Karnic as a genetic area

Claire Bowern

November 19, 2002

1 Introduction

The languages of the Lake Eyre Basin are an excellent set for case studies on diffusion and genetic linguistics. We know a lot not only about the languages themselves, but about the social and cultural life of the peoples that spoke them; thanks primarily to the work of Luise Hercus, Gavan Breen and Isobel McBryde, amongst others.[1]

I wish to argue here that despite recent claims to the contrary (for example Dixon 2001), the Karnic languages of the Lake Eyre Basin do form a discrete subgroup. I hope also to put aside the idea, prevalent in some quarters, that the use of the terms 'genetic relationship' or 'subgroup' necessarily implies the ability to model the languages on a neat family tree. Moreover, our lack of ability to model a given group of languages on a family tree does not necessarily imply a great time depth to the group.

Diffusionist models such as 'punctuated equilibrium' (Dixon 1997, 2001) are based on the assumption that linguistic genetic relationships should be able to be modelled on a family tree, and conversely that if one cannot model the relationships between given languages on a family tree, the relationship cannot be a genetic one. Indeed, punctuated equilibrium was invented in part to account for the lack of obvious binary-branching trees within the Pama-Nyungan family. In such a model, it is assumed that over time languages diverge through lack of contact and converge "on a common prototype" Dixon (1997:70) by borrowing back and forth from their neighbours. Dixon explicitly regards the Pama-Nyungan family as an ancient diffusion area with at least 40,000 years of relatively stable multilingualism, and any genetic boundaries have long been erased. "The time depth is so great that there is insufficient evidence to help us decide whether all the languages come from one genetic ancestor, or whether there were several genetic origins..." Dixon (2001:64)

While I do agree with Dixon regarding the importance considering language contact in reconstruction, especially in Australia, I do not agree with all aspects of Dixon's model. Some of the important points of disagreement include: languages and their speakers in Australian

[1]The ideas in this paper were first developed while I was writing my Honors thesis at the Australian National University in 1998, under Harold Koch and Luise Hercus. The ideas have continued to ferment since then, especially during my various periods as an visiting fellow at the ANU. I would like to Harold and Luise, and the rest of the ANU Centralian group for the many friendly and useful discussion we've had about these and other topics; thanks also to Jay Jasanoff, William Hawkins and Calvert Watkins. Thank you also to Pat McConvell for allowing this paper to be circulated with the other Arcling II papers, even though I am stuck on the other side of the world and unable to attend the conference in person.

have remained largely static since the mid-late Pleistocene, and there have been no wide-scale migrations (in Dixon's terms, 'major punctuations') in Australia since initial colonisation, up to 60,000 years BP; that in Australia basic vocabulary is borrowed as readily as non-basic; that languages in contact reach an equilibrium figure of approximately 50% of common vocabulary; and that languages with a close and well-defined genetic relationship can be modelled by a family tree. Dixon should, however, be thanked for pointing out some of the lazy assumptions that Australianists have been making regarding the historical linguistics of the continent, or that they have stopped making but have never got around to publishing retractions of. This paper owes as much to Dixon's work as to anyone else's.

This paper is already long enough and I have only scratched the surface of the issues involved; therefore my aim here is not to present a complete analysis of the Lake Eyre Basin or the ideas behind my views of the group in particular, and Pama-Nyungan in general. Rather, I propose to outline my objections to punctuated equilibrium in Australia, and to sketch an alternative view. I leave writing up the details and the implications of the model as another project.

2 The Lake Eyre languages

2.1 Geographical area

The languages which form the basis of this case study are those which were formally spoken in the Lake Eyre Basin of Eastern Central Australia, straddling the Queensland, Northern Territory, New South Wales and South Australian borders. The area where Karnic languages are spoken is broadly the area that comprises the Lake Eyre drainage basin; mostly rather flat semi-arid country, subject to occasional inundation. A map of the area, showing the languages under consideration, is given in Figure 1.

2.2 Classification

The classification of the Lake Eyre languages has not been stable. Researchers have vacillated between recognising a series of low-level groups with no closer higher relations between them, and grouping the languages into a larger family; the composition of this family, however, has also varied over time. This section briefly surveys the most widely-known classification schemata.

Schmidt (1919:43-44) defines a Karna group, which refers to Pitta-Pitta, Mithaka, Kunggari, and related dialects, as part of his *Süd-Zentral-gruppe*. He does say, however, that "[e]s ist sicher, daß die Kana-Sprachen nicht einheitlich sind, sondern aus verschiedenen Gruppen bestehen" [it is clear that the Kana languages are not uniform, but exist as different groups].[2] He also defines two *Untergruppen*; the Nulla-Untergruppe (Arabana-Wangkangurru) and the Dieri-Yarrawurka-Wonkamarra-Evelyn Creek-Untergruppe, whose name is self-explanatory. He places Badjiri in a Darling group.

[2]Schmidt's practice of naming groups after the word for 'man' has persisted in classifications to this day; Schmidt himself is responsible for the name *karna* for this area.

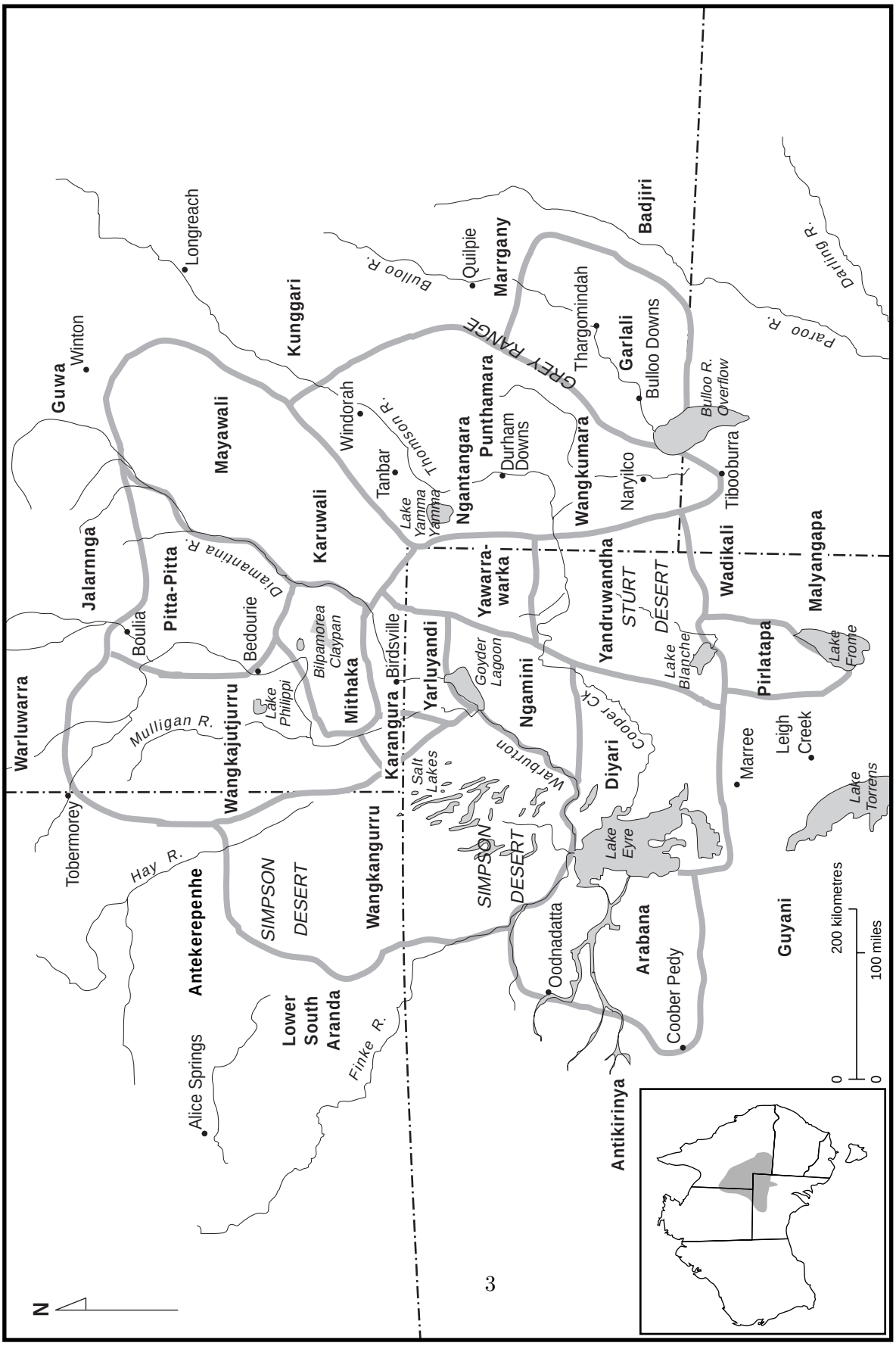

Figure 1: The languages of the Lake Eyre Basin; from Bowern (2001a).

O'Grady *et al.* (1966b) also recognise an entity referred to as 'Karnic', although their Karnic was considerably smaller than the family subgroup generally recognised now.[3] Other publications recognise several independent groups, implicationally no more closely related to each than to any other subgroup, such as Arandic or Tangkic. Breen (1971) recognised the wider relations of O'Grady, Voegelin and Voegelin's 'Karnic' group and related it to 'Mitakudic' and others in the Lake Eyre Basin.

Pretty much all the subgroupings have been based primarily on lexico-statistical data, whether as part of a wider preliminary survey of languages (such as O'Grady *et al.* 1966b and Wurm 1972) or whether involving a more detailed comparison, as in Breen 1971). Two of the later classifications, Austin 1990 and Bowern 2001a, both involve methods other than cognate inspection; both of these classifications also take morphological and lexical reconstruction into account, but because the reconstructions of the two authors are different they come to rather different conclusions regarding the subgrouping of the languages of the Lake Eyre Basin.

O'Grady *et al.* (1966b:42-44) give the following classification, which is very similar to that given by Wurm (1972):

(1) Language groups in the Lake Eyre Basin: (O'Grady *et al.* 1966b)

- a. Dieric Group
  - i. Karna Subgroup
    - Dieri – Tirari – Jandruwanta – Ngameni – Karangura – Yelyendi
    - Pilatapa
    - Jauaraworka (Jawaraworka)
    - Karendala – Kunkadutji (Gungadidji) – Kulumali – Bidia – Marulta – Karuwali
  - ii. Ngura Subgroup
    - Punthamara – Ngadangara
    - Kalali (Garlali) – Bidjara – Tereila – Wangkumara – Ngurawola
    - Badjiri (Badjari, Baddjeri)
  - iii. Yalyi Subgroup
    - Karenggapa – Nadikali – Malyangapa
- b. Pittapittic Group
  - Ulaolinya
  - Wangkadjera – Pittapitta – Kungkalenya – Karanya – Rakkaia – Ringuringu – Mayuli
- c. Mitakudic Group
  - Mitakudi
- d. Arabanic Group
  - Arabana – Wongkanguru – Wongkamala

[3]Their 'Karnic' comprised the languages immediately to the east and north of Lake Eyre - that is, Diyari, Ngamini, Yarluyandi, Yandruwandha and Yawarrawarka.

The classification given in Wurm (1972:131-3) is almost identical. The differences are that Karengapa is removed from the Yalyi subgroup (the name does not seem to appear anywhere in the classification), and the Ngura subgroup is defined as comprising two languages, Punthamara (with Ngadangara, Kalali (Garlali), Bidjara?, Tereila?, Wangkumara and Ngurawola?) and Badjiri (Badjari, Baddjeri).

Breen (1971) is a major survey of the languages of Western Queensland and in this book Breen modifies the groupings of O'Grady *et al.* (1966a) and Wurm (1972) in a number of ways. While Breen does not include a number of obviously Karnic languages, such a Pirlatapa and Arabana (since the scope of his study is a geographic region (Western Queensland), not a group of languages), the study is detailed and includes a careful study of the lexico-statistical material and brief consideration of linguistic factors other than shared vocabulary.

(2) Karnic Group (Breen 1971)
- a. Narla Subgroup
  - Arabana, Wangkangurru
- b. Karna Subgroup
  - i. Diyari, Ngamini, Yarluyandi,
  - ii. Mithaka
  - iii. Yawarrawarka, Yandruwandha
  - (and others outside the scope of the publication)
- c. Palku Subgroup
  - i. Lanima, Wangka-jutjuru
  - ii. Pitta-Pitta, Ringa-Ringa, Maiawali, Ngulubulu, Kunkalanja
- d. Ngura Subgroup
  - i. Punthamara, Thiraila, Mambangura, Wangkumara, Kungatutji, Karendala (Ngurawola)
  - ii. Galali, Wangkumara, Pitjara, Minkabari
  - iii. Badjiri

The greatest difference between Breen's classification and O'Grady *et al.* (1966b) is the recognition of a Karnic group, comprising the subgroups of the Lake Eyre Basin which were not previously regarded as particularly close. Mithaka is placed in the Karna subgroup. Wangkama(d)la, the third language of O'Grady, Voegelin and Voegelin's Arabanic group, is shown to be, in fact, the Wangkangurru name for Lhanima, and not a separate language. The Narla subgroup (the former Arabanic group) thus contains only two languages - Arabana and Wangkangurru. The details of the Ngura subgroup have also been altered. Kungatutji and Karendala are now part of this subgroup, and not the Karna subgroup as hypothesised by O'Grady, Voegelin and Voegelin. Badjiri remains distinct, but the other languages of the Ngura subgroup are split into two clusters. Note that Wangkumara appears twice, grouped both with Punthamara and with Garlali. The first is the "(lower) Wilson River language", described by Breen (nd.c); the second is Wangkumara as described by McDonald and Wurm (1979).

Austin (1990) represents a big departure in the classification of Lake Eyre languages. Unlike Breen (1971) and classifications following it, Austin excludes Arabana-Wangkangurru from Karnic, placing it as a primary subgroup on its own, although he does argue that the other languages of the area form a subgroup. His grouping is reproduced below in Figure 2. Note that Austin (1990) makes no mention of the Yardli languages or the subgroup to the East containing Wangkumara, Punthamara and Garlali.

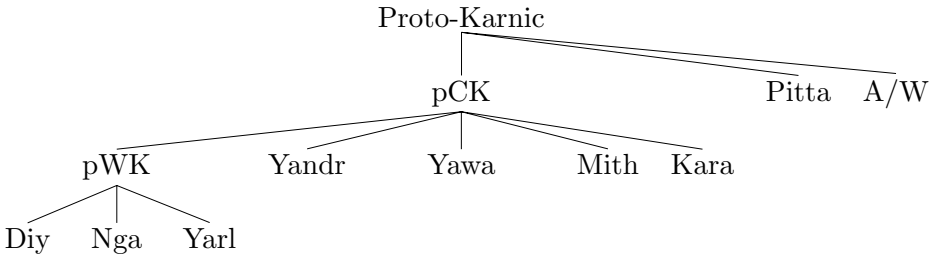

Figure 2: Classification of Lake Eyre languages: Austin (1990)

This is broadly in agreement with Breen (1971). The most significant point of difference between Austin and the other subgroupings is the exclusion of Arabana-Wangkangurru from the Karnic subgroup. Austin does this largely on the evidence of the pronominal forms in the relevant languages (Austin 1990:182ff).

In Bowern (2001a) I derive yet another classification, using primarily data from nominal morphology (but also making use of lexical innovation and sound changes). Like Austin's, my classification is based on the innovation method (for which see further Bowern and Koch 2002a), where subgroups are defined in reference to common changes that they undergone, and not simply by what they have in common. My classification differs from Austin's in several ways, however, mostly because I reconstruct different items to Proto-Karnic. I also consider the evidence for the inclusion or exclusion of the Yardli languages (for which see further Hercus and Austin 2002). Figure 3 below reproduces the family tree in Bowern (2001a).

In addition to the genetic groupings summarised in the tree in Figure 3, I also postulate two minor areas in the Lake Eyre Basin where there has been intense diffusion between groups of related languages; between Arabana-Wangkangurru and Pitta-Pitta and Wangkajutjuru (especially between Wangkangurru and Wangka-jutjurru) and between the Yandruwandha dialect ocmplex and Wangkumara.

This family tree belies the fact, however, that there is often rather little dividing the languages, and there is very little reconstructible at this stage in the way of shared innovations between higher level subgroups within Karnic; I am happy to admit that this family tree is much neater than it should be, and even the solution in Bowern (2001a), which shows several of the most important diffusion areas within Karnic, leaves something to be desired. Bowern (1998:Chapter 7) discusses the various family tress that one could draw for Karnic.

Dixon (1997:87-93) advocates a rather different view of the Karnic group, that languages with similarities to one another comprise linguistic areas rather than genetic groups, and the family tree model is inappropriate. Therefore the reconstruction of intervening proto-

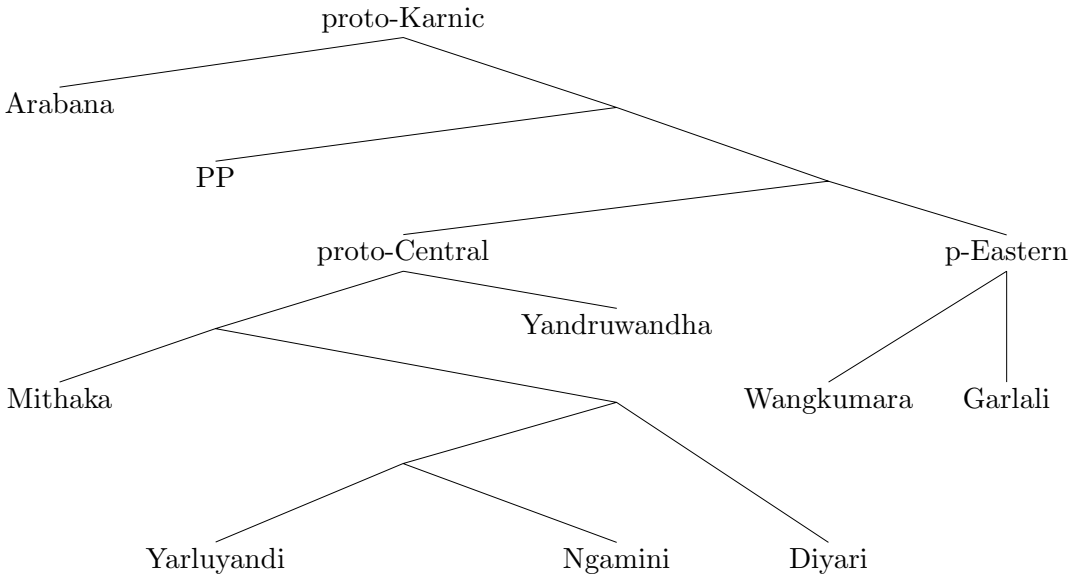

Figure 3: Family tree of Karnic after Bowern (2001a)

languages is either impossible or misguided.

2.3 Language data

Information on the language sources used are given in Table 1 of Bowern (2001a:248).

3 Overview of punctuated equilibrium

To my knowledge no detailed critique of punctuated equilibrium specifically applied to Australian languages has yet appeared in print. The reviews of Dixon 1997 were mostly favourable, mostly concerned with the second part of the book, which deals with endangered languages, and mostly did not address any of the specific claims of the theory (such as Joseph 1998). The only real exceptions to this are Kuteva (1999), which provides instances of mismatches between punctuation and family tree-like splits and equilibrium and linguistic diffusion, and Campbell (in press b), which provides a general critique.

In this section I will be concentrating on the major features of punctuated equilibrium as a model of language change. I will discuss the notions of 'punctuation' and 'equilibrium' and the problems with applying these notions to language change, and linking language change to non-linguistic events in a systematic way. I will address Dixon's criticisms of the comparative method and the family tree model, and the 'special situation' of Australian languages. If you like, I am taking up the challenge offered by Dixon (1997:fn 6, p74) to find counterexamples to the expected results of punctuated equilibrium.

Note that Dixon in no way claims that the family tree model is inapplicable to the whole of Australia, indeed at the end of Dixon 2001 there is a list of geographical, areal and genetic

groups of Australian languages. Note also that this is not specifically an attack on punctuated equilibrium or on Dixon himself; I feel, however, that there are several major aspects of punctuated equilibrium that should be addressed, and also that there are alternative explanations for many Australian phenomena that are not peculiar to the Australian continent.

3.1 General Summary

Dixon adapted the term 'punctuated equilibrium' from evolutionary biology (and specifically Eldredge and Gould 1972) to refer to his model in which language change can occur at vastly different rates over different periods of time.

In this model, language groups are assumed to have coexisted for most of human history without a great deal of disruption. During this time, languages continue to change of course, largely, Dixon (1997:71) implies, by borrowing from one another. In this way traits can diffuse across an area. Thus during equilibria local as well as larger linguistic areas are created, across boundaries of genetic families. The long periods of equilibrium are, in the punctuated equilibrium model, interspersed with much shorter periods of punctuation. Punctuation could be caused by several factors, either natural or manmade. The introduction of new technology, invasion, mass migration, fire or flood are some examples of punctuation. Dixon argues that it is only in times of punctuation that recognisable splits, of the kind well modelled by a family tree, occur (Dixon 1997:73). Note that in Dixon's model, linguistic areas can be 'punctuators', so the off-shoots of a family tree in a period of punctuation may not go back to a single proto-language: compare Figure 4, reproduced from Dixon (1997:101).

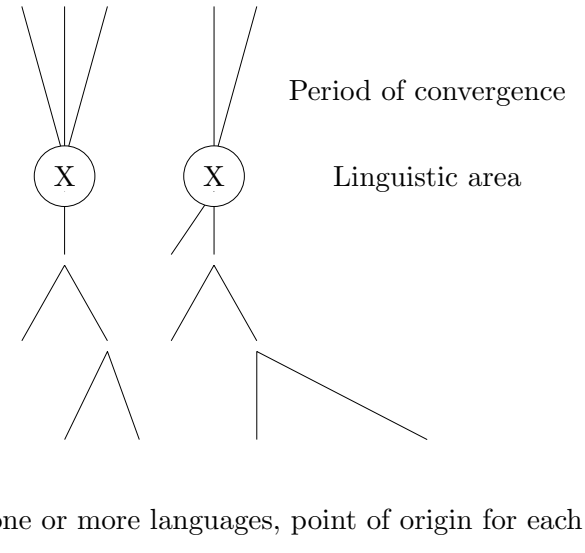

Figure 4: Schema of equilibrium and diffusion, from Dixon (figure 7.1 of 1997:101)

Dixon (1997:87-93) hypothesises a situation in Australia of extended equilibrium, with a number of minor punctuations due mostly to natural causes, such as fire and flood, and the

last, disastrous punctuation which began in January, 1788. The implication of this is that in this model, similarities between neighbouring languages are most likely to be the result of diffusion from one language to another, and not of common genetic inheritance.

The main aspects of Dixon's theory of language change are given in (3):

(3)
i. There is a direct correlation between patterns of linguistic relationship and non-linguistic, socio-historical, occurences;
ii. Language change is characterised by periods of stasis (or 'equilibrium') and much shorter periods of punctuation;
iii. Equilibrium has been the natural state for most of the world's languages for most of human history;
iv. The family tree model is applicable only in the time following a punctuation and language spread;
v. Punctuations are usually nonlinguistic phenomena (such as fires, floods or wars);
vi. Languages in an equilibrium situation will 'converge to a common prototype'; that is, they will tend to borrowing lexical items and grammatical structures until they reach approximately 50% in common;
vii. Any lexical or grammatical component of a language can be borrowed;
viii. There is no universal that basic vocabulary is borrowed less than non-basic (although this may be the case in Europe);
ix. Therefore,
    a. adjacent languages that have little in common are likely to have come into contact relatively recently;
    b. the family tree model is only applicable in some situations; those that result from punctuations (and by implication, in almost all punctuations);
    c. long-distance genetic comparison is not possible because diffusion has clouded all reliable genetic markers;
    d. similarities between adjacent languages are likely to be the result of shared diffusion, not shared genetic inheritance
x. Greater weight in the model is given to convergence than divergence (divergence typically happens when languages move apart; convergence happens when they move back together); more weight is given to linguistic change through external factors (i.e., diffusion) than to internally motivated changes;
xi. Genetic relationship is only one type of linguistic relationship.

3.2 Features of Punctuation

Dixon (1997:75) equates punctuation with primarily non-linguistic events, such as wars or natural disasters. Dixon also states that it is only in periods following punctuations that a family tree may accurately model the language split. However, there are punctuations, of broadly the same sort as Austronesian and Indo-European, which do not resemble a family tree. One of these is Turkic, which fits Dixon's time period of punctuation of 'a few hundred

to a few thousand years'. A review of the Turkic data is given in §7 below. For other examples of the non-correlation between socio-historical punctuation and family tree-like splits in languages, see Kuteva (1999) and Campbell (in press b) on agriculture.

Another problem of the notion of 'punctuation' is that it does not distinguish between the languages of the 'punctuators' and those of the 'punctuatees'. While all the examples of punctuations and family trees that Dixon gives are language spreads and therefore punctuators, Dixon does not provide a means of differentiating the two within his model. Punctuator-languages, for example, can show substratum influences (see, for example, the development of a set of retroflex consonants in Indic languages, or some of the features of Afrikaans, like double negation, which can be attributed to substrate influence from Khoi languages. The languages of the 'punctuatees' seem to be assumed to be swallowed up by those of the punctuators, but this does not always happen; not every conquest results in language shift.

I would suggest that there is an alternative explanation for the applicability of the family tree in Indo-European and Austronesian, rather than the rapid spread of these groups or their 'punctuation' *per se*. It is obvious that one of the major features of language change that will produce nice splits in family trees is the degree to which speakers of related languages remain in contact while and after they diverge. Language families that evolve with minimal contact between speakers after the split have fewer crossing isoglosses and fewer shared innovation that do not correspond to subgroups. This can be illustrated particularly well within Austronesian, where expansion was in many cases by island-hopping, with subsequent limited contact. Note crucially that in one of the first areas to be colonised by speakers of Oceanic languages, the Admiralty islands, Ross (1988) draws no clear subgroups, and refers instead to 'networks' of old dialect chains with a few major divisions It is probable in this case that the modern languages result from old dialect areas; that is, the languages have diversified *in situ*, and one cannot draw a family tree because it reflects a complicated dialect area (and note, in passing, the general inapplicability of family trees to dialects, as discussed in Hock 1991:450). An argument such as this is especially appealing when we compare the Admiralty situation with that of Fiji (as comprehensively described in Geraghty 1983). As shown below (§3.5), the relation of the type of language split and the (in)applicability of the family tree is much more a feature of the type of speech community and its linguistic relation to its neighbours than the result of non-linguistic factors such as war, conquest, flood or famine.

One of Dixon's arguments in favour of equilibrium in hunter-gatherer societies is the lack of perceived impetus for punctuation and spread. While Dixon himself does not explicitly link the spread of agriculture and punctuation, it is clearly a type of punctuation which he has in mind, and others, such as Bellwood (2001), discuss the spread of agriculture explicitly as a type of punctuation that produces a particular type of language split. Dixon links prestige and technological advantage to language spread and infers that because in hunter-gatherer communities such as one finds in Australia these features (major war machinery, agriculture, technological advances, etc) do not exist, one does not get these types of punctuation. Peter Bellwood does explicitly link punctuation and the initial migrations and split of families such as Niger-Congo and Indo-European with the spread of agriculture (see for example Bellwood 1997, 1998, 2001). The link between the migration of speakers of Proto-Indo-European and the spread of agriculture (due, as far as I am aware, primarily to Renfrew, for example

Renfrew 1987) has been comprehensively discredited, with agriculture spreading from the fertile crescent beginning towards the end of the 8th millennium BC, long before the spread of Indo-European, which must be dated to after the invention of secondary products such as cheese and weaving, some several thousand years later (cf Gimbutas 1963, Mallory 1989). The link of the spread of agriculture with Niger-Congo also has series problems, especially since it seems that Niger-Congo had already starting splitting before the introduction of agriculture. McConvell (2000) gives other examples of where the spread of agriculture (or a shift in the primary means of production for a community) did not involve language shift, such as the Tanzanian click language Sandawe. Campbell (in press b) shows that language spread is not tied directly to means of production, with 'spread' and 'non-spread' languages and language families being represented by both agriculturalists and hunter-gatherers.

3.3 Features of equilibrium

As Kuteva (1999) discusses in detail, Dixon (1997:75) equates punctuation and equilibrium with primarily non-linguistic events. She shows, however, using examples primarily from Africa, that socio-historical events of both punctuation and equilibrium can produce family tree-like splits, and conversely that periods of intense linguistic diffusion can be seen in both kinds of socio-historical situation.

A second point of consider, especially in relation to Indo-European, is the role of population movement in language split. Dixon makes no difference between 'punctuation' accomplished by the arrival of a new people to a previously occupied area, and language shift, accomplished by no large migration. The difference is illustrated nicely for the Caucasus in Nichols (1992:13-15), where the two occur in opposite directions. That is, languages tend to 'migrate' up mountains, where the language of a lower village will be known by a higher one, but not vice versa, whereas people themselves tend to migrate down mountains, in search of work, and so on. Dixon does not appear to distinguish between the spread of languages and the spread of people. Surely this would have an effect on his model. Presumably the spread of people over a large area, accompanied by a large migration, is a punctuation. But small-scale migration and the spread of language does not clearly fit into Dixon's model.

A third and very important point regarding the correlation between punctuation, equilibrium, and type of language change, is made in Watkins (2001:55 et passim) and his discussion of ancient Anatolia as a linguistic area. He points out that while Anatolia is a very good example of a linguistic area, with Indo-European and non-Indo-European languages in contact and borrowing from each from phonology, morphology and syntax, the languages in question are *at the same time* undergoing regular (ie, family-tree modelable) change and speciation. Watkins sees no reason for separating areal developments and genetic developments.

Finally, we may be able to counter some of these objections by assuming multiple types of equilibrium situations. Frankel (1988:41a), for example, quotes Butzer (1982:22) as defining five different notions of equilibrium and change in relation to the study of archaeology and prehistory in Australia. While his purpose is rather different, perhaps such notions could be adapted (but that is not a topic for this paper).

3.4 The current number of languages

> Approximate dates have been assigned for proto-languages ... no date earlier than around 10,000 BP is generally accepted. Yet archaeologists and human biologists believe that humankind developed language at least 100,000 years ago ... What happened between 100,000 years ago – or whenever language developed – and the proto-languages of modern families, 6,000 or 10,000 years ago?
>
> Dixon (1997:2)

This argument accounts for Dixon's assumption that for most of human history languages have been in a state of equilibrium rather than punctuation. He associates the creation of language families and rapid splits with non-linguistic punctuation, and much slower change and development with periods of equilibrium.

He does this partly to account for the current number of languages in the world (approximately 6,000); he argues (p 29) that a rate of splitting equivalent to that of Indo-European would have given about 10^{34} languages over the last 100,000 years, and even allowing for language death, this is still many orders of magnitude higher than we find. Therefore Indo-European splitting must be abnormal, and therefore that for most of history languages have not been splitting like this, but have rather existed in equilibrium.

Dixon appears to have vastly underestimated the factor of language death in reducing the number of currently spoken languages. He allows for the 'occasional death along the way'. Take Dixon's example of a family splitting into 2 every 6,000 years. If no languages had become extinct over that period, we would expect approximately 2^{17} (or about 132,000) languages to be spoken in the world today, assuming that Language has been spoken for approximately 100,000 years. If we assume, however, that the rate of language death is approximately 10% per generation, and use the same rate of splitting, we end up with a figure of approximately 22,000 spoken languages (assuming a constant rate of death across time/generations and, for the sake of simplicity in calculation, allowing for 'fractional' languages, a death rate of 10% gives us a figure of 1.8^{17}, or approx 21,860). Just as growth is exponential, so is language death, because in this model when a language 'dies' we remove from the calculation any offspring that the node would have had. Assuming a death rate of 20% with that rate of splitting brings us down to a modern figure of just under 3,000 (or $1.6^{17} \approx$ 2,950) languages.[4]

Moreover, language split in the last 10,000 years is not an accurate reflection of the previous 90,000 because of the role of agriculture in boosting population numbers and the social, economic and migrationary possibilities that this changes (see the papers in Harris 1996 for examples and discussion of the impact of population figures on language viability).

[4]There are several reasons why this experiment is not very meaningful. For example, assuming a rate of split of Indo-European as given in Dixon (1997:29) over the whole world would require a language death rate of 41% per generation to leave us with 6,000 modern languages. That in itself is perhaps not unreasonable, although I have no data on rates of language attrition for the relevant periods, and I do not know how such data would be obtainable. Furthermore, the model in this simplistic form is meaningless. We know, even from looking only at Indo-European, that language families split and diversify at different rates, and as far as I am aware we do not know what causes different rates of speciation; population numbers and density and migration are obviously two factors, but they are not the only factors.

3.5 The Family Tree Model

> Finally, it is appropriate to make an anti-assumption, that is, to question an assumption that is frequently made (if only implicitly). That is that all language development, and all types of proof of genetic relationship, must be like what happened in the Indo-European family. As Benveniste (1966:103) explained: 'il n'est pas certain que le modèle construit pour l'indo-européen soit le type constant de la classification génétique.' The family tree model, which so clearly explains most of the development of—and the relationships between—languages of the Indo-European family, is an important one, which has a measure of applicability in every language situation. Dixon (1997:14)

I agree with Dixon on his comments on the lack of possibility of reconstruction beyond a certain time depth, but not entirely for his reasons. Dixon (1997) claims that these stem from a wish to continue applying the comparative method to proto-languages *ad infinitum*. 'If the family tree model is all one has, there is a temptation to apply it, apply it again, and yet again. Why stop?' (Dixon 1997:37) Dixon argues that the failure of long-distance genetic comparison stems mostly from the results of diffusion, which make the comparative method inapplicable to much of human history. 'It is neither sensible nor provident to look for a family tree of family trees.' (Dixon 1997:152) However, the discussions of Nostratic in Campbell (1998) and Comrie (1998) gives the methodological flaws in the arguments presented in favour of these genetic groupings (which could be summarised by 'comparing languages is not the same as applying the comparative method'). The failure of the recognition of deep genetic relations through the comparative method has little to do with either the family model or punctuated equilibrium; whether language change is through diffusion/contact or internal change, at some point regular correspondences will be eradicated due simply to the accumulation of regular phonological and irregular or quasi-regular analogical or morphological changes which divide the languages concerned.

Finally, I wish to make a methodological point that has relevance to Dixon's qualms about genetic classification in Australia. Dixon (2001, in press) argues against the validity of a reconstructed language Proto-Pama-Nyungan partly on the grounds that one can reconstruct no innovations from a higher proto-language that set it apart from other Australian groups. One must realise, however, that demonstrating genetic relatedness and demonstrating membership of a subgroup are different things; for example, one does not need to believe in and reconstruct Proto-Nostratic in order to define Proto-Indo-European! It is true, of course, that if one views Proto-Pama-Nyungan as a subgroup of 'Proto-Australian', to justify this one must provide cognates from Proto-Australian and show innovations to Proto-Pama-Nyungan, but it is not necessary to do this to show the cohesiveness of a Pama-Nyungan family of itself. Demonstrating a subgrouping relationship and demonstrating genetic relatedness are two different processes.

3.6 On the Comparative Method

The comparative method is based on the assumption that internal language change is regular, and that regular, systematic correspondences between languages are meaningful and highly unlikely to be due to chance. This is the Neogrammarian hypothesis of the regularity of sound

change. What one does with this inference, however, is not strictly part of the comparative method.[5] The comparative method only relies on our assumption that languages employ regular sound change and exceptions to this require special pleading (as loan words, as words that have undergone sporadic change due to cliticisation, and the like). If Australian languages could be shown to change sporadically, then this would mean that the comparative method could not be applied in Australia, although nowhere in *The Rise and Fall of Languages* (or elsewhere, to my knowledge) does Dixon challenge the idea of regularity of sound change in principle.

Dixon (1997:149-152) attacks the comparative method as a 'discovery procedure', lambasting those (such as Hoenigswald (1960) and Anttila (1989/1972)) that equate the comparative method to an automatic synchronic phonological analysis. He argues that there are many places where a strict application of the comparative method will yield the wrong result, and thereby seeks to discredit the method. He is of course right to attack such reconstructions, although one would be hard put to find a thorough, detailed and plausible reconstruction of any family using the comparative method that applied it blindly, without taking into consideration the likelihood of the changes to be reconstructed, the possibilities of loanwords, and so on. In reality, neither synchronic nor diachronic phonology is practised as a mechanical discovery procedure. Problems like opacity are encountered, recognised, and dealt with in both synchronic and diachronic linguistics.

Dixon further criticises the comparative method because it will over-generate in some cases. For example, if a change occurs in all the languages of a group independently, it may well be reconstructed to the ancestor of the group, giving a false common innovation. We cannot reconstruct states for which we have no evidence. This is, of course, a weakness of the comparative method; however, it is a problem with all reconstructions of the past — if evidence has been obliterated, we won't see it, no matter what the method used (carbon-dating, stratigraphy, thermo-luminesce, punctuated equilibrium, glottochronology). This isn't a valid criticism of the comparative method *per se*, it's a fact of prehistory.

Dixon's third criticism of the comparative method is that proto-languages look neat and uniform, whereas 'natural languages are rather like an old garment that has been patched and mended' (Dixon 1997:45). Dixon implies here and elsewhere that the comparative method can only reconstruct regularity and that we can never reconstruct all aspects of a proto-language. I do not entirely agree. To take some established examples from Indo-European of areas we where can reconstruct irregularity: we can reconstruct certain irregularities in paradigms, such as the presence of full grades in locatives of ablauting root nouns despite the expected zero grade for oblique cases. We can reconstruct a vowel **a* distinct from that of h_2e despite their merger in all daughter languages: compare the behaviour of the Greek root *δαμάζω* 'I tame', and its ø-grade passive participle *δμᾱτός* < **damH-/dmH-tó-*, which shows a 'real' *a*, with the reflex of the root **steh₂-* 'stand' which does not lose the vowel but gives *στατός*, from **sth₂-tó-*, with vocalisation of the laryngeal. Another examples of reconstructing irregularity involves the remodelling of forms of the nominative

[5]Nor, incidentally, is the comparative method necessarily related to the family tree; this point will be taken up further below. The comparative method can still be used, for example, within linguistic areas and in the analysis of loan words (for an Australian example see Koch 1997 on loan words between Arandic and Warlpiri), which is what makes it such a powerful tool.

case in some paradigms, due to Szemerenyi's Law (involving compensatory lengthening over a resonant as the result of the loss of final *-*s*). There are numerous other rules to deal *precisely* with the irregularities which we can reconstruct, including Stang's Law (which deals with the unexpected non-vocalisation of word-final *m* in sequences of *VHm# (not the expected VHm̥)), Lindeman's Law (the variation in roots between *TRV(T)# ∼ *T(*R̥*)RV(T)), and the regularisation of certain proterokinetic paradigms in s-stem neuters in late Indo-European, such as the root **ĝenh*$_1$-, which we would expect to show ablaut between a nominative stem ***ĝenh*$_1$*-s-* and a genitive *ĝn̥h*$_1$*-és-s*, but instead appears to reflect a regularised, non-ablauting stem **ĝénh*$_1$-$^{e}/_{o}$ *s-*.

Another 'reconstruction of irregularity' which Dixon does not mention is loan-word analysis. It is precisely because these forms are irregular that we can identify them as being odd and worthy of separate explanation. It is also possible to reconstruct loans into proto-languages, using with case the same methods which we use to identify loans in modern languages. For example, it may be the case that the loan from B to A show a different reconstructible accent pattern, or an odd consonant cluster, or a morpheme otherwise unidentifiable, or an odd root shape. Such criteria allow us to infer that Proto-Germanic **rīk(i)ja-* (Go reikeis, OHG rihhi, Eng rich) is a loan from Celtic; comparison with Latin *rex* and Skt *rāj-* imply a reconstruction **rēĝ-*, and Celtic exhibits a regular sound change **ē* > *ī*, but this is not a regular change in Germanic. The irregular correspondence identifies the word as a probable loan. Thus instances where the comparative method fails can be as instructive as instances where it succeeds.

4 Punctuated equilibrium in Australia

Dixon (1997:91) regards Australia as a 'prototypical example of a long-time diffusion area'. Indeed, he states that his punctuated equilibrium model is the *only* way to account for the current distribution of Australian languages.

Dixon's views of several aspects of Australian languages are at variance from those of most other Australianists (as I understand them) and go some way to explaining his insistence on the lack of applicability of the family tree model in Australia. The main difference is the view that non-Pama-Nyungan languages are archaic in relation to Proto-Australian, and the non-Pama-Nyungan languages are innovative. This follows in part from the fact that in Dixon (in press) reconstruction proceeds according to general typological principles – that is, that languages tend to go from isolating to agglutinative to inflectional and back again; this is the basis for his view of the non-Pama-Nyungan languages as innovative and the Pama-Nyungan 'area' as an archaic diffusion area. The Pama-Nyungan type is assumed to be the original type, and the non-Pama-Nyungan languages are assumed to have innovated from that type by cliticising pronouns or catalysts to verbs to form inflectional bundles, as has also happened sporadically in Pama-Nyungan languages such as the Southern dialects of Baagandji (Hercus 1986).

Dixon assumes that Australian languages have been evolving fairly much *in situ* from early on since the initial expansion into Australia (probably around 60,000 years ago). His reasons are partly to do with origin myths (cf Dixon 1996) and partly because, as he states,

if Proto-Australian had been spoken only a few thousand years ago, the split and expansion would have happened rather recently, and a family tree model should be applicable' (Dixon 1997:92).

It is worth pointing out in general terms some of the problems with using origin myths and other types of myths to date population expansions and other aspects of prehistory. Take, for example, the following story about a volcanic eruption, told in Dyirbal to Dixon in the 1960s:

> It appears that beneath the veneer of fantasy some myths may provide accurate histories of events in the distant past of the people. There is, for instance, a Ngadyan myth that explains the origin of the three volcanic crater lakes Yidyam (Lake Eacham), Barany (Lake Barrine) and Ngimun (Lake Euramoo). It is said that two newly-initiated men broke a taboo and so angered the rainbow serpent, the major spirit of the area (...). As a result, 'the camping-place began to change, the earth under the camp roaring like thunder. The wind started to blow down, as if a cyclone were coming. The camping-place began to twist and crack. While this was happening there was in the sky a red cloud, of a hue never seen before. The people tried to run from side to side but were swallowed by a crack which opened in the ground ...' This is a plausible description of a volcanic eruption. After telling the myth, in 1964, the storyteller remarked that when this happened the country round the lakes was 'not jungle - just open scrub'. In 1968, a dated pollen diagram from the organic sediments of Lake Euramoo by Peter Kershaw showed, rather surprisingly, that the rain forest in that area is only about 7,600 years old. The formation of the three volcanic lakes took place at least 10,000 years ago. All this points to the story of the volcanic eruptions, and of the spread of the rain forest, having been handed down from generation to generation for something like ten millennia. This is perfectly possible: recent archaeological word suggests that aborigines have been in Australia for at least 25,000 years,[6] and the Dyirbal could well have been in more or less their present territory for 10,000 years or more. Dixon (1972:29)

There are several problems with the assumption that Dixon draws. First, just because people tell a story now and claim it as their own, it doesn't mean that they are the direct descendants of the people that the story happened to. Dixon gives no dates for population settlements in the areas he talks about, but from Mulvaney and Kamminga (1999:334) the area was settled around 5000 years ago, and extensive exploitation began from around 2000 years ago. That is after the savannah became rainforest and the sea level rose. In fact, it looks like exploitation of the islands came around with better marine technology (for examples of the borrowing of myths in Australia see Worms 1952).

Furthermore, the events in such stories are often very general and do not need to have been true (for example, seeing an island off the coast, it is very easy to make up a story that it was once possible to walk to it). The events, even if true, might not refer to the same area. Australia is very big. People migrate.

[6] The figure is now 55,000 years +.

More importantly for methodology, for every element in a story that might be true, there is usually a chunk of the story that one discards (for example, the rainbow serpent in the Dyirbal myth above). This very selective use of data leads us to ignore almost *everything* which does not fit our ideas of the prehistory of the area and only pick up on the aspects of the myth that fit.

4.1 Convergence in the lexicon revisited

An important component of Dixon's punctuated equilibrium model is the idea that Australian languages borrow back and forth until they reach an equilibrium with approximately 50% of their vocabulary in common; this figure and the algebra behind it was first published in Dixon (1972:332-337). It is asserted as fact ('Australian languages have an interesting property . . . ') in Dixon (1997:26) and is used as further evidence for the punctuated equilibrium model.

The 50% figure was derived like this: Suppose there are three languages in a row, A, B and C and each has 20% common vocabulary with its neighbours. In time T language B replaces 1% of its vocabulary by borrowing from its neighbours. Supposing that each language borrows about equally frequently from each of its neighbours, about 20% of the 1% vocabulary lost from B will be held in common with neighbour A (and similarly for neighbour C), but about half of the new vocabulary will have been borrowed from A (and similarly for C); thus after time T the new percentage of common vocabulary will be

(4) $20 - (0.2 + 0.2) + 0.5 + 0.5 = 20.6\%$

Thus the percentage of common vocabulary will slowly rise. Now, assuming that instead of 20% in common the languages had 70%, using the same reasoning will produce

(5) $70 - (0.7 + 0.7) + 0.5 + 0.5 = 69.6\%$

That is, that the figure of common vocabulary will slowly drop; the figures will converge at approximately 50%.[7]

Such as figure has been discredited by Alpher and Nash (1999), who point out (Alpher and Nash 1999:33) that lexical replacement is likely to be largely a factor of *internal* replacement rather than *external* replacement by borrowing. They studied replacement rates in Guugu Yimidhirr, analysing the replacement in lexical items from the wordlist collected by Captain Cook in 1770 and borrowings versus other replacements for 'sames' and 'not-sames' between several Cape York languages. Alpher and Nash (1999) show that replacement through mechanisms other than borrowing (which accounted for more than 50% of 'not-sames' in their sample) was likely to be an important factor which changes the equilibrium figure to something more than 25% common vocabulary.

5 An alternative model

I have gone into a discussion of punctuated equilibrium in some detail, in order to describe the model and to summarise the relevant issues in accounting for the current distribution of

[7]Further calculations are made for languages with multiple neighbours but Dixon (1972) does not model replacement due to mechanisms other than borrowing.

Australian languages. In this section, I outline my alternative scenario and the accompanying model of linguistic differentiation as it applies to the Australian situation.

I have called the model, for want of a better term, 'genetic areal linguistics', to imply that languages may be related genetically (that is, through the common inheritance of lexicon, morphology and syntax) while being also areally related; furthermore, following the speech-community network theory developed by Malcolm Ross, speech communities may undergo common innovation without us being able to classify them into subgroups.

The model that I am outlining here is not earth-shattering, indeed, the ideas in it have been around for quite a while, and it seems that various authors, including Evans and Jones (1997), McConvell and Evans (1997) and others assume something like what I am proposing here. The model of that describes linguistic differentiation is due almost entirely to Ross (1996, 1988, 1997). However, I think it is worth spelling out the scenario in detail, and making a number of assumptions explicit, precisely for the reason that this model does seem to be assumed or subsumed under a family tree model; no one to my knowledge is advocating something like this as a model for Australia—that is, genetic relatedness without a splitting family tree. Evans (e.g. Evans forthcoming), McConvell, and others who have written on the question still write in terms of a splitting family tree.

Another important point to make is that this model is a model of language change and classification/genetic relatedness; it is not a model encompassing socio-historical factors (other than the point that speakers of different linguistic allegiances must remain in contact); thus while I subscribe to the view (following Thomason and Kaufman 1988) that the history of a group of languages is a function of the history of its speakers, and does not exist in isolation, I do believe that socio-historical factors should be modelled independently of their linguistic consequences; as Kuteva (1999) and Campbell (in press b) (and Campbell in press a) have convincingly shown, there is no one-to-one correspondence between non-linguistic events and types of language change; thus while linguistic history is a function of non-linguistic events, the two can, and should, in my opinion, be modelled separately.[8]

Finally, what I am suggesting here is one possible way in which the languages of Australia, particularly of the Pama-Nyungan group, could be related, and a scenario to account for their current distribution. I am not claiming by any means that it is the only way, but it does fit our current knowledge not only of the Australian prehistoric situation, but our knowledge of mechanisms of language change elsewhere in the world — in my model, Australia is an unusual case, but it does not require special pleading.[9]

This is an outline of a path of inquiry that might be fruitfully investigated in detail for other parts of the continent.

[8]Of course, this is not to say that a unified treatment is impossible, just that it is impossible on the current data; this raises an important point about the design of linguistic models and what they do; they do not have to provide reasons for the effects exhibited in the data. There is a crucial difference between descriptive models and explanatory models, and many authors appear to confuse the two, or provide one for the other. This is (almost) purely a descriptive model.

[9]Note that although, especially in the following sections, I am directly contrasting my model with Dixon's; I do not mean this to be an adversarial paper, far from it! Rather, I wish simply to point out that there are strikingly different ways of viewing the same data, depending on one's initial assumptions.

5.1 The problem

The problem confronting us is how do we account for the recurrent similarities across the languages of a large proportion of the Australian continent? A neat family tree has remained elusive.

5.2 The scenario

The scenario that I propose to account for the lack of clear binary splits in Pama-Nyungan has its basis in dialect geography.

The principles of dialect geography are well-known from works such as Chambers 1998 and will only be briefly repeated here. The three most important for my purposes are:

- Speech communities are differentiated by isoglosses; isoglosses tend to bunch along natural barriers to communication but do not necessarily do so;
- Speech communities in the same geographical area often form chains of mutual intelligibility; that is, adjacent languages may differ minimally but either ends of the continuum may be mutually unintelligible;
- Speech varieties at the centre of an area tend to be innovative, while isolated conservative pockets exist on the fringes.

Now, consider the relationship between dialects and the family tree model; it has long been a paradox of genetic linguistics that the relationships between languages may be modelled on a family tree, but the relationships between languages do not fit a tree so easily. See, for example, Hock (1991:432ff), especially p 450:

> The linguistic relationship between neighbouring dialects of the same language very commonly cannot be stated in terms of tree diagrams. This is a consequence of the fact that these speech varieties remain mutually intelligible, stay in close contact, and therefore continue to interact with each other on a day-to-day basis, with shifting realignments as political and social circumstances change. It is therefore unrealistic to expect clear, 'tree-diagram' splits in such dialect continua.

One further point of note about the dialects of languages usually studied by dialectologists (and of which Hock is discussing) is that they are often subject to two normative pressures. One is that of a standard language (eg, RP, Hochdeutsch or Nynorsk) and the other is that of writing. I do not wish to go into this point in detail here but one should bear in mind that these two normative pressures did not exist for indigenous speech communities in Australia.[10]

Now, consider what we would expect to happen if dialectal speech communities continued to diverge, *in situ* and remaining in linguistic and social contact with one another. I argue that in such a situation we would expect to find a series of dialectal epicentres, with the speech communities around those epicentres speaking varieties more similar to one another's

[10]That is not to say that people were not proscriptive about language, of course.

than to varieties of other epicentres; we would expect to find some isoglossic scarps; that is, divisions where isoglosses have accumulated and bunch; we would also expect to find conflicting subgroupings. That is, we would find languages that exhibit traits that would in family tree terms lead them to be classified in multiple groups.

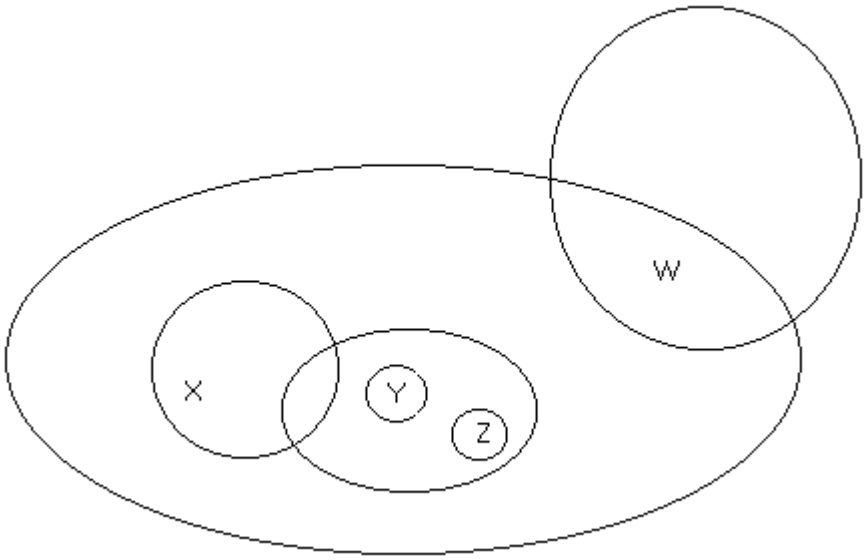

Figure 5: Schematic representation of an isogloss area. W, X, Y and Z represent languages and the lines represent isogloss boundaries. The relationship between languages X, Y and Z can be modelled on a family tree; language W, however, appears to belong both to the subgroup that contains X, Y and Z and to another group.

In this scenario, the possibilities for subgrouping would be quite limited, and often conflicting. However, crucially, the languages are still *genetically related*, in the same way that dialects without bunching isoglosses are. They will still probably exhibit regular correspondences and they are still descended from a single parent.

There seem to be three major advantages to this model over punctuated equilibrium to account for the distribution of languages in Australia. Firstly, there is a place for both divergence and convergence as processes of language change; punctuated equilibrium stresses convergence as the main mechanism of language change. Secondly, it makes Pama-Nyungan look much more similar to other areas of the world (surveyed very briefly in §7). We no longer have to assume that Australia is a special case. Thirdly, and related to this, we do not have to assume in this model that there has been intensive diffusion of many linguistic elements that in other parts of the world are resistent to borrowing (such as shared irregularities).

My model does assume that the dispersal of speakers of Pama-Nyungan is relatively recent; not, that is, of the order of 50,000 years or more ago, perhaps an order of magnitude more recent. This is in accordance with the relevant archaeological evidence, although, of course,

linking the spread of Pama-Nyungan with the series of (re)colonisations of abandoned or previously uninhabited areas is an unsupported assumption.

Note, incidentally, that my reasoning is similar to that in Golla (2000), but I draw different conclusions. Golla (2000:60ff) argues that languages and language families can be parameterised according to *split* versus *compact*; compact languages, he argues, form when speech communities are adjacent and in frequent intimate contact. Spread languages, by contrast, are spoken over large areas with infrequent contacts between members of different groups. Likewise families; spread families develop under the same conditions that are conducive to the formation of spread languages, while compact families develop out of compact languages, with all the social and isoglottal intertwining that one expects from such an area. Note further that Golla (2000:62-63) equates the 'classic model of linguistic relationships', that is, for families such as Indo-European and Austronesian, with spread families.

Pama-Nyungan, in Golla's terms, really have to be a spread family; its languages cover approximately 80% of the Australian continent. Karnic alone has an area of around 900 km^2. Pama-Nyungan would thus appear to provide a counterexample to Golla's claim that spread families are the classic Indo-European type.

These claims should not be considered original, and others have claimed what amounts almost to the same thing. The work of Malcolm Ross, in particular, discusses cases of conflicting subgrouping in Oceanic (for a brief summary see Lynch *et al.* 2002:92-94) and in particular 'innovaton-linked' subgroups; such groups share a common set of innovations, which are not shared outside the subgroup, yet the languages themselves provide us with no evidence for a single common proto-languages from which all the languages in the subgroup are descended.

5.3 Some more relevant parameters

Whatever one's model of language change in Australia, there are a series of questions that one must consider, and a set of parameters which must be addressed in order to define the model appropriately. I flag these as issues for further discussion, and do not address them in detail here.

- What constitutes genetic relatedness?
- What sort of social networks (in the sense of Milroy 1987) do hunter-gatherers have?
- What is the role of a standard (and/or written) language as a normative factor?
- What do we need to know about patterns of language contact, borrowing, and multilingualism?
- What are the relative roles of divergence and convergence in language change?
- What is the role of non-linguistic factors in determining events.

As a working definition I take the term 'genetic relatedness' following Thomason and Kaufman (1988:9-10) as meaning that two (or more) languages are descended from the same parent language and have been passed from parent generation to child generation, 'with

relatively small degrees of change over the short run'. Thus for two languages to be genetically related, they must have the same parent, and they must each have a single parent.

6 The Lake Eyre basin in detail

The languages of the Lake Eyre Basin are an excellent case with which to illustrate this model. Most of the languages are quite well described; we have adequate data for synchronic analysis and reconstruction at almost all levels of lexicon, morphology and grammar. Secondly, we have good archaeological data from a number of sites in the region; we also have extensive data on the trade and mythological networks of the area and their relations with other groups, especially to the West (my data are from McBryde 1987 and the summary in Mulvaney and Kamminga 1999, as well as unpublished work by Luise Hercus not quoted here; cf also Hercus 1980). Thirdly, given the lack of consistency in classification (see §2.2 above), and the fact that Dixon (2001) classifies Karnic as a small linguistic area, the Lake Eyre Basin should be a good area to test the ideas of genetic relatedness in this paper.

The evidence that I see in favour of considering Karnic a genetic subgroup of *something* (if not Pama-Nyungan in the way it is currently formulated) is presented in full in Bowern (1998) and in summary in Bowern (2001b). The evidence is summarised in the following sections. Due to limitations of space the supporting evidence is not presented in detail here, but is argued in full in earlier publications.

6.1 Contact and trade

We have very good evidence of well-defined and long-standing contact through trade and ceremony which cross language boundaries but which do follow major isoglosses. Hercus (1980), for example, describes one of the mythological tracks which cuts through the Simpson Desert and goes (amongst other places) between Arabana country and Lower Southern Aranda. McBryde (1987:260) gives the major sources of raw materials and trade networks in the Lake Eyre Basin. If Karnic were simply a diffusion area, we would not expect recurrent and regular similarities between languages which are not in frequent direct or semi-direct contact (such as Arabana-Wangkangurru and Wangkumara).

6.2 Lexical innovation

There are innovations in vocabulary between Proto Pama-Nyungan and Proto Karnic. That is, there are forms that are widespread in Pama-Nyungan languages (whether a genetic or an areal term) that are not shared by Karnic languages, but Karnic languages share another term. These are presented in Table 1. The sources for Proto-Pama-Nyungan lexical reconstructions are Alpher (forthcoming), Koch (1996), O'Grady (1990), Dixon (1980) and Capell (1963). The source for Proto-Karnic is Austin (1990). A number of doubtful items have been omitted from both lists.[11]

[11]The orthography is that used in Hercus (1994), which is a practical orthography most suited to the phonemic contrasts in Karnic languages. The velar nasal is written *ng*, stops are written as voiceless (except, of course, in the Karnic languages with phonemic voicing contrasts) and there are three rhotics—the trill is written *rr*, the flap *r* and the glide ʀ.

PPN	PK	English
kumpu	**purra*	urine
kami	**kanyini*	mother's mother
kutharra	**parrkulu*	two
muka	**pampu*	egg
nga-	**thayi*	eat
ngalirna	*no category*	1dl excl.
ngatyi	**kami*	mother's father
nguna	**parri*	lie down
nhu-	**nguntyi*	give
nhumpVlV	**nhula*	2dl
nyiina	**ngama*	sit
pangV	**paku*	dig
parnta	**marda*	stone
parnti	**panthama*	smell
patha	**matha*	bite
purlka	**pirna*	big
tharrV	**tharrka*	stand
thalany	**tharli*	tongue
paka	**paku*	dig
partu-	**karlathurra*	turkey
kunka	**kimpa*	alive, raw
pangkarra	**kalta*	blue-tongue lizard

Table 1: lexical innovations between Proto-Pama-Nyungan and Proto-Karnic

6.3 Shared morphological irregularity

Karnic languages share irregularities in certain pronominal paradigms, and these irregularities are not shared by other Pama-Nyungan languages. These include first person *ngathu : nganyi : ngantya* (erg:nom:dat); third person masculine singular *nhulu : nhinha* (erg:acc) (only Karnic languages have the vowel change); 2nd person dual *nhula* (compare Proto Pama-Nyungan **NHuNpal-*).

6.4 Shared phonological aberrancy

Karnic languages are the only languages in the area to reflect a voicing contrast in apical stops. This evidence is difficult, since not all languages preserve it clearly. Arabana-Wanggangurru and Pitta-Pitta reflect it as a three-way contrast in rhotics (tap ≠ trill ≠ glide); whereas other Karnic languages, in addition to showing the rhotic distinction also show a voicing contrast in apicals in clusters (nt ≠ $nd^{(r)}$) as in the Central languages, or at all points of articulation, as in Yandruwandha and Wangkumara. Further details can be found in Bowern (1998:39-40,43).

6.5 The diffusion scenario

Conversely, if we assume (as Dixon does) that Karnic is a typical diffusion area with no language more closely related in the group than to any other and most similarity (that not inherited from Proto-Australian), we have to assume the following:

(6) That the following characteristic 'Karnic' features have <u>all</u> diffused:
   i. Suppletive first person pronouns: ergative *ngathu*, nominative *nganyi*, accusative *nganha*, dative *ngantya* (*nganyi* and *ngantya* being found nowhere else in the area;[12]) - this is wrong; Malyangapa has *nganyi*... will fix.
   ii. a shift of the locative case to the meanings associated with the dative;
   iii. the change of **d* to a tap intervocalically and voicing contrasts in apicals were innovated in the same way in different languages.

(7) That, despite good archaeological evidence to the contrary, trade routes and dreaming tracks have shifted in the near past, bringing Arabana and Lower Southern Aranda speakers closer than in previous periods.

6.6 Diffusion patterns

We may contrast this with the patterns of features and general similarities that criss-cross the Lake Eyre Basin. The major ones are given below:

(8) i. Prestopping of sonorants – nasals and laterals in Arabana-Wangkangurru and Adnyamathanha; nasals only in Arandic; laterals and *n* in Yandruwandha; apical and lamino-dental nasals and laterals in Diyari and Ngamini;

[12] the only *nganyi* form I know of is in Nhanta (Blevins 1999) from the other side of the continent and I have not seen a form *ngantya* anywhere else, although *ngatyu* and *ngatya* are both quite widespread.

ii. l > r (a tap) in Arabana-Wangkangurru and Wangkajutjurru (but not Pitta-Pitta).

iii. inclusive pronouns have been borrowed from Yandruwandha (only in this language do the phonetic developments make sense) – Diyari borrowed the dual *ngaldra* and Wangkumara the plural *ngandra*

iv. trilled release of apical stops in homorganic nasal clusters in Diyari and Yandruwandha (Yandruwandha also has this in initial position; cf *rdriya* 'tooth' [dr̥ríyə])

v. Personal pronouns marked for kin terms.

6.7 Summary

Karnic appears to be a fairly typical area within Pama-Nyungan; a large amount of recent diffusion, some old dialect areas and some more major isogloss lines which can be used as the basis for genetic classification. I believe that there is no need to assume convergence over 50,000 years (as Dixon must do to explain the similarities). It is, in fact, possible to provide a concise explanation for the situation of Karnic by assuming that the modern languages developed out of dialect areas. Such a proposal has been made for the Admiralty Islands by Ross (1988) and for Fiji by Geraghty (1983). In each case the overall pattern looks similar; in an examination of the relative chronology of the area, one finds many loanwords, some recent irregular phonological borrowings and structural calquing (Ross's (1998) metatypy), and many old features that, all else being equal, one would reconstruct to a common proto-language, some of which cross language boundaries. This may well be simply what modern languages that have evolved *in situ* and in continued contact with their neighbours look like, and there is no need to posit types of changes any different from those familiar from 'normal' family trees (although this requires much more investigation).

We could thus consider the following stages:

(9)

Stage I:	'Proto-Karnic'
Stage II:	Isoglosses appeared as speech communities diversified; 'Common Karnic' developed dialects. There was movement of speakers, but different dialects maintained in contact.
Stage III:	More isoglosses built up. Isoglosses show 'bundling' but still a great deal of overlapping.
Stage IV:	(Modern attested stage) 'Dialects' are mutually unintelligible; certain 'family-tree-like' splits are found, but there are also a large number of isogloss boundaries which reflect old dialect areas.

This would result in some major dialectal divisions (a certain number of 'splits') but not necessarily intervening proto-languages, and not necessarily clear splits that the family tree implies. This is, indeed, exactly what we find. The difficulties in drawing family trees for dialects are well attested. Why, then, should it be thought a catastrophe for the method if the family tree does not work well in languages which have evolved from a complex dialect area? Indeed, why should we expect it to work as well in such areas as it does in areas where

continued contact has been much less. This casts no doubt on the genetic relationship between the languages, or, ultimately, on our ability to reconstruct the changes in the languages, but it does make it difficult to draw well-defined family trees.

7 Other potential areas of application: Turkic and beyond

Turkic is a family with several distributional characteristics that are strikingly similar to that of Pama-Nyungan. It is spoken over a large area, with structure and a certain amount of lexical homogeneity over the area as a whole, but with local diversity. There are some main bunching isoglosses but a lot of overlapping isoglosses which cut through otherwise quite homogeneous areas, and the geographical fringes of the family are quite diverse (Csató and Johansson 1998). Moreover, there has been suggestion that Turkic (like Dixon says for Pama-Nyungan) is not a genetic unity. '... a merely "genetic" classification is not sufficient. We have good reasons to assume that all the contemporary Turkic languages have not *directly* developed from a common ancestor.' (Schönig 1997a:118)

Note here that Schönig uses the term 'genetic heritage' in a way rather different from Thomason and Kaufman (1988:9). Schönig uses the term in the sense of neat splits in a family tree (that is, the appearance of diversification), whereas Thomason and Kaufman use 'genetic' to apply to the means of transmission of a language between one generation and the next (genetic being uninterupted, 'normal' acquisition).

From Dixon's point of view, Turkic would seem to be a paradigm case of a family that has undergone punctuation. Yet it also shows very few of the characteristics of an articulated family tree. It is, rather, a complex network of isoglosses that bunch in certain ways.

Schönig (1997a:119ff) (and also Schönig 1998, 1999) classify Turkic into Common Turkic, Norm Turkic, Central Turkic and Border Turkic. Common Turkic features are found in all languages classified as Turkic, and include a reflex of the verb **al-* 'to take' and a preterite in **-DI*. Norm Turkic is defined by the loss of word-initial **h-*, and nominal plural **-lAr*, and a few other innovative features.[13] The non-Norm Turkic languages, Chuvash, Khalaj and Lena Turkic show a mixture of innovations from the Norm-Turkic features (eg replacement of *-lAr* in Chuvash) and non-development of Norm features (Schönig 1997a:121-122). Similar critieria are used for Central Turkic (Oghuz, Kipchak and South East Turkic (including New Uighur)). There are languages in the group, such as Altaï, which show some Central Turkic features but not others (Schönig 1997a:128-129), and considerable overlap between Eastern Oghuz languages and other Central Turkic languages. We also see at the periphery of the family languages like Uzbek (which has lost vowel harmony under contact with Ossetic (Nichols 1998)).

I do not see how, in Dixon's model, we are to differentiate archaic dialect areas (perhaps thousands of years old) from archaic equilibrium areas (perhaps 50,000 years old, in the Australian example). Perhaps Dixon's point is that we cannot distinguish them without

[13] I ignore the common use in Turkic linguistics of relying on shared archaisms, such as in Tekin (1991), Tekin and Ölmez (1995) and others. In Schönig (1997b:272-273), for example, of the eleven characteristics used to argue for a common connection with Oghuz and Khalaj, only three of them are innovations (and one, the cliticisation of PT **bi(r)lä(n)* as a postposition meaning 'with' has happened in other Turkic languages as well and could be seen as a parallel innovation.

external linguistic evidence. However, I find this very unsatisfactory. If Dixon's model cannot distinguish between the results of a punctuation a few thousand years ago (which he says himself should produce a family tree-like set of splits) and the results of 50,000 years of equilibrium, perhaps his model needs refining. Dixon would perhaps rebut this by saying that Turkic is an exception to his model, but the Turkic expansion fits Dixon's stated views of a punctuation in such an exact way that it is a serious counter-example that needs addressing. Dixon could also argue that the Turkic expansion is too recent for a proper family tree to develop; however if in the next thousand years the Turkic languages continue to diversify, I would predict that we would be likely still to be able to reconstruct overlapping isoglosses. In that case, in a thousand years time, how different would Turkic look from a linguistic area like Karnic or Pama-Nyungan?

In summary, I would like to suggest that the data from Turkic point us to an analysis of Pama-Nyungan that does not require a great time depth with all similarities due to diffusion. Turkic seems to be a family where the languages have diversified while remaining in contact with one another, and therefore the innovations which one associates with dialectal differentiation have carried over to languages which have now ceased to be mutually intelligible.

8 Conclusions and further directions

I have hopefully shown in this paper that the punctuated equilibrium model and the classical family tree models are not the only way to view the history of Pama-Nyungan expansion in Australia; that we can assume a date of expansion more recent than initial colonisation and still account for the lack of an articulated family tree, and that we can reconcile diffusion areas with recent migration and still talk in terms of genetic relationship.

Of course, this model requires intensive testing, both at the level of Pama-Nyungan but also at the level of subgroups of Pama-Nyungan. Work on this is progressing, including the results published in Bowern and Koch (2002b) and Dench (2001); other work is in progress. Without detailed, step-by-step reconstruction and plotting of isoglosses, hypotheses of innovations, and the like, however, these ideas remain untestable.

References

Aikhenvald, Alexandra and R.M.W. Dixon (eds.) (2001). *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*. Oxford: Oxford University Press.

Alpher, Barry (forthcoming). Pama-Nyungan: Phonological reconstruction and status as a phylogenetic group. In Harold Koch and Claire Bowern (eds.), *Subgrouping and reconstruction: Case studies from Australian languages*, chap. 5, p. c 180 pp.

Alpher, Barry and David Nash (1999). Lexical replacement and cognate equilibrium in Australia. *Australian Journal of Linguistics* 19: 5–56.

Anttila, Raimo (1989/1972). *Historical and Comparative Linguistics*, vol. 6 of *Current Issues in Linguistic Theory*. Amsterdam: John Benjamins.

Austin, Peter (1990). Classification of Lake Eyre languages. *La Trobe University Working Papers in Linguistics* 3: 171–201.

Bellwood, Peter (1997). Prehistoric cultural explanations for widespread language families. In McConvell and Evans (1997), chap. 8, pp. 123–134.

Bellwood, Peter (1998). Agriculture and language spread. Talk at ANU (part of workshop on diffusion and genetic linguistics).

Bellwood, Peter (2001). Archaeology and the historical determinants of punctuation in language-family origins. In Aikhenvald and Dixon (2001), chap. 2, pp. 27–42.

Benveniste, Emile (1966). *Problèmes de linguistique générale*. Paris: Gallimard.

Blench, Roger and Matthew Spriggs (eds.) (1998). *Archaeology and Language II*. London: Routledge.

Blevins, Juliette (1999). Nhanta and its position within Pama-Nyungan. *Oceanic Linguistics* 38(2): 297–320.

Bowern, Claire (1998). *The Case of Proto-Karnic*. BA Honours Thesis, Australian National University, Canberra, Australia.

Bowern, Claire (2001a). Karnic classification revisited. In Jane Simpson, David Nash, Mary Laughren, Peter Austin and Barry Alpher (eds.), *Papers in Australian Linguistics*, Canberra: Pacific Linguistics, chap. 17, pp. 245–261.

Bowern, Claire (2001b). Root/affix faithfulness in Bardi, talk given at the 75th LSA annual meeting, Washington, DC.

Bowern, Claire and Harold Koch (2002a). Introduction. In Bowern and Koch (2002b), ms.

Bowern, Claire and Harold Koch (eds.) (2002b). *Subgrouping and reconstruction: case studies from Australian languages*. Cambridge, Mass and Canberra: ., ms.

Breen, Gavan (1971). Aboriginal languages of Western Queensland. *Linguistic Communications* 5: 1–88.

Butzer, K W (1982). *Archaeology as Human Ecology*. Cambridge: Cambridge University Press.

Campbell, Lyle (1998). Nostratic: a personal assessment. In Salmons and Joseph (1998), pp. 107–151.

Campbell, Lyle (in press a). Beyond the comparative method? In Barry Blake and Kate Burridge (eds.), *Selected papers from the Fifteenth International conference on historical linguistics*, Amsterdam: John Benjamins.

Campbell, Lyle (in press b). What drives linguistic diversity and language spread? In Colin Renfrew and Peter Bellwood (eds.), *Language-Farming dispersals*, Cambridge: McDonald Institute for Archaeological Research, p. 22 pp.

Capell, Arthur (1963). *Linguistic Survey of Australia*. Canberra: Australian Institute of Aboriginal Studies.

Chambers, J K (1998). *Dialectology*. New York: Cambridge University Press, 2nd edn.

Comrie, Bernard (1998). Regular sound correspondences and long-distance genetic comparison. In Salmons and Joseph (1998), pp. 271–276.

Csató, Eva and Lars Johansson (1998). *The Turkic languages*. London: Routledge.

Dench, Alan (2001). Descent and diffusion: the complexity of the Pilbara situation. In Aikhenvald and Dixon (2001), chap. 5, pp. 105–133.

Dixon, R.M.W. (1972). *The Dyirbal language of North Queensland*, vol. 9 of *Cambridge Studies in Linguistics*. London: Cambridge University Press.

Dixon, R.M.W. (1980). *The Languages of Australia*. Cambridge: Cambridge University Press.

Dixon, R.M.W. (1996). Origin legends and linguistic relationships. *Oceania* 67: 127–139.

Dixon, R.M.W. (1997). *The Rise and Fall of Languages*. Cambridge: Cambridge University Press.

Dixon, R.M.W. (2001). The Australian linguistic area. In Aikhenvald and Dixon (2001), chap. 4, pp. 64–104.

Dixon, R.M.W. (in press). *Australian languages: their nature and development*. Cambridge: Cambridge University Press.

Durie, M. and M Ross (eds.) (1996). *The comparative method reviewed; regularity and irregularity in language change*. New York: Oxford University Press.

Eldredge, Niles and Stephen Jay Gould (1972). Punctuated equilibria: an alternative to phyletic gradualism. In T J M Schopf (ed.), *Models in paleobiology*, San Francisco: Freeman, Cooper, pp. 82–115.

Evans, Nicholas (ed.) (forthcoming). *Studies in comparative Non-Pama-Nyungan*. Canberra: Pacific Linguistics.

Evans, Nicholas and Rhys Jones (1997). The cradle of the Pama-Nyungans: archaeological and linguistic speculations. In McConvell and Evans (1997), chap. 22, pp. 385–417.

Frankel, David (1988). Characterising change in prehistoric sequences: a view from Australia. *Archaeology in Oceania* 23(2): 41–48.

Geraghty, Paul (1983). *History of the Fijian languages*, vol. 19 of *Oceanic Linguistics special publications*. Honolulu: University of Hawai'i Press.

Gimbutas, Marija (1963). The Indo-Europeans: archaeological problems. *American Anthropologist* 65: 815–836.

Golla, Victor (2000). Language families of North America. In Colin Renfrew (ed.), *America past, America present: genes and languages in the Americas and beyond*, Papers in the Prehistory of Languages, Cambridge: McDonald Institute for Archaeological Research, chap. 5, pp. 59–75.

Harris, David R (ed.) (1996). *The origins and spread of agriculture and pastoralism in Eurasia*. London: UCL Press.

Hercus, Luise and Peter Austin (2002). the Yardli languages. In Bowern and Koch (2002b), ms.

Hercus, Luise Anna (1980). How we danced the mudlunga: Memories of 1901 and 1902. *Aboriginal History* 4(1): 5–32.

Hercus, Luise Anna (1986). *The Baagandji language*. Canberra: Pacific Linguistics.

Hercus, Luise Anna (1994). *A Grammar of the Arabana-Wangkangurru Language. Lake Eyre Basin, South Australia*, vol. C-128. Canberra: Pacific Linguistics.

Hock, Hans Henrich (1991). *Principles of historical linguistics*. Berlin: Mouton de Gruyter, 2nd edn.

Hoenigswald, Harry M (1960). *Language change and linguistic reconstruction*. Chicago: University of Chicago Press.

Joseph, Brian (1998). Review of Dixon (1997). *Diacrhonica* x(x): x–y.

Koch, Harold (1996). Reconstruction in morphology. In Durie and Ross (1996), chap. 8, pp. 218–263.

Koch, Harold (1997). Comparative linguistics and Australian prehistory. In McConvell and Evans (1997), chap. 3, pp. 27–43.

Kuteva, Tania (1999). Language and societes: the "punctuated equilibrium" model of language development. *Language and Communication* 19: 213–228.

Lynch, John, Malcolm Ross and Terry Crowley (2002). *The Oceanic languages*. Curzon language family series, Surrey: Curzon.

Mallory, J P (1989). *In search of the Indo-Europeans: language, archaeology and myth*. London: Thames and Hudson.

McBryde, Isobel (1987). Goods from another country: Exchange networks in the lake eyre basin. In *Australians to 1788*, Sydney: Fairfax, Syme and Weldon Associates, vol. III of *Australians: a historical library*, pp. 252–273.

McConvell, Patrick (2000). Language shift and language spread among hunter-gatherers, chapter 8.

McConvell, Patrick and Nicholas Evans (1997). *Archaeology and linguistics: Aboriginal Australia in global perspective*. Melbourne: Melbourne University Press.

McDonald, M A and S A Wurm (1979). *Basic materials in Wangkumara (Gaḷali)*, vol. C-? Canberra: Pacific Linguistics.

Milroy, Lesley (1987). *Language and social networks*, vol. 2 of *Language in society*. Oxford: Blackwell, 2nd edition edn.

Mulvaney, John and J. Kamminga (1999). *The prehistory of Australia*. ??: ??

Nichols, Johanna (1992). *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press.

Nichols, Johanna (1998). The Eurasian spread zone and the Indo-European dispersal. In Blench and Spriggs (1998), chap. 10, pp. 220–266.

O'Grady, Geoffrey, Kenneth Hale and Stephen Wurm (1966a). Languages of Australia. Map, university of Victoria, BC.

O'Grady, Geoffrey, C. Voegelin and F. Voegelin (1966b). *Languages of the World: Indo-Pacific XXX*. Anthropological Linguistics.

O'Grady, Geoffrey N. (1990). Prenasalisation in Pama-Nyungan. In Philip Baldi (ed.), *Linguistic Change and reconstruction methodology*, Berlin: Mouton de Gruyter, vol. 45 of *Trends in Linguistics: Studies and Monographs*, pp. 451–477.

Renfrew, Colin (1987). *Archaeology and language*. London: Jonathon Cape.

Ross, Malcolm (1996). Contact-induced change and the comparative method: cases from Papua New Guinea. In Durie and Ross (1996), chap. 8, pp. 180–217.

Ross, Malcolm D. (1988). *Proto-Oceanic and the Austronesian languages of Western Melanesia*, vol. C of *98*. Canberra: Pacific Linguistics.

Ross, Malcolm D. (1997). Social networks and kinds of speech-community event. In Roger Blench and Matthew Spriggs (eds.), *Archaeology and Language I*, London: Routledge, chap. 13, pp. 209–261.

Ross, Malcolm D. (1998). Sequencing and dating linguistic events in Oceania: the linguistics/archaeology interface. In Blench and Spriggs (1998), chap. 6, pp. 141–173.

Salmons, Joseph C and Brian D Joseph (eds.) (1998). *Nostratic: sifting the evidence*, vol. 142 of *Current Issues in Linguistic Theory*. Amsterdam: John Benjamins.

Schmidt, W. (1919). *Die Gliederung der Australischen Sprachen*. Vienna: Mecharisten-Buchdruckerei, english translation by D. Clark: Classification of the Australian Languages. 1972 ms. AIAS.

Schönig, Claus (1997a). A new attempt to classify the Turkic languages (1). *Turkic languages* 1(1): 117–133,.

Schönig, Claus (1997b). A new attempt to classify the Turkic languages (2). *Turkic languages* 1(2): 262–277.

Schönig, Claus (1998). A new attempt to classify the Turkic languages (3). *Turkic languages* 2(1): 130–151.

Schönig, Claus (1999). The internal division of modern Turkic and its historical implications. *Acta Orientalia* 52(1): 63–95.

Tekin, Talat (1991). A new classification of the Turkic languages. *Türk Dilleri Arastırmaları* 1: 5–18.

Tekin, Talat and M Ölmez (1995). *Türk Dilleri/Les langues turques*, vol. 20 of *Türk Dilleri Arastırmaları Dizisi*. Ankara.

Thomason, Sarah Grey and Terrence Kaufman (1988). *Language contact, creolization and genetic linguistics*. Berkeley and Los Angeles: University of California Press.

Watkins, Calvert (2001). An Indo-European linguistic area and its characteristics: ancient Anatolia.

Areal diffusion as a challenge to the comparative method? In Aikhenvald and Dixon (2001), pp. 44–62.

Worms, Ernst A. (1952). Djamar and his relation to other culture heroes. *Anthropos* 47: 540–560.

Wurm, Stephen A. (1972). *Languages of Australia and Tasmania*. the Hague: Mouton de Gruyter.

Department of Linguistics
Harvard University
305 Boylston Hall,
Cambridge, MA, 02138
USA

email: `bowern@fas.harvard.edu`