darling (Nash 1992:130).
instance, when other suffixes are added. In other languages, however, the old suffix has been frozen as part of the root but is clearly a suffix etymologically. It alternated with other suffixes, particularly -ji/-ju (also -li/-lu) occasionally in the same languages, often in different languages (discussed further in the section 3 and McConvell 2007b).

The root discussed above *kami also occurs with this suffix as gamidi in Narangga, a language of the Spencer Gulf in South Australia, which has the meaning 'mother's mother' as posited for the PPNy root. As with this case, on most occasions where this suffix occurs, there is no obvious change in meaning brought about by the addition of the suffix, when compared with the bare root in the same language or the imputed proto-form. In Gurindji in an old fused inseparable form like kurturtu (woman’s child; Nash cites it from Mudburra) possibly -rtu originally marked a kinship term as opposed to an unmarked form for ‘child’ kurtu (no longer used in Gurindji but retained in Warlpiri as kurdu).

Some other forms in Gurindji have optional -rti (-rnti) suffixes (not cited by Nash): ngaji(rti) ‘father’ ngamarnti ‘mother’, where the normal word for mother is ngamayi < *ngamaji by regular intervocalic lenition of j-y. My Gurindji consultants indicated that the difference was one of added politeness or formality in the register, with the -rti forms. They also tend to be used more as address forms, although not exclusively.

Apart from Pama-Nyungan, Nash finds solid evidence of what seems to be a related kin suffix -rdi only in Mara, a NPNy language of the Gulf of Carpentaria (1992:133). In this language it marks the first person referential form (‘my X’). Nash floats the hypothesis that this could be the origin of the suffix more generally across PNY. One advantage is that this is a parallel evolution to that outlined by Koch (1983, cited by Nash 1992:138, and discussed further in the next section) in which the first person possessed form of a kin is relatively unmarked and tends to be reanalysed as the base form. Disadvantages of the Mara-origin hypothesis are many, a number recognised by Nash himself (1992:133). He notes that languages in the Maran subgroup do not have this suffix, and it clearly seems an innovation in Mara, but there is no NPNy source in neighbouring languages. For me that fact in itself clearly rules out the hypothesis since it is impossible to imagine that such a recent innovation could be the source of a suffix found over such a wide area of Pama-Nyungan. Borrowing in the opposite direction from PNY into Mara with a shift in function could be an option or a hypothesis that Mara -rdi and PNY -rti are unrelated is also possible, as Nash suggests.

In any case, because of its wide distribution in PNY, and the unlikelihood that it has diffused, we are looking at a very old meaning of the suffix as used with kinship terms. At the moment we cannot be sure what that meaning was and whether it was a broader type of suffix that acquired a special type of meaning (often weak or non-existent eventually) or a specialised kinship suffix from the start.
Other apparent examples of the suffix -rti include its suffixation to a PP Ny root *piimu FZ in forms like Warlpiri pimirdi, although the Walmajarri form pimiri is anomalous in this light. A connection with CYP forms in -r is possible like pemr (Yir Yoront; Koko Minjena Thomson 1972:30). This raises the possibility of a link between the -rtu/-rti suffix and suffixes with liquids including -l, -ri, -rlu, etc., which is explored in McConvell (2007b).

4. The suffix -ji/-ju

Another problem with the hypothesis mentioned above deriving -rti from a first person possessive suffix is that such a suffix with such a form is unknown, either in the putative source language Mara, or indeed anywhere in Australia in PP Ny or NP Ny. However with the kin suffix -ji/-ju, which is often found in languages close to those with the -rti suffix or even in the same language, the prospects are much brighter because a first person singular dative and even kin-possessive suffix is attested in a number of Pama-Nyungan languages widely distributed across the continent.

It may be possible to reconstruct such a first person singular oblique enclitic to PP Ny either as a suffix on nouns like kinship terms or as a more general sentential enclitic, or both. The source of such a clitic could be a dative form of the first person pronoun *ngaju (H. Koch p.c.).

Koch has not discussed this particular kinship suffix in print but has examined the general phenomenon of first personal singular affixed forms of kinship terms becoming the unmarked lexical forms in Australian languages, and in particular the Arandic suffix -ye. Koch (1995:51-4) cites examples from northern Australian languages where the first person singular propositus form (‘my X’) is unmarked (the same as the citation form) and quotes Croft (1990: 145) as calling this an example of ‘local markedness’ where there is ‘third person/first person markedness reversal’ (third person being assumed to be the unmarked member of a paradigm generally).

In the above cases in Arnhem Land, the first person singular propositus forms cited by Koch do not appear to include any old first person affixes. In the case of Arandic however there is evidence that a suffix -ye has become a ‘zero’ – has been reanalysed as part of the stem. In Kaytete the suffix -ye is used to signal first singular propositus in such forms as arrenge-ye ‘my FF’ but other persons have the stem arrenge with kin prefixes. In Alyawarre, on the other hand, while a form like aknge-ye with the same prefix can be used to mean ‘my father’, it can also be used simply to mean ‘father’ with other propositus persons (e.g., akngeye ngkwenhe ‘your father’). Some Kaytete kin terms also contain the suffix but it has been reanalysed as part of the stem (e.g., arlweye < *kurla + -ye ‘father’) (Koch 1995:52-53).
Returning to the question of -ji/-ju, it may be that Arandic -ye is a cognate, but I will not attempt to probe into that further here. The kinship suffix of this form is segmentable in many north-western PNy languages since it is dropped when some other suffixes are used. For instance in Gurindji with ngamayi ‘mother’ (< *ngamaji by regular lenition, cf. Walmajarri ngamaji M) the dyadic form is ngama-rlang ‘mother and child’ (rlang is the regular dyadic suffix). Many of the kinship terms with -ji/-ju suffixes do not have any person signification but are personless forms of the kinterms, synchronically. However in Jiwarli and other western Pilbara languages -ju is the regular first person singular propositus suffix for kinterms – e.g., kantharri ‘MM’ kantharri-ju ‘my MM’ (Jiwarli, Austin 1992:94).

Other examples appear to have used -ji to form a kinship term from a non-kinship term, as with *kurri ‘woman’ + -ji > ‘mother-in-law’. This may be parallel to the function of -rti mentioned above as changing ‘child’ (non-relational) to ‘woman’s child’. There is an additional affinal sense added. In understanding the semantic shift here it is important to take account of the traditional ‘promise’ or betrothal system which was actually in many cases ‘mother-in-law bestowal’ (Maddock 1972) whereby a man (or a boy) was promised the future daughter of a girl or young woman as his future wife. In this context the most salient referent of ‘my woman’ could well be ‘my bestowed mother-in-law’. Data below is from Alpher 2006 (update of Alpher 2004).

**Warlpiri** kirriji ‘mother-in-law’ (KLH)

**Guurrindji** kurriji ‘mother-in-law’ [actually mother-in-law of woman] Syn. mali < *mayili and maliyi < *malici. (PMcC)

**Warumungu** kirriji ‘woman’ (JS&JH)

**NGarluma** kurri ‘girl (adolescent)’ Also kurrirri ‘Pleiades’. (KLH, GNOG)

**Panyjima** kurri ‘marriageable girl’ (AD)

Additionally kurri ‘woman’ is found in Western Desert dialects and Mudburra has kirri ‘woman’.

Elsewhere in the western Pilbara there is some indication that a parallel kin term can be shifted to an affinal meaning by the addition of -ji.

**Panyjima** mari ‘younger sister’ (AD)

**Watjarri** marici ‘brother’s wife’ (WD)

A form -ju is reconstructable as a first person oblique member of a paradigm of pronominal enclitics at least to some higher-level subgrouping of northwestern Pama-Nyungan. This proposal follows the position taken by McConvell and Laughren (2004) that there is a widespread old inherited pronominal enclitic system at least within western Pama-Nyungan and contrary to the ideas
of Dixon (2002; partially echoed by Sharp 2004) that pronominal enclitic systems diffuse and are recent innovations in this region. The subgroups within this putative larger sub-group in which this set of pronominal enclitics can be reconstructed include at least Ngumpin-Yapa, Marrngu, Wati, Kartu, and Ngayarda. Even in those languages in which most enclitic forms have been lost, the 1sg dative/benefactive -ju is retained e.g. the Ngayarda language Panyjima (Dench 1991:160).

There is a differentiation between the languages with -ju (which is also the proto-form *ju) and the Ngumpin languages which have diverged to -ji in the sentential pronominal enclitics, and in the pronominal kin suffixes, even when they are not productive or recognised by speakers, such as in the form *ngama-ji already discussed. This parallelism tends to support the idea that these elements are linked, and come from a common source morpheme.

However, this suffix is more widespread than just this western part of Pama-Nyungan. Umpila in central eastern CYP has kami-cu alongside kami ‘mother’s mother’ perhaps specially focussed on the reciprocal, wDC. In the initial-dropping language Uw-Olkola in southern CYP the reflex of the same root is found with a reflex of the *-ju suffix: abmi-dh ‘parallel grandparent’. Kaanytju pula-thu ‘son’s son’ (Hale 1976:58) has an allomorph -thu on the FF root *puula (proto-Paman with some probable cognates in PNy outside CYP) in the reciprocal meaning ‘son’s son’.

These are a few among quite a large number of examples in different Paman languages where -ju/-thu is apparently a (younger) reciprocal suffix. Umpila reciprocal ngaci-cu ‘daughter’s son’; cf. ngacimu ‘MF’ cf. PNy *ngaji MF Yintjingga (Thomson 1972:28 retranscription by Rigsby, called Ayapathu) kami FF,MM, etc., kamindhinhu, kamithu mSC, wDC and a number of others.

One can imagine perhaps that an address form may be used in the junior reciprocal function. If there was an original meaning ‘my’ this would fit with a change to an address form. However this scenario remains speculative and needs to be supported with evidence of such connections and pathways actually in the North Queensland area.

-ju is also a suffix on non-reciprocal kinterms such as many terms reflecting *yapuju, built on *yapu ‘younger brother’, such as Kuku Yalanji and Dyirbal yapuju ‘younger brother’. While the reciprocal aspect is missing here, the ‘younger’ element is present.

In the absence of evidence of actual links one may have to remain agnostic and treat the suffixes here (first person possessive and junior reciprocal) as accidental homophones rather than cognates. For those suffixes which currently have no particular effect on the meaning of the term it must be determined case by case whether they are more likely to be connected to the possessive pronoun or junior reciprocal.

In south-eastern parts of Australia there do not appear to be reflexes of -ji/-ju either as an archaic residue, or a current nominal possessive morpheme,
or a more general dative/benefactive enclitic which could lie behind this.\footnote{Ngiyambaa in central NSW has a second-position clausal enclitic for 1 sg oblique -dhii/-jii but this is fairly obviously a recent development from the free pronoun ngathii so could be counted as part of the story of the kinship suffix elsewhere.} In the Kulin languages of Victoria there are obligatory pronominal possessive suffixes on kinship terms and these too are the same as the clausal object pronouns at least in some languages (Hercus 1986: 51, 88, 129-131). The form of the 1\textsuperscript{st} singular possessive is -(ng)eg or -(ng)ai, however, apparently unrelated to -ji/-ju. In Buwandik, on the Victoria/South Australian border, again there are pronominal possessive suffixes which are the same as oblique pronominal enclitics, but the 1\textsuperscript{st} singular is -ngayn perhaps related to the Kulin forms but not to -ji/-ju.

One might speculate that formation of kin possessives from encliticised oblique pronominals is an old feature of Pama-Nyungan which was expressed by different forms in the far south on the one hand and the north and west on the other, but absorbed as ‘zero morphs’ into some kinterms in some parts of the north-east or entirely lost in the south-east. The reflexes of *-ju are widely distributed in Pama-Nyungan but other forms of first person possessive occurring on kin-terms in the areas of south-eastern Australia where *-ju reflexes are absent (-ngai, -ngeg, -ngi) may descend from the initial part of the first person singular pronoun rather than its second syllable, the probable origin of *-ju.

5. \textit{Kin reciprocals – from a dyadic/dual suffix?}

In some languages kin reciprocals – such as MM and woman’s DC – use the same term. In others there is a different reciprocal and this may be built on the complementary (usually senior) term by the addition of a suffix which may cover a number of or all reciprocals. We have already looked at examples – the use of the suffix -thu as a reciprocal.

Another reciprocal suffix already mentioned is the -ny-jarr suffix on the forms reflecting *kaminyjarr which provide such cogent proof of the widespread nature of the PNy family. As will be argued in the next section the -ny- here is probably a separate suffix found commonly with *kami and some other terms. The focus here is on the origin of -jar; there is no space for an extensive treatment of dyadic and reciprocal morphology.

Evans (2007) has argued that there is a common “type of extension, [is] from a morpheme originally meaning ‘pair’ [dyadic expression] to ‘one of a pair’”. He refers to Evans (2003) for exemplification from Australian kinship. In that paper (2003:5) he discusses “how areal diffusion can lead to parallel semantic patterning … how a single semantic target can be reached through widely varying patterns of semantic extension”. Clearly in our efforts to connect
kinship affixes of differing function whether by inheritance, formal borrowing or semantic replication, such hypotheses are of great value.

Evans (2003: 32-33) cites the suffix -nyjarr 'kin reciprocal' from McConvell (1997, discussed above in relation to the root *kami) as a possible example of such a semantic from dyad or dual to kin-reciprocal, but concludes ‘the deeper origins of this suffix are not known’.

There are certainly a number of possibilities here, some of which are discussed separately by Evans in the same paper.

(a) dyad suffixes in several languages of form -ntjerre (Evans 2003:32-33) could be connected (cf. also Kalkatungu grandparent–child dyad suffix -wanyjirr, Blake 1979:82);

(b) dyad forms like -jirr in Dyirbal (Evans 2003:15-16; Dixon 1972: 234-5)

(c) Gugadj dyad -tyira and following Breen (1976: 294) back to a (possibly Pama-Nyungan) proto-form for ‘two’ *kujarra. Dixon (2002:77) also supports this kind of etymology, including for Dyirbal -jarran ‘a pair’.

(d) as Breen mentions following Hale, the ‘two’ root could also be connected with the proprietive ‘having’ affixes -jarra and similar found in Pama-Nyungan, and the latter then to the dyadic morphemes of similar form.

For the semantics of such putative transitions see Evans (2003). The issue here is whether there could be a formal connection between the morphemes involved. One problem for a unitary account of most or all of these forms is the presence of an initial nasal in the -nyjarr and (a) forms above but not the others. However recall that even the more convincing of the cognates of the maternal grandchild root with *-nyjarr has a form without a nasal in Gidabal.

Two leading hypotheses here would be: 1) the nasal is a separate kin suffix; this is discussed in the next section; 2) there is some kind of alternation between homorganic nasal-stop clusters and plain oral stops which manifests itself in one form or the other between different languages.

I have discussed the synchronic phonological rule of Nasal Coda Dissimilation (NCD initially called Nasal Cluster Dissimilation, McConvell 1988) which has a patchy distribution across Australia and whose effects can also be seen in diachronic change. There are also a number of PNy roots which appear to alternate between nasal and oral consonants in initial position including kinship terms. It is possible that some early Australian proto-language (maybe

---

3 Evans (2003: 32-33) canvasses the idea that this suffix may be a cognate of verbal reciprocal morphemes in various Australian languages but finds major problems with it.
PNy) had a systematic distinction between oral and prenasalised stop phonemes which was subsequently largely lost (Harold Koch p.c.)

However, I pursue the simpler hypothesis 1 here, that the suffix is -jarr related to dual and pair morphemes and to *kujarra ‘two’ and that -ny is a separate kin suffix preceding it – this is discussed in the next section.

6. The suffix -ny

Another suffix which is found with our first example *kami is -ny. As suggested in the last section this may also make up the first part of the reciprocal suffix -nyjarr the second part of which may be derived from a dyad/dual suffix. The suffix -ny is found in Margany-Gunya gaminy MM, and has variants -n in the Mayi languages (possibly -nh?) -nya in Kungkari and -na in Yiningayi and Biri. A final -n in some words continues an earlier *-nh or *-ny (Alpher 1997:12).

In fact to some extent this set of suffixes is found on reflexes of all the PPNy grandparent terms and some others. The PPNy grandparent reconstructions (McConvell & Keen 2006, to appear) are as follows. It is assumed that the system was Kariera and gender of referent played a strong role in the early meaning of these roots but was demoted in many of the later reflexes:

- *kami mother’s mother & father’s father’s sister
- *ngaji mother’s father & father’s mother’s brother
- *papi father’s mother & mother’s father’s sister
- *mayi-ri/li father’s father & mother’s mother’s brother

There is an alternate root for *ngaji MF found mainly in the west *jamV, and for *papi MF also in the west *ngapV-. The FF root above is less certain as a PPNy reconstruction than the others and does not appear to take the -ny suffix or its variants.

The reflexes of the *ngaji and *papi roots with the -ny, -n, -na, etc., suffixes include:

- Margany-Gunya ngadhiny ‘mother’s father’ (GB)
- GBadhun ngajina ‘father’s father’ (PS)
- Margany babiny ‘father’s mother’ GB
- GBadhun babina ‘father’s mother’ PS
- Ngawun papin ‘father’s mother; son’s child (to woman)’

One might argue that this suffix has the same source as the -ny(a/-nh(a)) suffix which is an accusative/dative in early or proto-PNy (Blake 1987) but also
turns up as a marker of names or proper nouns, as in Western Desert. A variant
-n- has also been argued to account for the -n- which precedes case-markers etc
for animate nouns in Kuku Yalanji: Dative *-/nka (animates) */-ka (elsewhere);
‘having’ */ntji (animates); */-tji (elsewhere). Heath discusses an old accusative */-n
or inverse */N (Heath 1984, 1987) and its relation to accusative */-nha and dative
/*-na

Another possible source of */-ny could relate to the old PNY Feminine
pronoun base */nya(a) which is the source of disappearing -n. (Alpher 1987: 17).
However, while */kami and */papi are reconstructed in this paper as feminine in
proto-PNY, */ngaji MF is not, so the hypothesis that the nasal suffixes are gender
markers would not be supported.

Alpher further discusses (1987: Fn.14 p.17) the forms K-Thaayorr thu(u)wn
‘woman’s child’; Gugu-Badhun thuwana ‘nephew’; Bidjara dhuway ‘son’; Gugu-
Yimidhirr dhuway ‘woman’s child’; Kuku-Yalanji thuway ‘man’s sister’s child’;
Yuulingu dhuway FZC; with various suffixes compared to forms without a suffix
like Wik -Mungkan thuw. He comments that the Queensland languages vary as to
whether the form ends in n(a), y or nothing in a way that cannot be accounted
for by phonological conditioning. Languages that have */-na also have */-na in other
kinship terms – Gugu-Badhun kamina MM (cf. */kami); papina FM (cf. */papi) and
pimuna FZ (cf */piimu). Warlpiri attaches */-na to some kinship terms.

The root */thuwa is found in CYP generally as ‘woman’s child’ as in
Wunumara but as part of an Omaha skewing shift which becomes frozen,
changes meaning to FZC, marriageable male cross-cousin for woman in Yolngu
and further west (McConvell & Alpher 2003; McConvell & Keen 2006, to appear).
The meaning in Thalanyji far to the west is ‘the sister of such a man’. In
Wunumara the putative archaic suffix is */-ni rather than the */-na forms cited
above: thuwaNi ‘child (to woman); sister’s child’. Position of ‘N’ unknown. In
Thalanyji in the western Pilbara the form thuwayi ‘sister-in-law’ also shows
*/-ni, and the meaning seems to relate to the skewing found in Yolngu.

For the */-N(a) forms first person has been suggested as a source (particularly
for */-na) but perhaps more plausible is a connection with the proto-Pama-
Nyungan accusative */my(a) which also does duty as a marker of proper names in
several areas, perhaps also going back to PPNy. A proper name marker would
also be more likely where the form is used as an address form.

It is proposed here that this */-N suffix forms the first part of suffixes like
*/nyjarr which form junior reciprocals. This is not in contradiction with the
statement above that N is more associated with senior reciprocals. The second
part of the suffix is proposed to derive from a dyadic/dual morpheme */jarr and
this type of formation – which looks as if it goes back to PPNy or close – is
known to be usually formed from the senior of the pair. The meaning change
which then occurs is that which converts a term for a pair to the term for one of
a pair.
7. Conclusions

Harold Koch has argued consistently that etymology is at the core of our research, not only in lexical reconstruction, but also in morphology: we must search for the origins of the morphemes and patterns and the paths by which they were inherited or borrowed into recently attested languages. Because, as in this paper, we are often dealing with monosyllabic affixes, whose meaning and function can change over time, the dangers of confounding one affix with another are great and demand a rigorous and conservative approach. Consequently in this paper I have not been able to reach cast-iron determinations about the history of these bits of morphology, often now bleached of meaning and incorporated into lexical items.

Nevertheless I believe progress is being made and this is not the end of the road. Further detailed study of the distribution of these morphemes which are found with kinship terms, of the sub-grouping within Pama-Nyungan family and of the sound changes affecting these groups of languages will certainly enable us to get closer to the goal of tracking their paths accurately from ancient proto-languages.

Understanding of semantic change is also a key element, not only global and general tendencies, but also particular transitions which are embedded in the Australian Indigenous cultures – there are a number of examples in this paper where this has played a role.

Whether or not we can be sure of the ultimate provenance and function of these obscure ‘grand-daddy morphs’ at this stage, we can be sure that they are widely distributed within the Pama-Nyungan family. This leads us to the conclusion that they are inherited either from proto-Pama-Nyungan itself, or from some fairly high-level proto-language within it, the nature of which is sometimes unclear at our present stage of knowledge. Some kinship term roots from time to time have been borrowed with their affixes from or into non-Pama-Nyungan, and between sub-groups of Pama-Nyungan, but this is relatively minor phenomenon compared to inheritance.

One of the proto-affixes in which Harold Koch has been interested is *-ju (yielding *-ji in proto-Ngumpin-Yapa). It seems probable that this was originally an enclitic form of *ngaju the first person dative pronoun.

The suffix -rti/-rtu is widely distributed in Pama-Nyungan also with gaps in the east. The ideas that this was a borrowing from non-Pama-Nyungan or that it is also a form of an old first person suffix have not been supported here. Naturally there would be a corresponding sound to the retroflex stop in the eastern languages where it is absent, and my hunch has been that this is r. Consequently forms with -ri, and -r in the east could be connected, and also forms with lateral-initial suffixes -l, -li, -rlu may be related (McConvell 2007b). The meaning of *r(t)(i) seems to revolve around ‘junior reciprocal’ but there is also a ‘respect’ register affiliation in many occurrences.
A number of the *r(i)/*rti forms seem to form pairs with terms with nasal suffixes -ny, -n, -na leading to a conjecture that perhaps in some early system the *-N- forms may have formed senior reciprocals as a counterpart to the *-r- junior reciprocals. Partial reduplication, found in scattered locations from the tip of Cape York Peninsula to Adelaide at least (McConvell 2007a) could have been another way of marking senior reciprocals.

As with much of the content of this paper, this is work to be further pursued, both in its linguistic aspects and its anthropological implications. The etymological quest for which Harold Koch has been our guide requires painstaking rigour, but also yields rich rewards in the understanding of how the languages and the kinship systems of Australia evolved.

References


MORPHOLOGY OF THE EGGS, AND WHAT IT CAN TELL US ABOUT ROMANIAN NOMINAL INFECTION

KIM SCHULTE
University of Exeter

1. Introduction

In several Romance languages, the plural of reflexes of Latin ŒVUM ‘egg’ shows some degree of morphological and morpho-syntactic irregularity. After a brief presentation of the diachronic process that has led to this situation, this study focuses specifically on the genitive plural of Romanian ou ‘egg’, examining present-day variation between two forms; it is argued that what can be observed is a symptom of a currently ongoing, more general process of reanalysis of nominal inflectional morphology in Romanian.

2. Latin neuter nouns in Romance

The disappearance of the Latin neuter gender across the Romance-speaking world1 was caused, at least in part, by phonological changes (erosion of word-final consonants, loss of the opposition between long and short vowels) that increased the degree of syncretism already present in the Classical Latin declensional paradigms. Most neuter nouns became formally indistinguishable from masculine nouns in the nominative and accusative singular (-US, -UM > /-o/) and were therefore generally incorporated into the masculine category, while for those ending in a consonant or -E “the new gender appears to have been arbitrarily assigned” (Penny, 1991:107). Polinsky & van Everbroeck (2003:362-363) show that the decrease of neuter nouns was a gradual process that was already well underway in Late Latin.

In the plural, on the other hand, Latin neuter inflections differ significantly from their masculine and feminine counterparts, consistently ending in -A in the nominative and accusative plural, as exemplified in Table 1.

A simple phonologically motivated merger of the neuter with masculine or feminine forms is not possible in the plural, as the final -A does not coincide with any non-neuter plural endings, either before or after the regular sound changes undergone by Late Latin. However, the survival of neuter plural forms when the neuter had disappeared in the singular and arguably no longer existed as a distinct category at all appears to have presented speakers with an asymmetrical and possibly cognitively problematic paradigm that was dealt with in different ways; a number of strategies, either eliminating or integrating

---

1 Some Romance varieties do preserve traces of a neuter gender (cf. e.g. Lausberg, 1956:26-27).
these forms into a new two-term gender system, can be observed in different Romance varieties.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Nom./Acc. Singular</th>
<th>Nom./Acc. Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>neuter</td>
<td>LABIUM/LABRUM ‘lip’</td>
<td>LABIA/LABRA ‘lips’</td>
</tr>
<tr>
<td></td>
<td>(SUPER)CILIUM ‘eyebrow’</td>
<td>(SUPER)CILIA ‘eyebrows’</td>
</tr>
<tr>
<td></td>
<td>GENUCULUM ‘knee’</td>
<td>GENUCULA ‘knees’</td>
</tr>
<tr>
<td></td>
<td>ŒVUM ‘egg’</td>
<td>ŒVA ‘eggs’</td>
</tr>
<tr>
<td></td>
<td>BRACCHIUM ‘arm’</td>
<td>BRACCHIA ‘arms’</td>
</tr>
<tr>
<td></td>
<td>LIGNUM ‘wood’</td>
<td>LIGNA ‘pieces of wood’</td>
</tr>
<tr>
<td></td>
<td>TEMPUS ‘time’</td>
<td>TEMPORA ‘times’</td>
</tr>
<tr>
<td></td>
<td>MARE ‘sea’</td>
<td>MARIA ‘seas’</td>
</tr>
</tbody>
</table>

|        | VUM 'egg' | VA 'eggs' |
|        | GENUCULUM 'knee' | GENUCULA 'knees' |
|        | ŒVUM 'egg' | ŒVA 'eggs' |
|        | BRACCHIUM 'arm' | BRACCHIA 'arms' |
|        | LIGNUM 'wood' | LIGNA 'pieces of wood' |
|        | TEMPUS 'time' | TEMPORA 'times' |
|        | MARE 'sea' | MARIA 'seas' |

|        | LABIUM/LABRUM 'lip' | LABIA/LABRA 'lips' |
|        | (SUPER)CILIUM 'eyebrow' | (SUPER)CILIA 'eyebrows' |
|        | GENUCULUM 'knee' | GENUCULA 'knees' |
|        | ŒVUM 'egg' | ŒVA 'eggs' |
|        | BRACCHIUM 'arm' | BRACCHIA 'arms' |
|        | LIGNUM 'wood' | LIGNA 'pieces of wood' |
|        | TEMPUS 'time' | TEMPORA 'times' |
|        | MARE 'sea' | MARIA 'seas' |

Table 1

2.1 Ibero-Romance

In Ibero-Romance, the neuter plural forms were generally replaced by a regular plural, i.e. by suffixing the plural marker /-(e)s/ onto the singular form, as seen in the Castilian Spanish examples (reflexes of some of the Latin words in Table 1) in Table 2.

<table>
<thead>
<tr>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>labio 'lip'</td>
<td>labios 'lips'</td>
</tr>
<tr>
<td>huevo 'egg'</td>
<td>huevos 'eggs'</td>
</tr>
<tr>
<td>brazo 'arm'</td>
<td>brazos 'arms'</td>
</tr>
<tr>
<td>leño 'log'</td>
<td>leños 'logs'</td>
</tr>
<tr>
<td>tiempo 'time'</td>
<td>tiempos 'times'</td>
</tr>
<tr>
<td>mar 'sea'</td>
<td>mares 'seas'</td>
</tr>
</tbody>
</table>

Table 2

In some cases, the neuter plural form has survived into Spanish, after being reanalysed as a feminine singular noun. One such example is feminine singular ceja 'eyebrow' from Latin (SUPER)CILIA, the plural of the neuter noun (SUPER)CILIUM 'eyebrow, eyelash'.

The dual pathway (a) [neuter singular] > [masculine singular] and (b) [neuter plural] > [feminine singular] has given rise to the emergence of doublets such as leño 'log' vs. leña 'firewood' and huevo 'egg' vs. hueva 'spawn, roe', in which the feminine form, derived from the neuter plural, has taken on a lexically more specific collective meaning, referring to objects that, though possibly countable, are difficult to count and unlikely to be counted under normal circumstances. The fact that collective and mass nouns tend to be
morphologically singular is likely to have contributed to the reanalysis of plurals in /-a/ as feminine singulars.

In Portuguese, sobrancelha ‘eyebrow’ and lenha ‘firewood’ show the same development as in Spanish, whilst ovas ‘spawn’ is a plurale tantum in which the plural-marking morpheme /-s/ has been added to ova after its reanalysis as feminine singular, thereby re-establishing the plural notion that was lost in the process of reanalysis.

2.2 Italian

In Italian, a series of nouns that were neuter in Latin follow a somewhat unusual pattern in which the singular form behaves like a typical masculine singular noun, i.e. ending in /-o/ and triggering masculine agreement. In contrast to Spanish, the plural of this set of nouns is, however, not formed by extending the masculine morphology and forming an unmarked masculine plural in /-i/; instead, the original /-a/ is retained, triggering feminine plural agreement, presumably because /-a/ is cognitively closely linked to the feminine gender in Italian.2 Table 3 exemplifies this pattern.

<table>
<thead>
<tr>
<th>singular (masculine)</th>
<th>plural (feminine)</th>
</tr>
</thead>
<tbody>
<tr>
<td>il labbro bello</td>
<td>le labbra belle</td>
</tr>
<tr>
<td>il ciglio bello</td>
<td>le ciglia belle</td>
</tr>
<tr>
<td>il ginocchio bello</td>
<td>le ginocchia belle</td>
</tr>
<tr>
<td>l’uovo fresco</td>
<td>le uova fresche</td>
</tr>
<tr>
<td>il braccio lungo</td>
<td>le braccia lunghe</td>
</tr>
</tbody>
</table>

Table 3

For other nouns that were neuter in Latin, the masculine gender is extended to the plural form in a similar way as was shown, above, to have taken place in Spanish as seen in Table 4.

<table>
<thead>
<tr>
<th>singular (masculine)</th>
<th>plural (masculine)</th>
</tr>
</thead>
<tbody>
<tr>
<td>il tempo</td>
<td>i tempi</td>
</tr>
<tr>
<td>il mare</td>
<td>i mari</td>
</tr>
</tbody>
</table>

Table 4

Whether the Latin neuter plural ending /-a/ is preserved or replaced by an analogical masculine plural in /-i/ is lexically determined, and intimately related to the relative usage frequency of the singular and plural of the respective item. For items that occur more frequently in the plural than in the singular, such as paired body parts and eggs, the plural form in /-a/ survived,

2 The fact that /-a/, which otherwise marks the feminine singular, was reanalysed as feminine, but not as singular, suggests that gender may have (had) cognitive primacy over number in Italian; the simultaneous existence of a corresponding singular form is also likely to have influenced the reanalysis in favour of a non-singular reading.
while a predominance of the singular form in actual language usage led to the creation of an analogical masculine plural in /-i/.

The lack of a sufficiently large text corpus of early Italo-Romance makes it difficult to extract reliable usage data from the relevant period to confirm this frequency-based hypothesis. However, as the relative singular/plural distribution of these items is primarily determined by real-world situations, i.e. by the number of contexts or occasions on which it is relevant for speakers to refer to either one or more than one of the respective objects, the proportion of singular and plural forms is essentially extra-linguistic and should therefore be expected to be language-independent. This is confirmed by the modern-day frequency data presented in Table 5, drawn from three large corpora: the Corpus del español, the Corpus do português, and the British National Corpus.

<table>
<thead>
<tr>
<th>Spanish plural</th>
<th>Portuguese plural</th>
<th>English plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>labios</td>
<td>lábios(s)</td>
<td>lip(s)</td>
</tr>
<tr>
<td>cejas</td>
<td>sobrancelhas</td>
<td>eyebrow(s)</td>
</tr>
<tr>
<td>rodillas</td>
<td>joelhos</td>
<td>knee(s)</td>
</tr>
<tr>
<td>huevos</td>
<td>ovos</td>
<td>egg(s)</td>
</tr>
<tr>
<td>brazos</td>
<td>braços</td>
<td>arm(s)</td>
</tr>
<tr>
<td>tiempos</td>
<td>tempos</td>
<td>time(s)</td>
</tr>
<tr>
<td>mar(es)</td>
<td>mar(es)</td>
<td>sea(s)</td>
</tr>
</tbody>
</table>

Table 5

Table 5 shows some clear cross-linguistic trends regarding the relative frequency of singular and plural forms of relevant items. Crucially, those items forming their plural in /-a/ in Italian occur more frequently in the plural, whilst those that form their plural in /-i/ do not.

2.3 Romanian

In Romanian, a pattern similar to that of the small set of Italian nouns with a masculine singular and a feminine plural emerged, but rather than remaining limited to those Latin neuter nouns that occur predominantly in the plural, it was extended to a vast number of nouns, including loanwords from non-Romance source languages. These Romanian nouns are referred to as ‘neuter’, though according to their agreement patterns they are more accurately described as belonging to the masculine class in the singular and to the feminine class in the plural.

Morphologically, the plural of these ‘neuter’ nouns is either formed in the same way as that of feminine nouns (ending in either /-e/ or /-i/) or by suffixing the inflection /-uri/, which has its origin in the plural forms of certain Latin neuter nouns with stem allomorphy such as TEMP-US (nom./acc. sg.), TEMPOR-A (nom./acc. pl.); the final section of the longer stem was reanalysed as part of the plural inflection and subsequently extended to other nouns.
Table 6 shows masculine agreement (with the masculine indefinite article un) of neuter nouns in the singular, but feminine agreement (with două, the feminine form of 'two') in the plural. The Latin neuter plural marker -A does not survive into Romanian in the way it does in Italian. It appears that the feminine agreement of the plural forms of these nouns gave rise to an analogical extension of feminine plural morphology.\(^3\)

There is, however, a single noun with the plural desinence /-ǎ/: ouă, the plural of ou 'egg'. Whether this plural marker /-ǎ/ is a retention of the Latin neuter plural -A\(^4\), similar to the case of Italian uova 'eggs' (cf. 2.2.), or whether it is due to a regular sound change that turned /-e/ into /-ǎ/ after a labial glide (as in nouǎ < NOVEM) and which would have affected an earlier regular plural form oue, cannot be conclusively determined. Irrespective of the origin of this morphological isolate, the following section will show that it provides some valuable insights into the cognitive reality of nominal morphology in Romanian.

3. Remapping of meaning across morpheme boundaries

3.1 A brief outline of Romanian nominal morphology

Romanian distinguishes three cases, the nominative/accusative (N/A), the genitive/dative (G/D), and the vocative (which will not be discussed further here). Except in the feminine G/D singular, morphological case marking surfaces only in the preposed indefinite article or in the suffixed definite marker, as exemplified in Table 7.

---
\(^3\) An alternative explanation, reconstructing an earlier neuter plural in */-aec/ based on analogy with the neuter plural article illaec as the origin of /-e/, is proposed by Lausberg (1956:27).
\(^4\) Unstressed /a/ regularly changed to mid-central 'ǎ' in Romanian. (Rosetti 1978:370)
Except for some cumulative inflections in the feminine singular, the morphological structure of the noun is highly transparent, adhering to a structure that can be described as follows:

\[ \text{[stem]} + \text{[optional plural marker]} + \text{[optional definite+case marker]} \]

Whether the singular forms are morphologically unmarked or marked by a zero-morpheme is a theoretical issue that is not relevant here. There are four different plural markers, the choice of which is only partly predictable from the gender of the individual noun: /-i/ (m.,f.,n.), /-e/ (f.,n.), /-le/ (f.) and /-uri/ (n.).

As seen in Table 7, the word-final definite marker does not only encode information about case, but also about number and to some extent gender.

Definite plural forms are thus constructed by suffixing a definite marker to the indefinite plural form: /-i/, /-le/ in the nominative/accusative, /-lor/ in the genitive/dative.

### Table 7

<table>
<thead>
<tr>
<th>Case</th>
<th>masculine (lup 'wolf')</th>
<th>feminine (capr- 'goat')</th>
<th>'neuter' (lac- 'lake')</th>
</tr>
</thead>
<tbody>
<tr>
<td>N/A sg.</td>
<td>un lup</td>
<td>o capr-à</td>
<td>un lac</td>
</tr>
<tr>
<td>G/D sg.</td>
<td>unui lup</td>
<td>unui capr-e</td>
<td>unui lac</td>
</tr>
<tr>
<td>N/A pl.</td>
<td>(uni) lup-i</td>
<td>(une) capr-e</td>
<td>(une) lac-uri</td>
</tr>
<tr>
<td>G/D pl.</td>
<td>unor lup-i</td>
<td>unor capr-e</td>
<td>unor lac-uri-lor</td>
</tr>
</tbody>
</table>

3.2 The definite genitive/dative plural of 'ou'

In section 2.3 it was noted that ou 'egg' is exceptional because it is the only Romanian noun that has a plural form ending in /-ã/. According to the morphological rules briefly outlined above, the definite genitive/dative plural should be expected to be constructed by suffixing the definite genitive/dative morpheme /-lor/ to the indefinite plural ouã, yielding ouãlor 'to/of the eggs'. This is, indeed, the normatively correct form, but in actual language usage there is a frequently occurring alternative form, ouãlelor. In a small, informal survey⁶, nine out of 18 native speakers (i.e. 50%) from different areas across Romania used the form ouãelelor. There was no significant difference between male and female respondents. All respondents had a relatively high degree of formal education, most of them having completed a university degree; it is likely that among less educated respondents with a lower degree of exposure to the prescriptive standard, the proportion of respondents using the non-standard form is higher.

The /-le-/ in ouã-le-lor is not a direct continuation of anything found in Latin; it is most likely an analogical formation on the basis of a series of nouns

---

⁵ For neuter nouns, the plural in /-i/ is largely limited to loanwords with a singular in /-iu/.
⁶ Thanks to Alexandra Stavinschi and Benjamin Fagard for their help in collecting the data.
with a N/A sg. ending in /-e/, which regularly form their plural in /-le/, and therefore their definite G/D plural in /-le-lor/ (Table 8).

<table>
<thead>
<tr>
<th>Indefinite N/A sg.</th>
<th>Indefinite pl.</th>
<th>definite G/D pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>stea 'star'</td>
<td>ste-le</td>
<td>ste-le-lor</td>
</tr>
<tr>
<td>măsea 'molar'</td>
<td>măse-le</td>
<td>măse-le-lor</td>
</tr>
<tr>
<td>cafea 'coffee'</td>
<td>café-le</td>
<td>café-le-lor</td>
</tr>
<tr>
<td>perdea 'curtain'</td>
<td>perde-le</td>
<td>perde-le-lor</td>
</tr>
</tbody>
</table>

Table 8

In nouns inherited from Latin, this pattern is due to the preservation of an etymological /-l-/ in the plural that has been deleted in the singular (e.g. Lat. STÉLLA > Rom. stea 'star', Lat. STÉLLAE > Rom. stel-e 'stars'). This led to the reanalysis of /-le/ as a morphological unit, an allomorph of the feminine plural inflection: stel-e > ste-le. Following this reanalysis and the redrawing of the morpheme boundary, the resulting pattern has been extended to a number of nouns borrowed from Turkish, such as ‘curtain’ perdea (sg.), perdele (pl.) < Tk. perde (sg.), perdeler (pl.), in which the Turkish forms were assimilated to the phonologically most similar existing Romanian pattern.7

Whilst the sequence /-le-lor/ is the regular definite G/D plural ending for those items that have an indefinite plural in /-le/, there is no such clear motivation for the /-le-/ in the case of ou. In the following section, it will be shown that this form is a symptom of a process of cognitive restructuring of the definite plural morphology in Romanian.

3.3 Cognitive restructuring of Romanian definite plural morphology

In Section 3.1, the morphology of the definite plural forms of Romanian nouns was presented as sequential, with separate morphemes marking number and definiteness+case, respectively. The possible definite G/D plural sequences for feminine/neuter nouns are /-i-lor/ and /-e-lor/, with the variants /-uri-lor/ and /-le-lor/, as explained above, /-ã-lor/ occurs with a single lexical item, ou.

If the traditional analysis of these forms as consisting of a root plus two separate inflectional morphemes was an accurate representation of the cognitive structure of these forms, the form ou-ã-lor, following the normal

7 The choice of this pattern may be based on the similarity between the Turkish and Romanian singular forms, the plural forms, or both. Maiden (1999:329-330) believes it is the singular form, arguing that of all word-finally permitted vocalic sounds in Romanian, stressed /-e/ was the most similar to Turkish word-final stressed /-e/. This analysis rests on the assumption that the corresponding Turkish words have, or were perceived to have, word-final stress. However, the relevant nouns are perhaps more accurately described as having level stress or "gleichschwebende Betonung" (Steuerwald 1972:ix; 745 for perde), which makes the assimilation of /-e/ to stressed /-e/ as the sole trigger for their integration into this particular morphological pattern less convincing. The similarity of Romanian /-le/ and Turkish /-ler/, the latter with an /-r/ that is frequently realised weakly as a postalveolar fricative-approximant, makes it equally likely that the plural form played a role in this process.
pattern, would be the expected regular outcome, as there is no doubt or hesitation among native speakers concerning the simple (indefinite) plural form ou-ă. However, the fact that the sequence /-ălor/ is avoided by many informants is a strong indication that the definite G/D plural endings have the cognitive status of a single morphological unit for these speakers. Rather than constructing these endings from a plural marker and a definite case marker, these speakers have an inventory of four commonly occurring monomorphic allomorphs marking the definite G/D plural, /-ilor/, /-urilor/, /-elor/ and /-lelor/.

On the one hand, this reanalysis of the two morphemes as a single morphological unit is favoured by the fact that /-lor/ only ever occurs in this limited number of environments, and sets of several allomorphs are not unusual in Romanian. This is compounded by the fact that the inflections of the definite singular and all indefinite forms are monomorphemic, a fact that is likely to exert some analogical pressure on the bimorphemic definite plural. Furthermore, the fact that each of the two inflectional components in the definite plural forms implies plurality means that the plural is, in these forms, marked redundantly; their reanalysis as a single plural morpheme is thus also a process leading to increased cognitive economy.

All these factors facilitate a process which, following Croft’s terminology (2000:117-20), can be subdivided into two steps: (1) a non-overt form-function reanalysis, in which the frequent co-occurrence of the two morphemes leads to a mental remapping of their separate meanings as being expressed by the sequence as a whole, and (2) the actualisation of this reanalysis, the production of a previously unattested structure based on the reanalysis. In this particular case, it might be said that we are, in fact, dealing with an instance of ‘negative actualisation’, i.e. the avoidance of a previously attested structure, based on the reanalysis. The result of this negative actualisation is the avoidance of ouălor in favour of ouălelor, using the monomorphemic structure in which the meaning has been remapped onto the whole of /-lelor/.

The variation among speakers of Romanian suggests that this process is a currently ongoing one, one that has taken place in the grammar of those speakers considering ouălelor to be the correct form. To show that the reanalysis has not taken place in the grammar of speakers who use ouălor is perhaps impossible, as this particular form may, for them, have the status of a lexically specific exception, or they may have /-ălor/ as a fifth allomorph of the monomorphemic definite G/D plural inflection.

---

8 Lor also occurs as a separate lexical item, the dative/possessive plural pronoun. Though the semantic and phonological link between the pronoun and the inflection is obvious, their different morpho-syntactic status means that they have to be considered entirely distinct elements in modern Romanian.
Parallel to the definite G/D plural development, a less widespread definite N/A plural form *ouălele* instead of *ouăle* also occurs, as seen in the following example.

(1) Se adaugă **ouă-lele** bătute, unt-ul, drojdia dizolvată, sare-a REFL add-3SG egg-PL.DEF beaten butter-DEF yeast-DEF dissolved salt-DEF și zahăr-ul vanilat.
and sugar-DEF vanilla-flavoured
"The beaten eggs, the butter, the dissolved yeast, the salt and the vanilla sugar are added."

(Bucătăria Gabrielei, http://www.serbanfamily.net/bucatarie/)

Whether *ouălele* is formed by analogy with *ouălelor*, or whether a separate but comparable process of morphological restructuring is taking place in the definite N/A plural cannot be easily determined. However, *ouălele* is currently a marginal form, generally rejected by respondents who use *ouălelor* as the definite G/D plural.

3.4 Incorporation of adjoining material into the inflectional morpheme

According to Koch’s (1996:237-240) classification of changes in the place of morpheme boundaries, we can distinguish (a) loss of a morpheme boundary, in which two adjoining morphemes are reanalysed as a single one, (b) creation of a morpheme boundary, in which part of the phonological material of a morpheme is reanalysed as not belonging to the respective morpheme, and (c) shift of a morpheme boundary, in which a boundary is simultaneously lost in one place and created in another.

Incorporation of adjoining phonological material into the inflectional morpheme, leading either to the loss or shift of morpheme boundaries, has numerous precedents in Romanian. In Section 2.3. it was explained how the neuter plural inflection /-*uri*/ is the result of a morpheme boundary shift resulting from the reanalysis of Latin stem-final /-*or/- as part of the inflectional morpheme, as in Lat. TEMPOR-A > Rom. timp-uri. The new morpheme was then extended on a large scale to other Romanian ‘neuter’ nouns.

A similar process can also be observed in the masculine singular definite suffixes /-*ul/ and /-*ului/, in which the initial /-*u/- is historically derived from the vowel of the defunct Latin inflectional morpheme, whilst the rest of the suffix is derived from forms of the demonstrative pronoun ILLE. Thus, porcu-l ‘the pig’ was reanalysed as porc-ul; this reanalysis then gave rise to the generalisation of /-*ul/ as the definite marker for masculine and neuter singular nouns, as demonstrated by its use with nouns that do not have an etymological /-*u/-, such as lift-ul ‘the lift’. The meaning of /-*l/ (i.e. definiteness) having been remapped to the sequence /-*ul/, the /-*l/ of the definite suffix can, in fact, now
optionally be deleted in colloquial speech" (cf. Iscrulescu, 2006:146), yielding forms such as porc-u ‘the pig’ and McDrive-u ‘the McDrive restaurant’ which provide evidence that the incorporated /-u-/ has become the primary meaning-bearing element in the definite suffix.

4. Conclusion

The processes briefly outlined in Section 3.4 resemble the currently ongoing morphological restructuring of Romanian definite plural morphology discussed in this study to a great extent; in all three cases, meaning is remapped across morphological boundaries.

What sets the currently ongoing process apart from the other two is its actualisation. Whilst reanalysed /-uri/ and /-ul/ spread to large numbers of lexical items, such a visible actualisation is not possible for monomorphemic /-elor/ and its allomorphs because there is little scope for analogical extension that would lead to the emergence of previously impossible structures. Instead, the evidence for this reanalysis comes from what I have termed ‘negative actualisation’, i.e. the avoidance of a form that is no longer compatible with the new morphological structure. In the present case, such incompatibility arises only with a single lexical item, ou, due to its exceptional neuter plural ending.

The form ouāelor ‘of the eggs’ thus provides insights into a cognitive process affecting the structure of the Romanian nominal inflectional paradigm.

References

BNC Consortium. 2007. The British National Corpus, version 3 (BNC XML Edition). Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: http://www.natcorp.ox.ac.uk/


9 I am grateful to Martin Maiden for pointing out that in toponyms, e.g. Târgu Mureș, the deletion of the /-l/ of the definite article is an old phenomenon and certainly not restricted to the colloquial register. Rosetti (1978:466) provides some examples of final /-l/ having been apocopated from the definite article in texts from the 15th century. We may, thus, well be dealing with an old phenomenon that has, however, not (yet) become part of the less colloquial registers of modern Romanian.


THE REFUNCTIONALISATION OF FIRST-PERSON PLURAL INFECTION IN TIWI

JOHN CHARLES SMITH
University of Oxford (St Catherine’s College)

1. Introduction
Lass (1990) argued that a morphological form or opposition which had lost its original value might be reduced to vacuous ‘junk’, but could then emerge from this phase with a new linguistic function. He co-opted the term ‘exaptation’ from evolutionary biology to describe the process involved. In earlier work (Smith 1999; 2005; 2006; 2007; see also Donohue & Smith 1998), I have examined several instances of morphological refunctionalisation in the light of Lass’s hypothesis and suggested that, in all these cases, the oppositions in question are not at any stage ‘junk’. They may have been evacuated of concrete functional content (exponence), but a residual, more abstract, dichotomy remains (which may be discernible in frequency or markedness relations) — in short, an identity which, however diminished, is not yet ‘junk’, but is more akin to the notion of ‘skeuomorphy’, familiar in art history, whereby features which were originally functional lose their functionality, whilst retaining a content which is ‘decorative’ or part of a traditional or conventional design (Smith 2006:191-193).

If the opposition is refunctionalised, its refunctionalisation will be guided by this residual dichotomy — a process which I describe as ‘core-to-core mapping’. I have also drawn a parallel between this type of change and insect pupation (see especially Smith 2006:200): one coherent system (the larva) turns into another coherent system (the imago) via a transitional state (the chrysalis) which may appear incoherent, but which crucially mediates the relationship between the ‘input’ and the ‘output’. Although most studies of refunctionalisation are — almost of necessity, given the nature of the data — metachronic, dealing (in terms of the above metaphor) with direct comparisons between caterpillars and butterflies, one of the goals of a research programme in this area should ideally be to investigate the ‘chrysalis’ phase and the nature of the transition between one form–function relationship and another.

---

1 An earlier version of this paper was presented at the 38th Annual Meeting of the Societas Linguistica Europææa, held in Valencia (Spain) in September 2005. I am grateful to Paul Black, Jenny Lee, Elizabeth Traugott, and an anonymous referee for helpful comments and suggestions. Errors and omissions are, of course, my own.
In the present paper, I apply the approach outlined above to the evolution of first-person plural inflection in Tiwi, a non-Pama-Nyungan language spoken on Melville Island and Bathurst Island, off the coast of Australia’s Northern Territory. Tiwi exhibits a ‘minimal–augmented’ person system (Thomas 1955; Osborne 1974), and hence necessarily distinguishes between inclusive and exclusive forms of the first-person plural. However, since, and possibly as a result of, contact with English, Tiwi has been shifting to a ‘singular–plural’ person system (Lee 1987). Although it is possible for such systems to exhibit an inclusive vs. exclusive opposition in the first-person plural, the varieties of Tiwi which have shifted to the new system do not maintain this distinction. None the less, rather than disappearing or remaining in the language with a purely stylistic value or as stable sociolinguistic variation, the morphological opposition which originally encoded ‘clusivity’ (for this term, see Filimonova 2005) has been refunctionalised to encode a distinction of tense in the first-person plural, with the original inclusive form now acting as the exponent of the non-past, and the original exclusive form as that of the past. I shall suggest that the ‘core’ terms in these oppositions are, respectively, the inclusive person and the non-past tense, and the ‘non-core’ terms the exclusive person and the past tense. When the opposition is refunctionalised, it is therefore the ‘core’ form of the original opposition which takes on the ‘core’ function in the new opposition, with the result that this linguistic change, like other examples of morphological refunctionalisation, involves ‘core-to-core’ mapping.

2. The Tiwi person system

Viewed in traditional terms of person and number, Tiwi exhibits the following person system.

<table>
<thead>
<tr>
<th>SINGULAR</th>
<th>DUAL</th>
<th>PLURAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 inclusive</td>
<td>1 inclusive</td>
</tr>
<tr>
<td></td>
<td>1 exclusive</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1: A traditional description of the Tiwi person system

As Dixon (1980:351) observes, whilst this is a possible description of the system, it is an asymmetrical one. Osborne (1974:54) points out that such a system can be given a more satisfying and more economical representation in terms of the features [±Speaker] and [±Addressee]. These yield four possibilities for what is termed the ‘minimal’ system (effectively the singular and dual). If the terms of this ‘minimal’ system are taken to exclude other parties, the feature [±Other] will distinguish between this system and an ‘augmented’ system, in which other parties are included, as follows:
This type of analysis had earlier been proposed for the similar pronoun system of the Philippine language Ilocano by Thomas (1955). It accounts for Greenberg's observation (Greenberg 1988:1) that 'the first person inclusive is a favored category among duals'.

3. **The refunctionalisation of first-person plural inflection in Tiwi**

Lee (1987) examines the changes which have been taking place in Tiwi since, and possibly as a result of, contact with English. One of these changes is that 'the person–number system is being changed from a minimal–augmented system to a singular–plural one' (ibid.:176). In other words, many speakers of Tiwi now exhibit a person system which is comparable to that of English and many other Western European languages:

<table>
<thead>
<tr>
<th>SINGULAR</th>
<th>PLURAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

Like most non-Pama-Nyungan languages of Australia, Tiwi encodes the person of the verb by means of a subject prefix (for a detailed account of Tiwi verb morphology, see Capell 1967). In 'Traditional Tiwi' — defined by Lee (ibid.:12,18n14) as the language typically spoken by men aged 55 and over — which exhibits the system described in Figure 2 above, the first-person plural inclusive subject prefix is *nga-* and the first-person plural exclusive subject prefix is *ngi-*, in both past and non-past tenses. However, Lee notes (ibid.:176) that in 'Modern Tiwi' (the language of adult speakers under 30 (ibid.:10,12),
which exhibits the person system described in Figure 3 above), the prefix *nga*—‘is no longer used to mean “we (incl.)” as distinct from “we (excl.)”. In general, *nga*, or the changed form, *a*—[resulting from the loss of the initial velar nasal, a regular phonological development in “Modern Tiwi” (*ibid.*:39-40)], is used for “we” for non-past. The prefix *ngi*—[or the changed form *yi*—[again a regular development (*ibid.*:39-40)]) marks first-person plural in the past tense, regardless, once again, of whether or not the addressee is included. In other words, the original inclusive/exclusive opposition has been refunctionalised as an opposition of tense. Schematically:

<table>
<thead>
<tr>
<th>‘Traditional Tiwi’</th>
<th>‘Modern Tiwi’</th>
</tr>
</thead>
<tbody>
<tr>
<td>first-person pl. incl.</td>
<td>first-person pl. non-past</td>
</tr>
<tr>
<td><em>nga</em></td>
<td><em>nga</em>, <em>a</em></td>
</tr>
<tr>
<td>first-person pl. excl.</td>
<td>first-person pl. past</td>
</tr>
<tr>
<td><em>ngi</em></td>
<td><em>ngi</em>, <em>yi</em></td>
</tr>
</tbody>
</table>

4. Discussion — ‘Coreness’ vs. ‘non-coreness’

Is this change random, or is it motivated in some sense — in other words, does it exhibit ‘core-to-core mapping’? In Smith (1999; 2006), I suggested that the notion of ‘core’ value is associated with one or more of at least the following: qualitative unmarkedness; quantitative unmarkedness (higher frequency); default status. Often, these criteria will yield identical results; but not always. For this reason, I prefer not to use ‘unmarked’ as a cover term, leaving open the possibility that a particular criterion may dominate in particular circumstances.

Qualitative unmarkedness is defined by a number of well-known criteria, summarised by Battistella (1990:26) as ‘optimality, breadth of distribution, syncretisation, indeterminateness, simplicity, and prototypicality’. Optimality refers to the fact that ‘When certain segments or certain feature values imply others in language after language, those values are taken to be unmarked’ (*loc. cit.*). As far as distribution is concerned, ‘[u]nmarked terms are distinguished from their marked counterparts by having a greater freedom of occurrence and a greater ability to combine with other linguistic elements’ (*loc. cit.*) — the characteristic referred to by Croft (1990:77) as ‘versatility’. The unmarked term is also the one that occurs in positions of absolute neutralisation. Syncretisation means that ‘[u]nmarked categories tend to be more differentiated than marked ones’ (Battistella 1990:27). By the criterion of simplicity ‘unmarked elements are less elaborate in form than their [marked] counterparts’, and by that of prototypicality, they are ‘experientially more basic’ (*loc. cit.*).

Higher frequency is generally assumed to be a quantitative indicator of unmarkedness (see especially the discussion in Greenberg 1966:64). Bybee (1985:117ff.) further suggests that items which occur more frequently in texts or discourse have greater ‘lexical strength’ — that is, they are more firmly entrenched in the mental representation of the lexicon.
A default form is the one which occurs when there are no obvious criteria for selecting a particular item. I turn now to the applicability of these notions to the Tiwi data.

4.1 Quantitative ‘coreness’ of inclusive person

Donohue and Smith (1998) presented evidence from Indonesian and Classical Malay demonstrating the greater frequency of the first-person plural inclusive form in these languages (whilst recognizing that the ratio of the inclusive and exclusive forms will vary according to text type). Of course, these findings cannot simply be carried over without comment into a completely different language and culture, and we would need figures from Tiwi itself before we could draw any firm conclusions. However, as Tiwi is essentially an unwritten language, frequency data are hard to come by. The oral myths collected and transcribed by Osborne (1974:77-114) constitute far too small a corpus to be of significant statistical value, although some heuristic conclusions can be drawn from an analysis of these texts, in which the first-person plural inclusive subject prefix is used more frequently than its first-person plural exclusive counterpart, as shown below.

<table>
<thead>
<tr>
<th>Source: Osborne (1974:77-114)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corpus: 2907 words of oral Tiwi myths</td>
</tr>
<tr>
<td>Date corpus collected: 1967</td>
</tr>
<tr>
<td>First person plural inclusive</td>
</tr>
<tr>
<td>22 tokens</td>
</tr>
<tr>
<td>0.757% of total words</td>
</tr>
<tr>
<td>First person plural exclusive</td>
</tr>
<tr>
<td>17 tokens</td>
</tr>
<tr>
<td>0.585% of total words</td>
</tr>
<tr>
<td>Ratio of inclusive to exclusive = 1.29:1</td>
</tr>
</tbody>
</table>

Figure 4: Inclusive and exclusive first-person plural forms in Tiwi

The notion that the first-person plural inclusive is more frequent in Tiwi than is its exclusive counterpart is therefore plausible but unproven.

4.2 Qualitative ‘coreness’ of inclusive person

Qualitative arguments may be made for the ‘core’ status of the first-person plural inclusive relative to the corresponding exclusive form (see Donohue & Smith 1998). ‘Exclusive’ is arguably marked with respect to ‘inclusive’, in terms of both compositionality and discourse function. Compositionally, an exclusive first-person plural may be defined as ‘first person + third person’, whilst an inclusive first-person plural may be defined as ‘first person + second person (+ third person)’. Although Greenberg (1966:84-85) examines data from a number of languages which ‘lead one to posit, tentatively at least, a hierarchy in which the third person [is] the least marked, and the second person the most marked, with the first person intermediate’, an alternative, discourse-based, view might be put forward, in which discourse participants are less marked than non-
participants, and in which the speaker, as the necessary participant in every utterance, is less marked than the hearer, yielding a hierarchy 'first person > second person > third person'. Such a view underlies the work of Bühler (1934:79-148), and Benveniste (1956), and is explicitly articulated by Dixon (1994:84-90). Possibly the most detailed statement is by Silverstein (1976), who draws attention to the fact that some languages may invert the order of first and second persons in this hierarchy, apparently treating the second person as more animate or agentive than the first. None the less, in this framework, the inclusive always combines the two least marked persons, whilst the exclusive at best combines the least marked with the most marked, excluding the middle term, and at worst combines the two most marked — a reflection of the fact that, in discourse terms, a form which includes both (or all) discourse participants may be regarded as less marked than one which excludes one (or some) of them. The exclusive form therefore emerges as the marked term.

4.3 Quantitative and qualitative ‘coreness’ of non-past tense

As far as the tenses are concerned, Greenberg (1966:87) rehearses a number of reasons for regarding the past as quasi-universally marked with respect to the non-past; these include its lower frequency and its greater morphological complexity (see also Croft 1990; Battistella 1990). Although, once again, we have no statistically significant data for Tiwi, it seems reasonable to assume that this language is not exceptional with regard to the frequency relationship which normally seems to hold between the past and the non-past. Qualitatively, it is clear that the past tense is morphologically marked with respect to the non-past in ‘Traditional Tiwi’; Osborne (1974:41) states that “[n]on-past verbs have one of the non-past subject forms [...], but are not otherwise marked for tense. Past tense verbs are formed by (a) selecting one of the past tense set of subject forms, and (b) inserting the past tense marker r´- wherever the subject form carries no tense meaning”. See also Capell (1967).

4.4 Another aspect of ‘continuous’ mapping

I am indebted to Elizabeth Traugott (personal communication) for the observation that the mapping of ‘clusivity’ on to tense in the change under discussion may involve a further type of continuity. In an obvious sense, the interlocutor is clearly more proximal to the speaker and a third party more distal (a point already made in a different context above). The inclusive first-person plural therefore encodes a more proximal relationship than its exclusive counterpart. Likewise, the past is more remote from the moment of speech than the non-past (particularly given the default interpretation of the non-past as a present). In this sense, therefore, we are dealing with a ‘proximal to proximal’ and ‘distal to distal’ change, which can also be seen as ‘core-to-core mapping’.
5. Conclusion

It seems clear that, according to the relevant criteria, the inclusive first-person plural and the non-past tense are unmarked, more frequent (although statistically significant data for the forms of Tiwi itself are lacking), and more ‘proximal’, whilst the exclusive first-person plural and the past tense are marked, less frequent (subject to the same caveat), and more ‘distal’. We may tentatively conclude that the refunctionalisation of the first-person plural inflections of Tiwi, as outlined above, is in keeping with the principle set out (with respect to Romance data) in Smith (1999; 2005; 2006; 2007), in that it involves ‘core-to-core’ mapping. The change displays an element of continuity, and arguably at no stage involves ‘junk’.

References


1. Introduction

Changes in vowels are common, both historically as well as synchronically. However, little is known about their incidence in Chinese. This paper describes the first push-chain vowel-raising that occurred in the transition from Old Chinese to Middle Chinese, about 1600 years ago. It then explores two possible explanations for this type of common sound change, one speaker-driven and the other listener-oriented. Harold, one of my PhD thesis supervisors, taught me ‘Historical Linguistics’ at the ANU in 1990. I remember I was so excited when I got back my term paper with his comment ‘EXPERT’ (I still keep that apprentice article). I knew I was not an expert eighteen years ago, but that encouraged me to apply what I learnt in his class to the field of Chinese historical linguistics. Now I am pleased to present this little, but – I hope! – EXPERT, paper to Harold on his 65th birthday. Thank you, Harold.¹

2. Vocalic chain shifts

Vocalic chain shifts are common in many European languages. Among them, it is probably “the Great Vowel Shift” (Jespersen 1954: 232), which occurred in Middle-English long vowels, that is the most well-known:

<table>
<thead>
<tr>
<th>OC vowel</th>
<th>MC vowel</th>
</tr>
</thead>
<tbody>
<tr>
<td>a:</td>
<td>e:</td>
</tr>
<tr>
<td>e:</td>
<td>i:</td>
</tr>
<tr>
<td>i:</td>
<td>ai:</td>
</tr>
<tr>
<td>o:</td>
<td>o:</td>
</tr>
<tr>
<td>u:</td>
<td>au:</td>
</tr>
</tbody>
</table>

The incidence of chain shifts in the Chinese language, however, is not generally well known. This paper takes reconstructed vowels from two periods in the early history of Chinese – Old Chinese (OC) and Middle Chinese (MC) – and shows that the transition between them involves a chain raising of vowels. I also suggest possible explanations for the shift. The chain raising of long vowels has actually occurred several times in history of the Chinese language. This paper is concerned with the first: the Chinese Great Vowel Shift.

3. Rhyme groups and reconstruction of Old Chinese

Reconstructions of early states of Chinese are able to make use of often extensive philological data. In the case of the reconstructions of Old and Middle Chinese, recourse is had to the Rhymes of the Shi Jing – an anthology of poems

¹ This paper was supported by a grant from the Hong Kong Research Grant Council (HKUST 6426/06H). Thanks also to Dr. Phil Rose for discussing with me crucial concepts in the paper and editing the English version for me. I remain of course responsible for its contents!
from the period of Shang and Zhou Dynasties, said to have been edited by Confucius 2500 years ago. Morphemes (characters, really) that rhyme in the Shi Jing are grouped under different rhyme groups (RG) which are given names, like the RG zhī 貳. The indexing of vowels by their philological rhymes is actually very useful in charting the overall historical trajectory of vowels. In fact, in traditional Chinese historical phonology before the 20th century, there were no phonetic symbols available, and all phonological developments were described by reference to philological rhymes. The phonetic value of the vowel in the rhyme group, reconstructed after the introduction of the comparative method into China in the early 20th century, is derived from many sources: Sanskrit Buddhist classics, Old Chinese transliteration, phonetic-compound characters, internal reconstruction, and comparison of cognates between Chinese and Tibeto-Burmese languages, etc. There are several major reconstructions of the rhymes of Old and Middle Chinese. Our discussions in this paper will be based on the Old Chinese rhymes reconstructed by Zhengzhang Shangfang (2006), and the Middle Chinese rhymes reconstructed by Pan Wuyun (Pan 2000). Old Chinese, according to Zhengzhang, had six monophthongal vowels: *i (RG zhī 貰), *e (RG zhī 貰), *a (RG yú 魚), *o (RG hóu 侯), and *u (RG yōu 幽).

4. The First Great Chinese Vowel Shift

Several years ago I showed that historical changes in three of the OC vowels and a diphthong, *a (RG yú 魚), *o (RG hóu 侯), *u (RG yōu 幽), and *ai (RG gē 歌), actually constituted a chain vowel raising (Zhu 2002). This took place during the transition from Old Chinese to Middle Chinese (or between ca. 300 AD and the beginning of the 5th century). The shift can be represented by the schematic in (1).

\[(1) \quad *ai > *a \Rightarrow *a > *o \Rightarrow *o > *u \Rightarrow *u > *ou,\]

The rhyme groups of Old Chinese were divided among four grades (děng 等), with the first, second and fourth děng 等 having long vowels, and the third having short vowels. Grade three vowels were later lengthened and an [i] medial was epenthesised before the vowel chain shift began.

Within each of the rhyme groups there were usually more than one rhyme that could rhyme with each other. Not all of the rhymes in these rhyme groups underwent the first vowel shift; those undergoing the shift did not necessarily take the same path. To make things simple, I will only briefly mention which rhymes have gone through the vowel shift.

[*al > *ai] Within the gē 歌 rhyme group, only first děng 等 syllables without a medial, reconstructed as *-al (which later became *-ai) underwent the first chain shift. The gē 歌 RG was *al in early Old Chinese with the consonant coda
A CHAIN VOWEL RAISING IN THE EARLY HISTORY OF CHINESE

*-*l, however sometime during the Hàn 漢, Wèi 魏 and Jìn 晉 dynasties the coda vocalised to an *-i offglide, forming the diphthong *ai.

[*a > *o, *ā > *iā > *iə, *uā > *iua > *io] Within the yú 魚 group, three rhymes underwent the shift: *a (which later developed into the Middle Chinese mó 模 rhyme), the short vowel *ā (later the MC yú 魚 rhyme), and *uā with a rounded medial *-u- (later the MC yú 處 rhyme). Grade two rhymes of the OC yú 魚 group did not undergo the vowel shift.

[*o > *u, *ō > *io > *iu] The OC hóu 侯 rhyme group was comprised mainly of *o (later the MC hóu 侯 rhyme), and *ō (later the MC yú 處 rhyme).

[*u > *ou > *au, *ru > *ruau, *ū > *ieu, *rū > *eu] Lastly, the OC yòu 瘇 rhyme group included *u (later MC hào 䛟 rhyme), *ru (later the MC yao 雲 rhyme), *ū (later the OC yòu 瘇 rhyme), and *rū (later the MC yòu 瘇 rhyme). The vowels in all these rhymes underwent the first chain shift.

There are three parts to this shift. At one end of this chain of sound change was the monophthongisation of the diphthong *ai to *a. In the middle of the chain was a series of raising of the vowel *a to *u. Finally, at the other end was the diphthongisation of the vowel *u into *ou. There are two possible explanations for the initiation of the chain shift: either *a or *ai was the starter.

For the sake of convenience, let us assume that the vowel shift began with the low diphthong *ai (< *al). The rhyme ending *-i of the OC gē 歌 rhyme group was dropped around the beginning of the Northern Dynasties period, and this caused a conflict with the vowels in the yú 魚 rhyme group. There are two ways of resolving this kind of conflict: one is for a chain shift to occur in the yú 魚 rhyme group, and the other is a merger between the two rhyme groups, which is more common in the development of languages. In this case, however, the shift in the gē 歌 rhyme group from *ai to *a caused the *a in the yú 魚 rhyme group to raise to *o/ : *a > MC *o, *wa > MC *iwo, and *ā > MC *i. Later in the paper, however, I will consider another (more likely) possibility that the chain-shift began with the monophthong *a, which on the one hand evinced a series of vowel raisings, and on the other hand pulled out the diphthong *ai to fill the gap.

The chain shift occurred, without doubt, after the Western Jin Dynasties (ca. 300 AD) but before the famous rhyming dictionary Qièyùn 切韵 (601 AD). That it occurred after the Western Jin is pretty certain, given that the yú 魚 rhyme group was pronounced as *a before the Wei and Jin dynasties (Wang, 1923). That it occurred before Qièyùn is obvious because in Qièyùn phonology, the gē 歌, yú 魚, hóu 侯, yòu 瘇 rhyme groups were already *a, *o, *u, *ou, respectively (Karlgren 1915/1995; Pan 2000; Li 1956; Zhu 2002/05).

5. Motivations for the shift

According to the first of Labov’s (1994: 116) three principles, long vowels tend to undergo chain-raising. However, Labov does not explain why only raising of long vowels occurs and not lowering. Furthermore, why does chain-raising only
occur in long vowels and never short vowels? Below I will discuss two possible motivations for the chain raising of long vowels. Before that, however, we need to define a concept: the initial and the most natural articulatory configuration/position of articulation, or the default configuration, which produces the most natural, most unmarked, default sounds such as: vowel [a], or the second level [2] of a tone in a four-level scale (Zhu 2005). Based on this concept, I propose a hypothesis that many of the sound changes, including the vowel chain shifting, are involved in an articulatory movement called ‘returning to the default position’.

According to the listener-initiated sound change hypothesis, the first chain shift was initiated by the gē 歌 rhyme group *ai. That the first Chinese Great Vowel Shift may have originated from the gē 歌 rhyme group *ai is easy to understand. However, if you understand it as a listener-initiated sound change, more explanation is necessary. There is not always a one to one correspondence between the physical signal of a sound and the listener’s perception. If a listener hears a non-corresponding sound, they may easily misclassify it. The *i at the end of the gē 歌 rhyme group *ai is not a fully realised *i, but represents an upward movement of the tongue. Gliding is part of the speaker’s phonological target: they know what sound they mean to pronounce. But the listener may hear it as the natural ‘return’ to the default position typical in pronouncing an [a], due to the natural inclination to pronounce an [a] as [aa] or [a†]. This is what’s known as “hypo-correction” (Ohala, 1981). This kind of sound change is initiated by the listener, who unintentionally treats (i.e. corrects) the [ai] pronounced by the speaker as an intended [a].

(2) Sequence of Sound Change Initiated By Listener Hypo-Correction

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Listener</th>
</tr>
</thead>
<tbody>
<tr>
<td>ai</td>
<td>intentional</td>
</tr>
<tr>
<td></td>
<td>{1} speaker says: [ai]</td>
</tr>
<tr>
<td>a</td>
<td>unintentional</td>
</tr>
<tr>
<td></td>
<td>{2} listener hears: [a...a†]</td>
</tr>
<tr>
<td>a</td>
<td>{3} listener interprets: [a]</td>
</tr>
<tr>
<td></td>
<td>{4} listener repeats: [a]</td>
</tr>
</tbody>
</table>

The four steps in (2) outline the sequence of listener-initiated sound change. In terms of the speaker, they can pronounce /ai/ both as [ai] or as [a] (step 1). In terms of the listener, they may hear both sounds perfectly clearly, or may hear an [a...a...a†] (step 2). Under normal circumstances, the listener may correctly interpret it as [ai] or [a]. However, they may also quite possibly misinterpret as [a] what the speaker intended to pronounce as [ai] (step 3). When the listener repeats the sound, they will repeat the misinterpreted [a] (step 4). It is very easy for a non-native speaker of the language not to hear the distinction: to treat non-distinctive natural features of the final gliding as the linguistic target. This ignites the sound change. The fourth step in this case reflects free variation among individual speakers, which can happen every day among different speakers. Typically, this kind of variation will get corrected by
other speakers, or the social norm. This is not the case with non-native contact speakers of a language, who are unaware of ‘correct’ forms and may therefore easily pronounce the hypo-corrected forms initiated by the listener.

In terms of the vowel-chain-shift in the history of the Chinese language, it is quite possible that it did not start at the end of the chain, but on a link in the middle; the chain shift started with the raising of *a, which brought about the push chain *a > *o > *u and pull chain *ai > *a. Still, chain shifts in long vowels in other languages are always raising, and whether or not there is interference from a foreign/outside diphthong, it is quite probable that the first chain raising began with the low vowel *a.

The reason that a long, low vowel [aː] is raised so easily has to do with the fact that it is difficult to keep one’s mouth open and/or tongue depressed over a long period of time. After /a/ is maintained over a certain period, the speaker will, due to mis-timing and efficiency of articulation, tend to return to the default position, and consequently glide to a schwa [ə]. By ‘mistiming’ I mean variation in the timing coordination of the vocal apparatus. Generally, a sound is targeted, and then the vocal configuration either returns to an initial, default position, or moves on to a subsequent target. However, in an [a] for example, if the sound has not stopped but the vocal configuration nevertheless returns to the default position, it will create an [aː] or an [aːt] with an offglide.

6. Summary
To sum up, vowel chain raising is common cross-linguistically. Therefore, there must be a universal theory to account for this kind of linguistic phenomenon. Raising chain shifts are generally push chain shifts. The reason may be that maintaining an open mouth and depressed tongue in low vowels is difficult over a long period of time. Thus, it is quite easy for a mistiming to occur in the vocal apparatus, the result being a tendency for the vowel to centralise. Thus, the low vowel [a] may often be pronounced with an off-glide [aːt], which may in the longer term lead to vowel raising. This can explain why chain raising 1) only occurs in long vowels and not short vowels, and 2) is typically a push-chain shift rather than a pull chain shift. The concept of ‘returning to the default configuration’ is not only applicable to Chinese, but can be used to explain raising chain shifts in a lot of other languages. It can not only be used to explain the first Great Vowel Shift in Chinese, but can also be used to explain vowel shifts that happened later on. In fact, I (Zhu 2002) found and briefly reported another two vowel chain shifts occurred between the rhyme book Qieyun (601 AD) and the end of the Tang Dynasty (907 AD), though it is still difficult to pinpoint the exact time when they occurred due to conflict nature of relevant philological data. That will be another story in my next paper.
References


Li Rong 李荣. 1956. [The phonological system of the Qiéyún]. Beijing: Kexue Chubanshe.


INDEX OF LANGUAGES

A
Adnyamathanha. 21, 29, 109, 121, 139, 151, 153, 212
Aghu-Tharrnggala.........................212
Aiwong.................................148
Alawa..............................2, 59-69, 214
Alyawarr...............9, 71, 77, 82, 167, 173-176, 317
Anatolian..........................157, 158, 206, 209
Anadajin .....................................................187
Anmatyerr .......................................... 167, 173, 228
Arabana ............... 17, 20, 21, 24, 29, 139, 151, 152
Arabana-Wangkangurru ........... 21, 29, 139
Aranda ..................... See Arrernte
Arcadian ............................................. 156, 163, 165
Arrernte ...7, 8, 109, 121, 167, 174, 176-183, 226, 228, 231, 234, 321
Artuya (subgroup) .....................167
Atampaya.................101, 102, 212, See also Uradhi
Attic ...................... 93, 94
B
Baagandji.................................................231
Baham ......................................................4
Balto-Slavic................................. 156, 157, 159
Bardi ........................................ 7, 214
Basque ..............................................1, 6, 269-278
Bauzi ............................................. 44, 53, 54, 56
Bidyara..........................16, 19, 28, 212, 323
Biri ..................................................322
Boonwurrung ...................... 140, 146
Boraipar ........................................ 148, 149
Bundjalung dialects .............315
Western ........................................ 69, 147, 153
Bunuba ........................................... 214
Buwandik ........................................ 320
C
Cantonese ............................. 249
Chinese 5, 6, 235, 236, 245-252, 348-353, See also Mandarin, Cantonese, Hakka, Min
Middle ......................... 236-238, 348, 349
Chuqu ........................................... 236, 238
D
Diyari ...17, 22, 139, 151, 171, 172, 176, 181, 227, 230
Djabugay ................................ 22, 29, 135, 137, 212
Djadjawurrung ...................... 141
Djangadi ........................................ 214
Djapu ............................................. 126, 135, 137
Doric ............................................ 161
Dyirbal ..................... 18, 106, 133, 181, 182, 319, 321
Eastern Kulin (subgroup) .... 139, 140, 146
English .... 29, 38, 39, 40, 58, 69, 86, 140, 143, 148, 153, 164, 168, 180-182, 204, 205, 234, 297, 331, 341, 342, 348, 352
Euskara Batua ...................... See Basque
F
Forest River (language)........... 190, 192
French ........................................... 298-309, 338
Old .................................... 298-301, 305-308, 310
G
Gadhang ............................................. 231
Gajirrabeng .................................. 214
Gambera ........................................ 187, 192
Gamilaraay ......................... 4, 107, 108, 115, 120
Gan ............................................... 236
Germanic ..................................... 156, 163
Gidabal ...................................... 321
Gija ............................................... 214
Giraymalli ...................................... 101
Gooniyandi .................................. 199, 200
Gothic ........................................... 156, 216
Greek ............................................ 3, 4, 91-97, 156-165, 205, 206
Homeric ................................ 92, 93
Gugadji ......................... 321
Gugu-Badhun ...................... 323
Gugu-Yimidhirr .......... 15, 123-136, 315, 323
Gumbaynggirr ................. 125, 136
Gun-djeimi .................................. 214
Gunibidji ...................................... 102
Gunin/Kwini ...................... 187, 188, 192-195, 214
Gupapuyngu ..................... 147, 153
Guindji .. 71, 81, 82, 221, 224, 232, 315-316, 318, 326
Gurr-goni ........................................ 214
Guwij ........................................... 187, 192
<table>
<thead>
<tr>
<th>Language</th>
<th>Page Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>HAKKA</td>
<td>236</td>
</tr>
<tr>
<td>HATAM</td>
<td>44, 48, 49, 56, 57</td>
</tr>
<tr>
<td>HITITTE</td>
<td>3, 6, 157, 166, 201-209</td>
</tr>
<tr>
<td>Hoava</td>
<td>287-291, 295, 296</td>
</tr>
<tr>
<td>IBERO-ROMANCE</td>
<td>329</td>
</tr>
<tr>
<td>ILHA</td>
<td>44, 55, 56</td>
</tr>
<tr>
<td>INDO-EUROPEAN</td>
<td>1, 3, 6, 7, 155, 166, 201, 208, 212, 216, 285, 297</td>
</tr>
<tr>
<td>INDO-IRANIAN</td>
<td>156, 208</td>
</tr>
<tr>
<td>INNER-SLAVIC</td>
<td>160</td>
</tr>
<tr>
<td>IRISH</td>
<td>205, 285</td>
</tr>
<tr>
<td>ITALIAN</td>
<td>330-332</td>
</tr>
<tr>
<td>ITALO-ROMANCE</td>
<td>331</td>
</tr>
<tr>
<td>IWAITJA</td>
<td>214</td>
</tr>
<tr>
<td>JAMINJUNG</td>
<td>214</td>
</tr>
<tr>
<td>JARU</td>
<td>221, 224-226, 232</td>
</tr>
<tr>
<td>JAWOYN</td>
<td>214</td>
</tr>
<tr>
<td>JERSEY FRENCH</td>
<td>298</td>
</tr>
<tr>
<td>JINGILU</td>
<td>71</td>
</tr>
<tr>
<td>JIWARLI</td>
<td>34, 35, 40, 318, 325</td>
</tr>
<tr>
<td>KAKANYTU</td>
<td>319</td>
</tr>
<tr>
<td>KALA LAGAW YA</td>
<td>See Western Torres</td>
</tr>
<tr>
<td>Kalaw Kawaw Ya</td>
<td>See Western Torres</td>
</tr>
<tr>
<td>KANYARATI</td>
<td>31, 33, 34, 36, 37, 40</td>
</tr>
<tr>
<td>KARAJARRI</td>
<td>232</td>
</tr>
<tr>
<td>KARLAMAYI</td>
<td>212</td>
</tr>
<tr>
<td>KARNIC</td>
<td>139, 152, 313</td>
</tr>
<tr>
<td>KARTU</td>
<td>319</td>
</tr>
<tr>
<td>KATTING</td>
<td>See Gadhang</td>
</tr>
<tr>
<td>KAUNA</td>
<td>139, 147, 149, 151, 152, 227, 231</td>
</tr>
<tr>
<td>KOJ-KAPER</td>
<td>See Koko-Bera</td>
</tr>
<tr>
<td>KOKO MINJENA</td>
<td>316</td>
</tr>
<tr>
<td>KOKO-BERA</td>
<td>19, 103, 104</td>
</tr>
<tr>
<td>KOK-PAPRISK</td>
<td>103-105</td>
</tr>
<tr>
<td>KUGU-UWANH</td>
<td>22</td>
</tr>
<tr>
<td>KUKATA</td>
<td>150</td>
</tr>
<tr>
<td>KUKATJA</td>
<td>232</td>
</tr>
<tr>
<td>KUKU YALANJI</td>
<td>137, 319, 322</td>
</tr>
<tr>
<td>KULIN</td>
<td>140-148, 320</td>
</tr>
<tr>
<td>KUNBARLANG</td>
<td>214</td>
</tr>
<tr>
<td>KUNGKARI</td>
<td>322</td>
</tr>
<tr>
<td>KURRAMA</td>
<td>33, 34</td>
</tr>
<tr>
<td>KURTJAR</td>
<td>100-105</td>
</tr>
<tr>
<td>Kuuk-Thaayorl</td>
<td>323</td>
</tr>
<tr>
<td>Kuuk-Thaayorre</td>
<td>20, 21, 22, 29</td>
</tr>
<tr>
<td>KUYANI</td>
<td>139, 150-153</td>
</tr>
<tr>
<td>LARRIKIYA</td>
<td>214</td>
</tr>
<tr>
<td>LATNIATI</td>
<td>See Boraipar</td>
</tr>
<tr>
<td>LATVIA</td>
<td>160, 161</td>
</tr>
<tr>
<td>LETYLETYI</td>
<td>143, 146, 148</td>
</tr>
<tr>
<td>LIMILGAN</td>
<td>214</td>
</tr>
<tr>
<td>LINGUTHIGH</td>
<td>212</td>
</tr>
<tr>
<td>LITHUANIAN</td>
<td>156, 160-162</td>
</tr>
<tr>
<td>LONGGU</td>
<td>287-289, 296</td>
</tr>
<tr>
<td>LURITJA</td>
<td>176, 226-229, 231, 234</td>
</tr>
<tr>
<td>MAI</td>
<td>44, 48-49, 52, 56</td>
</tr>
<tr>
<td>MALAYANGAPA</td>
<td>152</td>
</tr>
<tr>
<td>MANAM</td>
<td>281, 296</td>
</tr>
<tr>
<td>MANDARIN</td>
<td>236</td>
</tr>
<tr>
<td>MANGALA</td>
<td>232</td>
</tr>
<tr>
<td>MANGARRAYI</td>
<td>59-65, 67, 69, 214</td>
</tr>
<tr>
<td>MANSIM</td>
<td>44, 48, 49</td>
</tr>
<tr>
<td>MANTHARTA</td>
<td>31-37, 40</td>
</tr>
<tr>
<td>MARA</td>
<td>2, 59-69, 214, 316</td>
</tr>
<tr>
<td>MARAN</td>
<td>316</td>
</tr>
<tr>
<td>MARAWARA</td>
<td>147, 148</td>
</tr>
<tr>
<td>MARGANY-GUYNA</td>
<td>322</td>
</tr>
<tr>
<td>MARRAMANINJSI</td>
<td>214</td>
</tr>
<tr>
<td>MARRGANY</td>
<td>23</td>
</tr>
<tr>
<td>MARRINGARR</td>
<td>214</td>
</tr>
<tr>
<td>MARRINGAYA</td>
<td>214</td>
</tr>
<tr>
<td>MARRINGAY</td>
<td>214</td>
</tr>
<tr>
<td>MARRINGAYU</td>
<td>319</td>
</tr>
<tr>
<td>MATHIMATHI</td>
<td>143, 145, 146</td>
</tr>
<tr>
<td>MAUNG</td>
<td>102, 214</td>
</tr>
<tr>
<td>MAYI</td>
<td>21, 28, 322</td>
</tr>
<tr>
<td>MAYI (SUBGROUP)</td>
<td>28, 322</td>
</tr>
<tr>
<td>MERYAM MIR</td>
<td>See Eastern Torres</td>
</tr>
<tr>
<td>MEYAH</td>
<td>44, 48, 49</td>
</tr>
<tr>
<td>MIDDLE CHINESE</td>
<td>236-238, 348, 349</td>
</tr>
<tr>
<td>MIDDLE HITITTE</td>
<td>202, 204</td>
</tr>
<tr>
<td>MIDDLE PAMAN</td>
<td>99, 106, 325</td>
</tr>
<tr>
<td>MIN</td>
<td>236, 237</td>
</tr>
<tr>
<td>MIRIWUNG</td>
<td>214</td>
</tr>
<tr>
<td>MIRMINY</td>
<td>149</td>
</tr>
<tr>
<td>MIWA</td>
<td>187</td>
</tr>
<tr>
<td>MOI</td>
<td>44, 48, 49</td>
</tr>
<tr>
<td>MPAKWITHI</td>
<td>103, 105</td>
</tr>
<tr>
<td>MPUR</td>
<td>44, 48, 49</td>
</tr>
<tr>
<td>MUDBURRA</td>
<td>71, 221, 224, 232, 316, 318</td>
</tr>
</tbody>
</table>
INDEX OF LANGUAGES

Munumburru ......................................................... 187, 192
Murrinh-patha .......................................................... 214
Muruwari ..................................................................... 119, 120

N

Na-kara .................................................................... 214
Narangga ................................................................. 139, 315
Nauo ........................................................................... 139, 149
Ndjebbana ................................................................. 102
New Hittite ................................................................ 202, 206
Ngadjuri .................................................................. 139, 152
Ngalakgan ................................................................. 131, 214
Ngan’gityemerri ........................................................ 215
Ngandi ....................................................................... 214
Nganyaywana .......................................................... 102, 104, 105, 214
Ngarinyin ................................................................. 187-199
Ngarla ....................................................................... 33
Ngarluma ................................................................. 32, 33, 229, 318
Ngarna subgroup ....................................................... 85
Ngarnawu .................................................................. 187
Ngawun ..................................................................... 21, 322
Ngayarda ................................................................. 32, 33, 37, 234, 319, 326
Ngayawang ............................................................... 148, 149
Ngiyamba ................................................................. 107, 120, 320, See also
Wangaaybuwan
Ngumpin ................................................................. 226, 231, 319
Ngumpin-Yapa 4, 71, 84, 221-227, 231-234, 315, 319, 324, 326
Nhuwala .................................................................... 33
Njeebana .................................................................. 214
non-Pama-Nyungan .................................................. 59, 62, 102, 194, 195, 199, 218, 232, 324, 341, 342
Norman French .......................................................... 300, 311
North-East Ambae .................................................... 281, 287-289, 296
Nukunu ...................................................................... 139, 151, 153
Nungali ...................................................................... 214
Nungubuyu ............................................................... 60-64, 214
Nyangumarta ........................................................... 85, 133, 134, 137, 212, 228-230, 326
Nyawaygi ................................................................. 101, 106
Nyikina ................................................................. 101, 178
Nyungar ............................................................... 228

O

Old Chinese .............................................................. 348, 349, 352
Old Church Slavonic ................................................. 156-162
Old French ............................................................. 298-301, 305-308, 310
Old Hittite ............................................................... 202
Old Irish ................................................................. 205, 285
Ojiiang ................................................................. 235
Oykangand ............................................................... 313

P

Pakantyi ................................................................. 147, 148, 153
Pagu ................................................................. 44, 47, 48, 51, 58
Paman
Middle ................................................................. 99, 106, 325
Paman (subgroup) ................................................... 15, 71, 99, 103, 319
Panyijima .............................................................. 37, 318, 319, 325
Parnkalla ............................................................. 29, 139, 147, 149, 151, 153, 228, 229, 231
Patjimalh ................................................................. 214
Payungu ................................................................. 34, 35, 37, 229
Pererapaperapa ....................................................... 141, 142, 144, 146, 148
Pikkolatpan ............................................................. 144, 146
Pingyang ............................................................... 244, 245
Pinikura .................................................................... 2, 31-40
Pintupi ................................................................. 226-234
Pitjantjatjara ............................................................ 176, 179, 183, 212
Pitjantjatjara-Yankunytjatjara ................................... 212
Pitta-Pitta ............................................................... 147, 152
Poraipar ................................................................. 146, 148
Portuguese .............................................................. 330, 331
pre-Arantic .............................................................. 77, 79
pre-Attic-Ionic .......................................................... 163
pre-Greek ............................................................... 162, 163
pre-Hittite ............................................................... 207
pre-Slavic ............................................................... 160
Proto-Arantic .......................................................... 77, 79, 80
Proto-Gunwinyguan .................................................. 129-131, 135, 136
Proto-Kulin ............................................................. 141, 146
Proto-Mirndi ............................................................ 77, 86
Proto-Ngarna .......................................................... 77, 82, 84
Proto-Oceanic ........................................................ 297
Proto-oujhang ......................................................... 235, 248, 249
Proto-Pama-Nyungan 15, 71, 77, 80, 84, 136, 174, 226, 323, 324
Proto-Warluwarric .................................................... 77
Proto-Western Kulin .................................................. 146-149
Proto-Worroran ........................................................ 192-199
Proto-Wu ................................................................. 235, 236, 237, 246
Proto-Yapa .............................................................. 77, 84
Proto-Yolngu ............................................................ 18
Punhamara ............................................................. 18
Purduna ................................................................. 32, 34, 35, 38

Q

Qingjiang ................................................................. 242, 243

R

Ramindjeri ............................................................... 17
INDEX OF LANGUAGES

Rembarrnga .......................................................... 131, 214
Ritharrnga ... ...................................................... 21
Romance ..........6, 278, 328, 331, 337, 346, 347
Romanian ........6, 328, 331-337
Roviana ........287-297

S
Salvador .... 44, 48, 58
Sanskrit ....... 155, 216, 349
Saweru ........44, 45, 46, 47, 53, 57
Shanghai .... 236, 243, 247, 249, 250, 352, 353
Sinitic .... 1, 236
Slavic .... 159-161
Spanish
Castilian ....... 329

T
Taihu .......... 236
Taizhou .......... 236
Tartitarti ...... 140, 143, 145, 146
Tarunggare ....... 44, 49, 57
Tehit ........... 44, 48, 49, 57
Thagungwurrung ......... 140, 146
Thalanyji .. 33-37, 40, 323
Thinn .............. 35
Thirrari ........ 151
Thura-Yura (subgroup) ....... 21, 29, 106, 139, 147, 149, 151-153
Tibeto-Burinese (family) .... 349
Tidore .......... 44, 47, 48, 57
Timor-Alor-Pantar ....... 43-45, 46, 54-56
Tiwi ............ 6, 170, 175, 182, 214, 341-346
Tjapwurrung .......... 141
Tocharian .......... 156-162, 166
Trans New Guinea (Phylum) .... 43, 45, 56
Turkish .......... 334
Tryantalla .......... See Werkaya

U
Umiiida ............. 187, 189, 192, 197
Umpila ............. 212, 229, 319
Ungarinyinyi ... 214, See Ngarinyinyi
Unggarrangu ....... 187, 189, 192, 197
Unggumi ........... 186-190, 192-198
Urudhii ........... 24, 100-105, 135, 136, 212, See also Atampanya
Urwa (subgroup) ......... 167
Uw-Olkola .......... 319

V
Vedic .......... 156, 158, 161

W
Wadigali ............. 152
Wagiman ........... 214
Wakaya ............ 214
Walmajarri ... 212, 221, 224, 232, 316, 318
Wambaya .......... 71
Wangaargguywan .... 107, 108, 113-120
Wangkajunga .......... 232
Wangkangurru ...... 17, 151, 152
Wangkumara ...... 18, 109, 119, 152
Wardaman ........ 214
Wargamay .......... 29, 101
Warlumbana .......... 71, 81, 82, 221, 224-232
Warlpiri .......... 9, 19, 71, 75, 128, 131, 133, 134, 177, 183, 212, 221-234, 316, 318, 323
Warlpirri (language) ... 71, 86
Warlpirri (subgroup) .... See Ngarna
Warnandargi ........ 2, 59-69, 214
Warray ............ 135
Warrgai .......... 214
Warriyangka ........ 35
Warumungu ....... 2, 9, 20, 71-86, 135, 137, 226, 232, 318
Wathawurrung .......... 140, 152
Wati (subgroup) .......... 148, 319
Watiwati ............ 143, 146, 148
Watjarri .......... 20, 125, 134, 136, 318
Wayi ............ 107, 114, 115, See also Ngiyambaa
Welsh .......... 205
Wembawemba .......... 139-144, 146, 153
Wencheng .......... 237, 245-246, 248
Wenzhou .......... 236-244, 250, 353
Werkaya .......... 140-143, 146
West Makuian ...... 44, 47, 48, 51, 53, 58
West Papuan (phyllium)... 2, 43, 44, 49-56, 58
Western Bundjalung ... 69, 147, 153
Western Desert (dialect chain) ... 19, 20, 85, 128, 149, 150, 232, 233, 318, 322
Western Kulin .......... 140, 141, 146
western Pilbara languages .......... 318
Western Torres (language) .... 2, 15-28
Wit-Mungknh ....... 212
Wilawila .......... 187, 192
Winyjarrumi ........ 189, 192
Wiradjuri ............ 107, 115
Wirangu .......... 21, 29, 139, 149, 151, 153
Woiwurrung .......... 140, 146, 152, 227
Woyami ........ 187, 192
Worrora .......... 186-200, 214
Worrora (family) ... 3, 185-200
Wu ............. 5, 235-238, 243, 247, 249, 250
Wunambal .......... 187-198, 200
Wunumara .......... 323
INDEX OF LANGUAGES

Wurla ............................................................. 187, 192
Wuzhou ............................................................. 236

X

Xiang ............................................................. 236
Xiangyang ............................................................. 236
Xuanzhou ............................................................. 236

Y

Yabulayabula ............................................................. 140, 144, 146
Yak Kumban ............................................................. 147, 149
Yakko-Yakko ............................................................. 140
Yakkumban ............................................................. 143, 145, 149
Yalarngga ............................................................. 143, 152
Yaluyandi ............................................................. 139, 152
Yandruwandha ............................................................. 18, 109, 120, 151, 152
Yanyuwa ............................................................. 99, 105
Yardilyawara ............................................................. 152
Yarli ............................................................. 152, 153
Yarri-yarri ............................................................. 145, 146
Yawâ ............................................................. 44, 45, 46, 53, 57
Yawijibaya ............................................................. 187, 189, 192, 197
Yawuru ............................................................. 214

Yidiny ............................................................. 22, 29, 131, 132, 135, 136, 212
Yijji ............................................................. 187, 192
Yiningyi ............................................................. 322
Yintjinga ............................................................. 319, See also Ayapathu
Yir-Yoront ............................................................. 19, 104
Yithayitha ............................................................. 140, 145, 146
Yolngu (subgroup) ............................................................. 17, 19, 24, 127, 137, 228, 315, 323
Yongia ............................................................. 238, 243-245, 248
Yortayorta ............................................................. See Yolngu
Yotayota ............................................................. 140, 144, 146
Yue ............................................................. 236
Yueqing ............................................................. 242
Yugambal ............................................................. 214
Yulparija ............................................................. 232
Yuulngu ............................................................. See Yolngu
Yuwaalaraay ............................................................. 4, 107, 108, 114-116, 119-121
Yuyu ............................................................. 140

Z

Zhenhai ............................................................. 236, 2
# INDEX OF SUBJECTS

| A | abduction ............................................ 304, 307, 310 |
|   | ablative ........................................... 18, 62, 63, 64, 68, 101, 207 |
|   | accusative ........................................ 17, 32, 33, 34, 44-47, 50, 51, 54, 63, 65, 95, 101, 103, 201, 206, 207, 272, 275, 322-323, 328, 332, 333, 347 |
|   | of extent ........................................... 206 |
|   | actualisation ....................................... 6, 335, 337 |
|   | address ............................................... 6, 335, 337 |
|   | term of ............................................. 199, 201, 270, 271, 312, 316, 319, 323 |
|   | affinal ................................................ 318 |
|   | allomorphy ........................................... 76, 80, 83, 134, 193, 195, 255, 256, 258, 260-262, 281, 282, 284, 332, 334, 336 |
|   | analogical change ................................... 5, 96, 130, 285, 298, 302, 307, 309-310 |
|   | analogy ............................................... 36, 47, 52, 68, 94, 96, 162, 164, 171, 217, 286, 298, 303-304, 307-308, 332, 336 |
|   | extension ............................................ 130, 332, 337 |
|   | apical ................................................... 74, 174 |
|   | dissimilation .................................... 78, 79, 83 |
|   | archaism .............................................. 145, 157, 171-173, 178, 181 |
|   | in song ............................................... 181 |
|   | augmented ........................................... 45, 56, 262, 341-342 |
| B | benefactive ........................................... 86, 108, 119, 319, 320 |
|   | Benveniste, Emile .................................. 201, 205, 208, 345, 346 |
|   | Bomberai ............................................. 44, 55, 56 |
|   | borrowing ........................................... 1, 2, 4, 24, 28, 34-36, 43, 55, 60, 82, 83, 148, 149, 152, 186, 194, 195, 211-213, 216, 227, 233, 262, 313, 320, 324 |
| C | Cape York Peninsula ............................ 20, 24, 29, 99-100, 103, 105, 212, 313, 315, 324-326 |
|   | Capell, Arthur ...................................... 15, 29, 31, 37, 43, 54, 56, 61, 188-192, 197, 200, 212, 218, 225, 226, 234, 342, 345-346 |
|   | China .................................................. 235, 250, 349, 353 |
|   | chain raising ........................................ 348, 350-352 |
|   | Central Australia ................................ 5, 7, 108, 120, 167-168, 170, 176, 180-183, 234 |
|   | construction ......................................... 269, 271 |
|   | marking ............................................... 6, 28, 33, 37, 63, 178, 252, 268, 270, 271, 275, 276, 278, 332 |
|   | cognitive restructuring ........................... 334 |
|   | Comparative Method ................................ 2, 3, 10, 11, 28, 85, 105-106, 136, 153, 182, 211, 214-218, 234, 325 |
|   | conjugation .......................................... 4, 19, 20, 22, 35, 36, 123, 125-127, 130, 132-136, 157, 161, 222, 230, 299, 301, 304 |
|   | marker ................................................ 123, 126, 127, 130, 135, 156 |
|   | consonant transfer .................................. 176 |
|   | continuum ............................................ 127, 129, 135 |
|   | coverbs .............................................. 60, 62, 68, 221, 223, 225, 233 |
|   | Curr, Edward ...................................... 143-145, 148, 150, 153 |
# INDEX OF SUBJECTS

**D**
- dative case ..................................178, 275, 276, 277
- definiteness ........................................333, 336
- dialect ........................................7, 32, 120, 137, 147, 148, 163, 225, 236-237, 243, 247, 249-250, 296, 300, 301, 305-309
- discourse .......................... 58, 183, 280, 285, 288, 293, 295, 343-345
- Dreaming .......................... 9, 168-170, 176, 179, 180, 182
- dual number ..................................23, 46, 49-51, 53, 55, 60, 63, 64, 66, 71, 73, 75, 80-85, 92, 95, 96, 99-104, 281, 314, 321-323, 329, 341, 346
- neutralisation ............................ 73, 80
- dyad ........................................312, 314, 318, 320-323, 325

**E**
- enclisis ...............................................318-320
- ergative case ..................................63, 273, 277
- suffixes ...........................................63
- Evans, Nicholas ................................2, 5, 10, 15, 16, 29, 129, 130, 136, 185, 195, 200, 213, 215, 218, 229, 288-290, 296, 312, 320, 321, 325, 326
- exaptation ........................................6, 340, 346
- extraparadigmatic ..........................206

**F**
- F0 ..................................................237-248
- feminine gender ..........................17, 45, 50, 60, 63, 65, 187, 192, 196, 323, 328-335
- fortis/lenis alternation ..............74, 75, 81
- four-place predicates ..................268-270, 275, 277
- France ...........................................302
- frequency .....................................93, 237, 249, 286-288, 291, 292, 295, 298, 302, 304, 308-311, 331, 340, 343-345

**G**
- Gascoyne ..................................31-33, 37, 38, 40
- genitive/dative case ..................332, 333

**H**
- h2e-conjugation ................................157
- Hale, Kenneth .............................3, 6, 7, 9-11, 22, 28, 29, 71, 73, 86, 99, 100, 106, 115, 120, 164, 166, 176, 182, 222, 234, 319, 321, 325
- Halmahera ..................................43-44, 47-58
- Harrison, Sheldon ...................3, 5, 11, 215-218
- Harvey, Mark ..........................4, 71, 129, 130, 136, 166, 196, 200, 211, 213, 218
- Heath, Jeffrey ..................20, 21, 29, 59-67, 69, 71-74, 86, 130, 136, 322, 325, 326
- Indoensia ..................................43, 56, 57, 58
- inflection...........................................34, 108, 301
- inheritance ..............................4, 159, 163, 200, 212, 213, 216, 227, 234, 320, 324
- initial dropping ..........................24, 103
- internal reconstruction ...........2, 46, 72-78, 80, 83, 211, 222, 349
- inversion marker ..........................72, 83, 86
- irregularity ................................10, 11, 57, 75, 216, 218, 305, 309-310, 325, 328, 338

**K**
- Kariera (kinship system) ..................322, 326
- kinship .................................6, 24, 82, 178, 182, 312-326
- Kurylowicz, Jerzy .........................155
- language contact ..........................16, 23, 29, 217, 234
- lenition ...................................22, 75, 80, 81, 100, 103, 147, 193, 257, 316, 318
- liaison ......................................299, 300, 305-307

**M**
- masculine gender ..........................17, 18, 45, 60, 63, 65, 187, 192, 328-333, 336
- Meillet, Antoine ..........................2, 11, 212, 215-218
- metathesis ..................................144, 163-166, 257
- metrical requirement .................5, 167, 170-172, 175, 178, 180
- monomorphic ................................115, 221, 232, 335, 336, 337
- morpheme boundary 
  - redrawing of ..........................74, 127, 199, 334, 336
- morphological similarity ..................7, 216
- multilingualism ..............................59, 62
### INDEX OF SUBJECTS

| N  | negation markers | 139-151 |
| N  | Neogrammarians | 3, 163, 166 |
| N  | New Guinea | See Papua New Guinea |
| N  | Omaha skewing | 323, 326 |
| P  | Papau New Guinea | 15, 281, 287 |
| P  | levelling | 74, 298 |
| P  | past tense | 20, 108, 203, 341-345 |
| P  | perfect | 15, 93, 155, 157, 159, 161, 203, 298-310 |
| P  | analytic | 203 |
| P  | person | first | 17, 50, 55, 71, 77-78, 95, 99-104, 281, 304, 316-320, 324, 342-345 |
| P  | second | 46, 50, 55, 60, 71, 75, 82-83, 86, 95, 99-104, 197, 281, 303, 306, 344, 345 |
| P  | philology | 2, 155 |
| P  | phonological change | 85, 133, 165, 207, 211-213, 252, 286, 309, 328 |
| P  | phonology | 2, 31, 36, 37, 57, 120, 128, 181, 234, 253-254, 263-264, 306, 310, 325, 337, 348, 350 |
| P  | Pokorny, Julius | 155, 156, 166 |
| P  | possessive case | 47, 65, 72-76, 78, 80-84, 103, 146, 281, 291, 312-316, 319-320, 335 |
| P  | prestopped nasal | 173 |
| P  | prestopping | 173, 313 |
| P  | preverbs | See coverbs |
| P  | pronoun | bound | 71, 72, 77, 85 |
| P  | free | 54, 71, 72, 75, 197, 320 |
| P  | possessive | 72, 73, 319 |
| P  | propositus | 317-318, See also kinship |
| P  | proprietive case | 34, 35, 321 |
| P  | prototypicality | 343 |
| P  | purposive inflection | 34, 35, 63, 101, 103, 116, 118, 152 |
| P  | push-chain | 348, 352 |
| R  | reanalysis | 3, 32, 34, 72, 73, 162, 197, 233, 284-285, 291, 295, 328-330, 334-337 |
| R  | reciprocal | kin | 320, 321 |
| R  | senior | 323, 324 |
| R  | reconstruction | 1, 2-6, 11, 16, 24, 28, 46, 51, 72, 73, 75, 99, 102, 105, 123, 128, 156, 167, 180, 185-186, 195, 199, 208, 211-218, 226, 227, 235, 237, 244, 248-254, 257, 260-264, 285, 312-341, 322, 323, 327 |
| R  | reduplication | 173, 175, 251, 256-257, 314, 324 |
| R  | reflexive | 20-22, 71, 75, 76, 78-84, 108, 119 |
| R  | refunctionalisation | 6, 340, 341, 346, 347 |
| R  | register | avoidance | 3, 179 |
| R  | regularity | 72, 85, 128, 298, 304, 309, 310 |
| R  | restructuring | 3, 6, 43, 254, 334, 336, 337 |
| R  | rhyme | 255, 256, 348-350, 352 |
| R  | rhyme groups | 348-350 |
| R  | root noun | 201, 206, 207 |
| R  | rounding | 173, 174 |
| S  | Saibai | 16 |
| S  | Schmidt, Wilhelm | 15, 29, 140, 161-163, 166, 251, 264 |
| S  | semantic | alignment | 44, 57 |
| S  | change | 3, 175, 186, 314, 324 |
| S  | extension | 180, 320 |
| S  | shift | 346, 351 |
| S  | shared aberrancy | 215, 216 |
| S  | shared song | 168, 169 |
364 INDEX OF SUBJECTS

song ......................................5, 9, 141, 144, 167-183
series.............................. 9, 168-170, 172, 178, 183
small ...................................................168, 169
text......................................................167-180
verbs in ..............................................176, 177
song language .........................5, 171, 178
sound change ..................... 1, 3, 15, 24, 93, 100, 103,
     128, 163, 166, 173, 186, 217, 225, 260, 261,
     290, 291, 299, 303-310, 313, 315, 324, 328,
     332, 348-352
sound correspondence .. 2-3, 15, 47, 54, 211-217
sound patterning .........................175, 180
stative ..........21, 45, 54, 109-116, 119, 161-162,
     260-262
stem17, 18, 21-23, 67, 68, 73-79, 81-84, 92, 93,
     108, 109, 112-115, 124, 127, 130-136, 146,
     155-157, 159, 161-164, 175, 177, 189, 205,
     207, 208, 216, 226, 229, 230, 252, 255, 258,
     261, 285, 290, 315, 317, 332, 333, 336
subgrouping.................................28, 53, 318
substrate ..........................................15, 28
superstrate ........................................15
syncretism .......................................298, 300, 305, 308, 328

T
targets..............................................239
tense..............19-23, 33-35, 45, 59, 71, 92, 93,
    107-109, 123, 132-135, 162, 172, 176, 177,
    203, 222, 281, 285, 304, 341-346
tone
    acoustics..............236, 238, 243, 246, 248, 249
    homogeneity ......................250
transitive .........19-22, 36, 45, 53, 71, 76, 101,
    108, 156-162, 165, 180, 187-188, 193, 197,
    262, 268, 270, 276, 280-283, 289-293, 295
U
unmarkedness ..................................343
usage frequency .........................330

V
variation.............. 2, 24, 27, 29, 74, 81, 86, 129,
    132, 133, 135, 140, 170-173, 185, 189, 235-
    237, 243-244, 246-249, 270, 276-278, 281,
    283, 300, 328, 335, 341, 351, 352
verb inventory .......................226, 232
versification ......................... 171, 172
voice ............. 20, 21, 91, 93, 156, 162, 171, 262
vowel
    assimilation ............74, 75, 80, 83
    chain shifts ....................352
    length ..................26
    shift ..................349, 350, 352
CURRENT ISSUES IN LINGUISTIC THEORY

E. F. K. Koerner, Editor
Zentrum für Allgemeine Sprachwissenschaft, Typologie und Universalienforschung, Berlin
efk.koerner@rz.hu-berlin.de

Current Issues in Linguistic Theory (CILT) is a theory-oriented series which welcomes contributions from scholars who have significant proposals to make towards the advancement of our understanding of language, its structure, functioning and development. CILT has been established in order to provide a forum for the presentation and discussion of linguistic opinions of scholars who do not necessarily accept the prevailing mode of thought in linguistic science. It offers an outlet for meaningful contributions to the current linguistic debate, and furnishes the diversity of opinion which a healthy discipline must have. A complete list of titles in this series can be found on the publishers' website, www.benjamins.com

Expected August 2008


290 MUGHAZY, Mustafa A. (ed.): Perspectives on Arabic Linguistics XX. Papers from the twentieth annual symposium on Arabic linguistics, Kalamazoo, Michigan, March 2006. 2007. xii, 247 pp.


287 CAMACHO, José, Nydia FLORES-FERRÁN, Liliana SÁNCHEZ, Viviane DÉPREZ and María José CARRERA (eds.): Romance Linguistics 2006. Selected papers from the 36th Linguistic Symposium on Romance Languages (LSRL), New Brunswick, March-April 2006. 2007. viii, 340 pp.

286 WEIJER, Jeroen van de and Erik Jan van der TORRE (eds.): Voicing in Dutch. (De)voicing – phonology, phonetics, and psycholinguistics. 2007. x, 186 pp.


282 PRIETO, Pilar, Joan MASCARÓ and Maria-Josep SOLÉ (eds.): Segmental and prosodic issues in Romance phonology. 2007. xvi, 262 pp.


HERRING, Susan C., Pieter van REENEN and Lene SCHØSLER (eds.): Textual Parameters in Older Languages. 2001. x, 448 pp.


AUTHIER, Jean-Marc, Barbara E. BULLOCK and Lisa A. REED (eds.): Formal Perspectives on Romance Linguistics. Selected papers from the 28th Linguistic Symposium on Romance Languages (LSRL XXVIII), University Park, 16–19 April 1999. 1999. xii, 272 pp.


SANKOFF, David: Diversity and Diachrony. 1986. xii, 430 pp.


BALDI, Philip (ed.): Papers from the XIIth Linguistic Symposium on Romance Languages, University Park, April 1–3, 1982. 1984. xii, 611 pp.


