# The diachrony of complex verb structures in northern Australia

Geordie Kidd,[1] Mark Harvey,[2] Robert Mailhammer[3]

[1]The University of Melbourne
[2]The University of Newcastle
[3]Western Sydney University

ALS 2025 · 4 December

# Complex verbs in northern Australia

► In the prefixing languages of northern Australia, there are commonly two kinds of verb structures: a simple verb, and a complex verb (Baker & Harvey, 2010; Bowern, 2004, 2014; McGregor, 2002; Schultze-Berndt, 2000; Wilson, 1999)

► Simple verb has single verb root with prefixes and suffixes

► Complex verbs involve two roots: finite verb root and 'coverb' root

► These complex verbs come in two main morphosyntactic configurations: **phrasal** (common) and **compound** (rare)

## Examples from Mangarrayi

▶ Simple and complex verbs can both be illustrated in Mangarrayi:

(1) a. [ŋar-**pu**-p]_Wd (Simple)
1TR>3SG-**hit**-PP
Prefix-**Verb.Root**-TAM
'We (three) hit it.'

   b. [**tum**]_Wd [ŋa-**pu**-p]_Wd (Phrasal)
      **split** 1SG>3SG-**HIT**-PP
      **Coverb** Prefix-**Verb.Root**-TAM
      'I split it.'

   c. [ŋawujan-**cawaŋ+pu**-p]_Wd (Compound)
      1SG>3PL-**answer+HIT**-PP
      Prefix-**Coverb+Verb.Root**-TAM
      'I answered them.'

## Aims of the study

- ▶ In all prefixing languages (except Mangarrayi) only one complex verb structure predominates and the other is either absent or restricted to minor functions (e.g., Kidd, 2024; Mansfield, 2019)
- ▶ Harvey and Mailhammer (2024, pp. 153–155) reconstruct the compound type to Proto-Australian, based on evidence of remnantal compounds in languages where only phrasal verbs predominate synchronically (see also Merlan, 1979)
- ▶ Where do phrasal verbs fit in? What is the diachronic relationship between phrasal verbs and compounds?

In languages where phrasal verbs are the predominant structure synchronically, but which have evidence of compounds historically, what process(es) led to the loss of compounds?

# Mangarrayi

► Mangarrayi (non-Pama-Nyungan isolate) (Merlan, 1982) has substantial quantities of phrasal verbs (N = 763) and compound verbs (N = 336)
► Offers a good testing ground for hypotheses about the diachrony of these structures
► Data drawn from grammar (Merlan, 1982), texts (Merlan, 1996), and unpublished dictionary (Merlan, n.d.)
► Examination of synchronic state of the grammar to infer likely historical changes, i.e., INTERNAL RECONSTRUCTION (Hock, 2021; Joseph, 2010)

Mangarrayi verbs and hypotheses

Data and evidence

Diachronic pathways

Conclusion

# MANGARRAYI VERBS AND HYPOTHESES

## Breakdown of verb types in Mangarrayi

► 36 simple finite verbs: /pu/ 'hit', /mi/ 'get', /ka/ 'take', etc.

► 336 Compound Verbs (CVCs)

► 763 Phrasal Verbs (PVCs)

(2) ka-ŋa-**kuɻ+wa**-n
PR-1SG>3SG-circle+VISIT-PR
'I'm circling (the tree).' (Merlan, n.d.)

(3) **taʔma** wur-**ka**-ŋiɲ
lift.up 3PL>3SG-TAKE-PP
'They lifted it up.' (Merlan, n.d.)

Quantities of parts

|  | CVC | PVC |
|---|---|---|
| N lexemes | 336 | 763 |
| Coverbs | 319 | 614 |
| Finite verbs | 15 | 44 |

## Differing characteristics between types

▶ PVCs are productive, responsible for adopting loans, causative alternations, and inchoative/factitive derivations, etc.

(4) Loaned verbs in Mangarrayi

    a. **ʈɻʌŋk** wuḻa-ma
       drunk 3PL-DO.PR
       'They're drunk.' (adapted from Merlan, 1982, p. 129)

    b. **klinim?** ŋa-ma-ɻi-wa      patɻum
       clean    1SG>3SG-DO-PC-FOC  bathroom
       'I cleaned the bathroom.' (Merlan, 1996, p. 65)

▶ CVCs are unproductive, not active in any domain of lexeme formation
(unlike compound-dominant languages; see Baker, 2008; Evans et al., 2004;
Saulwick, 2003)

## How did phrasal verbs come to predominate?

► So: PVCs are predominant and productive, while CVCs are far less numerous and unproductive

► There are two plausible hypotheses about how PVCs came to predominate in Mangarrayi:

H1 Compounds were inherited as a productive structure; phrasal verbs were acquired later on, and their expansion displaced compounds

H2 Both compounds and phrasal verbs were inherited as productive structures; compounds became lexicalised, subsequently closing and allowing phrasal verbs to predominate

# Hypotheses and predictions

▶ H1, the INNOVATION HYPOTHESIS, predicts that lexicalisation of compounds occurred **after** the expansion of phrasal verbs

▶ H2, the INHERITANCE HYPOTHESIS, predicts that this lexicalisation occurred **before** the expansion of phrasal verbs

▶ The synchronic verb system of Mangarrayi is simply explained under H2, but difficult to explain under H1

# DATA AND EVIDENCE

# Three sources of evidence

▶ There are three pieces of evidence that support H2:

1. CVCs are largely non-compositional, whereas phrasal verbs are much more compositional
2. Compound stems pattern like monomorphemic finite verbs
3. Compound stems can be used as independent coverbs

# 1. Compounds are non-compositional

▶ The compound system is largely non-compositional, whereas the phrasal system is largely compositional

(5) ca-**man+pu**-n
    3.PR-run+HIT-PR
    'He is running.' (Merlan, 1996, p. 10)

(6) **tuŋkurtuŋkur** ka-ŋa-**pu**-n
    crack            PR-1SG>3SG-HIT-PR
    'I am cracking it.' (Merlan, n.d.)

▶ Excluding tokens with 'say/do' finite verb, there are 102 potentially compositional CVCs, and 545 such PVCs

|  | Compound | Phrasal |
|---|---|---|
| Potentially compositional | 102 | 545 |
| Actually compositional | 33 | 335 |
| Proportion | **32.4%** | **61.5%** |

▶ The greater degree of lexicalisation in compounds indicates longer period of unproductivity, supporting the hypothesis that lexicalisation preceded the expansion of phrasal verbs (H2)

## 2. Compounds used as finite verbs

▶ In PVCs, 19/44 finite verbs (43.2%) are formally identifiable as compounds

(7) a. ca-**man+pu**-n
   3.PR-run+HIT-PR
   'He is running.'

   b. **pulkuc**   ca-**man+pu**-n
   rush.out   3.PR-run+HIT-PR
   'He is rushing out.' (Merlan, n.d.)

▶ Examples like (7b) are not different to 'regular' phrasal verbs (i.e., don't have distinct pragmatic interpretation)

## 2. Compounds used as finite verbs

(8) a. **men** ŋaɻa-**jiri+wa**-n ciku-nawu
watch 1PL.EX>3SG-look+VISIT-PR road-3SG.POSS
'We're watching his road.' (Merlan, n.d.)

  b. **naɲcar** ŋiɻa-**paʈ+namti**-wa
save 1PL.IN>3SG-put+HOLD.PC-NAR
'We saved it.' (Merlan, n.d.)

▶ The use of compounds in this capacity is enabled by their lexicalisation and loss of internal morphological structure; they are effectively like monomorphemic finite verbs

▶ This favours H2, since it requires compounds to have been lexicalised at the time new phrasal verbs were being coined

## 3. Positional alternations in coverbs

▶ Further evidence for the lexicalisation of compounds comes from patterns of alternation between compound and phrasal structures

▶ About 60 coverbs can occur either internally as part of a CVC structure, or externally as part of a PVC structure

(9) a. wuɭa-**tet+mi**-ji-ni
   3PL-paint+DO-RR-PC
   'They used to paint themselves.' (Merlan, 1996, p. 28)

   b. **tet**   wuɭa-**mi**-ji-ni
   paint   3PL-DO-RR-PC
   = (9a)

▶ Unproductive or frozen, sometimes with unpredictable differences in meaning (Merlan, 1982, p. 128f.)

## 3. Compounds used as coverbs

▶ In some cases, it is not the coverb alone that appears excorporated; rather, it is the entire [Coverb+Verb] compound:

(10) a. wur-**paɹap+ma**-ɹi
       3DU-look.back+DO-PC
       'They looked back.' (Merlan, 1996, p. 82)

    b. **paɹapma**   ŋan-jiri+wa-w!
       look.back  2SG>1SG-look+VISIT-IMP
       'Look back at me!' (Merlan, n.d.)

    c. *__**paɹap**__   ŋan-jiri+wa-w!

▶ No synchronic suffix /-ma/ (cf. Wagiman, Wilson, 1999, p. 51; Wardaman, Merlan, 1994, p. 263)

▶ There are several others (~20), some of which don't have corresponding compounds

| | | | | |
|---|---|---|---|---|
| [kalʔma]$_{Cov}$ [ni-]$_V$ | 'climb' | < | [kalʔ+ma-] | 'climb' |
| [caʈma]$_{Cov}$ [cajki-]$_V$ | 'go back' | < | [caʈ+ma-] | 'go back' |
| [jurmi]$_{Cov}$ [jak-]$_V$ | 'pass by' | < | [jur+ma-] | 'pass by' |
| [pelʔma]$_{Cov}$ [ni-]$_V$ | 'set (sun)' | < | — | |
| [cawʔma]$_{Cov}$ [ka-]$_V$ | 'grab' | < | — | |
| [waʔma]$_{Cov}$ [waji-]$_V$ | 'crouch' | < | — | |

▶ We can infer a historical morpheme boundary, since the glottal stop is otherwise restricted to morpheme edges (especially in coverbs)

▶ Thus, speakers must have reanalysed [Coverb+Verb] compounds as monomorphemic prior to coining new phrasal verbs with them, which affirms the predictions of H2 but not H1
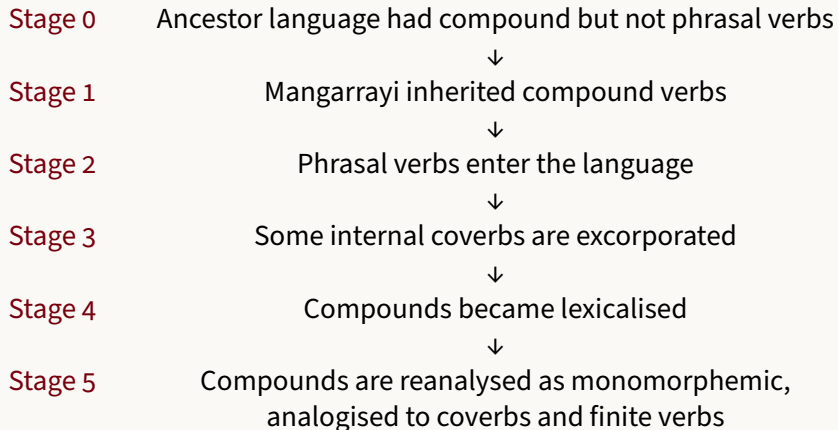
# Summary of evidence

▶ Compounds have a much lower degree of compositionality than phrasal verbs

▶ Certain compounds get used like monomorphemic light verbs in phrasal verb constructions

▶ Excorporation of compounds as coverbs also shows loss of morphological complexity

▶ Compounds behave synchronically like monomorphemic predicating elements: they can be light verbs (like simple finite verbs), and they can be external coverbs
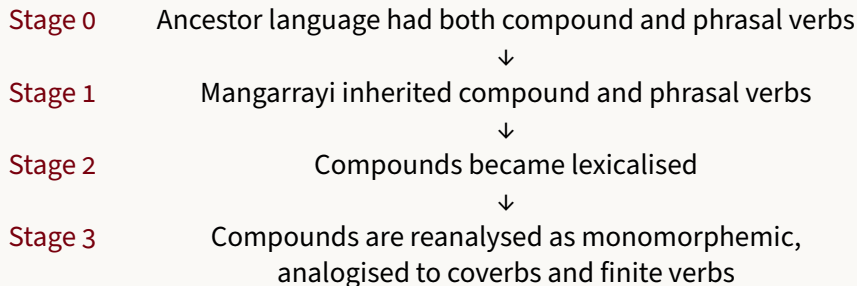
# DIACHRONIC PATHWAYS

## Evaluating the hypotheses

► To evaluate the hypotheses, we should see how well they account for the data just presented

H1 Compounds were inherited as a productive structure; phrasal verbs were acquired later on, and their expansion displaced compounds

H2 Both compounds and phrasal verbs were inherited as productive structures; compounds became lexicalised, subsequently closing and allowing phrasal verbs to predominate

► Let's compare the diachronic pathways implied by H1 and H2

# The innovation pathway (H1)

| | |
|---|---|
| Stage 0 | Ancestor language had compound but not phrasal verbs |
| | ↓ |
| Stage 1 | Mangarrayi inherited compound verbs |
| | ↓ |
| Stage 2 | Phrasal verbs enter the language |
| | ↓ |
| Stage 3 | Some internal coverbs are excorporated |
| | ↓ |
| Stage 4 | Compounds became lexicalised |
| | ↓ |
| Stage 5 | Compounds are reanalysed as monomorphemic, analogised to coverbs and finite verbs |

Stage 0     Ancestor language had both compound and phrasal verbs

↓

Stage 1     Mangarrayi inherited compound and phrasal verbs

↓

Stage 2     Compounds became lexicalised

↓

Stage 3     Compounds are reanalysed as monomorphemic,
analogised to coverbs and finite verbs

The inheritance pathway has several advantages:

1. Relies on fewer, and only widely-attested, historical changes (i.e., lexicalisation, analogy)
2. Alternating coverbs that aren't formally compounds are retentions from Stage 1 (compounds and phrasal verbs were in alternation)
3. 'Compound' coverbs are more recent coinages, explainable as reanalysis of stem boundaries in frozen compounds

# CONCLUSION

## Conclusion

▶ Phrasal verbs in Mangarrayi appear to be a retention from an immediate ancestor

▶ {INHERITANCE → LEXICALISATION → REPLACEMENT} is a possible pathway to a predominantly phrasal system

▶ For future work: assuming Proto-Australian is the immediate ancestor of Mangarrayi (Harvey & Mailhammer, 2024, pp. 290–294), then reconstruction of phrasal verbs is supported

▶ If Mangarrayi had fully lost compounds, it might look much more like other phrasal verb languages

# Thank you!*

## References

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

Baker, B., & Harvey, M. (2010). Complex predicate formation. In M. Amberber, B. Baker, & M. Harvey (Eds.), *Complex predicates: Cross-linguistic perspectives on event structure* (pp. 13–47). Cambridge University Press.

Bowern, C. (2004). *Bardi verb morphology in historical perspective* [Doctoral dissertation, Harvard University].

Bowern, C. (2014). Complex predicates in Australian languages. In H. Koch & R. Nordlinger (Eds.), *The languages and linguistics of Australia: A comprehensive guide* (pp. 263–294). De Gruyter Mouton.

Evans, N., Merlan, F., & Tukumba, M. (2004). *A first dictionary of Dalabon (Ngalkbon)*. Maningrida Arts and Culture.

Harvey, M., & Mailhammer, R. (2024). *Proto-Australian: Reconstruction of a common ancestor language*. De Gruyter Mouton.

Hock, H. H. (2021). *Principles of Historical Linguistics* (3rd ed.). De Gruyter Mouton.

# References

Joseph, B. D. (2010). Internal reconstruction. In S. Luraghi & V. Bubenik (Eds.), *The continuum companion to historical linguistics* (pp. 52–58). Bloomsbury.

Kidd, G. (2024). *Excorporation* [BA thesis]. University of Melbourne.

Mansfield, J. (2019). *Murrinhpatha morphology and phonology*. De Gruyter Mouton.

McGregor, W. (2002). *Verb classification in Australian languages*. De Gruyter Mouton.

Merlan, F. (1979, Summer). On the Prehistory of Some Australian Verbs. *Oceanic Linguistics*, *18*(1), 33. https://doi.org/10.2307/3622837

Merlan, F. (1982). *Mangarayi*. North-Holland Publishing Company.

Merlan, F. (1994). *A grammar of Wardaman: A language of the Northern Territory of Australia*. De Gruyter Mouton.

Merlan, F. (1996). *Big river country: Stories from Elsey Station. Told by Amy Dirngayg, Betty Lardy, Jackeroo Lirrawi, Jessie Roberts, Lulu Jilimbirrnga, Mary Nurniyn, Rita Morgan, Sheila Conway, Splinter Harris*. IAD Press.

## References

Merlan, F. (N.d.). *Mangarrayi dictionary draft.*

Saulwick, A. (2003). *Aspects of the verb in Rembarrnga, a polysynthetic language of northern Australia: Grammatical description, texts and dictionary* [Doctoral dissertation, University of Melbourne].

Schultze-Berndt, E. (2000). *Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language* [Doctoral dissertation, Katholieke Universiteit Nijmegen].

Wilson, S. (1999). *Coverbs and complex predicates in Wagiman*. CSLI Publications.