AUSTRALIAN JOURNAL OF LINGUISTICS
https://doi.org/10.1080/07268602.2020.1848796

The genetic position of Anindilyakwa

Marie-Elaine van Egmond and Brett Baker

Anhalt University of Applied Sciences and University of Melbourne

ABSTRACT

In this paper, we demonstrate that Anindilyakwa, spoken on Groote Eylandt, East Arnhem Land, is genetically closely related to Wubuy (Gunwinyguan). Anindilyakwa has long been believed to be a family-level isolate, but by a rigorous application of the Comparative Method we uncover regular sound correspondences from lexical correspondence sets, reconstruct the sound system of the proto-language, and suggest how the proto-phoneme inventory derives from the proto-Gunwinyguan system through phonological innovations. Although it has been hinted before that Anindilyakwa and Wubuy are related and together with Ngandi form a subgroup, this hypothesis is not borne out here: while Wubuy and Ngandi have been shown to share a significant amount of core vocabulary and irregular verbal paradigms, Anindilyakwa and Wubuy appear to have undergone separate development for a considerable length of time. Moreover, Anindilyakwa has independently undergone extensive further sound changes, resulting in a language that is phonologically (though not lexically or grammatically) quite unusual in Australia.

ARTICLE HISTORY
Accepted 30 October 2020

KEYWORDS
Non-Pama-Nyungan; Gunwinyguan; Anindilyakwa; Wubuy; Nunggubuyu; Comparative Method; linguistic reconstruction; North Australia

1. Introduction

The genetic position of Anindilyakwa [ɛniɲṯi'ʎakwa], the traditional language of the whole of Groote Eylandt and neighbouring Bickerton Island (Gulf of Carpentaria, NT; see Figure 1), has long presented a puzzle to Australian linguistics. O'Grady et al. (1966) classified both Anindilyakwa and its nearest geographical neighbour Wubuy (a.k.a. Nunggubuyu: Heath, 1984) as language isolates, while O'Grady (1979) excluded Anindilyakwa (and the Tasmanian languages) from membership in the "Original Australia" lineage (Evans, 2003, p. 9). Alpher et al. (2003; henceforth AEH) argued, based on verbal inflectional suffixes, that Wubuy should be added to the Gunwinyguan family, but "because of the difficulties of the [Anindilyakwa] data" (p. 308, fn. 4), they did not include this language in their reconstruction.

However, Jeffrey Heath, a linguist who did fieldwork on Anindilyakwa as well as Wubuy and Ngandi (and other languages), has claimed since 1978 (1978a, 1981, 1990, 1997, n.d.) that Anindilyakwa and Wubuy are related and form a subgroup with Ngandi, though without providing much formal evidence to support his claim. His view is not accepted by most other linguists (except Dixon, 2002). For example, AEH feel that the genetic relation of Anindilyakwa with Wubuy and/or Ngandi has not "been demonstrated

CONTACT Marie-Elaine van Egmond mevanegmond@gmail.com

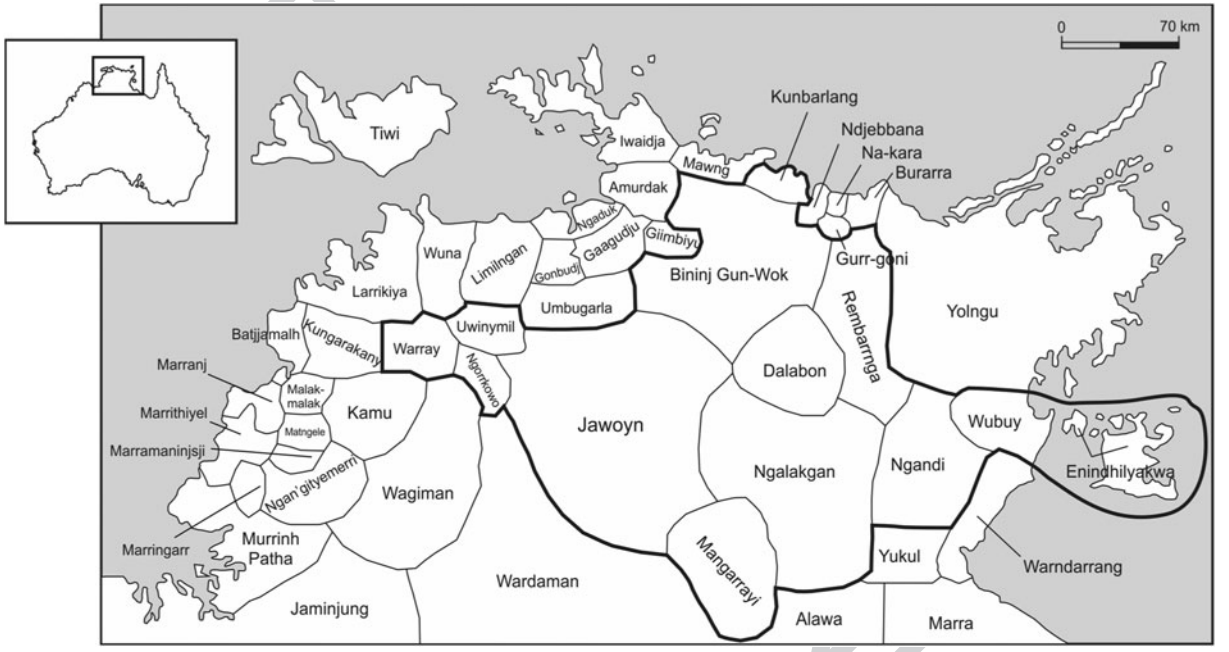

Figure 1 The Gunwinyguan family (based on Alpher et al., 2003; Evans, 2017; Harvey, 2003; Harvey & Mailhammer, 2017; Van Egmond, 2012)

conclusively at this point with any significant body of cognate lexical items or grammatical morphology" (p. 308). Evans (2005, p. 250), in his review of Dixon's (2002) book, concludes that "the evidence here is slender and in my view one should, for the moment, maintain the conservative position that [Anindilyakwa] is a family-level isolate". And in Baker's (2004) reconstruction of the Eastern Gunwinyguan group (consisting of Wubuy and Ngandi only), he notes that "Anindilyakwa retains so little in the way of recognizable monomorphemic roots and inflections that it is impossible to say if it was once related to Ngandi and [Wubuy]" (fn. 25).

In this paper, we demonstrate that Anindilyakwa is without doubt immediately related to Wubuy, but to a much lesser extent to Ngandi. We apply the Comparative Method to uncover systematic sound correspondences based on lexical correspondence sets and we reconstruct the sound system of the proto-language. In terms of shared sound changes, however, there appears to be little reason to include Ngandi in the reconstruction of the phoneme system of the proto-language: the Ngandi inventory is virtually identical to the one reconstructed for proto-Gunwinyguan by Harvey (2003), whereas Anindilyakwa and Wubuy have both undergone phonological innovations, namely: (a) elimination of the proto-Gunwinyguan fortis-lenis stop contrast together with (b) lenition of proto-Gunwinyguan lenis stops to approximants, (c) introduction of new phonemes such as lamino-dental lateral /l̪/ and (in the case of Anindilyakwa) lamino-palatal lateral /ʎ/, and (d) reduction of the vowel inventory from five in proto-Gunwinyguan to three in Wubuy and four in Anindilyakwa. None of these changes applied to Ngandi. Therefore, we exclude Ngandi in our reconstruction of the segmental inventory of the proto-language and reconstruct the sound system and vocabulary of the immediate ancestor of Anindilyakwa and Wubuy only, which we call 'proto-Insular Gunwinyguan' for practical purposes (this name refers to the fact that Anindilyakwa is spoken only on islands, and the suggestion – from Anindilyakwa and Wubuy speakers – that Wubuy-speakers also migrated to these islands in the past).[1] We recognize that there is insufficient evidence presented herein to demonstrate a solid argument with respect to subgrouping, and that further research may well uncover reasons to include other languages within this group.

The considerable amount of vocabulary shared between Anindilyakwa and Wubuy raises the question of why linguists have failed for decades to notice the relatedness of the two languages. One likely reason is the presumed isolate status of Anindilyakwa, so that Van Egmond (2012) was the first scholar to take up the task of systematically comparing Anindilyakwa and Wubuy. Another reason, addressed in detail in this paper, may be the rather dramatic changes that Anindilyakwa phonology and phonotactics have undergone, resulting in a system that is not only atypical of the Gunwinyguan languages (or, indeed, of Arnhem Land languages more generally), but which may also have obscured cognate forms. The unusual phonology of the language may furthermore have hampered previous attempts at reconstruction, so that Anindilyakwa was long considered as "perhaps the most difficult of all Australian languages, with a very complex grammar" (Dixon, 1980, p. 84; see also Capell, 1942, p. 376; Yallop, 1982, p. 40), and the language was excluded from virtually all Gunwinyguan comparative studies (e.g. Alpher et al., 2003; Green, 2003b; Harvey, 2003; Verstraete, 2005).[2]

[1] See also Tu[illegible] 1974, 1999) for discussion of the mainland origin of some of the Groote Eylandt and Bickerton Island clans.

[2] Based on Van Egmond (2012), Evans (2017) and Harvey and Mailhammer (2017) also include Anindilyakwa in their discussion of the Gunwinyguan languages.

Van Egmond (2012) has attempted to show, however, that this somewhat mystical status is not fully deserved and that Anindilyakwa is typically Gunwinyguan in many ways.

AEH include the following languages in the Gunwinyguan family: Uwinymil, Warray, Jawoyn, Bininj Gun-Wok (Mayali, Kunwinjku, Kune), Dalabon, Rembarrnga, Mangarrayi, Ngalakgan, Ngandi, Wubuy, with Kungarakany and Kunbarlang being possible additional members. Uwinymil and Kungarakany are now extinct; most of the other languages (including Ngandi) are effectively moribund, while Wubuy is spoken by perhaps 60 speakers as their first language – the younger generation speaking mostly Kriol. Only Anindilyakwa and Bininj Gun-Wok are still being learnt by children as their first language.

We start by discussing several aspects of the phonological and morphological systems of Anindilyakwa, Wubuy and Ngandi which are relevant to reconstructing the ancestral sound system in §2: an outline of the phonologies of Ngandi, Wubuy and Anindilyakwa (§2.1), the nature of phonotactics in Wubuy and Anindilyakwa (§2.2), the nature of historical lenition in differentiating Wubuy and Anindilyakwa from proto-Gunwinyguan (§2.3), and the noun class prefix systems (§2.4). We then present our data and methodology in §3, and examine the shared core vocabulary in §4. Section 5 reconstructs the phoneme inventory of the ancestral language by applying the Comparative Method. Section 6 finishes the paper by summarizing the main findings and presenting conclusions.

Abbreviations and sources (unless otherwise noted) for the languages discussed here are:

proto-Gunwinyguan (pGN)	Alpher et al. (2003), Harvey (2003)
Ngandi (Ng)	Heath (1978b)
Wubuy (Wu)	Heath (1982, 1984)
Anindilyakwa (An)	Groote Eylandt Linguistics (1993), Leeding (1989), Van Egmond (2012), Waddy (1988), unpublished digital dictionary, first author's fieldnotes
Makassarese (Mkr)	Evans (1992)
Yolngu (Yol), Ritharrngu (Ri)	Zorc (1986)

2. Phonological and morphological issues relevant to reconstruction

2.1 Phoneme inventories

Tables 1–3 present the consonant inventories of Ngandi, Wubuy and Anindilyakwa, respectively. The Ngandi inventory is identical to the one reconstructed for pGN by Harvey (2003), except for the dental nasal, which occurs in Ngandi (though it is rare) but is not reconstructed for pGN. The lamino-dental stop /t̪/ is also absent

Table 1 PGN and Ngandi consonant phonemes (n̪ only occurs – though rarely – in Ngandi and is not reconstructed for pGN)

	Labial	Lamino-dental	Apico-alveolar	Apico-post-alveolar (retroflex)	Lamino-alveo-palatal	Dorso-velar	Glottal
Stop: fortis	pp	t̪t̪	tt	ʈʈ	cc	kk	ʔ
Stop: lenis	p	t̪	t	ʈ	c	k	
Nasal	m	(n̪)	n	ɳ	ɲ	ŋ	
Lateral			l	ɭ			
Tap			r				
Approx.	w			ɻ	j		

Table 2 Wubuy consonant phonemes (with rare phonemes in parentheses)

	Labial	Lamino-dental	Apico-alveolar	Apico-post-alveolar (retroflex)	Lamino-alveo-palatal	Dorso-velar
Stop	p	t̪	t	ʈ	c	k
Nasal	m	(n̪)	n	ɳ	ɲ	ŋ
Lateral		l̪	l	ɭ		
Tap			r			
Approx.	w			ɻ	j	

Table 3 Anindilyakwa consonant phonemes (with rare phonemes in parentheses)

	Labial	Lamino-dental	Apico-alveolar	Apico-post-alveolar (retroflex)	Lamino-alveo-palatal	Dorso-velar	Labialized velar	Labio-velar
Stop	p	t̪	(t)	ʈ	c	k	kw	kp
Nasal	m	(n̪)	n	ɳ	ɲ	ŋ	ŋw	ŋm
Prenasal	mp	n̪t̪	nt	ɳʈ	ɲc	ŋk	ŋkw	ŋp
Lateral		l̪	(l)	(ɭ)	ʎ			
Tap			r					
Approx.	w			ɻ	j			

in the other GN languages, save Wubuy and Anindilyakwa, but Harvey (2003) reconstructs it for pGN based on the divergent set of correspondences found in some reconstructed cognate sets of GN (i.e. /t/, /c/, /t̪/, /l̪/). Table 1 also shows a contrast between two series of stops (labelled 'fortis' and 'lenis') that has been reconstructed for pGN and preserved in Ngandi and other daughter languages. This stop contrast is absent in Wubuy and Anindilyakwa, which both only have a single stop series. Both languages also lack the glottal stop reconstructed for pGN and preserved in Ngandi. The sound that Anindilyakwa and Wubuy share and that is absent from all other GN languages (as well as other languages in the area) is the lamino-dental lateral /l̪/.

The Anindilyakwa consonant inventory differs from the neighbouring languages in a number of ways. Firstly, the alveolar stop and lateral are rare, as is the retroflex lateral.[3] Secondly, Anindilyakwa has a palatal lateral /ʎ/ absent in any of the neighbouring languages. Only the laminal laterals /l̪ ʎ/ occur with any frequency in Anindilyakwa, which is unusual in Australia.[4] And thirdly, Anindilyakwa has a series of contrastive labialized velar consonants /kw, ŋw, ŋkw/ plus a series of contrastive complex consonants, otherwise unattested in Australian languages (except perhaps Kugu-Nganhcara: Smith & Johnson, 2000). These complex segments consist of a simultaneous velar and labial constriction: there is a stop /kp/, a prenasalized stop /ŋp/ and nasal /ŋm/.

Tables 4–6 present the vowel inventories of the three languages. Ngandi (Table 4) has the five-vowel system typical of GN languages and reconstructed for pGN. Wubuy (Table 5) has a standard three-vowel system plus a length contrast for all three vowels. Van Egmond (2012) analyses Anindilyakwa (Table 6) with a highly unusual four-vowel system, with two contrastive front vowels and two central vowels, but no back vowels at all (the nature of the vowel system in Anindilyakwa, and indeed its phonology in general, has been the most controversial subject among researchers: c.f. Heath, n.d.;

[3]The Anindilyakwa standard orthography does not distinguish between dental and alveolar consonants, the orthographic symbol representing the most common sound: *d* = /t̪/, *n* = /n/ and *l* = /l̪/.

[4]As Hamilton (1996, p. 62) notes, Australian languages always have an apical lateral, if they have a lateral at all, and the presence of a laminal lateral implies the presence of an apical one. See §5.3.1 for further discussion.

Table 4 Proto-Gunwinyguan and Ngandi vowel phonemes

	Front	Central	Back
High	i		u
Mid	e		o
Low		a	

Table 5 Wubuy vowel phonemes (plus length)

	Front	Central	Back
High	i		u
Low		a	

Table 6 Anindilyakwa vowel phonemes

	Front	Central
High	i	
Mid	e	ə
Low		a

Leeding, 1989; Moody, 1954; Stokes, 1981; Van Egmond, 2012). The back vowel [u] that occurs at the surface can be predictably derived from either /i/ or /ə/ under the influence of neighbouring consonants with a [round] or [labial] feature, as in some analyses of Arandic languages (e.g. Eastern Arrernte: Henderson, 1998). That is, [u] only occurs preceding the labialized velars (e.g. *ṯu**k**^{w}a* 'maybe', *aku**ŋ**^{w}a* 'water') or /w/ (e.g. *ŋaj**uw**a* 'I'), or following phonetic dorso-velars for which an underlying representation as a labialized velar is possible (e.g. *aku**ŋu**-l̪aŋ^{w}a* 'water-ABL'; *cu**ku**-cu**k**^{w}a* 'RDP-chicken') or /w/ (e.g. *mama**wu**ɻa* 'sun'). We follow Van Egmond (2012) in assuming that [u] is not contrastive but obtains its rounding from contiguous rounded velars. Consequently, the An reflexes of *u in the proto-language are fully predictable, as we will see in §5.6.

2.2 Phonotactics

Wubuy phonotactics in roots are not significantly different from other GN languages such as Ngandi. In particular, a wide range of root-internal heterorganic clusters occur (see Heath, 1984, p. 20), though there are no clusters involving velars (/k/ or /ŋ/) as the first member (Heath, 1984, p. 21), and clusters of this kind arising at morpheme boundaries are subject to a rule deleting /k/ before other consonants and assimilation of /ŋ/ to the point of articulation of the following stop (Heath, 1984, pp. 70–73). Words can end in a number of consonantal segments (Heath, 1984, p. 19) – all apicals, lamino-palatals and velars, and all vowels.

Anindilyakwa has a radically different phonotactic system from any nearby language. It appears to have reshaped the phonology of the ancestral language, according to two constraints: a constraint against coda consonants, and a constraint militating against morpheme boundaries which do not align with syllable boundaries. These constraints have had the following effects (Van Egmond, 2012):

- a[redacted]ords are vowel-final at the surface (the final vowel is always [a]);

- heterorganic clusters are heavily restricted and many of these clusters in the proto-language have been eliminated;
- others have become complex segments: both homorganic NC clusters (mp, nt̪, …) and labio-velar complex segments (kp, ŋp, ŋm) are a result of this process; and
- epenthesis is an option for all remaining clusters, and at morpheme boundaries where C–C.

The only permitted clusters are apical liquids and nasals plus non-coronal nasals and stops or laminal stops. The velar+labial sequences are analyzed as complex segments in Van Egmond (2012) (as mentioned, such clusters are completely absent in Wubuy). Unlike true heterorganic clusters, these are never broken up by epenthesis or separated in reduplication (e.g. *ere**kp**-ere**kp**ici-* 'RDP-spit', *-i**ŋp**-i**ŋp**əṯa* 'RDP-strong'), whereas other heterorganic consonant clusters may be (e.g. *-mə**r**ə-mə**rk**+paɤ́a* 'RDP-soft'). Moreover, speakers syllabify labio-velar segments as onsets (e.g. *ma.mə.**ŋp**a* 'hair', *a.l̪a.**kp**a* 'leg', *ma.**ŋm**a* 'brain'), whereas they pronounce other consonant clusters hetero-syllabically (e.g. *a.mə**ŋ.ṯ**a* 'shoulder', *me**n.p**a* 'eye').

The labio-velar segment /kp/ appears to have several potential sources (Table 7). In some cases, the complex segment corresponds to a cluster in Wu (e.g. 'blame') or in Ngandi (e.g. 'cut-leaved palm', which has been simplified in Wu according to the /k/-deletion rule mentioned above). In others, the complex labio-velar stop segment /kp/ corresponds to a reconstructed fortis labial stop in pGN (e.g. 'plain, flat area'), to a morpheme boundary (e.g. 'Grewia orientalis'), or to a long vowel in Wu followed by /p/ (e.g. 'pelican').

The /ŋp/ cluster corresponds to /mp/ in Wu according to the /ŋ/-assimilation rule in Wu mentioned above: e.g. Wu **l̪a**ŋ**+**p**atca- > l̪a**m**+**p**atca-* : An *ɤ́a**ŋ**+**p**atca-* (pronounced *ɤ́a.**ŋp**at.ca*) 'head+hit'. There are no correspondences involving the cluster /ŋm/.

Moreover, the coronal contrasts in Anindilyakwa are much more weakly attested than in Wubuy (which maintains a four-way coronal contrast, even word-initially: Heath, 1984). According to Van Egmond (2012), retroflex consonants (apart from /ɻ/) do not occur root-initially. All coronals, including the prenasals, can occur in syllable onset (though /l/ and /l̪/ are rare and almost entirely confined to coda position and loanwords). Hence the contrast between apical and laminal laterals in onsets is weakly attested (which is rather unusual in Australia; see §5.3.1 for further discussion). Only apical sonorants can occur in syllable coda position (/l/, /r/, /ɻ/, /n/, /ɳ/). It appears that the apical contrast is largely neutralized in Anindilyakwa here, resulting in alveolar realizations of laterals corresponding to retroflex laterals in Wubuy (e.g. Wu *mu**l̪**k*ʷ*a* in *mu**l**k*ʷ*a* 'belly'; see §5.3). In codas, neither Wubuy nor Anindilyakwa allow lamino-dentals (in common with other Australian languages: Hamilton, 1996).

Table 7 Correspondences of Anindilyakwa /kp/

	Anindilyakwa	Wubuy	Ngandi	Others
BLAME	*l̪e**kp**a-*	*l̪a**cp**u-*		
CUT-LEAVED PALM	*ji+ŋa**kp**arŋa*	*ŋa**p**anŋa*	*ŋa**kp**anŋa*	Ri *ŋa**kp**anŋa*
NOW	*aɻa**kp**a*	*aɻa**p**a*		
PLAIN	*a**kp**al-*	*a**p**aɭa*		pGN **ka**pp**al*
OLD	*-enuŋkuɻa**kp**a*	*nuŋkuɻaa**p**a*		
Grewia orientalis	*ma+**kp**ija*	*ma**p**uju*	*(ma-)**p**uju?*	Ri ***p**uju?*
PELICAN	*ṯə+ma**kp**əl̪a*	*maa**p**ul̪u*		

Table 8 Anindilyakwa final /a/ epenthesis in sonorant-final roots

Gloss	Anindilyakwa	Wubuy	Ngandi	pGN / Yolngu
MANGROVE TREE	*ji+l̪erkəra*	*l̪alkur*	*t̪alkkurk*	Ri *t̪alkkurk*
BACKBONE	*məraŋa*	*muurŋ* 'shell'	*murŋ*	
WOOMERA	*ji+maŋala*	*maŋalŋ*		pGN **maŋal,* Yol *maŋalʔ*
HIGH	*karawaɻa*	*arwaɻ*		Yol *karwaɻ*

Table 9 Deletion of word-final stops in Anindilyakwa

Gloss	Anindilyakwa	Wubuy	Ngandi	pGN / Yolngu
SKIN	*[illegible]+kuɕa*	*makuɭa**k***	*kuɭa**ʔ***	pGN **kuɭa**k***, Ri *kuɭa**ʔ***
BLOODWOOD	*a+l̪amiɕa*	*l̪umuɭu**k***	*t̪umuɭu**ʔ***	Yol *t̪umuɭu**ʔ***
DRY, HOT	*aɻətara*	*ɻaɻar**k***		
LIVINGSTONE PALM	*ji+l̪erpa*	*(jii-)l̪alpi**c***	*(ku-)t̪alppi**ʔ***	Ri *t̪alppi**ʔ***
THIN	*-jarmijarma*	*carmajarma**c*** 'long, tall'		

Finally, the fact that all words end in /a/ in Anindilyakwa probably is an innovation due to the general tendency of the language to avoid codas. It appears that word-final /a/ has been attained by both deletion or epenthesis, depending on whether the word-final segment in the proto-language was a stop or a sonorant.[5] For cognates ending in a sonorant, the Anindilyakwa form usually involves a final epenthetic vowel /a/ (Table 8).[6] This strategy does not apply to bound nominal roots in An, which fail to show epenthesis (e.g. *l̪əkar-* 'track' Table 14, *ɕaŋ-* 'head' Table 26, below), presumably because these can never occur word finally and therefore do not conflict with the constraint against word-final consonants.

In stop-final roots, however, it appears that deletion of the final stop was the favoured strategy (Table 9). It may be noted that most of these examples involve final /k/. Note that in the two cases where we find a cognate in Ngandi, there is also replacement of final /k/ with /ʔ/. Word-final /k/ is unstable in Wubuy also; some words vary between forms with final /k/ and forms without (e.g. Wu *maraaɻa(k)* 'egret', *maaŋa(k)* 'heron'). This strategy also does not apply to bound nominal roots in An, where final stops are maintained (e.g. *calk-* 'ground' Table 15, below), again presumably because these can never occur word finally.

Since An word-final /a/ is a language-specific innovation, we reconstruct the Wu endings for the ancestral language.

2.3 Lenition

In this section we discuss the thorny problem of historical lenition, which has led to rather drastic differences between the sound systems of Wubuy and Anindilyakwa compared to other Gunwinyguan languages. We argue that lenition was a sound change which distinguished proto-Insular Gunwinyguan from the rest of the Gunwinyguan bloc.

In Ngandi and other GN languages, there are two series of stops, variously labelled 'fortis/lenis', 'geminate/singleton' or 'long/short', depending on the analysis. The acoustic

[5]We thank an anonymous reviewer for drawing our attention to this pattern.
[6]However, 'boil' **ɻulul* > *ji+ɻəla* is an exception to this. The Anindilyakwa dictionary notes this form as a loan from Wubuy. This may indicate a change in strategy over time.

Table 10 Operation of classic lenition chain shift in Ngandi and Wubuy (boldface marks corresponding segments; source: Harvey, 2012)

Ngandi	Wubuy	Gloss	Correspondences	Word class
*meɭe**pp**eʔ*	*maɭi**p**i*	shoulder blade	pp : p	body part
***p**urk-t̪a-*	***w**urt̪a-*	bury	p : w	verb
*ɻa**tt**ar*	*ɻa**t̪**ar*	harpoon	tt : t̪	noun
***t̪**a*	***l̪**a-*	mouth	t̪ : l̪	body part
***ʈ**a-**p**u-*	*ɻ**a**-**w**u-*	to test, to try out	ʈ : ɻ, p : w	verb
***k**ulk-t̪u-*	***w**ult̪a-*	to cut, to cut through	k : w	verb
***k**uc-ka-*	*icka-*	to go hunting with dogs	k : Ø	verb

Table 11 Reflexes of systematic correspondence sets in Wubuy and Ngandi (Harvey, 2003, p. 211; Heath, 1978a, pp. 37–41)

Ngandi	*pp*	*p*	*tt*	*t̪*	*tt*	*tt*	*ʈ*	*cc*	*c*	*kk*	*k*
Wubuy	*p*	*w*	*t̪*	*l̪*	*t*	*ʈ*	*ɻ*	*c*	*j*	*k*	*w*

Table 12 Probable retentions of stops in Wubuy (source: Harvey, 2012)

Ngandi	Wubuy	Gloss	Correspondences	Word class
ʈikku	*ʈiku-*	raw, uncooked	ʈ : ʈ	adjective
*ŋaʈʈu**k**u*	*ŋaʈu**k**u*	rope	k : k	artifact
*c**a**laʈi?*	*c**a**laʈi*	vine sp.	c : c, ʈ : ʈ	natural species
*c**ur**kupattu*	*c**ur**kupatu*	goanna sp.	c c, k : k, p : p	natural species
*p**uʈu**k**a*	*p**uʈu**k**a*	shrub sp.	p : p, ʈ : ʈ, k : k	natural species
*c**i**r?*	*c**i**r-*	hair; fur, body hair	c : c	body part
*k**a**laŋ*	*k**ak**alaŋ*	egg	k : k	body part

correlates differentiating these stops appear to be primarily duration (see e.g. Baker (2008) on Ngalakgan, the closest GN language to Wubuy and Anindilyakwa which has been analyzed acoustically). Based on cognates such as those shown in Table 10, Heath (1978a) proposes that in the process of Wubuy differentiating from the proto-language which was the ancestor of Ngandi and Wubuy, Wubuy underwent a chain shift, as schematized in Table 11, whereby the proto-language fortis stops are realized as simple stops in Wubuy, and proto-language lenis stops are realized as approximants at the same place of articulation, or zero, in both initial and medial environments (between non-nasal sonorants and vowels). The assumption is that Ngandi preserves the original system intact.[7]

However, the outcome of lenition in Wubuy is far from straightforward. Table 12 shows examples of probable cognates in Ngandi and Wubuy where a lenis or initial stop in Ngandi matches to a stop in Wubuy – in other words, an apparent 'failure' of lenition. Some of these are possible later loans, after the lenition rule ceased to operate in Wubuy. However, items such as *cir-* 'hair', which only exist as bound roots in Wubuy, are unlikely to be borrowings.[8] However, we have no straightforward way of demonstrating inheritance rather than borrowing in any case. The presence of lenition cannot be used as evidence for retention because of the numerous loans which are demonstrably

[7]Although hinting that Anindilyakwa might be closely related to Wubuy and Ngandi, Heath (1978a) leaves this question open a[illegible]es not include Anindilyakwa in his analysis.

[8]Howeve[illegible] is not ironclad evidence against borrowing; Wu bound modifier *ʈiku-* 'raw', which also has lenited forms *ɻiku-* and *ɻuku-* in Wu, the latter agreeing with the An adjective *aɻuk^{w}a* 'raw', appears to be a borrowing from Yolngu *ʈiku* 'raw', since the latter is not restricted to a particular variety of Yolngu according to Zorc (1986).

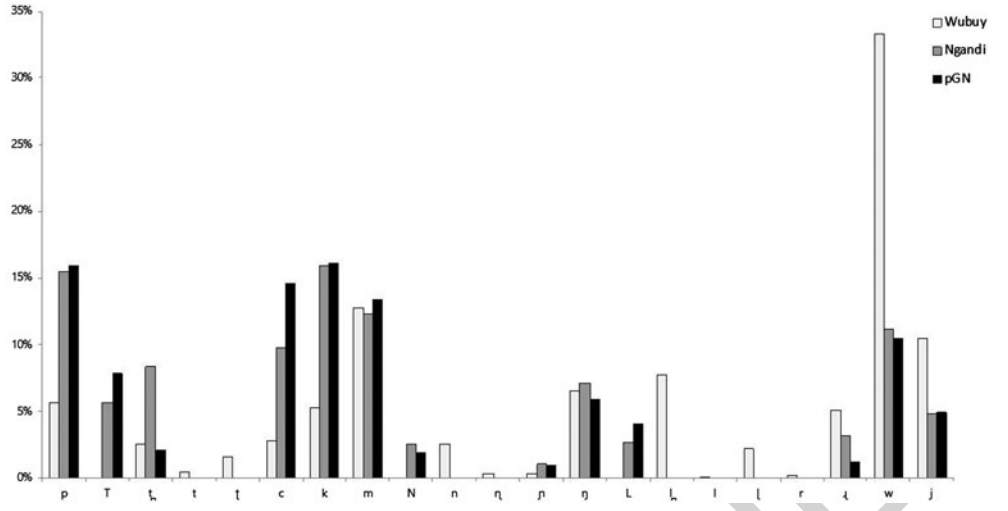

Figure 2 Frequencies of initial segments in Wubuy, Ngandi and Proto-GunwinyguanSources: Wubuy – Bundgaard-Nielsen and Baker (2014); Ngandi – Heath (1978b); pGN roots – Harvey (2003).
Note: T, N, L represent neutralized initial apical segments.

from neighbouring, unrelated Yolngu varieties which have undergone lenition in both Wubuy and Anindilyakwa (see Table 14).

We can gain an indirect measure of the degree of lenition in Wubuy by comparing the relative frequencies of different segments in initial position with the proportions found in the reconstructed pGN lexicon of 1,409 roots (Harvey, 2003), and those found in the synchronic lexicon of Ngandi, compared in Figure 2. All things being equal, we would expect an even distribution of segments in initial position. In reconstructed pGN, there are 13 possible segments in initial position, and so the predicted proportion would be around 8% each. The relative proportions of initial peripheral and laminal stops have roughly double the predicted frequencies (around 16%), a pattern typical of the frequencies of root-initial segments in Australian languages in general (Hamilton, 1996, p. 218). The synchronic lexicon of Ngandi preserves this distribution almost exactly, the biggest difference being the relative proportions of the two laminal stops.

The distribution of initial segments in Wubuy is quite different. There is a significant over-representation of roots beginning in /w/, and in other non-nasal sonorants. There are 16 possible initial consonantal segments (ignoring vowel-initial roots), so the predicted proportion for each is about 6%. The segment /w/ therefore has nearly six times the predicted proportion of the lexicon.[9] Roots beginning in stop segments constitute a correspondingly much smaller proportion than in pGN. Since, by hypothesis, pGN initial *p and *k both lenited to /w/ in Wu, we have an explanation for the difference in the relative proportions of these segments in the Wubuy lexicon compared to the pGN lexicon, largely preserved in Ngandi.

[9]In Hamilton's (1996, p. 218) survey of word-initial frequencies, there are no languages reported with this kind of distribution. The language with the highest frequency of word-initial /w/ is Yirr-Yorront, with a mere 13%. Similarly, there are no languages with such low frequencies of initial velar and labial stops as we find in Wubuy. In other languages, the range for these stops in initial position is 14%–25%. The distribution of initial segments in Wubuy is more similar to that in the Atampaya dialect of Uradhi (Crowley, 1983, p. 319), a language which also underwent historical initial lenition.

Table 13 Cognates showing lenition of initial stops in Wubuy and Anindilyakwa

Gloss	pGN	Ngandi	Anindilyakwa	Wubuy	Corr.
BLOODWOOD	*ṯumuḻuk	ṯumuḻuʔ	a+ḻamiɕa	ḻumuḻuk	ṯ : ḻ
Planchonia careya	*ṯaŋkiʔ 'tree sp.'	ṯaŋkiʔ	ḻaŋ(k)- 'long tall standing things, e.g. trees'	ḻaŋki	ṯ : ḻ
UPPER LEG	*ṯar	ṯar+ppic	ḻaṇṯar-	ḻar+pic	ṯ : ḻ
OLD PERSON		ŋacciwaŋku (FEM only)	-juwaŋkʷa	-jiwaŋku	c : j
SNEAK UP ON		kalkka-	walka-	waalka-	k : w
BITE	*pa-	pa(ŋa)-	(w)aŋa-	wa(ŋa)-	p : w
CHEST	*pere	pere	weri-		p : w

Table 14 Roots shared with Yolngu and Ngandi showing lenition of initial stops

	Yolngu	Ngandi	Anindilyakwa	Wubuy	
EEL-TAILED CATFISH	kaŋŋal	kaŋŋal	aŋalpija	(w)aŋŋalpija	k : An Ø : Wu w ~ Ø
CANOE	ṯamukku		a+ḻamukʷa	ḻamuku	ṯ : ḻ
LIVINGSTONE PALM	ṯalppiʔ	ṯalppiʔ	ji+ḻerpa	ḻalpic	ṯ : ḻ
MANGROVE TREE	Ri ṯalkkurk	ṯalkkurk	ji+ḻerkəra	ḻalkur	ṯ : ḻ
(1) PIPE; (2) Excoecaria parvifolia	Ri cilaraʔ (2)	cilaraʔ (2)	jiɕara (1, 2) (Scaevola sp.)	jilaari (1, 2) (Scaevola sp.)	c : j
WATER PYTHON	Ri kaḻcca		[illegible]lca	aḻca	k : An w : Wu Ø
TRACK, CROSSING	Ri ṯakar		[illegible]	ḻa(w)ar	ṯ : ḻ

Phonotactics in Anindilyakwa are complicated by the fact that all roots are obligatorily prefixed in Anindilyakwa (Van Egmond, 2012, p. 26). Nominals begin in one of the set of segments {a/e m n ŋ w j k ṯ}, and verbs take a similar, constrained set of segments, in both cases representing the initial segments of prefixes. Therefore, the root-initial position is not word-initial in Anindilyakwa most of the time (and the root-initial position in nouns is not identified in the Anindilyakwa dictionary). Secondly, the prefix–root boundary in verbs is difficult to identify in many cases, in particular in cases where the root appears to begin in a vowel. These factors militate against a straightforward assessment of root-initial phonotactics in Anindilyakwa. Nevertheless, if we accept the premise that Anindilyakwa is related to Wubuy (and thence to other GN languages), then the high frequency of the lateral /ḻ/ in Anindilyakwa argues for the operation of historical lenition from *ṯ, its only likely source apart from borrowing from Wubuy.

Table 13 shows the few cognates in Anindilyakwa and Wubuy with reconstructed initial stops from a source in pGN, and Table 14 shows roots which are shared with Yolngu. Both languages agree in showing lenitions of reconstructed (or attested) initial stops. Verbs with initial stops are relatively rare in pGN, and those which are shared by An and Wu are even fewer. The 'bite' and 'sneak up on' verbs (Table 13), which are cognate in all three languages, are lenited in both An and Wu.[10] The first three of the Yolngu roots listed in Table 14 are found in all, or a significant majority of, Yolngu varieties (according to Zorc, 1986), not just those which neighbour Wubuy and/or Anindilyakwa. This would strongly suggest that the direction of borrowing was from (a variety of) Yolngu into the latter languages. The others are

[10]See Van Egmond (2012, Ch. 9) for comparison of the inflectional paradigms of the verbs discussed here.

Table 15 Retention of [illegible] stops in Anindilyakwa and/or Wubuy

Gloss	pGN	Ngandi	Anindilyakwa	Wubuy	Corr.
MEAT	*ṯaŋku	ṯaŋku	-ma+ṯaŋkʷa	l̪aŋku	ṯ : An ṯ : Wu l̪
SKIN	*kuɭak	kuɭaʔ	-ma+kuʃa	-ma+kuɭak	k : k
MoMo	*ka(k)kak	kokkok	kaku	kaaku	k : k
OWL		(kurʔkur)	ju+kurkʷa	jikurku	k : k
GROUND	*colkko	colkko	calk-		c : c
INCHOATIVE	*+ṯi-	+ṯṯi-	+ṯə-	+ṯi-	ṯṯ : ṯ
THEMATIC	*ṯa-	+ṯa-	+ṯa-	+ṯa-	ṯ : ṯ
THEMATIC	*ka- 'carry'	ka- 'carry'	+ka-	+ka-	k : k
HIT	*patca- Yolngu	pacca-	+patca-	+watca-	p : An p : Wu w
RAW	ɭikku	ɭikku	aɻukʷa	ɭiku ~ ɻiku- ~ ɻuku-	ɭ : An ɻ : Wu ɭ~ɻ
SPEAR BUSH	Ri waɭawaɭa	waɭawaɭa	a+waɻuwaɻa	waɭawaɭa	ɻ : Wu ɭ
SEAGULL	carak	carak	ji+cara	carak	
MORNING STAR	ṯaampuɭ	ṯampul	ṯampul	ṯaampuɭk	ṯ : ṯ

found just in Ritharrngu, which neighbours Ngandi and Wubuy, but not Anindilyakwa. Given that the Wubuy and Anindilyakwa forms show lenition, but the Ritharrngu forms do not, this suggests that the roots are GN or pIGN, and that borrowing was from Ngandi or a similar non-leniting NPN language into Ritharrngu, or else the borrowing pre-dates lenition in pIGN, Wu or An. Many of these roots are also found in other NPN languages of the area, such as Ngalakgan. In two cases, Ri *kaɭcca* 'water python' and *ṯakar* 'crossing', the root is not found in Ngandi (which uses the term *puɻurcci* for 'water python', a term also used by Ritharrngu and Ngalakgan speakers, the word for 'crossing' in Ngandi is unknown).

Nevertheless, there are exceptions to the generalization that pGN cognates and Yolngu borrowings undergo lenition (Table 15), and Wubuy and Anindilyakwa do not always agree in whether a given reflex is lenited or not. Furthermore, the disagreement can be in either direction (lenited in Wu, unlenited in An, as with the 'meat' form, or the reverse, as with 'spear bush'). If we confine our attention to the finite verbs inherited from a higher ancestor (perhaps pGN), we can observe that these are, apart from Wu +*watca*- 'hit', unlenited in Wu and An. However, these verbs occur only in bound positions in both languages, i.e. following another lexical root. Thematic +*ṯa*- for instance only occurs in frozen complex stems (e.g. Wu, An *l̪ar+ka*- 'send'; Wu *ŋur+ka*-, An *ŋur+kʷa*- 'hunt'), which may descend from pGN **ka*- 'carry' (Heath, 1984, p. 419; Van Egmond, 2012, p. 349). The 'hit' verb only occurs in compounds, taking incorporated nouns or coverbs, e.g. Wu *l̪am+patca*-, An *ʃaŋ+patca*- 'head+hit' (where Wu initial /w/ has hardened to /p/, see below). Moreover, both languages use the meaningless finite verb +*ṯa*- to act as a host for tense–aspect–mood inflection for borrowed verbs (as from Kriol and English), as well as some native coverbs, found in examples such as An *pejiṉ+ṯa*-, Wu *pajiṉ+ṯa*- 'buy' (< Eng *buy* via Kriol *bayim*). The verb also occurs in Ngandi, as a bound verb taking coverbs, though it is no longer productive (the similarly meaningless finite verb+*ṯu*- is used instead for loanverbs; this verb is not found in either An or Wu but is reconstructible to pGN). The bound finite verb root+*ṯi*- (Wu), +*ṯə*- (An) derives intransitive verb stems from nominals/adjectives in An, e.g. *eni-juwaŋku+ṯə*- 'become old (for men)', from *n-eni-juwaŋkʷa* 'old man', including recent loanwords. This root is unproductive in Wu, but occurs in a few verb stems (Heath, 1984, p. 398), including *jiwaŋku+ṯi*- 'become old' (from *jiwaŋku* 'old man'). The lack of lenition in this

Table 16 Ngandi, Anindilyakwa and Wubuy noun class prefixes to nominal constituents

	Ngandi	Anindilyakwa	Wubuy (Oblique)
MASC	*ŋi-*	*ji-*	*jii-* ∼ *jiri-*
FEM	*ŋa-*	*t̪ə-*	*jii-* ∼ *jiri-*
COLL/PL	*pa-*	*wurə-*	*waa-* / __C ∼ *wara-* / __V
VEG	*ma-*	*m(a)-*	*ama-*
NEUT	*ku-*	*a-*	*a-*
ANIM	*a-*	*a-*	*a-*

and other bound finite verbs may be related to the fact that the cognate verb form has an initial fortis stop in Ngandi. Historical fortition of the initial stops of bound verb roots is a sporadic phenomenon in Ngandi, Ngalakgan and Rembarrnga.

From Table 15 we can infer that it is unlikely that lenition (and fortition) was a uniform process differentiating pIGN from pGN, and inherited by each daughter language. Rather, we assume (as seems reasonable) that pIGN was a language where lexemes and speakers varied in the presence and frequency of lenition. The descriptions of Iwaidja (Shaw et al., 2020) and Gaagudju (Harvey, 2002, pp. 28–43) provide accounts of exactly this kind, where there is both inter- and intra-speaker variation in the application of lenition, and furthermore, lexeme-specific differences in the frequency or presence of lenition. We therefore attribute the synchronic differences in the application of historical lenition to historical variation of this kind in the ancestor language from which the daughters sourced their lexicons.

To summarize, lenition in both Anindilyakwa and Wubuy has apparently been a long and complex process, with somewhat unpredictable outcomes, as argued for other nearby languages by Harvey (2012). Mailhammer and Harvey (2018) assume multiple waves of lenition for Iwaidjan. The evidence is however too scant to enable us to draw any firm conclusions about the relative timing of lenition in Wu and An with respect to the putative ancestor pIGN, except that it must have followed initial borrowings from neighbouring Yolngu varieties, but had ceased by the time of Macassan contact in the eighteenth century.[11]

2.4 Word structure

Finally, in this background section, we briefly discuss the nature of noun class prefixation in the languages in question, since this is relevant to the reconstruction of nouns. Wubuy, Ngandi and Anindilyakwa all have a noun class system of around five–six classes (Table 16).[12] The class membership of nouns is reflected in agreement on verbs, demonstratives and other constituents of clauses in all three languages. Unlike many other languages of Arnhem Land (e.g. Bininj Gun-Wok, Burarra, Alawa), all nouns also take an overt noun class prefix.

[11]Given that Makassarese and English loans consistently fail to lenite (e.g. An *t̪ɛripa*, Wu *t̪aaripa* 'trepang' < Mkr *taripaŋ*; An *puciketa*, Wu *(a-)pucikan* < *pussycat*), lenition is not a synchronically active process in either language.

[12]We use the Wu Oblique series (what Heath (1984) labels the 'Punctual' forms) in this paper, as these are formally most similar to the An class prefixes. Apart from being the only GN languages to show agreement for all noun classes on the verb, these three are also the only GN languages to have extended the use of a 3pl pronoun as a plural prefix for humans (also found in neighbouring Warndarrang, Marra and Alawa), and (in An and Wu) as an additional non-human noun class. While these patterns are suggestive, we do not have the space here to investigate the noun class systems in detail.

Table 17 An and Wu noun roots belonging to distinct noun classes

	Anindilyakwa	Wubuy
MUDWHELK	*ji+l̪aɻa* (MASC)	*(waa-)l̪aaɻu* (COLL)
SALMON	*mu+walŋaɻa* (VEG)	*(a-)walŋaɻi* (NEUT)
ROCK WALLABY	*t̪ə+l̪anta* (FEM)	*(a-)l̪aanta* (NEUT)
RIDLEY TURTLE	*ji+ciɻakamərа* (MASC)	*(a-)caɻakaamiri* (NEUT)

However, the status of the associated prefixal morphology differs. In both Ngandi and Wubuy, noun class prefixes are omissible, and nouns in citation form (as well as other contexts) frequently lack noun class prefixes (Heath, 1978b, p. 35, 1984, p. 160). In Anindilyakwa, on the other hand, with the exception of some nouns denoting human referents, the erstwhile noun class prefixes are frozen components of nouns, and cannot be omitted or replaced by other prefixes (Van Egmond, 2012, Ch. 3). The frozen status of noun class morphology becomes evident from the phonological classification of loanwords. Makassarese and English loans do not take an overt noun class prefix but may be classified based on the initial segment of the noun root (e.g. *t̪ə̪mpala* 'cloth, sail' (< Mkr *sombalaʔ*) and *t̪eripa* 'trepang' (< Mkr *taripaŋ*) are assigned feminine class due to their initial /t̪/, as are the English loans *t̪ə̪ɻacica* 'trousers' and *t̪aŋki* 'donkey').[13]

Evidence that the An class prefixes were formerly more flexible comes from cognate noun roots that belong to different noun classes, represented by distinct class prefixes (Table 17).

These cognates show that the class prefixes historically must have been a separate constituent from the noun root, as they still are in Wubuy.[14]

In citing Anindilyakwa noun roots in sets of cognates, we separate the (inferred) noun class prefix with the plus symbol '+' to indicate a lexicalized morphological boundary. We infer that in Anindilyakwa the noun class morphology was formerly more flexible, as it is in other GN languages, and reconstruct a system with an omissible noun class prefix to pIGN.

3. Data and method

The foundation for demonstrating genetic relatedness is the uncovering of systematic sound correspondences between words with similar meaning through the Comparative Method (Campbell & Poser, 2008). Our method can be divided into two general steps: first, we composed a 150-item Swadesh list (Alpher & Nash, 1999) of Anindilyakwa, Wubuy and Ngandi basic 'core' vocabulary from the dictionaries listed above (Appendix 1). The Swadesh list includes body part nouns, verbs, adjectival nominals, terms for common natural phenomena, manufactured items and basic human classification terms. This core vocabulary is considered to be more resistant to borrowing than, for instance, nouns belonging to the domain of ceremonies, natural species and some material objects (e.g. Baker, 2004; Harvey, 2003; Hock & Joseph, 1996, p. 257; Swadesh, 1952, 1956; Tadmor et al., 2010).

[13]In Wubuy and Ngandi, by contrast, loaned nouns are overtly prefixed for class: Wu *ana-pipa* 'NEUT-paper', Ng *ku*-bottle 'NEUT-bottle' (c.f. An *cura* 'paper' (NEUT) < Mkr *surat*, *puṯiɻa* 'bottle' (VEG)).

[14]Further evidence that the An prefixes were formerly flexible comes from nouns incorporated into verbs or adjectives, which do not take a class prefix: e.g. *l̪akpa-* 'leg' (< *a+l̪akpa* 'NEUT-leg'), *ŋampa-* 'groin' (< *ji+ŋampa* 'MASC+groin').

Second, a larger corpus of 244 potential cognate sets across the three languages was compiled (Appendix 2). The basis of this corpus was the thesaurus in Heath's Wubuy dictionary (1982, pp. 309–326), which includes flora and fauna terms, body parts and parts of objects, manufactured artifacts, nautical terms, and terms for the physical environment. For all these terms, the Anindilyakwa and Ngandi dictionaries were searched for potential cognate forms, as was Harvey's (2003) list of reconstructed pGN items. These word lists were used to establish regular sound correspondences between the three languages, with the aim to reconstruct the phoneme system of the proto-language and any sound changes that have taken place in the daughter languages.

For the reconstruction of the proto-phoneme inventory, we extracted forms with attestations in Anindilyakwa, Wubuy and Ngandi from this corpus, following the methodology proposed by Mailhammer and Harvey (2018), who propose a framework for evaluating reconstructions and etymologies objectively. Their method involves two basic factors that influence the quality of etymologies: (i) the regularity of the correspondences the reconstructions are based on, and (ii) the number of cognate sets supporting those correspondences. Given that the Comparative Method is probabilistic, they claim that:

> The more sets of cognates that support a given correspondence we can find, the more strongly our correspondence is supported, and the more certain it is that the best explanation for the sound correspondence is in fact inheritance as opposed to language contact or chance. This crucially depends on establishing that the correspondence is systematic, i.e. regular. (Mailhammer & Harvey, 2018, pp. 335–336)

To determine the degree of regularity between correspondences in quantitative terms, Mailhammer and Harvey propose a Regularity Coefficient (RC), defined as follows:

$$\mathrm{RC} = \frac{\text{instances supporting a pattern for one corresponding segment}}{\text{number of all correspondences of the relevant segment}}$$

For example, there are 56 cognate sets in our corpus that involve altogether 64 correspondence sets with /p/ (as one word may contain more than one instance of /p/). Hence 'number of all correspondences of the relevant segment' = 64. In all cases, Anindilyakwa /p/ corresponds with Wubuy /p/ (= 'instances supporting a pattern for one corresponding segment'), so RC = 1. That means in probabilistic terms, there is a 100% chance that An /p/ and Wu /p/ match. By contrast, there are 13 correspondence sets with /t/, but /t/ only matches in five cases (RC = 5/13 = 0.38). In the remaining eight instances, An /t/ corresponds to a different coronal stop in Wu, or to /r/. There are no apparent conditioning factors for these different sound correspondences, apart from the fact that /t/ is a marginal phoneme in An. So, there is a 38% chance that the alveolar stops match in the two languages. Mailhammer and Harvey use this RC as a means of objectively evaluating correspondences, because "a higher RC gives greater confidence in a correspondence to be able to rule out chance or borrowing as the reason for its existence" (2018, p. 336).

The second factor that influences the quality of etymologies, and thus the confidence we can place in correspondences being due to inheritance rather than chance or borrowing, is the number of cognate sets that support a given correspondence. The correspondence An /p/ – Wu /p/ is 100% regular and supported by 64 correspondence sets. By contrast, the correspondence An /ṉ/ – Wu /ṉ/ is also 100% regular but supported by

Table 18 Categories of etymological quality based on support for cognacy (source: Mailhammer & Harvey, 2018, p. 337)

Category	Cognacy	Criteria
1	Excellent support	Exact formal and semantic match in proposed cognates. All changes conform to regular patterns. RC = 1 for all segments in the entire stem
2	Good support	Exact formal and semantic match in only the roots of proposed cognates. RC = 1 for the root; remaining part of the word is transparent
3	Moderate support	Cognacy requires assumption of not fully regular formal changes (RC = 0.5–1); proposed semantic changes have close parallels or are common
4	Tentative support	Connection rests on *ad hoc* assumptions and correspondences are irregular (RC < 0.5)

only six correspondence sets. Clearly, the latter pattern is less reliable than the former, as the high 'regularity' could easily be due to the small dataset. In what follows, the number of cognate sets supporting a correspondence set is provided with each sound correspondence. Moreover, the etymological quality of cognate sets depends on the degree of support for cognacy, which Mailhammer and Harvey (2018) divide into four categories, as in Table 18.

In Category 1, all changes are regular: they conform to sound laws and the semantics of the compared lexical items are the same. For example, An *mulk^{w}a* ~ *muḻku* 'belly' is an instance of Category 1 cognacy: the correspondences An /l/ ~ /ḻ/ /__. and An /k^{w}a/ ~ /ku/ /__# are fully regular (RC = 1 and 0.96, resp.), and the meanings of the words are identical. Therefore, the reconstruction of pIGN **muḻku* 'belly' based on regular sound laws has excellent support.

Category 2 involves cases where the root matches but not the entire stem. For example, the An verb *kiru+waɲci-* 'emit smell' is a synchronically frozen compound structure. The corresponding verb in Wu *waɲca-* ∼ *paɲca-* 'emit smell' is clearly cognate but lacks the *kiru*+element present in An. Thus, reconstruction of pIGN **waɲca-* 'emit smell', which continues in An only as a compound (and which in Wu synchronically hardens to *paɲca-* after a stop/nasal), is very plausible and has good support.

Category 3 reconstructions involve additional assumptions that follow a less regular pattern. For example, the An verb *ṯarpə-* corresponds to Wu *ṯaṯarpi-*, both meaning 'to move'. The Wu verb shows reduplication of the first syllable /ṯa/, which is absent in An (the An /ə/ ~ /i/ correspondence is regular). Reduplication is common in both An and Wu, indicating repetition or prolongation. But inherent reduplication, where there is copying of form but no additional semantic component (as in the case of *ṯaṯarpi*), is also common among verbs (Heath, 1984, p. 342). As a result, the reconstruction of pIGN **ṯarpi-* is plausible but only moderately supported, given the reduplication in Wu that is absent in An.

Category 4 consists of reconstructions that only have tentative support and require *ad hoc* assumptions, such as formal changes that are irregular, or unexplained semantic differences. An example is An *ṯu+waʎa*, Wu *(jii-)wuwaḻuḻu*, Ng *(a-)kuwaḻuḻu*, pGN **kuwaḻu* 'curlew'. Evidently, the Wu and Ng forms are cognate and descend from pGN. The An word, however, is two syllables shorter than the Wu and Ng words, and an explanation for the presumably absent syllables rests on *ad hoc* assumptions, as it assumes either an unmotivated loss in An, or an addition in Wu and Ng. Consequently, the reconstruction of 'curlew' for pIGN only has tentative support.

Table 19 Cognate densities of Anindilyakwa, Wubuy and Ngandi

	Ngandi	Wubuy
Anindilyakwa	17.3%	34.4%
Wubuy	30%	

Following Mailhammer and Harvey (2018), we consider the cognates in Categories 1–3 as sufficiently supported to be used in our reconstruction of the pIGN phoneme system – but not those in Category 4. This resulted in a corpus of 233 reliable cognate sets. Finally, a note about lexical 'depth'. We assume, as is standard in historical linguistics (Meillet, 1967, pp. 41–42), that forms which are part of complex and irregular paradigms (for example, verbs), or are comparatively less independent, such as bound forms more generally, carry relatively greater evidentiary weight in assessing cognacy. Therefore, nouns such as An *ɾaŋ-* ~ Ng *ɭoŋ-* 'head' which exist only as bound roots are more indicative of a genetic relationship than independent nouns such as Wu *l̪umuɭuk* ~ An *a+l̪əmiɾa* 'blood-wood', since the latter are more easily borrowed than the former.

4. Shared core vocabulary: Swadesh list

Appendix 1 (https://doi.org/10.5281/zenodo.4045329) presents a 150-item Swadesh list of Anindilyakwa, Wubuy and Ngandi core vocabulary (Alpher & Nash, 1999). This list consists of body part nouns, verbs, adjectival nominals, terms for common natural phenomena, manufactured items and basic human classification terms. Inspection of the core vocabulary list of cognates belonging to Categories 1–3 gives the lexical cognacy rates shown in Table 19.

From this table we see that Anindilyakwa and Wubuy share about the same amount of basic vocabulary as Wubuy and Ngandi do, whereas Anindilyakwa and Ngandi share about half that amount. About 14% is shared by all three languages (and with few exceptions, most words that An and Ng have in common also occur in Wu). The lexemes shared by Anindilyakwa and Wubuy are not different in kind from those shared by Wubuy and Ngandi. In both cases, for example, there are finite verbs shared (12 shared by Wu and Ng – around 27% of Ngandi's finite verb inventory, and 14 shared by Wu and An). Such shared vocabulary levels are comparatively high for NPN languages (see Bowern et al., 2011; Harvey, 2011a); low levels of shared vocabulary between contiguous related languages are a well-known phenomenon in Australia. For example, Harvey (2011a) mentions Ngan'gityemerri and Murriny-Patha, which, based on shared irregular verbal paradigms, have been shown to be genetically related, belonging to the Southern Daly family (Green, 2003a), yet they share only 8% of their vocabulary. And within the GN family, Jawoyn and Ngalakgan only share 10% (Harvey, 2003, 2011a, p. 353).

5. Reconstructing the phoneme inventory of proto-insular Gunwinyguan

In this section, we focus on reconstructing the proto-phonemes of the common ancestor of Wubuy and Anindilyakwa only, for purely practical reasons: since we examine lexical cognacy and sound reconstruction, and the great majority of sound changes are limited to Wu and An, Ngandi cannot contribute to the reconstruction of the phoneme

Table 20 PIGN consonant phoneme inventory

	Labial	Lamino-dental	Apico-alveolar	Apico-retroflex	Lamino-alveo-palatal	Dorso-velar
Stop	p	t̪	t	ʈ	c	k
Nasal	m	(n̪)	n	ɳ	ɲ	ŋ
Lateral		l̪	l	ɭ		
Tap			r			
Approx.	w			ɻ	j	

Table 21 PIGN vowel phoneme inventory

	Front	Central	Back
High	i		u
Low		a	

inventory of the proto-language (recall that the Ngandi phoneme inventory is almost identical to the one reconstructed for pGN (Table 1); Harvey (2003, p. 225) lists only one sound change for Ngandi from pGN).

We call the immediate ancestor of Anindilyakwa and Wubuy 'proto-Insular Gunwinyguan' (pIGN) for practical purposes, without making a claim that these two languages form a subgroup. Nor do we wish to claim that Ngandi is only distantly related to Wubuy or Anindilyakwa; as mentioned, based on verbal paradigms, Baker (2004) has shown that Wubuy and Ngandi are closely related – although the corresponding verbal suffixes may well be retentions rather than innovations (see Green, 2003b). But Anindilyakwa and Wubuy are clearly united by phonological innovations absent in Ngandi. Further research on the morphology of Anindilyakwa and Wubuy may establish more secure reasons for subgrouping these two languages.

We apply the Comparative Method to reconstruct the sound system of the immediate ancestor of Anindilyakwa and Wubuy; if this reconstruction is successful, it shows that the assumption that the two languages are genetically related is warranted (e.g. Campbell & Poser, 2008). The pIGN phoneme inventory we reconstruct is given in Table 20 (consonants) and Table 21 (vowels). The pIGN consonant inventory is similar to the one reconstructed for pGN, apart from the absence of the stop contrast and the glottal stop, and the presence of the lamino-dental lateral and nasal. We thus assume here that lenition eliminated the inherited fortis–lenis consonant contrast and distinguished pIGN from pGN, in the process innovating a new phoneme /l̪/. The pIGN three-vowel inventory (*/i/, */u/, */a/) is distinct from the five vowels reconstructed for pGN (*/i/, */u/, */e/, */o/, */a/). Both the pIGN consonants and vowels are preserved exactly in Wubuy (apart from the vowel length contrast currently found in Wubuy, which we do not reconstruct for the proto language), but both have changed rather dramatically in Anindilyakwa, in particular the vowels. We present the general patterns in terms of correspondence sets grouped according to manner of articulation: stops (§5.1), nasals (§5.2), laterals (§5.3), the tap (§5.4), continuants (§5.5) and vowels (§5.6). Appendix 2 (https://doi.org/10.5281/zenodo.4045329) contains the full dataset including the cognate sets, as well as details on the reconstructions. Ngandi and/or pGN forms are included and used in the reconstruction where available: for example, based on An *-məɻa*: Wu *muuɻi* 'FaFa' only, we cannot confidently reconstruct the quality of pIGN vowels for this word. But when considering pGN **muɻi*, Ng *muɻi*, which have the same vowels as Wu, we can state with certainty that the

Table 22 PIGN stops correspondence sets

Meaning	pIGN	Anin.	Wubuy	Env.	# corr. sets	RC
	*/p/	/p/	/p/	#__, C__, V__V	63	1
SEA WASP	*ja**p**uŋuru	ja**p**uŋura	(waa-)ja**p**uŋuru			
TREE SP. (Sterculia quadrifida)	*(ma-)**p**a|**p**i	ma+**p**a|**p**a	(ama-)**p**a|**p**i			
	*/t̪/	/t̪/	/t̪/	#__, C__, V__V	33	0.97
WORK	*wa𝑙+**t̪**a-	*wa𝑙+**t̪**a-	*wa𝑙+**t̪**a-			
STAB	*a**t̪**uka-	a**t̪**uk^{w}a-	a**t̪**uka-			
	*/t/	/t/	/t/	n__, V__V, __ c	13	0.38
BUTTER-FLY FISH	*man**t**ara	man**t**ara	(jii-)man**t**ara			
HIT, KILL	*+pa**t**ca	+pa**t**ca-	+wa**t**ca- ∼ +pa**t**ca-			
KANGAROO	*(?-)l̪anta	t̪ə+l̪anta	(a-)l̪aanta			
MISS (THE MARK)	*kVcVr+pa**t**ca-	k^{w}ijer+pa**t**ca-	wacir+pa**t**ca-			
WIRE SPEAR	*t̪urapa**t**a	t̪ərapa**t**a	(a-)t̪urapa**t**a ∼ (a-)t̪utapa**t**a			
	*/𝑙/	/𝑙/	/𝑙/	ŋ__, V__V	33	0.58
GRASS	*(a-)ma𝑙a	a+ma𝑙a	(a-)ma𝑙a			
VENUS SHELL	*jiŋ𝑙irani	jaŋ𝑙ərena	(jii-)jiŋ𝑙irani			
	*/c/	/c/	/c/	#__, C__, V__V	25	0.92
MIST	*(a-)wu**c**i	a+wi**c**a	(a-)wuu**c**i			
PUSH	*c**u**𝑙a-	c**i**𝑙a-	c**u**𝑙a-			
	*/c/	Ø	/c/	__.	6	0.83
LIVINGST.PALM	*(ji-)l̪alpi**c**	ji+l̪erpa	(jii-)l̪alpi**c**			
PARROT	*wuruwa**c**pa	wuruwepa	(waa-)wuru[illegible]pa			
	*/k/	/k/	/k/	#__, C__, V__V	37	0.97
SACRED	*k**u**𝑙u**k**u𝑙u	-ak**u**𝑙u**k**u𝑙a	k**u**𝑙u**k**u𝑙u			
	*/k/	Ø	/k/	__#	12	1
CYRENE SHELL	*(ji-)wa𝑙kur**k**	ju+walkura	(jii-)wa𝑙kur**k**			
	*/ku/	/k^{w}a/	/ku/	__#	22	0.96
BARK CANOE	*(a-)l̪amu**ku**	a+l̪amu**k^{w}a**	(a-)l̪amu**ku**			
	*/ka/	/k^{w}a/	/ka/	u__#	5	1
STONE	*(a-)ŋ**uka**	a+ŋ**uk^{w}a**	(a-)ŋ**uka**			

Wu form must be original, and An has changed */u/ to /ə/ (this is a regular shift preceding retroflex consonants in An, as is the conversion of any word-final vowel into /a/).

5.1 Stops

We propose that pIGN had a similar stop inventory to pGN as reconstructed by Harvey (2003): bilabial, lamino-dental, alveolar, retroflex, palatal and velar stops, as set out in Table 22. The differences are that the pGN stop contrast and glottal stop (both preserved in Ngandi) were lost in pIGN (even though the lenition process responsible for this loss was probably not complete before Wu and An separated; §2.3). The reconstructed pIGN phonemes appear in column 2, together with two example reconstructed forms. Columns 3 and 4 contain the corresponding reflexes in Anindilyakwa and Wubuy, respectively, while column 5 presents the phonological environments of the relevant correspondence set, column 6 the number of cognate sets relevant to the reconstruction of the proto-phoneme, and column 7 the

Regularity Coefficient (RC) that indicates the degree of regularity between the correspondences, as discussed in §3 above.

Some general observations can be made from this table:

- As mentioned, all words end in /a/ in Anindilyakwa. Since this is an innovation probably due to the general tendency in the language to avoid codas, we reconstruct the Wubuy endings for pIGN. In §2.2 we discussed the evidence that word-final /a/ was attained by either deletion or epenthesis, depending on whether the word-final segment in the proto-language was a stop or a sonorant.
- For some noun roots starting with a segment that represents one of the class prefixes in An (a, m, j, ṯ, w), the Wu correspondence appears to have an extra class prefix, as in An *japuŋuru* : Wu *(waa-)japuŋuru* 'sea wasp'. The An noun is classified MASC because of the initial /j/, while the Wu noun is COLL, as represented by the *waa-* prefix. Since in such cases it is unclear whether An has lost or Wu has added the noun class prefix, we do not attempt to reconstruct it to pIGN.
- The reconstructed pIGN forms are almost always identical to the Wubuy forms, whereas Anindilyakwa has undergone major phonological changes. As we will see below, this observation holds more generally in the history of pIGN.

More specifically, regarding the stops, we can state the following:

- The evidence for reconstructing */p/, */ṯ/, */c/ and */k/ is strong: all are supported by at least 25 cognate sets and the relevant correspondence sets are > 90% regular.
- Wu /k/ corresponds to An /k/, /k^{w}/, or zero, but the distribution is fully predictable: word-final Wu /k/ is deleted in An (RC = 1); Wu /ku/ is realized as An /k^{w}a/ word-finally (RC = 0.96);[15] Wu word-final /ka/ is realized as An /k^{w}a/ when preceded by /u/ (RC = 1) (recall that An [u] obtains its rounding from contiguous rounded velars, so the Wu velars are analyzed as underlyingly rounded). Everywhere else, the velar stops match (RC = 0.97).
- Reconstruction of the alveolar stop is supported by 13 correspondence sets, which are only 38% regular (all shown in Table 22). We attribute this low match to /t/ being a marginal phoneme in An (as noted in §2.1): *t is reflexed variably by /t/, /ṯ/, /ʈ/, /r/, /ɻ/ ~ zero, with no apparent conditioning factors determining the outcome (Table 23). This would indicate that at some point in the history of An, *t was disfavoured and was (somewhat inconsistently) realized as other coronal segments, depending partly, but not entirely, on phonotactic position.

  For instance, the same reconstructed string */anti/ is reflexed by /aŋʈə/ in 'criticize' but by /aṉṯa/ in 'barramundi'. Since An generally appears to eschew alveolar stops as well as heterorganic clusters, combined with the available evidence from other languages (e.g. Ri *jatpuwa*), we reconstruct the Wu stops and clusters to pIGN.
- The relatively low degree of regularity of the retroflex stop (33 correspondence sets, which are 58% regular) is due to 11 cases where the Wubuy stop corresponds to a continuant in Anindilyakwa (e.g. Wu *(jii-)jimiʈaari* : An *jimaɻara* 'milkfish'). In §2.3 we

[15]The reason why RC < 1 here is the correspondence Wu *(a-)ku**ku*** : An *a+ku**ŋ^{w}a*** 'water', where (if these are in fact cognate) the velar stop appears to have become a nasal in An (**a+ku**ku*** > **a+ku**k^{w}a*** > *a+ku**ŋ^{w}a***). Evidence that the stop must be older comes from Bininj Gun-Wok *ku**kk**u* 'water'.

Table 23 Other reflexes of pIGN *t in An

Meaning	pIGN	Anin.	Wubuy	Env.	# corr. sets
	*/t/	/t̪/	/t/		
SHUT	*t̪i**t**a-	t̪i**t̪**a-	t̪i**t**a-	V__V	1
BARRAMUNDI	*jiŋku**l**pa**nt**i	jukulpa**nt̪**a	(jii-)jiŋku**l**pa**nt**i	N__	2
	*/t/	/ʈ/	/t/		
CRITICIZE	*a**nt**ira-	aŋʈ**ə**ra-	aa**nt**ira-	N__	2
	*/t/	/r/	/t/		
CUTTLEFISH (SHELL)	*ja**t**puwa	je**r**ipuwa	(jii-)ja**t**puwa	__.	2
	*/t/	/ɻ/ ~ Ø	/t/		
WASH	*a**t**cara-	aɻcira- ~ acira	a**t**cara-	__ c	1

proposed that although lenition appears to have been shared at the pIGN level, it was unlikely to have been a uniform process. Presumably, then, An /ʈ/ lenited (further) here whereas Wu /ʈ/ did not.[16]

- Syllable-final /c/ and word-final /k/ in Wu correspond to zero in An. This pattern is supported by five cognate sets (RC = 0.83) for /c/ and 10 cognate sets (RC = 1) for /k/. Since it is most likely an expression of the need in An to avoid codas, as discussed in §2.2, we can reconstruct the word-final consonants to pIGN.[17]
- */t/ and */ʈ/ are not definitely attested in initial position. Both are absent here in An and initial /t/ is rare in Wu (there are no correspondence sets involving initial /t/). Although /ʈ/ in initial position is not uncommon in Wu, there are only two correspondence sets with this sound in this position. Both have an initial vowel in An (Wu *ʈiku* : An *aɻukwa* 'raw' and Wu *ʈaʈark* : An *aʈətara* 'dry').[18] Because this contrast in initial position is not reconstructed to pGN (Harvey, 2003, p. 214) we do not reconstruct it in this position for pIGN. Ngandi, like most Australian languages, neutralizes the apical contrast in initial position: all initial apicals are underlyingly retroflex (Heath, 1978b).

5.2 Nasals

We reconstruct nasals at the same six places of articulations as the stops: bilabial, lamino-dental, alveolar, retroflex, palatal and velar, as set out in Table 24, below. We can most confidently reconstruct */m/, */n/, */ɳ/ and */ŋ/ (all have at least 15 correspondence sets and RC > 0.74). The dental nasal, on the other hand, is rare in both languages, represented by only seven cognate sets. Heath notes that /n̪/ occurs in only a handful of flora and fauna terms in Wu (1978a, p. 36), suggesting that, since this phoneme appears to be archaic in Yolngu, it may have entered Wu through loanwords from Yolngu. However, since /n̪/ also occurs in inherited words, however rare, we reconstruct it as a phoneme of pIGN. The lamino-palatal nasal is also rare and mostly (64%) occurs in homorganic clusters followed by a stop. However, since it occurs in a non-conditioning environment in two cognate sets, we reconstruct */ɲ/ as a phoneme to pIGN. For the

[16]The retroflex stop rarely lenited in Wubuy, whether in initial or medial position. In 25 roots reconstructible to pGN involving a medial retroflex stop, only two show lenited reflexes in Wubuy.

[17]A reviewer suggests that to make this constraint solid, it is necessary to spell out the range of possible options and factors influencing the chosen path, along the lines of the preference-law-based approaches developed in Vennemann (1988). However, limitations of space prevent us from going into detail here.

[18]As mentioned in footnote 8, the 'raw' root appears to be a borrowing from Yolngu *ʈiku* 'raw' (Zorc, 1986).

Table 24 PIGN nasals correspondence sets

Meaning	pIGN	Anindilyakwa	Wubuy	Env.	# corr. sets	RC
	*/m/	/m/	/m/	#__, C__, V__V, __C	85	0.99
CHASE	*l̪ar**m**a-	l̪ar**m**a-	l̪ar**m**a-			
PADDLE	***m**ijaca	***m**ijeca	(ama-)**m**ijaca			
	*/n̪/	/n̪/	/n̪/	__ t̪, V__V	6	1
ORGAN	*a**n̪**t̪i**l̪**i	a**n̪**t̪ə**l̪**a 'kidney'	(a-)a**n̪**t̪i**l̪**i 'heart'			
MANGROVE	*a**n̪**uma	a**n̪**əma	(a-)a**n̪**uma			
	*/n/	/n/	/n/	#__, V__V, __C	27	0.74
SPEAR	*a**n**uŋku	e**n**uŋk^{w}a	(a-)a**n**uŋku			
HAWK	*(ji-)**n**ikarka	ji+**n**ikarka	(jii-)**n**ikarka			
	*/ŋ̪/	/ŋ̪/	/ŋ̪/	__ t̪, __C, V__V, __#, #__	15	0.8
ELBOW	*a**ŋ̪**l̪a	a**ŋ̪**l̪a	(a-)a**ŋ̪**l̪a			
STONE	*(a-)**ŋ̪**uka	a+**ŋ̪**uk^{w}a	(a-)**ŋ̪**uka			
	*/ŋ̪/	/m/	/ŋ̪/	__p	5	1
FIG	*ma**ŋ̪**pa**ŋ̪**puru	ma**l̪m**pa**l̪m**pəra	(ama-)ma**ŋ̪**pa**ŋ̪**puru			
PLACE NAME (NE coast of Groote Eylandt)	*ama**ŋ̪**pa**ŋ̪**pina	ama**l̪m**pa**l̪m**pəna	ama**ŋ̪**pa**ŋ̪**pina			
	*/ɲ/	/ɲ/	/ɲ/	__c, C__	11	0.82
DOLPHIN	*(?)ŋa**ɲ**capani	t̪ə+ŋi**ɲ**capena	(jii-)na**ɲ**capana			
Grewia retusifolia	*(ma-)mur**ɲ**aN	ma+məri**ɲ**iɲa	(ama-)mur**ɲ**aŋ			
	*/ŋ/	/ŋ/	/ŋ/	#__, __#, V__V, C__, __C	36	0.97
FIRE	*(a-)**ŋ**u**l̪**a	a+**ŋ**u**l̪**a	(a-)**ŋ**u**l̪**a			
	*/ŋ/	/ŋ/	/m/	__p	3	1
GREENBACK MULLET	*(ji-)**l̪aŋ**pa**l̪aŋ**pa**t̪u**	ji+ʎa**ŋ**piʎa**ŋ**pa**t̪**a	(jii-)**l̪am**pa**l̪am**paa**t̪u**			
	*/ŋu/	/ŋ^{w}a/	/ŋu/	__#	4	1
EMPHATIC SUFF.	*-**l̪aŋu**	-**l̪aŋ^{w}a**	-**l̪aŋu**			
	*/ŋ/	Ø	/ŋ/	__#	3	1
OWNER	*apali-cu**ŋ**	-apalica	napali-cu**ŋ**			

alveolar nasal, only 74% of the correspondences match. Of the seven instances where the alveolar nasals do not match, five involve the cluster *nt. Above we saw that /t/ is a marginal phoneme in An, which also holds for the homorganic NC cluster, which is realized as either /ŋ̪t̪/ or /n̪t̪/ in An (e.g. Wu *mu**nt**uwa-* /n̪ *mə**ŋ̪t̪**uwa-* 'count', Wu *(a-)t̪ii**nt**u* /n *t̪i**n̪t̪**a* 'cane grass spear'). It appears that vowel conditioning may play a role, with dental clusters found following /i/ and retroflex following /u/, but following /a/ both dental and retroflex reflexes are found.

The */ŋp/ cluster is one of the few instances where we reconstruct the An form to pIGN, rather than the Wu form. As mentioned, this cluster is non-existent in Wu because the nasal assimilates to the point of articulation of the following stop, as we can observe in our data: the three correspondence sets with /ŋp/ are all realized as /mp/ in Wu (e.g. An *ji+ʎa**ŋp**iʎa**ŋp**a**t̪**u* /u *(jii-)l̪a**mp**a**l̪amp**aa**t̪**u* 'greenback mullet'). We can thus reconstruct */ŋp/ to pIGN, which continues in An (where it has become a complex phoneme) but changed to /mp/ in Wu.[19]

[19]Heath (1984, p. 70) explicitly states that this rule does not apply to /ɲ/ and /ŋ̪/, which retain their points of articulation when followed by a stop. Since An nasals frequently assimilate in point of articulation to a following stop, we reconstruct the Wu /ŋp/ cluster to pIGN for, e.g. 'fig' (*ma**ŋp**a**ŋp**uru* > An *ma**l̪mp**a**l̪mp**əra*).

Table 25 PIGN laterals correspondence sets

Meaning	pIGN	Anindilyakwa	Wubuy	Env.	# corr. sets	RC
	*/l̪/	/l̪/	/l̪/	#__, V__V	27	0.96
LIE DOWN	*murkuɭa-	murkuɭa-	murkuɭa-			
SEND	*ɭarka-	ɭarka-	ɭarka-			
	*/l/	/l/	/l/	V__V, __.	7	0.64
TURBAN SHELL	*juŋpalmi	jimpalma	(a-) juŋpalmi			
	*/l/	/r/	/l/	__stop	5	0.6
LIVINGST. PALM	*(ji-)l̪alpi	ji+l̪erpa	(jii-)l̪alpic			
	*l	Ø	*l	__#	1	1
BOIL	*(jii-)ɻulul	ji+ɻəla	(jii-)ɻulul			
	*/l̪/	/ɻ/	/l̪/	.__	8	0.89
HONEY	*(a-)ɭaku	a+ʂakwa	(a-)ɭaku			
SKIN	*-ma+kuɭak	-ma+kuʂa	(a-) makuɭak			
	*/l̪/	/l/	/l̪/	__.	9	1
BELLY	*muɭku	mulkwa	(ama-)muɭku			
WATER PYTHON	*(?-)waɭca	t̪u+walca	(jii-)aɭca			
	*/l̪/	/l/	/l̪/		2	1
CLAN NAME	*ɭaɭ[illegible]	lalaɻa	ɭal[illegible]			
EGRET	*(a-)[illegible]la(k)	a+marəla	(a-[illegible]aaɭa(k)			

5.3 Laterals

We reconstruct three laterals for pIGN: dental, alveolar and retroflex, as set out in Table 25; however, these three only contrasted in intervocalic position in the proto-language. The dental lateral is an innovation unique (among GN languages) to Anindilyakwa and Wubuy and is a reflex of pGN */t̪/. Anindilyakwa has a fourth – palatal – lateral, which is a language-specific innovation, corresponding to /l̪/ in Wubuy. Reconstruction of the dental lateral is well-supported (27 correspondence sets, RC = 0.96): the change pGN */t̪/ (preserved in Ngandi) > /l̪/ is one of the characteristic features of pIGN (e.g. pGN, Ng *t̪aŋkiʔ 'tree sp.' > Wu ɭaŋki, An ɭaŋ- 'tree'). This plosive > lateral change appears to be rare cross-linguistically (Kümmel, 2007, p. 248). The dental lateral is absent in all other GN languages and surrounding languages in the area, but it is common in An and Wu.[20]

The other two laterals are marginal phonemes in An. The alveolar lateral correspondence is only 64% regular, and Wu /l/ may correspond to /l̪/ in An (e.g. Wu *(mana-)wurculuku*: An *wuciɭukwa* 'barred long-tom') or to /ɻ/ (Wu *(jii-)jilaa[illegible] ji+ʂara* 'pine'). In coda position, Wu /l/ corresponds to An /r/ 60% of the time (e.g. Wu *(jii-)l̪alpic* [illegible] *ji+l̪erpa* 'Livingstone palm'), which is also a synchronic variation in An [illegible]*lkʷa* 'belly' was recorded as *murukʷa* in 2018. Since Wu /l/ is common, we reconstruct */l/ to pIGN.

The retroflex lateral is virtually absent in An, having shifted to a palatal lateral in onset position, and an alveolar in coda position.[21] The */ɻ/ > /ɻ/ shift is supported by eight cognate sets (RC = 0.89), which also include the language name: Wu *nun-t̪iɻaku* [illegible] *e+niŋ+t̪iʂakʷa* (where Wubuy *nun-* is the Gentilic prefix found on a number of ethnonyms, including the Wubuy ethnonym 'Nunggubuyu' (< *nun-wupuj* 'people who speak Wubuy')). PIGN */ɻ/ continues from pGN (e.g. pGN *kuɭa[illegible] GN, Wu *-ma+kuɭak* [illegible] *-ma+kuʂa* 'skin').

[20] A dental lateral is also reconstructed for Proto-Iwaidjan (Mailhammer & Harvey, 2018, pp. 348–349).

[21] It only occurs in a handful of loanwords, such as the Yolngu loans *maɻuwija* 'emu' (< *maɻwija*), *acarkaɻa* 'fish species' (< *waccarkkaɻi*) (Leeding, 1989, p. 30; Stokes, 1981, p. 143). The only inherited Anindilyakwa word listed in the dictionary with a retroflex lateral is *juwaɻkura* 'kneecap, Cyrene shell', but this word varies between [juwaɭkuɾa ~ juwalkuɾa], and which is in coda position, in any case. Only older speakers seem to use the retroflex lateral (Van Egmond, 2012, p. 32).

Table 26 Wubuy, Ngandi, pGN onset /l̻/ : Anindilyakwa /ɻ/

	Wubuy	Ngandi	Anindilyakwa	pGN
HEAD	*l̻aaŋ*	*l̻oŋ-*	*ɻaŋ-*	**L/ɻoŋ*
SKIN	*maku l̻a*	*ku l̻a?*	*-ma+kuɻa*	**ku l̻ak*
BLOODWOOD	*l̻umu l̻uk*	*t̪umu l̻u?*	*a+l̻əmiɻa*	**t̪umu l̻uk*
SUGARBAG	*l̻aku*		*a+ɻak^{w}a*	
[LANGUAGE NAME]	*nun-t̪i-l̻aku*		*e+niŋ+t̪i+ɻak^{w}a*	

The two cognate sets that do not support the /l̻/ : /ɻ/ correspondence are An /ɻ/ that corresponds to a different continuant in Wu: either /l/ (Wu *(jii-)jiɬaari* *jiɻara* 'pipe') or /ɻ/ (Wu *(jii-)jiniŋiɻa* *jiniŋiɻa* 'sandbar').

Finally, the */l̻/ > /l/ shift in coda position is 100% regular and supported by nine cognate sets.

5.3.1 History of laterals in Anindilyakwa

The change from retroflex to palatal lateral attested in An, through correspondence sets such as those below (Table 26), is one of the most interesting and surprising sound changes in the history of IGN. We are aware of no other Australian languages (or indeed, languages elsewhere) with this sound correspondence. A lamino-palatal lateral does not occur as a contrastive phoneme in Wubuy or Ngandi (or any other GN language), and is not reconstructed for the phoneme inventory of pGN by Harvey (2003). There are several factors arguing for the validity of the change l̻ > ɻ that we propose, however. One is that the forms in question are unlikely to be recent borrowings. As mentioned (§2.4), noun class prefixes (for non-human nouns) in Anindilyakwa are entirely frozen. Evidence from Makassan borrowings (which lack prefixes) suggests that the class system has been defunct for at least several hundred years. This in turn suggests that forms such as *a-ɻak^{w}a* 'honey' (Wubuy: *a-l̻aku*) are quite old, and presumably retained from the proto-language. This form is not found in nearby languages such as Yolngu (the equivalent term is *kukku*), nor in the nearest relative Ngandi (which uses *kuŋ* for this concept).

Similar considerations suggest the other items in Table 26 are also very old. The item glossed 'head' in this set is only found as an incorporated noun in both Ngandi and Anindilyakwa, not as an independent noun in either language; it is found as both an independent and an incorporable noun in Wubuy. A potential cognate is also found in Ngalakgan *ɻoŋ* 'face' (Harvey, 2003, p. 242).[22] In addition, Van Egmond (2012) argues that noun incorporation in Anindilyakwa is largely defunct, making recent borrowings of such forms highly unlikely. Unlike the other forms in this set, the form for 'skin' is highly represented among the GN languages (see Harvey, 2003), as well as in several frozen forms in Wubuy and Anindilyakwa. The existence of this form (if it is indeed a legitimate cognate) suggests that the /l̻/ : /ɻ/ correspondence is not limited to word/morpheme onsets.

If the correspondence were limited to morpheme-initial position, then we could argue on the basis of the well-established restrictions on word-initial apicals in Australian languages, discussed at length by Hamilton (1996), that Anindilyakwa palatalization is a result of the same constraint. The cognate appears to be genuine, hence we cannot use the constraints on initial position to explain the /l̻/ : /ɻ/ correspondence. However,

[22]Though Harvey (2003) does not list Ngalakgan /ɻ/ : Wu, Ng /l̻/ as a regular sound correspondence.

Table 27 Reconstruction of the phonotactic distribution of laterals in pIGN

Word onsets, post-consonantal	Intervocalic	Syllable coda
*l	l	l
ɭ	ɭ	ɭ
l̪	l̪	*l̪

apart from a few instances in the dictionary, apical laterals are extremely rare in syllable onsets in Anindilyakwa, suggesting that they were largely eliminated early on and replaced by lamino-palatal laterals, to differentiate them maximally from the other lateral: the lamino-dental.[23] This sound change was presumably facilitated by the lack of a pre-existing phoneme in this part of the inventory (as with the t̪ > l̪ change). The development of the lamino-palatal lateral must have followed the generation of the lamino-dental lateral through lenition (or other processes), a sound change that is shared by Wubuy and Anindilyakwa. Apical laterals were later re-introduced through loanwords from Makassarese and Yolngu varieties.

In particular, we propose the following path of development. In pIGN, there was a contrast between three laterals: l, ɭ and l̪ (from *t̪). This three-way contrast was neutralized in different ways depending on syllable position, however. The only position where all three laterals were in contrast was the intervocalic position, as is commonly the case in Australian languages (Dixon, 1972, p. 272; Hamilton, 1996, p. 76). In word-initial position, and following any preceding consonant (if indeed this was possible), the apical contrast was neutralized: only retroflex and dental laterals could occur. In syllable codas, only the apicals were licensed and dentals were illegal as is also commonly the case in Australian languages (Hamilton, 1996, p. 126) (Table 27). However, this kind of lateral inventory is extremely rare in Australia. Hamilton's survey of 116 Australian languages finds that laminal laterals are marked with respect to apical laterals (1996, p. 64), the presence of a laminal lateral in the inventory implies an apical one. Furthermore, if a language has one laminal lateral, it is the alveo-palatal; Wubuy and Yirr-Yorront are the only exceptions.

Among the coronals, the contrast between alveolars and dentals is the least differentiated overall (Hamilton, 1996, p. 49). In word-initial position, the primary cue is the burst (Bundgaard-Nielsen et al., 2016, p. 2806; Ladefoged, 2001, pp. 158–160, both reporting on Wubuy stops), which laterals lack; vowel formant transition cues are extremely weak to non-significant. This is presumably the reason for the rarity of initial contrasts in apical and dental laterals. We therefore propose that such an inventory is inherently unstable, particularly in initial position.

At some point in the history of Anindilyakwa, it seems that initial apical laterals were highly disfavoured, possibly for the reasons set out here. Retroflex/neutralized apical laterals were replaced by alveo-palatals in word-initial positions, possibly as a way of maximizing contrast in this position. However, retroflex laterals were disfavoured more broadly and were replaced by alveo-palatals in onset positions in general, and by alveolars in coda positions. Inherited alveolar lateral codas tended to be replaced by trills (presumably a further push-chain shift). The nature of the processes underlying these shifts are the topic of further research. We are aware of no

[23] Apart from Makassan loans and in coda position, the dictionary lists: *mejela* 'nervous', *wupərikila* 'spear', *jimaŋala* 'flat woomera' (< Yolngu *maŋal?*), *jiɻəla* 'boil' (< Wubuy).

Table 28 Makassarese, Wubuy /l/, Anindilyakwa /l̪/, /l/, /ɻ/

	Mkr /l/	Wu /l/	An /l̪/	An /l/	An /ɻ/
BOTTLE	*bóṯolo?*	*puṯ̪ili*	*puṯ̪il̪a*		
SWORD	*galéwaŋ*	*kal̪iwaŋa*	*kal̪iwaŋa*		
EUROPEAN (< 'Hollander')	*balánda*			*palanta*	
ANCHOR	*baláŋo*	*palaŋu*		*palaŋ^{w}a*	
SAIL	*sómbala?*	*ṯ̪umpala*		*ṯ̪umpala*	
KNIFE	*ladiŋ*				*ɻeɻiŋa*
AXE	*paŋkulu?*	*paŋkilicina*			*paŋkiɻa*

Table 29 Makassarese /l/, Anindilyakwa /la/, /l̪i/

Mkr /l/	An /la/	Mkr /l/	An /l̪i/
balánda 'European'	*palanta*	*bóṯolo?* 'bottle'	*puṯ̪il̪a*
baláŋo	*palaŋwa* ~ *pal̪aŋ^{w}a*		
pallajaraŋ ~ *palliaraŋ*	*palijara*	*gúliŋ* 'rudder'	*kul̪aŋa*
	ṯ̪umpala	*galéwaŋ*	*kal̪iwaŋa*
		karóro?	*kal̪iwura*

sound changes of the particular kind discussed here, although historical changes from retroflex to prepalatal articulations in Arandic languages may constitute a related phenomenon (Harvey, 2011b).

Apical laterals in onset positions were reintroduced primarily through loanwords. There are no words with word-initial apical laterals, except for the clan name *lalara* (An) : *l̺al̺ara* (Wu). This is a clan which originated on the mainland, in the Wubuy speaking area (Turner, 1999), and given this and the absence otherwise of apical laterals in word-initial position in Anindilyakwa, we assume this name is originally from Wubuy. Apical laterals in onsets were re-introduced into Anindilyakwa through loans from Wubuy, Makassarese and Yolngu varieties, nevertheless we often observe variation, indicating ongoing disfavouring of apical laterals in onsets. In Wubuy, Makassarese /l/ (described as an apico-alveolar lateral in Jukes (2006)) is uniformly realized as the corresponding Wubuy phoneme /l/. This is not the case in Anindilyakwa however. Here, we observe several alternant reflexes. Most of the time, reflexes are almost evenly split between the alveolar /l/ and the lamino-dental /l̪/. In just two cases, both of them dubious, Mkr /l/ is realized as a lamino-alveopalatal /ɻ/. It is difficult to discern any overwhelming conditioning factors involved in this alternation. It is true that all word-initial Mkr /l/s are dental in loans, and conversely, apico-alveolar /l/ is found only word-internally in Anindilyakwa (Tables 29 and 28). On the other hand, we find dental /l̪/ in this environment as well. All of the corresponding Wubuy forms show alveolar /l/ corresponding to Mkr /l/. This is therefore evidence that the contrast between the alveolar lateral /l/ and the lamino-dental lateral /l̪/ in Anindilyakwa was tenuous, at best, at the time when Mkr loans were borrowed by An speakers.

There appears to be some evidence for the distribution of these two laterals in An conditioned according to the following vowel. There is a strong tendency for alveolar /l/ when the neighbouring vowels are /a/, and /l̪/ when the following vowel is /i/. The form *puṯ̪il̪a* 'bottle' does not counter-exemplify this: all word-final vowels in An are synchronically /a/; the Wu form *puṯ̪ili* suggests the An form may have been /i/-final historically as well. But the tendency *is* counter-exemplified by the *palijara* form, which the dictionary specifically says is pronounced alveolar. The fact that many Mkr laterals were borrowed as lamino-

Table 30 PIGN tap correspondence sets

Meaning	pIGN	Anindilyakwa	Wubuy	Env.	# corr. sets	RC
	*/r/	/r/	/r/	__C, V__V	71	0.96
BLUE TUSK FISH	**jampi**r**ku*	*jempə**r**k^{w}a*	*(jii-)jampi**r**ku*			
DART	**ja**r**uwa**r**i*	*ja**r**uwa**r**a*	*(jii-)ja**r**wa**r**i*			

Table 31 PIGN approximant correspondence sets

Meaning	pIGN	Anindilyakwa	Wubuy	Env.	# corr. sets	RC
	*/w/	/w/	/w/	V__V, #__	55	0.95
HIT, KILL	****w**aɻa-*	**w**aɻa-	**w**aɻa-			
NEW	**kaɻu**w**a*	-kaɻu**w**a	kaɻu**w**a			
	*/ɻ/	/ɻ/	/ɻ/	__C, V__V, #__	22	1
BILLABONG	**(a-)wuɻuku*	a+wuɻuk^{w}a	(a-)wuɻuku			
VINE SP.	**(ma-)ɻica*	ma+ɻica	(ama-)ɻiica			
	*/j/	/j/	/j/	#__, V__V	39	0.87
BREAM	****j**iwunpiɳa*	**j**uwunpəɳa	*(jii-)**j**iwunpiɳa*			
THROAT	****j**ampi**j**a*	**j**ampi**j**a	*(jii-)**j**ampi**j**a*			

dentals, rather than apico-alveolars as in Wubuy, strongly suggests that at the time of Makassan contact, Anindilyakwa was a language with two laterals, both laminal. We further suggest that apical laterals were eventually re-introduced through loans in forms such as *palijara*, creating a contrast between what were formerly allophones conditioned by the following vowel or by position in the syllable.

5.4 Tap

A tap /r/ can be confidently reconstructed for pIGN, as set out in Table 30. The three cases where /r/ does not match involve a consonant cluster in Wu and a simple segment in An: Wu *(ama-)wu**rc**uluk* : An *wu**c**i̪uk^{w}a* 'barred long-tom', and Wu *(jii-)ɻa+mu**rl̪**umu**rl̪**u* : *ji* [illegible]*ŋ+mə**l̪**ə**l̪**a* 'stonefish'. The /rc/ cluster is uncommon and unstable in An (the dictionary lists two entries, both of which vary between /rc/ and /c/), and the liquid+dental cluster /rl̪/ is non-existent. Therefore, we reconstruct the Wu clusters to pIGN and propose that they have simplified to single segments in An.

5.5 Approximants

We reconstruct central approximants in three places of articulation: labiovelar /w/, retroflex /ɻ/ and palatal /j/, as set out in Table 31. The three sets where /w/ fails to correspond in the two languages involve an initial zero in either language (e.g. Wu *(a-)**w**aɻkara* : An *alkəra* 'herring', Wu *(jii-)aɻca* : An *t̪u+**w**alca* 'water python'). Three of the five sets where /j/ does not match involve An zero in initial position (e.g. Wu *(jii-)**j**ara̪n̪t̪i* : An *ara̪n̪t̪a* 'sweetlips'). The two other cases show hardening in Wu but not in An: e.g. Wu ***c**armajarmac* : An *-**j**armijarma* 'thin'. In compounds, however, the Wu form is *-**j**armajarmac*, so we can safely reconstruct the unhardened form to pIGN (recall that hardening of continuants to stops is a synchronically productive process in Wu).

Table 32 PIGN */a/ correspondence sets

Meaning	pIGN	Anindilyakwa	Wubuy	Env.	# corr. sets	RC
	*/a/	/e/	/a/	__Ci, __Ce, __C[+ high]	29	0.86
BLUE TUSK FISH	***ja**mpirku*	***je**mpərk^{w}a*	*(jii-)**ja**mpirku*			
BUSH	***wa**ri*	***e**ripe**a**pa*	*(a-)**a**ṯa**a**riɲ*			
PARROT	*(wara-)**wa**pa*	*wuru+**we**pa*	*(waa-)**wu**ruw**a**cpa*			
	/a/	*/e/*	*/a/*	__n	6	1
STRINGYBARK	***ma**nuŋku*	*e+**me**nuŋk^{w}a*	*(ama-)**ma**nuŋku*			
	/a/	*/ə/*	*/a/*	[-stress], __C[+retroflex, rhotic]	?	?
STINGRAY	*(a-)**ma**ɳiɳik*	*a+mərɻəɳa*	*(a-)**ma**ɳiɳik*			
	/a/	*/a/*	*/a/*	elsewhere	196	1
BARRACUDA	***a**ma**a**rŋ**a***	***a**ma**a**rŋ**a***	*(a-)**a**ma**a**rŋ**a***			
SWEETLIPS	***ja**kari*	***ja**kara*	*(jii-)**ja**kari*			

5.6 Vowels

Vowels are the most controversial topic in the available descriptions of Anindilyakwa, with the number of vowels in the language ranging from one primary vowel (/a/: Heath, n.d.), to two (/a/, /i/: Leeding, 1989), four (/a/, /i/, /e/, /u/: Stokes, 1981; or /a/, /i/, /e/, /ə/: Van Egmond, 2012) or five (/a/, /i/, /ə/, /u/, /ae/: Moody, 1954). The vowels are problematic because on the one hand, the quality of the mid and high vowels seems to be conditioned by the surrounding consonants, though on closer inspection, /i/, /e/ and /ə/ also occur in non-conditioning environments.

We reconstruct the common Australian three vowel system for pIGN, */a/, */i/, */u/, which continues unchanged in Wubuy (apart from the length contrast not reconstructed for pIGN), but which has altered quite dramatically in Anindilyakwa:

- */a/ > /e/ when adjacent to a palatal C or /i/ in the next syllable;
- */a/ > /e/ before /n/;
- */i/ > /ə/ preceding retroflex consonants or /r/; retraction of front vowels in a retroflex context is common cross-linguistically (Catford, 1977; Gafos, 1999, pp. 140–141, 213; Steriade, 2001, pp. 226–227) and also, within Australia, in the GN languages Rembarrnga and Dalabon (Baker, 2014, pp. 175–176);
- */i/ > [u] adjacent to labio-velars;
- */u/ > [u] only adjacent to dorso-velars for which an underlying representation as a labialized velar is possible (= [u]-conditioning environment);
- */u/ > /i/ adjacent to palatals; and
- */u/ > /ə/ preceding retroflex consonants or /r/, and in non-[u]-conditioning environments.

We start with */a/ (Table 32), followed by */i/ (Table 33) and */u/ (Table 34). There are in total 277 correspondence sets with /a/, but in only 76% of these does Wu /a/ correspond to An /a/. In the other instances, */a/ is raised to /e/ in An by following lamino-palatals, */i/ or */n/, and */a/ is reduced to /ə/ in unstressed environments followed by a retroflex or rhotic consonant. The raising of /a/ to /e/ is both a diachronic process in An, having created the phoneme /e/ (absent in Wu), and a synchronic process that varies per

Table 33 PIGN */i/ correspondence sets

Meaning	pIGN	Anindilyakwa	Wubuy	Environment	# corr. sets
	*/i/	/i/	/i/	some non-conditioning environments, but esp. adjacent to palatal C(s)	34
TO SHUT	*ṯita-	ṯiṯa-	ṯita-		
BUSH	*wariɲpariɲ	eripéripa	(a-)wariɲpariɲ		
	*/i/	/ə/	/i/	preceding retroflex C(s) or /r/	34
BREAM	*jiwunpiŋa	juwunpəŋa	(jii-)jiwunpiŋa		
INSTR. CASE SUFFIX	*-miri	-məra	-miri		
	*/i/	/u/	/i/	adjacent to labio-velars	17
TO CLIMB	*wiɻa-	wuɻa-	wiɻa-		
RAW	?*ɻiku	-aɻukwa	ɻiku ~ ɻiku+ ɻuku+		

Table 34 PIGN */u/ correspondence sets

Meaning	pIGN	Anindilyakwa	Wubuy	Environment	# corr. sets
	*/u/	/u/	/u/	adjacent to labio-velars only	54
BILLABONG	*(a-)wuɻuku	a+wuɻukwa	(a-)wuɻuku		
SEA WASP	*japuŋuru	japuŋura	(waa-)japuŋuru		
	*/u/	/ə/	/u/	adjacent to retroflex C(s) or /r/, and in non-cond. env.	21
BACKBONE	*murŋ	mərəŋa	(ama-)muurŋ		
BLOODWOOD	*(a-)l̪umuɭuk	a+l̪əmiʎa	(a-)l̪umuɭuk		
	*/u/	/i/	/u/	adjacent to palatal Cs only	8
BLOODWOOD	*(a-)l̪umuɭuk	a+l̪əmiʎa	(a-)l̪umuɭuk		
TO SCRAPE	*ɻuca-	ɻica-	ɻuca-		

speaker.[24] The historical raising of */a/ to /e/ conditioned by /i/ in the following syllable, or when followed by a palatal consonant, is supported by 29 cognate sets, which are 86% regular (e.g. **mijaca* > An *mijeca* 'paddle', **maɭiwu* > An *meɭuwa* 'razor clam'). The conditioning /i/ may have disappeared, as in the 'razor clam' example. However, raising is blocked when */i/:

- occurs word-finally: e.g. **jakari* > An *jakara* 'sweetlips' (not **jakeri*) (16 correspondence sets, RC = 0.94);
- is preceded by a retroflex consonant: e.g. **(a-)maŋiṉṯaŋu* > An *a+maŋəṉṯaŋʷa* 'eagle ray' (not **a+meŋəṉṯaŋʷa*);
- is preceded by a consonant cluster: e.g. **t̪arpi-* > *t̪arpə-* 'move' (not **t̪erpə-*).

The alveolar nasal also conditions the raising */a/ > /e/. The six cognate sets that support the Wu /an/ : An /en/ correspondence are 100% regular. Indeed, there are very few instances of /an/ in the An dictionary, all of which are loanwords from Makassarese (e.g. *pacananga* 'lantern', *paciɭana* 'fish hook'), as opposed to many examples of /en/ in inherited words (e.g. *emeniṯara* 'NEUT.fins', the demonstrative ***ena*** ~ ***enena*** 'this here', and in many grammatical affixes).[25] By contrast, the dental nasal, however rare, *can* be preceded by /a/, as in ***a***ṉ*əŋa* 'vegetable food' and *jiṉ**a**ṉa* 'fingernail'.

[24]For example, *ŋampe-na* 'bathe-NP' varies with *ŋempe-na*, and ***aja**rka* 'hand' varies with ***eije**rka* (van Egmond, fieldnotes).

[25]The language name itself is not an exception: although spelled *Anindilyakwa*, it is consistently pronounced [**en**iṇṭiʎakwa].

From these data it may seem that /e/ is not contrastive in An, as its distribution appears to be fully predictable. However, synchronically, this vowel also occurs in non-conditioning environments where it is contrastive, as is apparent from (near-)minimal pairs such as *m**a**ra* 'wattle tree'–*m**e**ra* 'blood', ***a**ra* 'forehead'–***e**ra* 'vomit', *m**a**n̪t̪a* 'heron'–*jim**e**n̪t̪a* 'turtle'. Hence, we treat /e/ as a synchronic phoneme of An. */a/ tends to be realized as /ə/ when occurring in a non-stressed environment *and* followed by a retroflex or rhotic consonant. Exact numbers cannot be given, however, since we do not reconstruct the prosody of pIGN. The reduction of any vowel to /ə/ in unstressed environments is a strong tendency in An in general (Van Egmond, 2012).

Table 33 sets out the correspondences with the reconstructed vowel */i/. Of the altogether 81 correspondence sets, only 42% shows an /i/ reflex in An. This is because pIGN */i/ is realized as /ə/ in a retroflex or rhotic environment in An, and as /u/ when adjacent to labio-velars. Although /i/ occurs in all environments in An, it is by far the most common adjacent to palatal consonants. Consequently, the An reflex of pIGN */i/ is consistently /i/ in this conditioning environment, but it frequently obtains a different quality in others: */i/ > /ə/ in a retroflex environment or preceding /r/, and */i/ > [u] when adjacent to labio-velars. Regularity coefficients cannot be given here because the quality of the An vowels is to a high degree predictable from the surrounding consonants.

Table 34 sets out the correspondence sets involving pIGN */u/, which in An continues in a fully predictable fashion: exclusively adjacent to velars and labio-velars. Proto-IGN */u/ only continues in An in a conditioning environment: when adjacent to /w, k^w, $ŋ^w$/ or to [k, ŋ] that could underlyingly be interpreted as labialized. This is because in An, [u] is not contrastive but either generated by absorbing the roundness of a preceding labio-velar (e.g. /kwV/ > [ku]) or by assimilating to a following labio-velar (e.g. /Vk^wa/ > [uk^wa]). For instance, both */u/ vowels in **(a-)w**u**ɻ**u**ku* 'billabong' continue in An because they occur in environments compatible with /u/: following */w/ and preceding */ku/. Examples such as **m**u**rk**u**n̪a-* > An *m**u**rk**u**n̪a-* 'lie down' and **m**u**ɻku* > An *m**u**lkwa* 'belly' show that an intervening consonant (r, ɻ) does not prevent the high vowel from assimilating to the following labio-velar. The vowels /i/ and /ə/ do not occur in this environment in An. For example, the An reflex of **w**i**ɻa-* 'to climb' is *w**u**ɻa-* (Table 33) because pIGN */i/ assimilates to the preceding /w/.

In other environments, */u/ cannot be maintained in An. When followed by a retroflex consonant or /r/, the An reflex is predictably /ə/ (e.g. **m**u**ŋɻuka-* > An *məŋɻukwa-* 'come together', and note that this rule is in a counter-feeding relationship with the preceding *i > u rule producing *w**u**ɻa-*). When adjacent to a palatal, the An reflex of */u/ is /i/ (e.g. **am**u**ɲ* > An *m+am**i**ɲa* 'firefly'). In non-conditioning environments, */u/ cannot be maintained either and continues as /ə/ (e.g. **(a-)n̪**u**muɻuk* > *a+n̪əmiʎa* 'bloodwood').

From Tables 32–34 it might seem that /ə/ is not contrastive in An, as its distribution appears to be fully predictable. However, synchronically, this vowel also occurs in non-conditioning environments where it is contrastive, as in the following (near-)minimal pairs: *aɻəmpa* 'NEUT.blister' *aɻ**ə**mpa* 'NEUT.stingray', *jiŋ**a**ŋa* 'MASC.nail': *təŋ**ə**ŋa* 'FEM.mosquito', *məɻəŋkəra* 'VEG.sinker' *əŋ**u**kura* 'NEUT.face'. Thus, we treat /ə/ as a synchronically contrastive phoneme in An.

In sum, the An four-vowel system has clearly developed from the pIGN three-vowel system preserved in Wu: whereas */u/ has lost its phonemic status, /e/ and /ə/ have gained it, having developed from */a/ by raising, and from */i, u, a/ by centralizing, respectively.

6. Conclusions

Using a dataset of 233 cognates with reliable etymologies and applying the traditional techniques of historical reconstruction, we have reconstructed the phoneme inventory of the immediate ancestor of Wubuy and Anindilyakwa, which we call proto-Insular Gunwinyguan (pIGN) in view of the island-based homelands of a significant proportion of the descendant speakers. To obtain as rigorous, transparent and objective a reconstruction as possible, we followed Mailhammer and Harvey's (2018) measure of regularity to evaluate the reliability of systematic correspondences, and their classification system to index the quality of reconstructions. This way, our reconstruction can be followed and evaluated objectively.

Our results can be summarized as follows:

- It can be demonstrated that Anindilyakwa and Wubuy are closely related genetically, as there are systematic correspondences between phonemes that are better interpreted as indicating an origin in an intermediate common ancestor than having occurred independently, by chance, or by borrowing.
- The pIGN consonant phoneme inventory is similar to the pGN consonantal inventory, and many reflexes are identical to their historical sources. The IGN languages have innovated two additional consonants, both lamino-dental: /n̪/ and /l̪/. The latter is a reflex of pGN */t̪/, while the former may have diffused from the Yolngu languages into pIGN. Both consonants also occur in Ngandi, but only in a handful of loanwords from Wubuy (Heath, 1978a, p. 7). Anindilyakwa independently developed a lamino-palatal lateral and two series of complex consonants.
- The stop contrast reconstructed for pGN was lost in pIGN (as described by Heath (1978a) for Wubuy). Heath's proposed lenition chain that is responsible for this change is problematic, however (§2.3). In particular, it is unlikely that lenition was a 'one-off', comprehensive change. Rather, it can be characterized as a highly variable, lexeme-specific and incomplete change. The factors behind this kind of sound change are the subject of ongoing research.
- The pIGN vowel inventory (three vowels) differs from the pGN system (five vowels). It is preserved in Wubuy but significantly altered in Anindilyakwa.

The following example illustrates some of these developments. An ***a***+*l̪əmiɻa* 'NEUT bloodwood' descends from pGN **t̪umuɻuk* 'bloodwood' through the following sound changes, which are all regular:

**t̪umuɻuk*	
(i) Lenition of lamino-dental stop	> ***l̪***umuɻuk (Wu, An)
(ii) Conversion of *u to /ə/ in non-[u]-compatible environment	> *l̪***ə**məɻ**ə**k (An)
(iii) Retroflex lateral > lamino-palatal in onset	> *l̪**ə**mi**ɻ**ək (An)
(iv) Loss of word-final coda, word-final vowel > [a]	> *l̪əmiɻ***a** (An)
(v) Lexicalization of NEUT noun class prefix	> ***a***+l̪əmiɻa (An)

While lenition of the type described in §2.3 is one of the few prevalent sound changes across Australia (Baker, 2014, pp. 170–171), Anindilyakwa has undergone more extensive changes from its reconstructed ancestor, pushing its phonemic inventory and

phonotactics into an area not shared by any neighbouring language. The dramatic phonological innovations of Anindilyakwa clearly deviate from the general tendency in Australia for most languages having undergone few sound changes (Koch, 2014, p. 26), so that cognates tend to be identical or near-identical (Miceli, 2008, p. 212).

We have not directly addressed the issue of subgrouping in this paper, so it remains unclear what the precise nature of the internal relationship between Anindilyakwa, Wubuy and Ngandi is. Based on the synchronic differences between them, it appears that Anindilyakwa and Wubuy have undergone separate development for a considerable period of time. Space precludes a consideration of their similarities in morphology here (but see Van Egmond, 2012, Ch. 9). Wubuy and Ngandi have previously been shown to have shared irregular morphology in their verb paradigms (Baker, 2004). However, it is unclear whether the verbal morphology shared by Wubuy and Ngandi (and likewise Anindilyakwa and Wubuy) constitutes retention or innovation (c.f. Green, 2003b), and we leave this question open for future research.

Finally, we do not concur with Heath's observation that Anindilyakwa shows a high degree of lexical divergence, that it shares relatively few lexical items with Wubuy and Ngandi – or indeed any other language (Heath, 1981, p. 338 fn. 2). In fact, Anindilyakwa and Wubuy, and even Anindilyakwa and Ngandi, share more basic vocabulary than do many other genetically related languages in Australia, including other neighbouring languages within the Gunwinyguan family (such as Ngalakgan and Jawoyn, which, as mentioned, share just 10%). The fact that this has gone unnoticed for so long could be due in part to the extensive sound changes that have taken place in Anindilyakwa after the split of pGN and pIGN, while Wubuy appears to have remained largely stable after this point. Such dramatic sound changes, which are an exception rather than the rule in Australia, have resulted in a language that is phonologically – though not lexically or grammatically – quite unusual.

Acknowledgements

The authors would like to thank Rikke Bundgaard-Nielsen, Nick Evans, Mark Harvey and Robert Mailhammer for useful input to the topic of this article over many years, though they should by no means be held to account for the conclusions we draw here. We would also like to thank Rikke Bundgaard-Nielsen for help with Figure 2. Marie-Elaine van Egmond wishes to thank the many Anindilyakwa speakers who have tried to teach her their language, especially: †Dugururru Lalara, †Caroline Wurramara, Judy Lalara, Letoria Yulidjirri and Corey Amagula. She furthermore would like to thank the Groote Eylandt Language Centre in Angurugu and Umbakumba, and the Anindilyakwa Arts Centre in Angurugu, for their invaluable help and support.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Marie-Elaine van Egmond's work was supported by fieldwork grants from the Australian Institute of Aboriginal and Torres Strait Islander Studies under Grant G2008/7358; the Endangered Languages Documentation Programme at SOAS, London under Grant FTG0152; the ARC Centre of Excellence for the Dynamics of Language; and the THEORIA Scientific Programme for the Humanities of the

Ministry of Education, Science and Culture, Mecklenburg-Vorpommern, Germany. Brett Baker's work Q7 Q6 was supported by the Australian Research Council's Discovery Program under Grant DP190100646 'One potato, two wotatoes, three otatoes: Lexical access in Australian languages' (with co-investi- Q8 gators Rikke Bundgaard-Nielsen, Mark Harvey and Robert Mailhammer).

Geolocation information

Groote Eylandt:
Latitude: −13° 57′ 59.99″ S
Longitude: 136° 34′ 59.99″ E
Numbulwar:
Latitude: −14° 16′ 60.00″ S
Longitude: 135° 43′ 59.99″ E

Notes on contributors

Marie-Elaine van Egmond completed her PhD at the University of Sydney on the phonology, morphosyntax and genetic position of Anindilyakwa. She then taught English Linguistics as a lecturer at the University of Greifswald in Germany and continued working on Anindilyakwa as a postdoctoral researcher in the THEORIA Scientific Programme, investigating the influence of English in the emerging varieties of the younger generations.

Brett Baker is an Associate Professor in Linguistics at the University of Melbourne. He has been researching the languages of south-eastern Arnhem land since the mid-1990s, in particular Ngalakgan and Wubuy, with a focus on word structure, phonology and more recently processing and perception. He has published books and articles on phonology, speech science, morphology, syntax, semantics, discourse analysis, historical linguistics and toponymy.

ORCID

Marie-Elaine van Egmond http://orcid.org/0000-0001-5045-3081
Brett Baker http://orcid.org/0000-0002-7012-4442

Q19 References

Alpher, B., Evans, N., & Harvey, M. (2003). Proto Gunwinyguan verb suffixes. In N. Evans (Ed.), *The Non-Pama-Nyungan languages of northern Australia: Comparative studies of the continent's most linguistically complex region*. Pacific linguistics 552 (pp. 305–352). Pacific Linguistics.

Alpher, B., & Nash, D. (1999). Lexical replacement and cognate equilibrium in Australia. *Australian Journal of Linguistics*, *19*(1), 5–56. https://doi.org/10.1080/07268609908599573

Baker, B. (2008). *Word structure in Ngalakgan*. CSLI Publications.

Baker, B. (2014). Word structure in Australian languages. In H. Koch & R. Nordlinger (Eds.), *The languages and linguistics of Australia* (pp. 139–213). Mouton de Gruyter.

Baker, B. J. (2004). Stem forms and paradigm reshaping in Gunwinyguan. In C. Bowern & H. Koch (Eds.), *Australian languages: Classification and the comparative method*. Current issues in linguistic theory 249 (pp. 313–340). John Benjamins.

Bowern, C., Epps, P., Gray, R., Hill, J., Hunley, K., McConvell, P., & Zentz, J. (2011). Does lateral transmission obscure inheritance in hunter-gatherer languages? *PLoS ONE*, *6*(9): e25195. https://doi.org/10.1371/journal.pone.0025195 Q20

Bundgaard-Nielsen, R. L., & Baker, B. (2014). Frequency in the input affects perception of phonological contrasts for native speakers. In *Proceedings of the fifteenth Australasian international conference on speech science and technology 2014*. [Published electronically.].

Bundgaard-Nielsen, R. L., Kroos, C., Baker, B. J., Best, C. T., & Harvey, M. (2016). Consonantal timing and release burst acoustics distinguish multiple coronal stop place distinctions in Wubuy (Australia). *The Journal of the Acoustical Society of America*, *140*(4), 2794–2809. https://doi.org/10.1121/1.4964399

Campbell, L., & Poser, W. J. (2008). *Language classification: History and method*. Cambridge University Press.

Capell, A. (1942). Languages of Arnhem Land, North Australia. *Oceania; A Journal Devoted To the Study of the Native Peoples of Australia, New Guinea, and the Islands of the Pacific*, *12*(4), 364–392. https://doi.org/10.1002/j.1834-4461.1942.tb00365.x

Catford, J. C. (1977). *Fundamental problems in phonetics*. Edinburgh University Press.

Crowley, T. (1983). Uradhi. In R. M. W. Dixon & B. Blake (Eds.), *Handbook of Australian languages* (pp. 307–428). John Benjamins.

Dixon, R. M. W. (1972). *The dyirbal language of north Queensland*. Cambridge University Press.

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

Dixon, R. M. W. (2002). *Australian languages: Their nature and development*. Cambridge University Press.

Evans, N. (1992). Macassan loanwords in top end languages. *Australian Journal of Linguistics*, *12*(1), 45–91. https://doi.org/10.1080/07268609208599471

Evans, N. (2003). Introduction: Comparative non-Pama-Nyungan and Australian historical linguistics. In N. Evans (Ed.), *The non-Pama-Nyungan languages of northern Australia: Comparative studies of the continent's most linguistically complex region*. Pacific linguistics 552 (pp. 3–25). Pacific Linguistics.

Evans, N. (2005). Australian languages reconsidered: A review of Dixon (2002). *Oceanic Linguistics*, *44*(1), 242–286. https://doi.org/10.1353/ol.2005.0020

Evans, N. (2017). Polysynthesis in Northern Australia. In M. Fortescue, M. Mithun, & N. Evans (Eds.), *The Oxford handbook of polysynthesis* (pp. 312–335). Oxford University Press.

Gafos, A. (1999). *The articulatory basis of locality in phonology* [Outstanding dissertations in linguistics] Garland.

Green, I. (2003a). The genetic status of Murrinh-Patha. In N. Evans (Ed.), *The non-Pama-Nyungan languages of northern Australia: Comparative studies of the continent's most linguistically complex region*. Pacific linguistics 552 (pp. 125–158). Pacific Linguistics.

Green, R. (2003b). Proto Maningrida within Proto Arnhem: Evidence from verbal inflectional suffixes. In N. Evans (Ed.), *The non-Pama-Nyungan languages of northern Australia: Comparative studies of the continent's most linguistically complex region*. Pacific linguistics 552 (pp. 369–421). Pacific Linguistics.

Groote Eylandt Linguistics. (1993). *Eningerribirra-langwa jurra 'book about all sorts of things'*. Groote Eylandt Linguistics.

Hamilton, P. J. (1996). *Phonetic constraints and markedness in the phonotactics of Australian Aboriginal languages* [Doctoral dissertation, University of Toronto]. https://twpl.library.utoronto.ca/index.php/twpl/article/view/6520/3491

Harvey, M. (2002). *A grammar of Gaagudju*. Walter De Gruyter.

Harvey, M. (2003). An initial reconstruction of Gunwinyguan phonology. In N. Evans (Ed.), *The non-Pama-Nyungan languages of northern Australia: Comparative studies of the continent's most linguistically complex region*. Pacific linguistics 552 (pp. 205–268). Pacific Linguistics.

Harvey, M. (2011a). Lexical change in pre-colonial Australia. *Diachronica*, *28*(3), 345–381. https://doi.org/10.1075/dia.28.3.03har

Harvey, M. (2011b). Prepalatals in Arandic. *Australian Journal of Linguistics*, *31*(1), 79–110. https://doi.org/10.1080/07268602.2011.532858

Harvey, M. (2012). Warndarrang and Marra: A diffusional or genetic relationship? *Australian Journal of Linguistics*, *32*(3), 327–360. https://doi.org/10.1080/07268602.2012.705578

Harvey, M., & Mailhammer, R. (2017). Reconstructing remote relationships: Proto-Australian noun class prefixation. *Diachronica*, *34*(4), 470–515. https://doi.org/10.1075/dia.15032.har

Heath, J. (1978a). *Linguistic diffusion in Arnhem Land*. AIAS research and regional studies 13. Australian Institute of Aboriginal Studies.

Heath, J. (1978b). *Ngandi grammar, texts, and dictionary*. Australian Institute of Aboriginal Studies.

Heath, J. (1981). A case of intensive lexical diffusion: Arnhem Land, Australia. *Language*, *57*(2), 335–367. https://doi.org/10.2307/413694

Heath, J. (1982). *Nunggubuyu dictionary*. Australian Institute of Aboriginal Studies.

Heath, J. (1984). *Functional grammar of Nunggubuyu*. Australian Institute of Aboriginal Studies.

Heath, J. (1990). Verbal inflection and macro-subgrouping of Australian languages: The search for conjugation markers in non-Pama-Nyungan. In P. Baldi (Ed.), *Linguistic change and reconstruction methodology* (pp. 403–417). Mouton.

Heath, J. (1997). Lost wax: Abrupt replacement of key morphemes in Australian agreement complexes. *Diachronica*, *14*(2), 197–232. https://doi.org/10.1075/dia.14.2.02hea

Heath, J. (n.d). *Anindhilyakwa grammatical sketch*. Australian Institute of Aboriginal Studies.

Henderson, J. (1998). *Topics in Eastern and Central Arrernte grammar* [Doctoral dissertation, University of Western Australia].

Hock, H. H., & Joseph, B. D. (1996). *Language history, language change, and language relationship: An introduction to historical and comparative linguistics*. Mouton de Gruyter.

Jukes, A. (2006). *Makassarese (basa Mangkasara'): A description of an Austronesian language of South Sulawesi* [Doctoral dissertation, University of Melbourne].

Koch, H. (2014). Historical relations among the Australian languages: Genetic classification and contact-based diffusion. In H. Koch & R. Nordlinger (Eds.), *The languages and linguistics of Australia vol. 2*. (The world of linguistics) (pp. 23–90). De Gruyter Mouton.

Kümmel, M. J. (2007). *Konsonantenwandel*. Reichert.

Ladefoged, P. (2001). *Vowels and consonants*. Blackwell.

Leeding, V. (1989). *Anindilyakwa phonology and morphology* [Doctoral dissertation, University of Sydney].

Mailhammer, R., & Harvey, M. (2018). A reconstruction of the Proto-Iwaidjan phoneme system. *Australian Journal of Linguistics*, *38*(3), 329–359. https://doi.org/10.1080/07268602.2018.1470455

Meillet, A. (1967). *The comparative method in historical linguistics*. Librairie Honoré Champion.

Miceli, L. (2008). Morphological reconstruction and Australian languages. In C. Bowern, B. Evans, & L. Miceli (Eds.), *Morphology and language history: In honour of Harold Koch* (pp. 211–219). John Benjamins.

Moody, M. L. A. (1954). *A descriptive statement of the phonemics and morphology of Anindilyaugwa, the language of Groote Eylandt, Northern Territory* [MA thesis, University of Sydney].

O'Grady, G. N. (1979). Preliminaries to a proto nuclear Pama-Nyungan stem list. In S. A. Wurm (Ed.), *Australian linguistic studies*. Pacific linguistics C54 (pp. 107–139). Pacific Linguistics.

O'Grady G, N., Wurm, S. A., & Hale, K. L. (1966). *[Map of] Aboriginal languages of Australia (a preliminary classification)*. Department of Linguistics, University of Victoria.

Shaw, J. A., Carignan, C., Agostini, T. G., Mailhammer, R., Harvey, M., & Derrick, D. (2020). Phonological contrast and phonetic variation: The case of velars in Iwaidja. *Language*, *96*(3), 578–617. https://doi.org/10.1353/lan.2020.0042

Smith, I., & Johnson, S. (2000). Kugu Nganhcara. In R. M. W. Dixon & B. Blake (Eds.), *A handbook of Australian languages* (Vol. 5, pp. 357–489). Oxford University Press.

Steriade, D. (2001). Directional asymmetries in place assimilation: A perceptual account. In E. Hume & K. Johnson (Eds.), *The role of speech perception in phonology* (pp. 219–250). Academic Press.

Stokes, J. (1981). Anindilyakwa phonology from phoneme to syllable. In B. Waters (Ed.), *Work papers of SIL-AAB series A volume 5. Australian phonologies: Collected papers* (pp. 139–181). SIL.

Swadesh, M. (1952). Lexico-statistic dating of prehistoric ethnic contacts: With special reference to North American Indians and Eskimos. *Proceedings of the American Philosophical Society*, *96*, 452–463.

Swadesh, M. (1956). Some limitations of diffusional change in vocabulary. *American Anthropologist*, *58*(2), 301–306. https://doi.org/10.1525/aa.1956.58.2.02a00540

Tadmor, U., Haspelmath, M., & Taylor, B. (2010). Borrowability and the notion of basic vocabulary. *Diachronica*, *27*(2), 226–246. https://doi.org/10.1075/dia.27.2.04tad

Turner, D. H. (1974). *Tradition and transformation: A study of Aborigines in the Groote Eylandt area, northern Australia*. Australian Institute of Aboriginal Studies.

Turner, D. H. (1999). An aboriginal outstation movement in Arnhem Land and the Perils of advocacy anthropology. *Nomadic Peoples*, *3*(1), 8–21. https://doi.org/10.3167/082279499782409514

Van Egmond, M.-E. (2012). *Enindhilyakwa phonology, morphosyntax and genetic position* [Doctoral dissertation, University of Sydney]. http://hdl.handle.net/2123/8747;oai:ses.library.usyd.edu.au:2123/8747

Vennemann, T. (1988). *Preference laws for syllable structure and the explanation of sound change*. Mouton de Gruyter.

Verstraete, J.-C. (2005). The semantics and pragmatics of composite mood marking: The non-Pama-Nyungan languages of northern Australia. *Linguistic Typology*, *9*, 223–268. https://doi.org/10.1515/lity.2005.9.2.223

Waddy, J. (1988). Classification of plants & animals from a Groote Eylandt aboriginal point of view, volume 2. In *Australian National University North Australia research unit monograph*. Darwin.

Yallop, C. (1982). *Australian aboriginal languages*. Andre Deutsch.

Zorc, R. (1986). *Yolngu-matha dictionary*. School of Australian Linguistics, Darwin Institute of Technology.