School of Linguistics and Applied Language Studies

Linguistic travels in time and space: Festschrift for Liz Pearce

Wellington Working Papers in Linguistics
Volume 23, 2017

School of Linguistics and Applied Language Studies

Linguistic travels in time and space: Festschrift for Liz Pearce

Wellington Working Papers in Linguistics

Volume 23, 2017

ISSN 1170-1978 (Print)
ISSN 2230-4681 (Online)

TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI
VICTORIA
UNIVERSITY OF WELLINGTON

Linguistic travels in time and space: Festschrift for Liz Pearce

Wellington Working Papers in Linguistics

Volume 23, 2017

Edited by Heidi Quinn, Diane Massam, and Lisa Matthewson

School of Linguistics and Applied Language Studies
Victoria University of Wellington
P.O. Box 600
Wellington
New Zealand

Published 2017

Back cover photo: Diane Massam

ISSN 1170-1978 (Print)
ISSN 2230-4681 (Online)

Linguistic travels in time and space: Festschrift for Liz Pearce

Wellington Working Papers in Linguistics
Volume 23, 2017

CONTENTS

Editorial note

Volume 23 of the WWPL is dedicated to Liz Pearce on her retirement from many years of service in the School of Linguistics and Applied Language Studies at Victoria University of Wellington. For this special occasion we decided to relax the WWPL policy guidelines and invite contributions from linguists around the world who have drawn inspiration from Liz's work. The resulting Festschrift brings together papers from the wide range of subject areas spanned by Liz's research interests: historical linguistics, Romance linguistics, syntactic theory, Austronesian linguistics, and language documentation. We realise that our decision to keep the Festschrift a surprise has meant that our invitation could only reach a small subset of those who have benefitted from interacting with Liz. Even so, the diversity of the contributions is a fitting tribute to the many ways in which Liz has enriched our lives and the field of linguistics, both as a scholar and a friend.

Thank you, Liz, from all of us who have signed the tabula congratulatoria and no doubt many others.

Heidi Quinn, Diane Massam, Lisa Matthewson

August 2017

Tabula congratulatoria

"Non sta scritto da nessuna parte che non ce la puoi fare."
— Elena Ferrante, *L'amica geniale*

The following colleagues and friends of Elizabeth Pearce would like to express their best wishes on the occasion of her retirement from Victoria University of Wellington. Thank you for all your inspiration, collegiality, and friendship – and Congratulations, Liz!

Ngā mihi nui me te aroha nui!

Toutes nos félicitations à l'occasion de votre retrait!

Complimenti per i tuoi contributi passati e i nostri auguri più sinceri per quelli futuri!

En joie et santé li ciels vous tienne!

Numkut ge xo re wi kala nene Elizabeth Pearce re honta reptu nikia naha lele Tabula Gratulatoria nge tele niye wi nahaine tuwa daras pesei pa. Dame. (Ninde)

Nimorot tieng ahr re noh halan Elizabeth Pearce re rong retu nehen ahr len Tabula Gratulatoria tieng gen nerorongien ti ahr gen ei. Ileh. (Nahavaq)

Batin re naxerr sen imavos ni buravi. (Unua: Pearce 2015, 478:50 (Mark 4:29))

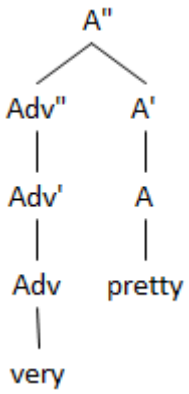
Edith Aldridge
Julie Barbour
Laurie Bauer
Winifred Bauer
Sigrid Beck
Adriana Belletti
Claire Bowern
Mary Boyce
Dave Britain
Andrea Calabrese
Sasha Calhoun
Henry Yungli Chang
Helen Charters
Sandy Chung
Guglielmo Cinque
Peter Cole
Nicola Daly
William Davies
Paul de Lacy
Laura Dimock
Robert Early
Alex Francois
Mark Hale
Ray Harlow
Jennifer Hay
Gabriella Hermon
Paul Hirschbühler
Hans Henrich Hock
Janet Holmes
Robin Hooper
Leina Isno
Dianne Jonas
Richard Kayne
Kate Kearns
Michael Kenstowicz
Alistair Knott
Marie Labelle
Paul Law
John Lynch
John Macalister
Meredith Marra
France Martineau
Leisel Masinglow
Diane Massam
Lisa Matthewson
Ross McKerras
Timothy Mckinnon
Miriam Meyerhoff
Claire Moyse Faurie
Yuko Otsuka
Jill Musgrave
Bill Palmer
Ileana Paul
Andrew Pawley
Matt Pearson
Maria Polinsky
Heidi Quinn
Martin Paviour-Smith
Eric Potsdam
Norvin Richards
Luigi Rizzi
Yves Roberge
Malcolm Ross
Mario Saltarelli
Cindy Schneider
Corinne Seals
Yono Sukarno
Jae Jung Song
Martin Takac
Lynsey Talagi
Helen Tamtam
Melenaite Taumoefolau
Nick Thieberger
Shizuka Torii
Lisa Travis
Michelle Troberg
Bernadette Vine
Elaine Vine
Dieter Wanner
Paul Warren
Yanti
Kie Ross Zuraw

How can you put *Liz* into a tree?

Laurie Bauer

In an X'-grammar, the metarules on the formalisation of rules require that any phrasal node is rewritten as one node with the same number of bars as the dominating node or one less, and otherwise only phrasal categories. Since most X'-bar grammars also have a requirement of binarity, this means that any phrasal node must be rewritten as one (optional) phrasal category and a category (which may or may not be phrasal) which includes the head of the construction. That is, we find trees like that in (1):

1.



If we extend this to noun phrases, then any NP has to be rewritten with structures that eventually contain an N (if it was not, it could not be an NP). This means that if we want to draw a tree for the sentence in (2) we have to have an NP in the subject which dominates (directly or indirectly) the N *Liz,* which is precisely what is done by Jackendoff (1987) and, at least implicitly, by Longobardi (1994).

2. Liz is retiring.

The difficulty with this is that if we look at the distribution of *Liz*, it is not an N but an NP (Payne & Huddleston 2002: 516).[1] Consider the substitution tables in (3) and (4).

3.

Liz She The person who lectures in syntax A syntax lecturer	is retiring

[1] Payne & Huddleston (2002) do not use a DP analysis, but have determiners as modifiers within the NP; for those who do use DPs, it might be more accurate to say names are DPs.

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 1–4

4.

A This Our[2]	Ø valued fit[3]	colleague *she *Liz	is retiring

The problem is that there is no way in an X'-grammar that a lexical item like *Liz* can belong to a phrasal category, and yet this seems to be what is required.

There are alternative solutions. Perhaps *Liz* is not an NP but a NameP, so that *Liz* is a Name. This still does not resolve the problem that *Liz* seems to act as a phrase rather than as an element within a phrase, though it does help with the fact that, in English, personal names like *Liz* do not in general take determiners: a DP could allow an NP but not a NameP within it. But if we assume the word-class theory of Chomsky (1970), there are only four word-classes available: noun, verb, adjective and adposition, defined by the feature set [±N], [±V]. In such a model Name cannot be a new word-class (and neither, actually, can adverb, to which appeal was made in (1)).

So perhaps name is a feature on an N rather than a word-class in its own right. The feature would bar the use of modifiers with names like *Liz*. That might work if NP included determiners, but under a DP analysis it would have the presence or absence of a determiner being provoked by a feature on a node in the non-head position of the DP.

There is another set of facts, however, which might be taken to imply that the usage illustrated in (2) is atypical, and that items like *Liz* are really nouns after all (this is the conclusion drawn by Longobardi 1994: 636). The relevant constructions are like those in (5).

5. a. The Liz we know and love.
   b. There is no Liz in this house.
   c. Do you mean the Liz in linguistics or the one who was married to Richard Burton?
   d. A young Liz stared out of the photograph.
   e. To solve this problem, we need another Liz.
   f. I've never met a Liz before.

[2] In the English of the area in which I grew up, *Our Liz* is grammatical in the sense 'the Liz who is a member of a group to which the speaker also belongs, especially the family'. This is not standard usage. In any case, it probably changes the category of *Liz* – see below.

[3] It might be objected that some adjectives would be perfectly in order here provided there was no determiner: *Poor Liz*, for instance, would be possible. I suspect this is a different construction. Payne & Huddleston (2002: 520) term things like *poor* here 'embellishments'. It is not clear how embellishments differ from the type of example shown in (5) in terms of grammatical structure. The asterisk also assumes that there is only one person called 'Liz' in the universe of discourse; otherwise the sentence becomes like one of those illustrated in (5): acceptable, but a different construction.

Payne & Huddleston (2002: 520) talk here of 'secondary uses' of proper names, where the inherent definiteness normally associated with names is lost. Bauer et al. (2013: 558) talk about this being a case of type coercion, and say that *Liz* in examples like those in (5) is a common noun, no longer a name.

If that is the answer, the question becomes where the type coercion arises. One strong possibility is that it arises from the syntactic tree in which the item is placed. If *Liz* is placed in a tree where it has determiners or modifiers (pre- or post-), then it is a noun; if it arises in a tree with no other constituent of the DP/NP it is a name. This is clearly a workable solution, but it still says that a name is the unique constituent of an N' which is the unique constituent of an N" (which may be the unique constituent of a D"). It does not allow us to say that a name is an NP (or possibly a DP, depending upon your theoretical orientation).

There is dispute in the literature as to whether names like *Liz* have any sense, can have a definition or just have reference (see Searle 1971). To a certain extent a position taken on this issue has an implication for whether *Liz* in a sentence like (2) and *Liz* in sentences like those in (5) belong to distinct categories (Lehrer 1994) or more or less canonical types of the same category (Ullman 1957: 74). Jespersen (1924: 67) suggests that the forms in (5) are figurative interpretations of a form which has a literal meaning in (2).

The fact that *Liz* in (2) refers to a single individual, while *Liz* in most of the examples in (5) refers to a class of individuals who have in common that they can all be called 'Liz' ((5e) is rather different, and picks out a quality in Liz which we wish to replicate in another person), goes a long way to explaining the different grammar. But in English, at least, the distinction is an absolute one. The same is not true in, for instance, German, where names like *Liz* can be used with a definite article (with certain stylistic implications), as in (6).

6. Die Liz ist erst 1987 nach Wellington gekommen.
   the Liz is first 1987 to Wellington come
   'Liz didn't come to Wellington until 1987.'

In a language like German, then, there seems to be more reason to treat a form like *Liz* as a noun; in a language like English it is primarily something which substitutes for a category larger than N, and overlaps with N only in secondary uses.

I have presented this argument in terms of X'-grammar in particular, but in fact I know of no theory of phrase structure which allows single words to be phrasal categories. Rather, what happens in phrase structure is that a phrasal category gets rewritten as a single terminal category. But this fails to distinguish between an item like *Liz* which is, canonically, a whole phrase, and an items like *lecturers* which may be a whole phrase, but which can also have other items in the phrase that it heads. Only in dependency grammars – where phrases are derived notions: the head and all its dependents – are the two plausibly treated the same. As I have shown here, this is a problem for names. It is also a problem for pronouns. Given our theoretical assumptions, it seems that neither fit properly into phrase structure trees.

References

Bauer, Laurie & Lieber, Rochelle & Plag, Ingo. 2013. *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.

Chomsky, Noam. 1970. Remarks on nominalization. In Jacobs, Roderick A. & Rosenbaum, Peter S. (eds.) *Readings in English transformational grammar*, 184–221.Waltham, MA: Ginn.

Jackendoff, Ray. 1987. *X' syntax: A study of phrase structure*. Cambridge, MA: MIT Press.

Jespersen, Otto. 1924. *The philosophy of grammar*. London: Allen & Unwin.

Lehrer, Adrienne. 1994. Proper names: Linguistic aspects. In Asher, Ronald E. (ed.) *The encyclopedia of language and linguistics*, vol 6, 3372–3374. Oxford: Pergamon.

Longobardi, Giuseppe. 1994. Reference and proper names: A theory of N-movement in syntax and logical form. *Linguistic Inquiry* 25. 609–665.

Payne, John & Huddleston, Rodney. 2002. Nouns and noun phrases. In Huddleston, Rodney & Pullum, Geoffrey K. (eds.) *The Cambridge grammar of the English language*, 323–523. Cambridge: Cambridge University Press.

Searle, John A. 1971. The problem of proper names. In Steinberg, Danny D. & Jakobovits, Leon A. (eds.) *Semantics*, 134–141. Cambridge: Cambridge University Press.

Ullman, Stephen. 1957. *The principles of semantics*. 2nd edn. Oxford: Blackwell and Glasgow: Jackson.

Laurie Bauer
Victoria University of Wellington
laurie.bauer@vuw.ac.nz

An alternative semantic cycle for universal quantifiers

Sigrid Beck

Abstract

This squib investigates a diachronic semantic development concerning universal quantifiers. Haspelmath (1995) shows that a cross-linguistically frequent historical source of universal quantification is free choice relative clauses. This development necessitates a complex compositional semantic change whose individual steps and components are as yet unexplored. I propose a hypothesis, the AltSemCycle, according to which the free choice relative and the standard universal quantifier are two stages in a cycle (Jespersen 1917) or spiral (Gergel 2016), and I make some first proposals towards spelling out the compositional semantic steps that get us from one stage of the cycle to the next.

1. The issue

The diachronic development that we will explore is showcased in (1). A free choice relative clause FCR like (1a), which contains a wh-expression and an element like *ever*, is a primary source of standard universal quantifiers like (1b). Haspelmath (1995) collects an impressive sample of cross-linguistic and diachronic data that make this point. The diachronic development is well attested for English and German universal quantifiers like *each*, *every* and *jeder* 'every' (e.g. Kluge 1995). We concentrate on English here, for ease of exposition.

1. a. I will read [whichever book you recommend]. FCR
   b. I will read every book (that you recommend). universal DP

Under the view that the overall sentence meaning remains stable between the steps of a change (this is discussed concretely as *Constant Entailments* in Beck (2012) and Beck & Gergel (2015); see Eckardt (2009) for a broader perspective), this development is quite intuitive: both (1a) and (1b) universally quantify over a relevant set of books, as sketched informally in (2).

2. For all x, x a relevant book: I will read x

However, the way that this overall sentence meaning comes about is very different. Let's assume the standard meaning for *every* given in (3) (e.g. Barwise & Cooper 1981; Heim & Kratzer 1998). According to (3), the word *every* combines with two properties and says that every individual that has the first property also has the second. This lexical entry directly leads to the meaning paraphrased in (2) for sentence (1b) (with P

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 5–13

being identified as the property of being a book and Q being identified as the property of being read by me).

3. [[every]] = $\lambda P.\lambda Q.\forall x[P(x) \rightarrow Q(x)]$

Example (1a) is more complex. I adopt concretely a recent analysis by Hirsch (2015) which is sketched in (4).

4. For all alternative propositions p:
   If p is the proposition that you recommend x (x a book) and p is true,
   then I will read the book that you recommend.

The sketch in (4) is very informal, but it serves to illustrate the problem: in the FCR, there is universal quantification. But the quantification is over alternative propositions (not individuals), and it is not expressed by an overt morpheme. The wh-expression *which(ever) book* introduces alternatives (Hamblin 1973), alternative books in the example. Those alternatives project to the propositional level, yielding alternative propositions 'you recommend book x'. In an FCR these propositions go into the antecedent of a conditional, which gives the FCR its modal flavour. The overall semantics amounts to this: If you recommend *West Coast Walking - A Naturalist's Guide*, I will read *West Coast Walking - A Naturalist's Guide*. If you recommend *An Introduction to Modal Logic*, I will read *An Introduction to Modal Logic*. If you recommend *A Grammar of Unua*, I will read *A Grammar of Unua*. And so on. So the question is, how does the grammar have to change in order to get from covert universal quantification over alternative propositions (triggered by a wh-expression), as in (4), to overt universal quantification over individuals, as in (2),(3)?

2. The idea of the cycle

It is clear that one does not get from (4) to (2),(3) with one change in the grammar. Rather, a cascade of smaller changes is required which jointly bring about the development to standard universal quantification. We certainly expect intermediate stages, for which it would be good to identify present-day cross-linguistic counterparts.

Second, it is important to note that items can move away from the semantics in (3) as well. Present-Day English *every,* for example, is in fact showing signs of doing so. Consider (5). Standard universal quantification would lead us to truth conditions amounting to (6a), which make no sense. Instead, *everyone* in (5) plausibly has a group denotation, which combines with the collective predicate sensibly as in (6b). See e.g. Champollion (2010) for relevant discussion.

5. Everyone gathered under the rata.

6. a. #For each x, x a relevant person, x gathered under the rata.
   b. A group consisting of the relevant persons gathered under the rata.

Interestingly, Kluge (1995) identifies the *je* contained in German *jeder* 'every' as being derived from an older form **aiwin* 'always', so this would be another instance of a universal quantifier (namely 'always') losing its quantifier status (in this case, changing

to a dependent particle *je*, not a plural), and itself reentering the development in forming *jeder* 'every'.

These considerations lead me to propose, as a hypothesis, an alternative semantic cycle for universal quantification, the AltSemCycle sketched in (7):

7. stage 1: covert universal quantification over alternative propositions in FCRs
   $\forall p[p \in \{FCR(x) | x \in Alt(wh)\} \rightarrow \varphi(p)]$
   stage 2: lexical universal quantification over alternatives
   $\forall x[x \in Alt(wh) \rightarrow Q(x)]$
   stage 3: lexical universal quantification over individuals by universal DP
   $\forall x[P(x) \rightarrow Q(x)]$
   stage 4: group-denoting DP with possible universal distributive readings
   $max(\lambda x.P(x)) \in [\lambda z.\forall x[x \leq z \rightarrow Q(x)]$

The general model for the cycle is Jespersen's (1917) famous cycle for negation. Similarly to Jespersen's cycle, the AltSemCycle begins at stage 1 with a linguistic expression that does not denote a certain logical concept, but occurs in environments in which the concept is expressed by a complex composition of ingredients. In the case of Jespersen's cycle, the relevant linguistic expression is a negative polarity item; in our case, it is a wh-element. The logical concept is negation for Jespersen and universal quantification for us. Moving to stage 2, expressing the logical concept becomes a lexical property of the linguistic expression, but in conjunction with further ingredients (two negative particles in Jespersen's cycle, a wh-element plus a universal in the AltSemCycle). At stage 3, the semantics of the linguistic expression is the logical concept (*not, every*). In the final stage 4, that semantics is weakened. This opens the door to a new start of the cycle or spiral with potentially fresh linguistic expressions. In what follows, I explain and work out the stages of the proposed AltSemCycle (7) in some more detail.

3. Stage 1: A semantic analysis of FCR

Let me explain what stage 1, repeated in (8), is intended to convey. Essentially (8) is an informal sketch of the semantic analysis proposed for FCRs like (1a).

8. stage 1: covert universal quantification over alternative propositions in FCRs
   $\forall p[p \in \{FCR(x) | x \in Alt(wh)\} \rightarrow \varphi(p)]$

1. a. I will read [whichever book you recommend]. FCR

Let's look at the compositional interpretation of (1a). Following Hirsch (2015), I assume that the wh-expression, as usual, introduces alternatives ('Alt(wh)' in (8)), for example (9a). At the level of the FCR, we have the alternative propositions (9b) ('FCR(x)' in (8)).

9. a. [[whichever book]] = {*West Coast Walking - A Naturalist's Guide*, *An Introduction to Modal Logic*, *A Grammar of Unua*} = {W, M, U}
   b. [[whichever book you recommend]]
   = {you recommend W, you recommend M, you recommend U}

The complete structure of (1a) is spelled out in some more detail in (10). It contains a hidden conditional structure (represented as IF) and a covert universal quantifier over alternative propositions (represented as ALL-ALT). Covert material is indicated in capitals. Motivation for Hirsch's analysis comes from so-called unconditionals (Rawlins (2013)) which require just those ingredients. An example of an unconditional and its analysis is sketched in (11). Given (11), the meaning of (10) amounts to (12). Hirsch (2015) proposes an E-type semantics for the covert pronoun IT – [[IT]] = the book that you recommend. Hence we get (4) as the overall meaning of (10)/(1a), as anticipated.

10. [ALL-ALT$_{FCR}$ [[IF [$_{FCR}$ whichever book you recommend]] [I read IT]]]

11. a. Whichever book you recommend, I will read *Pride & Prejudice*.
    b. [ALL-ALT $_{FCR}$ [[IF [$_{FCR}$ whichever book you recommend]] [I read P&P]]]
    c. [[ALL-ALT$_C$ φ]] = 1 iff for all propositions p∈C, φ is true
    d. [[IF A B]] = 1 iff all worlds in which A is the case are worlds in which B is the case
    e. For all propositions p ∈{you recommend W, you recommend M, you recommend U}:
       all worlds in which p is the case are worlds in which I read P&P.

12. For all propositions p ∈ {you recommend W, you recommend M, you recommend U}: all worlds in which p is the case are worlds in which I read IT.

4. For all alternative propositions p:
   If p is the proposition that you recommend x (x a book) and p is true,
   then I will read the book that you recommend.

Crucially, the universal quantifier ALL-ALT is a covert quantifier over alternative propositions. Its domain of quantification is generated by the wh-expression, which triggers the introduction of alternatives into the semantics. The combination of these two things results in seeming universal quantification over books in our example (1a). This, I propose, is a compositional semantic analysis of the structures identified by Haspelmath (1995) as the source of universal DPs and conjectured to be the beginning of the AltSemCycle at stage 1.

4. Stage 2: Lexical quantification over alternatives

It is not at all obvious what intermediate steps could lead from this state of affairs to a standard universal quantifier (3). What precedents do the languages of the world provide for an intermediate stage on the way from (12)/(4) to (3)? I suggest looking at Japanese alternative sensitive universal quantification expressed by MO. An example is given in (13) (from Shimoyama (2001)).

13. a. dono gakusei-no okaasan-MO odotta
       which student-GEN mother-MO danced
    b. For all alternatives x such that x ∈ { y's mother | y a student }: x danced

According to Shimoyama (2001; 2006), MO expresses universal quantification. Unlike with Present-Day English *every*, quantification is over the alternatives provided by the sister of MO. The analysis is motivated by the fact that MO co-occurs with wh-expressions, and the right interpretation arises when the alternatives triggered by the wh-expression project up to the point where MO is encountered. Example (13) is analysed in (14).

14. a. [which student's mother]-MO danced.
    b. [[which student]] = {x: x is a student}
       e.g. {Linda, Julia, Saskia}
    c. [[which student's mother]] = { x's mother | x is a student }
       e.g. {Linda's mother, Julia's mother, Saskia's mother}
    d. If Z =[XP -MO] then [[Z]] = $\lambda Q.\forall x[x \in [[XP]] \rightarrow Q(x)]$
    e. [[[which student's mother] -MO]]
       = $\lambda Q.\forall x[x \in$ { Linda's mother, Julia's mother, Saskia's mother$\} \rightarrow Q(x)]$
    f. [[[which student's mother] -MO danced]] = 1 iff
       $\forall x[x \in$ {Linda's mother, Julia's mother, Saskia's mother$\} \rightarrow$ x danced]

The relevance of Shimoyama's analysis for present purposes lies in the fact that this analysis mediates between FCRs and standard universal quantification: universal quantification is conventionally tied to a lexical item (in the example, MO). Moreover, quantification can be over individuals rather than propositions. At the same time, quantification still happens at the level of alternatives, and the wh-expression still plays the role of the alternative trigger. This, I suggest, makes it a potential precedent for stage 2 of the AltSemCycle.

15. stage 2: lexical universal quantification over alternatives
    $\forall x[x \in Alt(wh) \rightarrow Q(x)]$

A couple of further remarks are in order. First, it is interesting that the etymological source of MO seems to be the meaning 'also'. Additive focus-sensitive particles like 'also' often show up in the FCRs of stage 1 (Haspelmath 1995); a Present-Day German example is given in (16a). The role of these particles in the transition from stage 1 to stage 2 needs to be explored further. Second, a more obvious candidate for an intermediate stage between FCR and universal DP might have been free choice DPs, for instance (16b). While there is a rich and lively research on free choice indefinites (e.g. Chierchia 2013; Menendez-Benito 2010), I have not been able to identify an analysis that I can take to provide a clear conceptual transition between stages 1 and 3 of the AltSemCycle.

16. a. Wer auch immer kommen mag ist eingeladen.
       who also always come wants is invited
       'Whoever wants to come is invited.'

    b. Choose whichever card!

I must leave this matter for future research and note that diachronic evidence has yet to be provided that a semantics similar to the one sketched for MO is plausibly

operative at some stage in languages like English or German. My suggestion here is based purely on theoretical reasoning.

5. Stages 3 and 4: The standard universal quantifier and its weakening

If stage 2 as sketched above is a relevant intermediate state of the grammar, it is fairly clear what has to happen to get us to stage 3:

17. stage 3: lexical universal quantification over individuals by universal DP
$\forall x[P(x) \rightarrow Q(x)]$

The wh-expression has to lose its status as an alternative trigger, thus allowing the universal quantification to switch from the level of alternatives to the ordinary semantics. The lexical semantic change in the quantifier and the loss of a genuine wh-element have to happen *in tandem*.

There is not much more to say about stage 3 since it is the standard case of quantification. Instead, let's consider how a change from stage 3 to stage 4 might be motivated.

18. stage 4: group denoting DP with possible universal distributive readings
$\max(\lambda x.P(x)) \in [\lambda z.\forall x[x \leq z \rightarrow Q(x)]$

Let me first explain what (18) intends to convey. Reference to a group of individuals is the standard semantics assumed for definite plural DPs (e.g. Link 1983), as shown in (19).

19. a. The sheep gathered under the rata.
    b. [[the sheep]] = $\max(\lambda x.\text{sheep}(x))$
    c. $\max(\lambda x.P(x)) = \text{the } x\text{: } P(x) \,\&\, \forall y[\, P(y) \rightarrow y \leq x]$
    d. [[the sheep]] is the largest x such that 'sheep' applies to x and everything that is a sheep is part of x, i.e. the largest group of individuals that consists of sheep

The hypothesized last stage in the AltSemCycle takes such a reference to a group to be a denotation of the universal quantifier of stage 3, cf. (20a), as motivated by (5), (6). This raises the question of how an expression might change from a universal quantifier denotation (20b) to reference to a group (20a).

20. a. [[everyone]] = $\max(\lambda x.\text{person}(x))$
    b. [[everyone]] = $\lambda Q.\forall x[\text{person}(x) \rightarrow Q(x)]$

I propose that a bridge might be provided by the equivalence of (21a,b) (see also Champollion (2010)):

21. a. Heidi read every book.
    b. Heidi read the books.

(21a) receives its truth conditions via the lexical entry in (3) as sketched in (22). The example with the group-denoting plural DP (21b) is analysed in (23). While the object DP denotes a group, universal quantification comes from the distributive interpretation of the predicate represented by the DIST operator (e.g. once more Link 1983). The truth conditions of (21a) and (21b) are equivalent and amount to (24), even though the meanings of the two object DPs (22b), (23b) are quite different.

22. a. [[every book] [Heidi read t]]
    b. [[every book]] = $\lambda Q.\forall x[book(x) \rightarrow Q(x)]$
    c. $\forall x[book(x) \rightarrow$ Heidi read x]

23. a. [[the books] [DIST [Heidi read t]]]
    b. [[the books]] = $max(\lambda y.books(y))$
    c. [[DIST]] = $\lambda Q.\lambda z.\forall x[x \leq z \rightarrow Q(x)]$
    d. $max(\lambda y.books(y)) \in [\lambda z.\forall x[x \leq z \rightarrow$ Heidi read x] =
       $\forall x[x \leq max(\lambda y.books(y)) \rightarrow$ Heidi read x]
       'For each x such that x is a part of the largest book group, Heidi read x.'

24. For each x such that x is a book, Heidi read x.

Suppose the meaning of *every book* is reanalysed as group denoting, like (23b). In distributive predication, the overall truth conditions remain stable (in line with *Constant Entailments*). This, I speculate, is permitting Present-Day English *every* to be thus reanalysed, bringing us on the way to stage 4.

6. Conclusions

I have made a suggestion regarding a compositional analysis of Haspelmath's finding that universal quantifiers may derive from free choice relative clause constructions. The finding, I propose, describes two stages in a diachronic cycle, the AltSemCycle, by which linguistic expressions change from an alternative semantics to ordinary universal quantifiers and later lose their quantificational nature, opening a gap for the cycle to begin anew.

I have based my discussion in very general terms on Haspelmath's generalisations. A lot of empirical and analytical work remains to be done. The details of the stages I have proposed need to be spelled out as well as the transitions between them. This will make fine-grained predictions about the diachronic development, which need to be followed up on for individual languages and time periods.

If this research project can be carried out satisfactorily, we may make progress regarding two important general questions in semantics. The first is, why is universal quantification very stable cross-linguistically compared to potential quantifiers like *most, exactly n, no more than n*? The AltSemCycle invites the hypothesis that universal quantification over alternative propositions is reliably available, and standard quantification has this as a source. The second question is, why does natural language burden itself with two separate quantificational systems, ordinary quantification and alternative semantic quantification? The AltSemCycle regards the two systems as not,

in fact, separate. Interestingly, the basic source of quantification is in the domain of alternatives, and what we have come to look at as standard is derived from that.

In sum, while this squib can only be the starting point of a comprehensive research project, I hope to have shown that it would be a very worthwhile project for semantics and diachronic linguistics.

References

Barwise, John & Cooper, Robin. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4(2). 159–219.

Beck, Sigrid. 2012. Pluractional comparisons. *Linguistics and Philosophy* 35. 57–110.

Beck, Sigrid & Gergel, Remus. 2015. The diachronic semantics of English *again*. *Natural Language Semantics* 23. 157–203.

Champollion, Lucas. 2010. Cumulative readings of *every* do not provide evidence for events and thematic roles. In Aloni, Maria & Bastiaanse, Harald & de Jager, Tikitu & Schulz, Katrin (eds.) *Proceedings of the 17th Amsterdam Colloquium,* 213–222. Berlin & Heidelberg: Springer.

Chierchia, Gennaro. 2013. *Logic in grammar*. Oxford: Oxford University Press.

Gergel, Remus. 2016. Modality and gradation. In van Gelderen, Elly (ed.) *Cyclical change continued,* 319–350. Amsterdam: John Benjamins.

Eckardt, Regine. 2009. *Grammaticalization and semantic reanalysis.* Oxford: Oxford University Press.

Hamblin, Charles. 1973. Questions in Montague English. *Foundations of Language* 10. 41–53.

Haspelmath, Martin. 1995. Diachronic sources of 'all' and 'every'. In Bach, Emmon & Jelinek, Eloise & Kratzer, Angelika & Partee, Barbara (eds.) *Quantification in natural languages,* 363–382. Dordrecht: Kluwer.

Heim, Irene & Kratzer, Angelika. 1998. *Semantics in generative grammar*. Oxford: Blackwell.

Hirsch, Aron. 2015. A compositional semantics for wh-ever free relatives. In Bade, Nadine & Berezovskaya, Polina & Schöller, Anthea (eds.) *Proceedings of Sinn und Bedeutung* 20, 341–358. (semanticsarchive.net)

Jespersen, Otto. 1917. *Negation in English and other languages.* Copenhagen: Høst.

Kluge, Friedrich. 1995. *Etymologisches Wörterbuch der deutschen Sprache.* 23. Auflage. Berlin & New York: De Gruyter.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In Bäuerle, Rainer & Schwarze, Christoph & von Stechow, Arnim (eds.) *Meaning, use, and interpretation of language*, 302–323. Berlin: De Gruyter.

Menéndez-Benito, Paula. 2010. On universal free choice items. *Natural Language Semantics* 18(1). 33–64.

Rawlins, Kyle. 2013. (Un)conditionals. *Natural Language Semantics* 21(2). 111–178.

Shimoyama, Junko 2001. Wh-constructions in Japanese. Amherst, MA: University of Massachusetts. (Doctoral dissertation)

Shimoyama, Junko. 2006. Indeterminate phrase quantification in Japanese. *Natural Language Semantics* 14. 139–173.

Sigrid Beck
Universität Tübingen
sigridbeck@me.com

Passive and movement of verbal chunks in a V/head-movement language

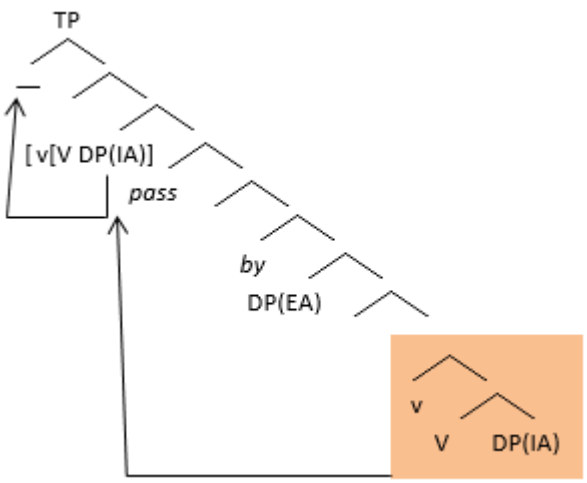
Adriana Belletti

The present squib raises the following question: Could movement of a verbal chunk be generally disfavoured in a given language? In particular, if the language is best described as displaying instances of head movement processes including V-movement ones in the domain of inflectional morphology, should this then entail that processes affecting (portions of) phrasal verbal constituents should be globally excluded?

If the influential analysis of passive in terms of movement of a verbal chunk/ *smuggling* à la Collins (2005) is assumed (cf. also Belletti & Rizzi 2012; Belletti 2014) the question above becomes particularly relevant, and amounts to asking whether passive should be excluded, or at least disfavoured, in a V/head movement language as compared to a language without (or with reduced instances of) V/head movement.

The analysis of passive in terms of movement of a verbal chunk has the important side-effect of overcoming a potential locality/Relativized Minimality (Rizzi 1990; 2004) violation. The violation of locality would inevitably arise if the DP internal argument (IA), the derived subject in the passive, were raised into the (relevant, EPP) subject Spec/TP position directly, across the intervening DP external argument (EA). As (1) illustrates, no such violation occurs if the process affects a chunk of the verb phrase, containing the verb (with the appropriate morphology) and DP(IA). The moved chunk does not contain the DP(EA), which is left behind in a lower position in the clause structure. The DP(IA) is thus raised into Spec/TP from the derived position, encountering no intervention problem. According to the analysis sketched out in (1), a component of the passive voice, labelled *pass* for convenience, attracts the relevant portion of the verb phrase into its specifier:

1.



Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 15–18

The derivation in (1) is thus motivated on principled grounds.

Māori is a V/head movement language. According to Pearce (2002), the head movement analysis simply accounts for complex basic word orders of the language and for the morphosyntax of adverbial particles, which are best analysed as building up the functional spine of the clause, as in (2):

2. E tupu tonu mai nei anō i te pari o taua whēnua.
   T/A grow still hither near1 still P the abundance of that land
   'Because of the fertility of that land, (they) are still growing there.' (Pearce 2002: ex.2)

Furthermore, the fact that a modified noun cannot undergo incorporation is also a clear indication that head movement-type processes are at work in the language in the nominal domain as well. This is illustrated by the contrast in (3a) vs. (3b) with respect to the position of the T/A marker *ana,* (which contrasts with the situation in Niuean as described in Massam 2001 , discussed in Pearce 2002):

3. a. *E kimi whare kia noho ai ia ana tērā tangata.
      T/A seek house SUBJNCT live there 3SG T/A that man

   b. E kimi whare ana tērā tangata kia noho ai ia.
      T/A seek house T/A that man SUBJNCT live there 3SG
      'That man is looking for a house for him to live in.' (Pearce 2002: ex.22)

Pearce's V/head movement analysis of Māori appears to account more adequately for the relevant observed distributions than a potential alternative analysis in terms of remnant phrasal movement. In the latter analysis, a chunk of the verb phrase (previously voided of its arguments and only containing the verb) would move (into the Spec of a given functional head), giving the impression of a head movement process (along the lines of e.g. Koopman & Szabolcsi's 2000 approach).

Going back to the question raised at the outset, if this question is given a positive answer, then passive should be absent or at least disfavoured in the relevant language, as it involves phrasal movement of a portion of the verb phrase. The use of alternatives to passive should then be expected, such as e.g. object topicalisation or use of arbitrary subjects in active clauses (a way of demoting the subject comparable to passive). Note that recent evidence from the acquisition of Italian indicates that Italian-speaking young children adopt precisely these alternatives when they still do not adequately master passive, specifically when they do not yet master movement of the relevant verbal chunk implicated in passive (cf. experimental results presented in Belletti & Manetti 2017; Snyder & Hyams 2015 for related discussion).

If, in contrast, the question raised at the outset is given a negative answer, this would mean that no "global" dis-preference for moving chunks of verb phrases should be expected in principle, but rather, the existence of processes of this type should be established construction by construction, in fact probing/attracting head by probing/attracting head. Specifically, one should expect that passive could be in principle available in the language, all other things being equal.

The latter conclusion should probably be preferred on conceptual grounds: there does not seem to be any formal, principled way to characterize a global ban against phrasal movement of portions of verb phrases. The availability of the process should be determined through the setting of different micro-parameters related to the (movement-attracting) properties of different functional heads (cf. Rizzi forthcoming for related considerations on micro-parametric variation, in the spirit of the Borer-Chomsky conjecture). Moreover, the *smuggling*-type movement illustrated in (1) appears to be different from the remnant phrasal movement of a verb phrase that mimics head movement: in the latter the verb phrase has been emptied of its arguments; in the former, in contrast, the moved chunk crucially contains both the verb and the DP(IA).

Italian is a language that is best described as instantiating the V/head-movement process, in both finite and non-finite inflectional morphology (Belletti 1990; Cinque 1999). Passive is a productive construction in Italian. This indicates that the conceptually preferable conclusion is met in Italian.

All other things being equal, one would expect passive to also be available in Māori, still maintaining Pearce's V/head movement analysis of the language. An example like the following from Pearce's (2002) article indicates that the conceptually preferable conclusion is met in Māori as well, as passive appears to be a productive construction in the language (a very productive one in fact; see Pucilowski 2006 for discussion of possible reasons linked to the analysis of the language in terms of split-ergativity).

4. Ā mahara-tia tonu-tia hoki e mātou ngā kupu o tāna waiata.
   and remember-PASS still-PASS also by 1PLEXCL the-PL word of her song
   'And we still remember the words of her song.'

In conclusion, even if the V/head-movement analysis seems preferable and more adequate in the domain of inflectional morphology, phrasal instances of movement of portions the verb phrase may still be present in a given language, as in the case of the derivation of passive addressed in this squib.

The question remains open whether phrasal movements of portions of the verb phrase are in fact always of the type implemented in passive, in which the verbal phrasal chunk that is moved into the specifier of the attracting functional head contains more than just the verb; in the case of passive, it contains the verb and its internal argument. Another case similar to passive in this respect is provided by Romance-type causatives (Belletti forthcoming).

References

Belletti, Adriana. 1990. *Generalized verb movement*. Turin: Rosenberg & Sellier.

Belletti, Adriana. 2014. Notes on passive object relatives. In Svenonius, Peter (ed.) *Functional structure from top to toe*, 97–114. Oxford: Oxford University Press.

Belletti, Adriana. forthcoming. Labeling (Romance) causatives. In Enoch Aboh & Haeberli, Eric & Puskas, Genoveva & Shonenberg, Manuela (eds.) *Elements of comparative syntax: Theory and description,* 13–46. Berlin: De Gruyter Mouton.

Belletti Adriana & Rizzi, Luigi. 2012. Moving verbal chunks in the low functional field. In Brugè, Laura & Cardinaletti, Anna & Giusti, Giuliana & Munaro, Nicola & Poletto, Cecilia (eds.) *Functional heads: the cartography of syntactic structures*, vol. 7*,* 129–137. Oxford: Oxford University Press.

Belletti, Adriana & Manetti, Claudia. 2017. Topics and passives in Italian speaking children and adults. Université de Genève & Università di Siena. (Submitted ms.)

Cinque, Guglielmo. 1999. *Adverbs and functional heads*. Oxford: Oxford University Press.

Collins, Chris. 2005. A smuggling approach to the passive in English. *Syntax* 8(2). 81–120.

Koopman, Hilda & Szabolcsi, Anna. 2000. *Verbal complexes.* Cambridge, MA: MIT Press.

Massam, Diane. 2001. Pseudo noun incorporation in Niuean. *Natural Language and Linguistic Theory* 19. 153–197.

Pearce, Elizabeth. 2002. VP-raising versus V-raising in Māori. In Rakowski, Andrea & Richards, Norvin (eds.) *Proceedings of AFLA VIII: the eighth meeting of the Austronesian Formal Linguistics Association* (MIT Working Papers in Linguistics 44), 225–240. Cambridge, MA: MITWPL.

Pucilowski, Anna. 2006. *Split-ergativity in Māori*. Christchurch, NZ: University of Canterbury. (MA thesis)

Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.

Rizzi, Luigi. 2004. Locality and left periphery. In Belletti, Adriana (ed.) *Structures and beyond: the cartography of syntactic structures*, vol. 3, 223–251. Oxford: Oxford University Press.

Rizzi, Luigi. forthcoming. On the format and locus of parameters. *Linguistic Analysis.*

Snyder, William & Hyams, Nina. 2015. Minimality effects in children's passives. In Di Domenico, Elisa & Hamann, Cornelia & Matteini, Simona. (eds.) *Structures, strategies and beyond,* 343–368*.* Amsterdam: John Benjamins.

Adriana Belletti
University of Geneva, University of Siena
adriana.belletti@unige.ch

A note on Romance and Germanic past participle relative clauses*

Guglielmo Cinque

Abstract

Some evidence will be presented for the existence in Romance and Germanic of an IP (though not of a CP) layer in past participle 'reduced' relative clauses, and for the PRO nature of their internal head (for at least some of them).

1. Introduction

Past participial relative clauses (RCs) are often considered to be reduced, in the sense that they are taken to lack the higher clausal layers IP and CP. The main properties adduced in support of this conclusion are: 1) the absence of overt complementizers, 2) the absence of tense, 3) the absence of external arguments, and 4) the fact that only internal arguments promoted to subject position can be relativized[1] (see, among others, Burzio 1981: §3.3; 1986: 150–152, 193–198; Chomsky 1981: 167; Hazout 2001; Siloni 1995; 1997: Chapter 4; Cecchetto & Donati 2015: §3.4; Harwood 2016; Douglas 2016: Chapter 6).[2] The limited goal of this article is to present some evidence for the existence in Romance and Germanic of an IP (though not of a CP) layer in past participle 'reduced' RCs, and for the PRO nature of their internal head.[3]

---

* To Liz, looking back to the early days of our acquaintance and friendship.

[1] Namely, the derived subjects of passive and pseudo-passive participles (*the students accepted t in the program*; *the rights infringed upon t; a man believed t to know the truth* (Burzio 1986: 151f, 190), *anyone given this opportunity*, and certain active past participles of unaccusative verbs (*the recently arrived letter* – Kayne 1994: 99; *?the leaf fallen from the tree is red* – Marvin 2002: 141), but not a subject raised from the complement of raising verbs: **a man seemed to know the truth* (Burzio 1986: 191). For restrictions on reduced relatives with active past participles of unaccusative verbs in English see Burzio (1986: 191), Stanton (2011: 61) and Douglas (2016: §5.1).

[2] As a matter of fact, as Belikova (2008) and Stanton (2011) note, the term "reduced" as a cross-linguistic characterization may be a misnomer, since in many languages participial RCs can be structurally complex and because the four properties mentioned above turn out to be independent from one another, as there are languages that have overt tense but no overt subjects, and languages that have no overt tense but have overt subjects (thus being able to relativize non-subjects as well). Also see, among others, Doron & Reintges (2007) for discussion of languages with temporally marked participles, and Krause (2001) and Alcazár (2007) for past participle RCs with overt subjects allowing relativisation of non-subjects.

[3] As in Cinque (2008; 2013), here I am assuming that relative clauses are merged in a specifier of the nominal extended projection and have both an internal and an external head. In Cinque

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 19–27

2. Evidence for the IP/TP nature of past participle relatives

Evidence that in Italian past participle RCs are larger than VP, plausibly as big as IP/TP, (but not larger – cf. Section 4), comes from two distinct considerations: 1) the possible occurrence of speaker and subject oriented IP adverbs (see the example in (1)):[4]

1. I film sfortunatamente/presumibilmente persi durante
   the films unfortunately/presumably lost during
   l'ultima guerra sono davvero molti.
   the.last war are really many
   'The films unfortunately/presumably lost during the last war are really many.'

and 2) the tense interpretation of the participles. Both unaccusative and passive participles have a deictic past tense (not just an anterior tense) interpretation. This can be seen from the fact that they are compatible with deictic past tense adverbs like *ieri* 'yesterday', *la settimana scorsa* 'last week', but not with deictic future tense adverbs like *domani* 'tomorrow', *la settimana prossima* 'next week'. See (2–3):

2. a. Gli ospiti arrivati la settimana scorsa sono stati tutti
      the guests arrived the week last are been all
      sistemati in albergo.
      accommodated in hotel
      'The guests who arrived last week have all been accommodated in a hotel.'

   b. Il film visto ieri sarà premiato lunedì.
      the film seen yesterday will.be prized Monday
      'The film seen yesterday will get a prize on Monday'

(2010: §4.2) participial RCs are argued to be merged lower than cardinals (and full finite restrictive RCs):

i. $[_{DP}$ the [two $[_{ParticP}$ PRO_i recently appointed $t_i]$ $professors_i]]$

Sleeman (2011) claims that prenominal participles in Germanic are structurally different from postnominal ones, the latter being bigger (CP) than the former (AspP). This is however dubious given the lack of evidence for a CP structure for either one (see below) and given that the differences between them in terms of presence of complements (for postnominal ones) and agreement with the head (for pre-nominal ones) seem ascribable to independent factors (cf. Cinque 2010: Chapter 4, and Appendix, A6).

[4] The same appears to be true of French (cf.(i)) and English (cf.(ii)):

i. L'étudiant probablement/malheuresement arrêté par la police est un étranger.
   the.student probably/unfortunately arrested by the police is a foreigner
   'The student probably/unfortunately arrested by the police is a foreigner.' (Douglas 2016: 241)

ii. the cakes fortunately eaten by the guests (Douglas 2016: 229)

3. a.* Gli ospiti arrivati la prossima settimana verranno tutti
the guests arrived the next week will.come all
sistemati in albergo.
accommodated in hotel
'The guests who will have arrived next week will all be accommodated in a hotel.'
(cf. *Gli ospiti che saranno arrivati la prossima settimana verranno*
the guests that will.be arrived the next week will.come
tutti sistemati in albergo.
all accommodated in hotel
'The guests who will have arrived next week will be accommodated in a hotel.')

b.* Il film visto domani sarà premiato lunedì.
the film seen tomorrow will.be prized Monday
'The film seen tomorrow will get a prize on Monday.'
(cf. *Il film che sarà visto domani sarà premiato lunedì*.
the film that will.be seen tomorrow will.be prized Monday
'The film which will be seen tomorrow will get a prize on Monday.')

If the participle involved a relative anterior tense (in Reichenbachian terms E – R: "the event time precedes a reference time") rather than a past tense (in Reichenbachian terms E,R – S "the event time coinciding with the reference time precedes the speech time"), (3a,b) should be compatible with a deictic future tense (as the corresponding finite sentences following them indicate). But they are only compatible with a deictic past tense.

3. Evidence for the PRO nature of the internal head of past participle relatives

Following Chomsky (1981: 167f) and Burzio (1986: §3.2), I take the subject of past participle reduced RCs in Italian to be PRO. German displays direct evidence for this conclusion. As noted in Fanselow (1986), 'floating' distributive phrases like *einer nach dem anderen* 'one after the other' agree in Case with the DP with which they are construed. See (4a–b):

4. a. Wir_{Nom} haben $\text{Maria}_{\text{Acc}}$ $\text{einer}_{\text{Nom}}$/*$\text{einen}_{\text{Acc}}$ nach dem anderen geküßt.
we have Maria one after the other kissed
'We have kissed Maria one after the other.'

b. $\text{Maria}_{\text{Nom}}$ hat die $\text{Männer}_{\text{Acc}}$ $\text{einen}_{\text{Acc}}$/*$\text{einer}_{\text{Nom}}$ nach dem anderen geküßt.
Maria has the men one after the other kissed
'Maria kissed the men one after the other.'

As Fanselow further observes, if such floating phrases are construed with the PRO subject of an infinitive, they invariably bear nominative Case. This is particularly evident in such cases as (5), where the controller of PRO bears a different Case:

5. Weil ich die Männer$_{Acc}$ überzeugte, PRO Renate einer$_{Nom}$/*einen$_{Acc}$
   as I the men convinced Renate one
   nach dem anderen zu küssen, ...
   after the other to kiss
   'As I convinced the men to kiss Renate one after the other, …'

Now, what we observe in the reduced relative clause case is that the floating distributive phrase also appears in nominative Case, irrespective of the Case borne by the head with which it is construed (example also from Fanselow 1986):

6. Wir sahen die [einer$_{Nom}$/*einen$_{Acc}$ nach dem anderen angekommenen]
   we saw the one after the other arrived
   Studenten$_{Acc}$
   students
   'We saw the students who had arrived one after the other.'

For past participle reduced RCs this clearly points to the presence of a PRO with which the floating distributive phrase is construed.[5]

Cecchetto & Donati (2015: §3.4) propose that past participle reduced relatives involve 'movement of N', which 'from its argument position to its derived position relabels the structure, conveniently providing the external determiner with the NP it needs to select' (p. 77). Cf. their (7):

7. the [$_{NP}$ [$_{N}$ philosopher] [$_{VP}$ admired [$_{N}$ ~~philosopher~~]]"

In addition to the evidence just reviewed for taking the internal argument moved to subject to be PRO, further evidence exists that such an argument has to be a phrase rather than a bare head N. This phrase can control a PRO (*un documento$_i$* [$_{IP}$ PRO$_i$ *archiviato* t_i [*dopo* PRO$_i$ *esser stato letto* t_i] 'a document filed after being read') and can enter an A-chain in restructuring configurations. See the cases of long passivisation in (8) and those involving raising of a lower subject or object in (9):[6]

[5] Siloni (1995: §4.3.2) argues against PRO in reduced RCs, but see Cinque (2010: 131, note 34). Restrictive past participle RCs possibly also allow a raising derivation, as the head can be an idiom chunk *(The headway made so far is not negligible* – Cinque 2010: 56), which raises the question whether the internal head remains in Spec,IP/TP or moves higher, and how it is Case-marked.

[6] There is also evidence that such participles are verbal and are to be kept distinct from adjectives morphologically derived from past participles. *Un uomo amatissimo da tutti* 'a man most loved by everybody' (Cecchetto & Donati 2015: 78) contains an adjective – witness the superlative morphology – which however becomes impossible if clausal adverbs are added: **un uomo da poco amatissimo da tutti* 'a man recently most loved by everybody'. See Cinque (2010: Chapter 6, note 1). Also past participles in relatives clauses with clitics attached to them (*le sole persone presentateci …* 'the only people introduced to us …') cannot be adjectives, as these resist cliticisation (Benincà & Cinque 1991: §2.3).

8. a. Le case [PRO$_i$ finite di costruire t_i negli anni '50] …
the houses finished to build in.the years 50
'The houses that they finished building in the '50's …'

   b. I bambini andati/venuti a prendere t_i a scuola …
the children gone/come to fetch at school
'The children who were fetched from school …'

9. a. Un vicino [PRO$_i$ venutomi a t_i chiedere un favore] …
a neighbour come-to-me to ask a favour
'A neighbour who had come to ask me a favour …' (Burzio 1986: 334)

   b. i soli quadri [PRO$_i$ andati t_i persi t_i]
the only paintings gone lost
'the only paintings that were lost'

4. Evidence for the lack of the CP (and other left periphery) layers[7]

As observed in Jacobson (1995: 460), 'headless' reduced RCs appear not to be possible. See (10):

10. a.* What(ever) displayed in this window will be sold by midnight.
    b.* Whoever invited to the party is supposed to bring a bottle.

The same seems to be true of Italian:[8]

[7] Also see Hazout (2001).

[8] Apparent exceptions, like (i), can be accounted for if, as suggested in Donati & Cecchetto (2011: Appendix), *qualunque cosa* and *chiunque* can be external heads (see (ii), where they stand by themselves):

i. a. Qualunque cosa persa da uno di voi non verrà ricomprata.
whatever thing lost by one of you not will.come rebought
'Whatever thing [is] lost by one of you will not be bought again.'

   b.? Chiunque sorpreso a rubare verrà multato.
whoever caught to steal will.come fined
'Whoever [is] caught stealing will be fined.'

ii. a. Farei qualunque cosa per aiutarti.
I.would.do whatever thing to help.you
'I would do anything to help you.'

   b. Parlerebbe con chiunque.
(s)he.would.talk with whoever
'(S)he would talk with anyone.'

The same appears to be true of cases involving *quanto* 'what', given the possibility of (iv):

ii. Quanto pattuito è senz'altro soddisfacente.
how.much agreed is certainly satisfying
'What [is] agreed upon is certainly satisfying.'

11. a.* Chi invitato alla festa dovrà portare una bottiglia.
who invited to.the party will.must bring a bottle
* 'Who invited to the party will have to bring a bottle.'

b.* Quanti intervistati hanno negato di conoscerlo.
how many interviewed have denied to know.him
* 'Who$_{pl}$ interviewed denied knowing him.'

The ungrammaticality of (10) and (11) could be blamed on the fact that the *wh*-phrase receives no Case within the 'headless' RC, due to the absence of finite Tense (cf. Kayne 1994: §8.4). Nonetheless, the fact that 'headless' reduced RCs appear not to be possible even where no Case licensing is plausibly at issue, as is the case in (12), suggests that the cause of the ungrammaticality is related to the fact that they contain no CP capable of hosting a *wh*-phrase (also see Chomsky 1981: 167):

12. a.* Quando addormentata Gianna si mise a russare.
when fallen.asleep Gianna herself put to snore
'When fallen asleep, Gianna started snoring.'

b.* Dove alloggiati non era certo il miglior posto che potevano
where housed not was certainly the best place that they.could
scegliere.
choose
'Where housed was certainly not the best place they could choose.'
(cf. *Dove siamo stati alloggiati non era certo il miglior posto*
where we.are been housed not was certainly the best place
che potevano scegliere
that they.could choose
'Where we have been housed was certainly not the best place they could choose.').

The absence of a CP in reduced RCs should also account for the impossibility of **the book which recently sent …* (cf. Kayne 1994: 98), as well for the ungrammaticality of (13) as opposed to (14), which involves pied piping of the whole participle phrase to the Spec of the matrix CP:[9]

---

iv. a. Ho fatto tutto quanto.
I have done all how.much
'I did everything.'

b. Ecco quanto.
here how.much
'Here is what (I did/said).'

[9] As apparent from (12) and from such ordinary adjunct participles as *Sdraiato sul divano, lui si sentirà meglio* 'Lain on the sofa, he will feel better', it is possible for such participles to have a controlled PRO subject.

13. a.* Questo è il divano sul quale sdraiato lui si sentirà meglio.
this is the sofa on.the which lain he himself will.feel better
'This is the sofa on which lain he will feel better.'

b.* That's the screen behind which sat he won't be able to see her.

14. a. Questo è il divano sdraiato sul quale lui si sentirà meglio.
this is the sofa lain on.the which he himself will.feel better
'This is the sofa lain on which he will feel better.'

b. That's the screen sat behind which he won't be able to see her.

In addition to the CP layer that hosts *wh*-relative phrases, there is evidence that past participle RCs also lack the Topic and Focus layers, which (to judge from finite restrictive or non-restrictive RCs – cf. (16)) are lower than the *wh*-relative layer:

15. a.* Il libro a Carlo regalato(gli) ieri.[10]
the book to Carlo given(to.him) yesterday

b.* Il libro A MARIA dato (non a Carlo) era molto costoso.
the book to Maria(FOCUS) given (not to Carlo) was very expensive

16. a. Il libro che a Carlo gli era stato dato solo ieri …
the book that to Carlo to.him had been given only yesterday …

b. Il libro che A MARIA era stato dato, non a Carlo, …
the book that to Maria(FOCUS) had been given not to Carlo

References

Alcazár, Asier. 2007. *A minimalist analysis of participial constructions: Towards a phase account of non-finite structures.* Los Angeles: University of Southern California. (Doctoral dissertation)

Belikova, Alyona. 2008. Syntactically challenged rather than reduced: Participial relatives revisited. In Susie Jones (ed.) *Proceedings of the 2008 annual conference of the Canadian Linguistic Association*, 1–15. Toronto: Canadian Linguistic Association. (http://homes.chass.utoronto.ca/~cla-acl/actes2008/CLA2008_Belikova.pdf)

Benincà, Paola & Cinque, Guglielmo. 1991. Participio presente. In Lorenzo Renzi & Gianpaolo Salvi (eds.) *Grande grammatica italiana di consultazione*, vol. II, 604–609. Bologna: il Mulino.

[10] Cases involving tonic pronouns, like (i), are only apparent cases of ordinary Clitic Left Dislocation/ Topicalisation or Focus movement, as the same pronouns (as opposed to full DPs) can also occur in finite relative clauses in between auxiliaries and past participles. See (ii):

i. Il libro a noi regalato ieri
the book to us given yesterday
'the book given to us yesterday'

ii. il libro che era stato a noi regalato ieri
the book that had been to us given yesterday
'the book that had been given to us yesterday'

Burzio, Luigi. 1981. Intransitive verbs and Italian auxiliaries. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation)

Burzio, Luigi. 1986. *Italian Syntax: A Government-Binding approach*. Dordrecht: Reidel.

Cecchetto, Carlo & Donati, Caterina. 2015. *(Re)labeling*. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 2008. More on the indefinite character of the Head of restrictive relatives. In Paola Benincà & Federico Damonte & Nicoletta Penello (eds.) *Selected proceedings of the 34th Incontro di Grammatica Generativa* (Special issue of the *Rivista di Grammatica Generativa*, volume 33)*,* 3–24. Padova: Unipress. (Reprinted in Cinque 2013: 167–180) (http://lear.unive.it/jspui/ bitstream/11707/133/1/cinqueRGG33-2008.pdf)

Cinque, Guglielmo. 2010. *The syntax of adjectives: A comparative study*. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 2013. *Typological studies: Word order and relative clauses*. London: Routledge.

Donati, Caterina & Cecchetto, Carlo. 2011. Relabeling heads: A unified account for relativization structures. *Linguistic Inquiry* 42. 519–560.

Doron, Edit & Reintges, Chris. 2007. On the syntax of participial modifiers. Hebrew University of Jerusalem. (Unpublished ms.) (http://pluto.huji.ac.il/~edit/edit/DORON&REINTGES2.pdf)

Douglas, Jamie. 2016. *The syntactic structures of relativization*. Cambridge: University of Cambridge. (Doctoral dissertation) (http://ling.auf.net/lingbuzz/003182)

Fanselow, Gisbert. 1986. On the sentential nature of prenominal adjectives in German. *Folia Linguistica* 20. 341–380.

Harwood, Will. 2016. *Reduced relatives and extended phases: A phase-based analysis of inflectional restrictions on English reduced relative clauses*. Leuven: KU Leuven. (Doctoral dissertation) (http://ling.auf.net/lingbuzz/002867)

Hazout, Ilan. 2001. Predicate formation: The case of participial relatives. *The Linguistic Review* 18. 97–123.

Jacobson, Pauline. 1995. On the quantificational force of English free relatives. In Bach, Emmon & Jelinek, Eloise & Kratzer, Angelika & Partee, Barbara (eds.) *Quantification in natural languages*, 451–486. Dordrecht: Kluwer.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Krause, Cornelia. 2001. *On reduced relatives with genitive subjects*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation)

Marvin, Tatjana. 2002. Past participles in reduced relatives. In van Koppen, Mario & Thrift, Erica & van der Torre, Erik Jan & Zimmermann, Malte (eds.) *Proceedings of ConSOLE IX*, 141–156. Leiden: Universiteit Leiden. (http://www.hum2.leidenuniv.nl/pdf/lucl/sole/ console9/console9-marvin.pdf)

Siloni, Tal. 1995. On participial relatives and complementizer D°: A case study in Hebrew and French. *Natural Language and Linguistic Theory* 13. 445–487.

Siloni, Tal. 1997. *Noun phrases and nominalizations: The syntax of DPs*. Dordrecht: Kluwer.

Sleeman, Petra. 2011. Verbal and adjectival participles: Position and internal structure. *Lingua* 121. 1569–1587.

Stanton, Tom. 2011. The reduced relative clause: A misnomer? *Début: The Undergraduate Journal of Languages, Linguistics and Area Studies* 2(2). 55–65. (http://www.studyinglanguages.ac.uk/sites/default/files/pages/106/Debut_Volume_2_Number_2_stanton.pdf)

Guglielmo Cinque
Università Ca' Foscari Venezia
cinque@unive.it

Teachers' understandings of the role of translation in vernacular language maintenance in Malekula: some early thoughts

Nicola Daly and Julie Barbour

1. Introduction

Existing research concerning literacy in the Pacific indicates three key considerations: (1) Traditional print literacy levels are low (Toumu'a 2016), (2) the development of culturally and linguistically relevant quality print resources is a key approach to raising these levels (Toumu'a 2016); and (3) the development of such resources can also contribute to maintenance of local knowledge (Paviour-Smith 2005).

Studies of traditional ecological knowledge (TEK) have also indicated that their inclusion in formal school curricula is an important aspect of the maintenance of this knowledge (McCarter & Gavin 2011). Paviour-Smith (2005) has worked with a community in the Aulua locale of Malekula to standardise orthography and develop materials in local languages with 26 villagers. Along the way he and his colleagues discovered the importance of paying attention to local discourse models. Thus, understanding the forms of local storytelling and the literacy practices surrounding these stories is an important first step in creating effective print resources using vernacular languages and containing TEK which aim to raise the print literacy levels in Malekula communities.

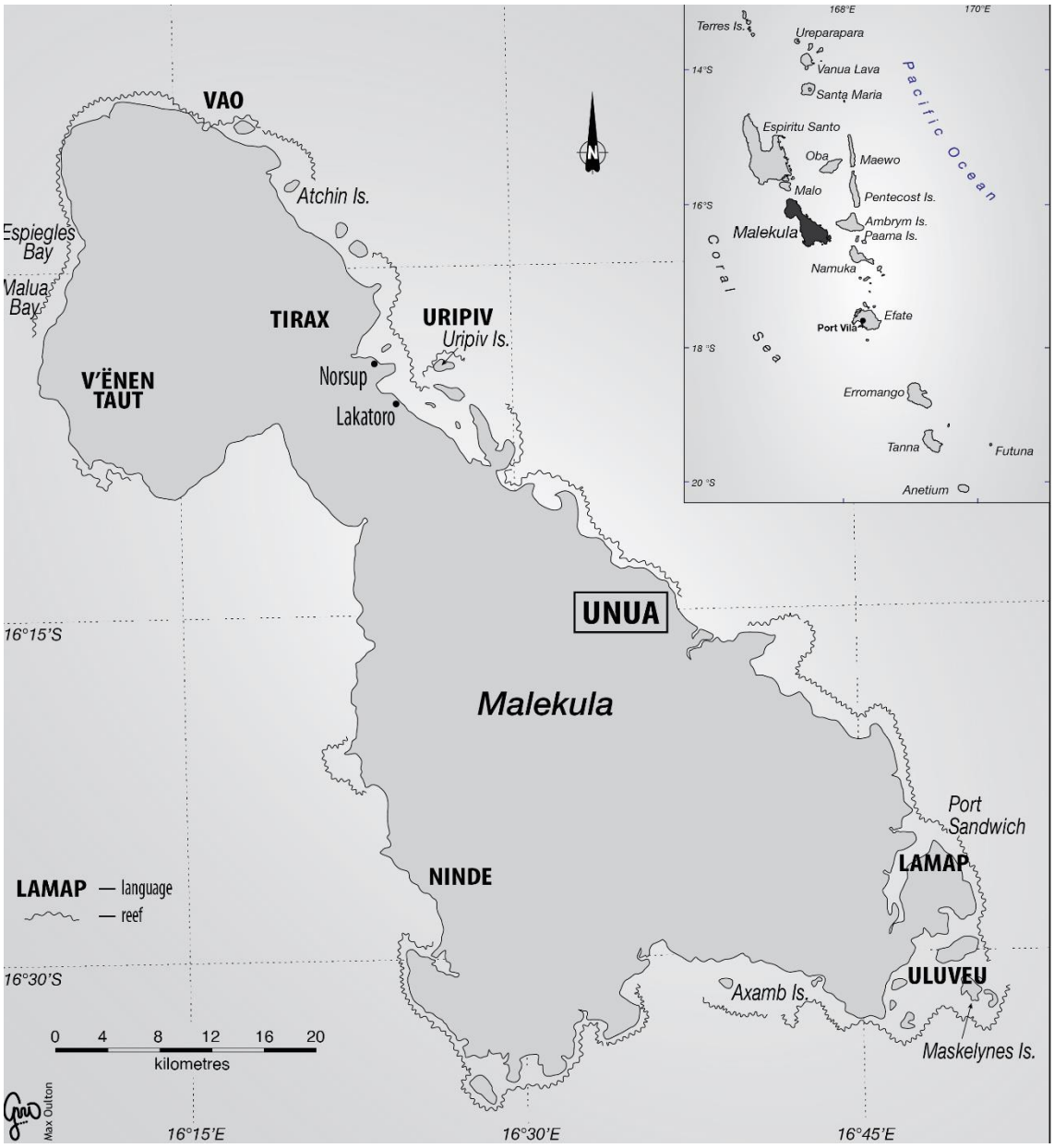
In 2012, a National Language Policy was endorsed in Vanuatu, accommodating the vernacular languages of Vanuatu within the formal education system for the first time in the nation's history (Ministry of Education 2012). To enact the new National Language Policy, the Ministry of Education developed a plan to create vernacular resources for around half of Vanuatu's 100+ languages. To support this process, linguists were approached, including one of the authors (Julie Barbour) who has been working with communities in Malekula, Vanuatu for the last 15 years (see e.g. Barbour 2010; 2012; 2013; 2015). In 2015, Julie facilitated the first regional workshop on Malekula Island to translate Year 1 reading materials into seven local languages with Nicola assisting. Nicola's research interest since 2007 in picturebooks which use more than one language (Daly 2007; 2008; 2016) and Julie's knowledge of local languages in Malekula through her own research, and the supervision of research students, have led to the development of the project introduced in this article documenting local literacy and storytelling practices in Malekula.

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 29–34

2. Setting

The setting for this work was in a Ministry of Education Building in Norsup on the island of Malekula in the Malampa Province of Vanuatu. Malampa is an administrative region including the islands of Malekula, Ambrym and Paama. Malekula is the second largest island in Vanuatu, and is known as the most linguistically diverse of all the islands, with more than 30 distinct languages (Lynch & Crowley 2001). In November 2017, a group of four postgraduate linguistics students and the two authors arrived to work with 14 teachers from around Malekula. Our main focus was to support local teachers to translate the Year 2 and 3 reading materials produced by the Ministry of Education in English into seven local languages: Lamap (spoken in southeast Malekula), Mae (spoken in east Malekula), Maskelynes (spoken on small islands south of Malekula), Ninde (spoken in the southwest of Malekula), Tirax (spoken in north central Malekula); Uripiv (spoken on the northeast coast of Malekula) and V'ënen Taut (spoken in northwest Malekula).

Map 1. Malekula Island showing locations of languages included in Ministry of Education vernacular education project, along with Unua, the language studied by Elizabeth Pearce.



3. Data collection progress

While the translation work was progressing, the 14 Malekula teachers present were invited to speak with Nicola using a semi-structured interview format regarding the place of storytelling in their communities, and to express their views on the importance of translating children's reading materials into local languages (see questions in the Appendix). There were nine volunteers: four men and five women, with at least one volunteer for each of the seven languages represented in the Year 2 and 3 Reader translation workshop. A local translator was present during the interviews, which were conducted in English. While most of the nine interviewees were proficient in English, some were more fluent in French and some translation was required using Bislama, a dialect of Melanesian Pidgin and the local *lingua franca*. These interviews lasted from between 15 and 30 minutes, and were recorded using Quicktime, and then transcribed by a transcriber with local knowledge. These transcriptions are currently with the participants for checking before analysis takes place.

4. Early thoughts

While the details of the interview transcriptions are yet to be confirmed by the nine Malekula teacher participants, the tenor of the interviews without exception indicated a strong awareness of the importance of literacy materials translated into vernacular languages. Vernacular language materials were positioned as being relevant both to the success of school children and to the maintenance of local languages. Participants commonly refer back to their own experiences of learning to read for the first time in a completely unknown language. The policy until 2015 had been for education to be delivered in the medium of either English or French. The teachers' early experiences of literacy were therefore in a foreign language, and they recalled how difficult that was.

We look forward to unpacking the detailed nuances of these transcripts to help us understand more about traditional literacy and storytelling practices. We believe that the analysis of ideas and comments made by the nine volunteers concerning the place of storytelling in a range of Malekula communities will be invaluable during the development of future literacy materials using local languages. Much like Toumu'a (2016), who argues for culturally and linguistically relevant literacy resources to be developed in the Pacific region more generally, we believe that the more aligned literacy materials can be with the discourse practices around oral storytelling in the communities, the more effective these materials will be.

We are honoured to be able to offer this brief introduction to our project to mark the retirement of our colleague (JB and ND) and former lecturer (ND), Liz Pearce, whose own work on the language of Unua in Vanuatu continues to be an important resource for us and our research students.

Appendix: Semi-structured interview questions

1. Can you tell me about storytelling in your community?
   a. When are stories told?
   b. How are they told? (books, song, dance?)
   c. How do people in your community learn to tell stories in these different ways?
   d. Who tells them? (adults/children?)
   e. What are the stories about?
   f. Are there stories which belong to your community?
   g. Are any of these stories written down?

2. Can you tell me about other times when stories are told? (prompts: in ceremonies? Singing? In dance? In church?)
   a. How do people in your community learn to use language in these ways?
   b. When is dance/singing/ceremonies done?
   c. How do community members learn ceremonies/songs/dances?
   d. Can anyone learn ceremonies/songs/dances?
   e. When people make mistakes, what happens?
   f. What stories do ceremonies/songs/dances tell?

3. Do any of the reading materials used in schools tell any of the stories or knowledge from your community? If so, which ones and what stories/knowledge?

4. How do the stories in the reading materials relate to the lives of the children in your communities?

5. How do you feel about the translation of the school reading materials into local languages?
   a. What do you see as the advantages of having reading materials in local languages for the students? The teachers? The community?
   b. What are the disadvantages of having reading materials in local languages for the students? The teachers? The community?

6. How does the community feel about the translation of the school reading materials into local languages?

7. How has the process of translating the reading materials into local languages been for you?
   a. What have been the challenges?
   b. What have you learned?

References

Barbour, Julie. 2010. Neverver: A study of language vitality and community initiatives. In Florey, Margaret (ed.) *Endangered languages of Austronesia,* 225–244. Oxford: Oxford University Press.

Barbour, Julie. 2012. *A grammar of Neverver.* Berlin: De Gruyter Mouton.

Barbour, Julie. 2013. To the field and back ... In Denzin, Norman K. (ed.) *Studies in symbolic interaction*, vol. 40, 247–268. Bingley, UK: Emerald Books.

Barbour, Julie. 2015. Small stories and associated identity in Neverver. In Gounder, Farzana (ed.) *Narrative and identity construction in the Pacific Islands,* 81–100*.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Daly, Nicola. 2007. Kūkupa, koro and kai: The use of Māori vocabulary items in New Zealand English picture books. *New Zealand English Journal* 21. 20–33.

Daly, Nicola. 2008. The narrative contract and the use of Māori loanwords in New Zealand English picture cooks*. Journal of Children's Literature Studies* 5(2). 1–17.

Daly, Nicola. 2016. Dual language picturebooks in Māori and English. *Bookbird* 54(3). 10–17.

Lynch, John & Crowley, Terry. 2001. *Languages of Vanuatu: A new survey and bibliography*. Canberra: Pacific Linguistics.

McCarter, Joe & Gavin, Michael C. 2011. Perceptions of the value of traditional ecological knowledge to formal school curricula: Opportunities and challenges from Malekula Island, Vanuatu. *Journal of Ethnobiology and Ethnomedicine* 7:38. (doi:10.1186/1746-4269-7-38)

Ministry of Education. 2012. *Vanuatu national language policy*. Port Vila: Ministry of Education.

Paviour-Smith, Martin. 2005. Is it Alua or education dressed up in Kastom?: A report on the ongoing negotiation of literacy and identity in a Ni Vanuatu community. *Current Issues in Language Planning* 6(2). 224–238.

Toumu'a, Ruth. 2016. Culturally and linguistically relevant resource development for literacy in the Pacific. In Toumu'a, Ruth & Sanga, Kabini & Johansson-Fua, Seu'ula Falelalava (eds.) *Weaving education theory and practice in Oceania: Selected papers from the second Vaka Pasifiki Education Conference,* 118–122*.* Suva, Fiji: Institute of Education, USP.

Nicola Daly
University of Waikato
nicolad@waikato.ac.nz

Julie Barbour
University of Waikato
jbarbour@waikato.ac.nz

Untangling multiple Madurese benefactives

William D. Davies

Abstract

A Madurese benefactive can be expressed as a prepositional object or in an applicative structure, and a single clause can include more than one benefactive. However, the interpretation-structure pairings are restricted with multiple benefactives. It is argued here that the restriction has an explanation in the derived preposition that expresses the benefactive relation.

1. Madurese benefactive applicatives

As in other closely-related Indonesian-type languages (i.e. Balinese, Javanese, Malay-Indonesian, Sundanese and others), Madurese includes an applicative construction that takes transitive predicates as input and derives a ditransitive structure with a benefactive meaning, as in (1).

1. a. Bhibbhi' ngerra' bhâbâng kaangghuy embu'.
   aunt AV.slice onion for mother
   'Auntie sliced onion for Mother.'

   b. Bhibbhi' ngerra'-**aghi** **embu'** bhâbâng
   aunt AV.slice-AGHI mother onion
   'Auntie sliced onion for Mother.'

In the benefactive applicative in (1b), the verb *ngerra'* 'slice' takes the applicative suffix *-aghi*,[1] and the beneficiary *embu'* 'mother' occurs as a core argument in immediate postverbal position. The applied object is the primary object and is the only internal argument that can occur as the subject of a passive.

The *-aghi* suffix is cognate with Indonesian *-kan*, and it is instructive to compare the Madurese benefactive applicative to the structure that has been described for Indonesian. Consider the Indonesian examples in (2) from Son & Cole (2008).[2]

[1] The initial vowel in the affix alternates between [a] and [ɣ] (represented as <â> in the *Ejaan 2004* orthography adopted here; cf. Roben 2015). The vowel alternation is conditioned by the Madurese vowel harmony processes discussed in Davies (2010: 30-35). Since *-aghi* is the elsewhere form that occurs in a wider range of phonological environments, it is given as the citation form in this paper. *–âghi* mainly occurs after voiced and voiceless aspirated stops, and after stems that end in a high vowel + glottal stop.

[2] The prefix *meN-* occurs primarily on transitive verbs (Peter Cole, p.c.).

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 35–40

2. a. Tika memanggang roti itu untuk Eric.
   Tika meN.bake bread the for Eric
   i. 'Tika baked the bread for Eric to have/to give to Eric.'
   ii. 'Tika baked the bread for the benefit of Eric.' (Son & Cole 2008: 5)

   b. Tika memanggang-**kan Eric** roti itu.
   Tika meN.bake-kan Eric bread the
   'Tika baked the bread for Eric (to have it).'

According to Son & Cole, (2a) is ambiguous, giving rise to both an interpretation in which the bread is to be possessed by Eric (i) and one in which Tika bakes the bread instead of Eric baking the bread (ii), the so-called 'deputative benefactive' of Van Valin & LaPolla (1997). On the other hand, Son & Cole state that (2b) is unambiguous, allowing only a change-of-possession interpretation.

The Madurese structures differ in this regard. Like the Indonesian, the sentence in (1a) with the prepositional beneficiary is ambiguous between change-of-possession and deputative interpretations. So, in (1a), Auntie may slice the onions for Mother to have, or she may slice them instead of Mother doing so. Where Madurese differs is with respect to the interpretation of the applied structure. The Madurese sentence in (1b) evinces the same ambiguity as (1a). While the change-of-possession interpretation seems to be preferred by speakers in a decontextualized presentation, the deputative interpretation is possible in an appropriate context. Further, with situations in which there is no possible change of possession of the theme argument, the applied structure is deemed acceptable, as in (3).

3. Hasan notob-**bhâghi** Siti labâng.
   Hasan AV.close-AGHI Siti door
   'Hasan closed the door for Siti.'

Son & Cole additionally note that it is possible to get the applicative *-kan* even when the beneficiary occurs as a prepositional object of *untuk* 'for', as in (4).

4. Tika memanggang-**kan** roti itu **untuk** Eric.
   Tika meN.bake-kan bread the for Eric
   'Tika baked the bread for Eric (to have it).'(Son & Cole 2008: 11)

Son & Cole assert that unlike (2a), when the beneficiary is a prepositional object and co-occurs with *-kan*, only the change-of-possession interpretation is possible in (4).[3]

2. Multiple benefactives

At first glance, Madurese provides a similar surface paradigm. It is indeed possible for a prepositional beneficiary to co-occur with the applicative morpheme *-aghi*, as in (5).

[3] Son & Cole (2008) further assert that the benefactive PP in (3) is a subcategorized argument. However, the argument or non-argument status of the beneficiary does not bear on the current discussion.

5. Bhibbhi' ngerra'-aghi bhâbâng kaangghuy embu'.
   aunt AV.slice-AGHI onion for mother
   'Auntie sliced onion for Mother.'

Like the Indonesian example, speakers report that the preferred interpretation for (5) is that Auntie is slicing the onion and Mother is going to possess it. However, peeling back the layers and inspecting a sentence like (5) in greater detail indicates that there is more to the structure than is apparent from the surface.

Speakers report that unlike (1a) or (1b), there is, in fact, more than a single beneficiary involved. Speakers report that in (5) there is the beneficiary that will likely possess the onion, Mother, but there is also another participant not explicitly identified who will benefit from the action. The sentence in (6) illustrates this as well.

6. Marlèna ngèbâ'-âghi kothak kaangghuy Siti.
   Marlena AV.carry-AGHI box for Siti
   'Marlena carried the box for Siti.'

The sentence in (6) does not simply mean that Marlena carried the box and Siti was the recipient of the box. Included in (6) is the notion that Marlena carried the box to Siti to help out someone else, possibly though not obligatorily Siti. That is, (6), and (5) as well, includes both a change-of-possession and deputative beneficiary. This becomes obvious when the deputative beneficiary is explicitly mentioned, as in (7).

7. Marlèna ngèbâ'-âghi Ita kothak kaangghuy Siti.
   Marlena AV.carry-AGHI Ita box for Siti
   'Marlena carried the box for Siti for Ita.'

In (7), it is clear that there are two beneficaries, both a deputative beneficiary and a benefactive participant who will possess the box. The fact is that some speakers find (6) — and (5) — of questionable acceptability, preferring a sentence like (7), in which the additional beneficiary is identified explicitly.

The Madurese examples in (5–7) are, in fact, open to yet another interpretation. With appropriate contextualization, each sentence can be interpreted as involving two deputative beneficiaries. This possibility becomes clearer when considering a sentence such as (8), in which neither benefactive participant can be a possessor as there is nothing tangible that can be transferred from one person to another.

8. Alè' nyawapo'-aghi embuk kamar kaangghuy èbhu'.
   younger.sibling AV.sweep-AGHI elder.sister room for mother
   'Little Brother swept the room for Big Sister for Mother.'

3. A puzzle and a solution

The Madurese data considered thus far present us with the paradigm in (9).

9. a. Marlèna ngèbâ' kothak kaangghuy Siti.
Marlena AV.carry box for Siti
'Marlena carried the box for Siti.'

b. Marlèna ngèbâ'-âghi Siti kothak.
Marlena AV.carry-AGHI Siti box
'Marlena carried the box for Siti.'

c. Marlèna ngèbâ'-âghi kothak kaangghuy Siti.
Marlena AV.carry-AGHI box for Siti
'Marlena carried the box for Siti for someone.'

d. Marlèna ngèbâ'-âghi Ita kothak kaangghuy Siti.
Marlena AV.carry-AGHI Ita box for Siti
'Marlena carried the box for Siti for Ita.'

In (9a), the beneficiary Siti occurs in a PP and there is no applicative affix on the verb. While the sentence is ambiguous between a change-of-possession and deputative benefactive, in the absence of context the change-of-possession interpretation is the default and Siti is likely the recipient of the box. In (9b), the beneficiary is a non-prepositional core argument and the applicative affix occurs on the verb. Like (9a), this sentence is ambiguous, but again the change-of-possession interpretation is preferred. The applicative suffix occurs in (9c) but a beneficiary occurs as a prepositional object. The sentence in (9c) includes both the applicative suffix and the benefactive preposition *kaangghuy*. Here there is a deputative interpretation for the unidentified beneficiary associated with the suffix *-aghi*, and the prepositional object Siti is likely the recipient of the box (although as discussed with respect to (7), a double deputative benefactive interpretation is possible though certainly not preferred). In (9d), where we find two beneficiaries, only the prepositional beneficiary Siti is a possible recipient. Ita is obligatorily interpreted as benefitting from the event, but is not a possible recipient. The same is true of the passive equivalent of (9d) given in (10).

10. Ita è-gibâ'-âghi kothak moso' Marlèna kaangghuy Siti.
Ita PV-carry-AGHI box by Marlena for Siti
'Marlena carried the box for Ita for Siti.'

The data present the following puzzle:

- Why is the PP benefactive more strongly interpreted as the recipient when it co-occurs with the *-aghi* suffix (or obligatorily so with *-kan* in Indonesian)?

- If both recipient and deputative benefactives can occur as NP and PP objects, why is only the PP object a possible recipient when both occur in one sentence?

Note that this is not a peculiarity of the data in (9) and (10), but is true when comparing (1a) with (5), and is also true of (11).

11. Bhibbhi' ngerra'-aghi embuk moso' bhâbâng kaangghuy embu'.
aunt AV.slice-AGHI elder.sister by onion for mother
'Auntie sliced onion for Mother for Big Sister.'

In (11), *embuk* 'Big Sister' must be interpreted as a deputative beneficiary and only *embu'* 'Mother' is a possible (and likely) recipient of the sliced onion.

It appears that the answer lies in the nature of the preposition *kaangghuy*, which marks beneficiaries. Like other Indonesian-type languages, Madurese includes only a handful of basic prepositions. Other prepositions are derived, largely from verbs. This is true of *kaangghuy*, which is derived from a verbal base, *angghuy* 'use', with the resultative/abilitive prefix *ka-*. This abilitive use is illustrated in (12).

12. a. Tina ng-angghuy buku rowa.
Tina AV-use book that
'Tina read the book.'
lit. 'Tina uses this book.'

b. Buku rowa ka-angghuy (bi') Tina.
book that ABIL-use by Tina
'The book can be read by Tina.'

The sentence in (12a) exemplifies the predicate *angghuy* in an active clause. (12b) is a passive structure in which *buku rowa* 'this book' is the subject and *Tina* is the agent, which can optionally take the preposition *bi'* 'by, with'.[4] In benefactive constructions, the preposition *kaangghuy* 'for' may carry the implication that its object will use the item that the agent acts upon, which necessitates that the object of *kaangghuy* will possess that item. Clearly, this is not a necessary implication inasmuch as we have seen that the object of *kaangghuy* can be interpreted in the deputative sense. However, inasmuch as *-aghi* introduces an additional beneficiary when the clause contains the prepositional beneficiary, it stands to reason that the beneficiary marked with a derivative of the verb 'use' will be the recipient of the item in question, as opposed to the unmarked beneficiary.

This does not mean, however, that in benefactive structures *kaangghuy* is verbal. In (12b), in which *ka-angghuy* is a verb, *bi'* 'by, with' can precede the postverbal agent. However, when *kaangghuy* is used prepositionally, its object cannot take *bi'*, (13).[5]

[4] As is the case with most Indonesian-type languages, there are a number of environments in which a preposition can be omitted in Madurese, as long as the object of the preposition occurs in immediate post-verbal position. See Davies (2010) for details.

[5] In Javanese, the preposition *kanggo* 'for' is also derived from the verb *anggo* 'use'. A similar state of affairs has been reported for Sasak (Asikin-Garmager 2013) and Balinese (Natarina 2015); in each case the preposition marking the beneficiary is deverbal, and the base verb ('get' and 'give', respectively) describes a state of affairs in which the participant involved takes possession of some object.

13. Buku rèya è-belli embu' kaangghuy (*bi') Tina.
book this PV-buy mother for by Tina
'This book was bought by Mother for Tina.'

The unacceptability of the agentive preposition *bi'* in (13) indicates that *kaangghuy Tina* in (13) cannot be analysed as a verbal phrase.

The benefactive constructions in Madurese evince a somewhat different set of properties from those reported by Son & Cole (2008) for Indonesian, and unlike Indonesian *-kan*, the suffix *-aghi* is in no sense optional in benefactive structures.

Acknowledgements

The work that this paper draws on was funded by the National Science Foundation through grant no. BCS 980944, the Toyota Foundation, and the Arts and Humanities Initiative of the University of Iowa.

References

Asikin-Garmager, Eli. 2013. The structure of /*-an*/ in Sasak: Relevance for analyses of the cognate suffix shared by languages of Java. University of Iowa. (Unpublished ms.)

Davies, William D. 2010. *A grammar of Madurese*. Berlin: De Gruyter.

Natarina, Ari. 2015. The lexical and pragmatic effects of the Balinese morpheme *-ang*. (Presented at the Fifth International Symposium on the Languages of Java. Bandung, Indonesia, June 7)

Roben, Thomas. 2015. *Two orthographies in written Madurese and their effect on reading performance of Madurese junior high school students in Pamekasan of Madura Island, Indonesia*. Penang: Universiti Sains Malaysia. (MA thesis) (http://eprints.usm.my/30368/1/Thesis_by_Thomas_Roben_on_2_Madurese_orthography.pdf)

Son, Minjeong & Cole, Peter. 2008. An event-based account of *-kan* constructions in Standard Indonesian. *Language* 84. 120–160.

Van Valin, Robert D. & LaPolla, Randy J. 1997. *Syntax: Structure, meaning, and function*. Cambridge: Cambridge University Press.

William D. Davies
The University of Iowa
william-davies@uiowa.edu

Circumscriptive haplologizing reduplicants

Paul de Lacy

Abstract

This paper argues that there are morphemes that circumscribe subparts of stems, creating word-internal boundaries that condition phonological processes, trigger allomorphy, and provide a locus for concatenation. Examples are found in Māori, a Polynesian language spoken in New Zealand. Circumscriptive morphemes are shown to trigger allomorphy in a number of morphemes, including the imperative *e*~∅ and tense marker *ka*~*ka*ː. They are also used to account for vowel lengthening in certain plural and passive forms, and motivate infixation in reduplication.

1. Introduction

Some morphemes have no underlying content of their own; they borrow it from the stems they attach to (McCarthy & Prince 1999). Such 'reduplicants' are ubiquitous. There are several in the Polynesian language Māori, which will be the empirical focus here (ISO: mri; Austronesian>Malayo-Polynesian>Polynesian>Central Eastern Nuclear Polynesian). For example, the base [pi.rau] *pirau* 'to be extinguished' reduplicates as [pi-pi.rau] <pipirau> 'to be decayed' and [pi.ra-pi.rau] <pirapirau> 'to be decayed' (Keegan 1996).

Some morphemes undergo 'morphological haplology': their segments merge with neighbouring material when it is sufficiently similar. For example, the genitive suffix in many dialects of English merges with the plural morpheme: e.g. /dɔg-z_1-z_2/ 'dog'+plural+possessive → [dɔgz] <dogs'>, *[dɔgzəz]. Similarly, French /ist/ *-iste* merges with preceding /i{s,z}(t)/: e.g. /analiz-ist/ 'analyse'+agentive→ [analist] <analyste>, *[analizist]. For other examples, and why morphological haplology involves merger rather than deletion, see de Lacy (1999).

Here, I will argue that there are morphemes that are *both* contentless and undergo morphological haplology. Such 'haplologizing reduplicants' merge with the phonological material of adjacent morphemes. After de Lacy (1996), I will call such morphemes 'circumscriptive' (CR) because of their phonological and morphological effects.

The behaviour of circumscriptive morphemes is schematized in (1) below. The string [taka] in the output [pataka] is an exponent of both the root morpheme PATAKA and the haplologizing reduplicant CR. Lines show morpheme affiliation, not association.

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 41–51

1. *Haplologizing reduplicants schematized*

Underlying: /pataka+CR/

Reduplication: [pataka-taka]

Morphological haplology: [pataka]

Circumscriptive morphemes can be detected by how they affect their environment. Phonologically, they can cause prosodic and segmental changes in the stem they merge with. Morphologically, they can condition allomorphy, and serve as the base of affixation.

I argued in de Lacy (1996) that circumscriptive morphemes are found in Māori. I go further here in arguing that Māori contains two distinct types of circumscriptive morpheme: 'coextensive' and 'non-coextensive'. Coextensive circumscriptive morphemes must align with the edges of the stems they merge with. Non-coextensive circumscriptive morphemes can merge with parts of stems.

Section 2 discusses non-coextensive circumscriptive morphemes, focusing on the Māori kinship plural, lengthening in the passive, and infixing reduplication. Section 3 deals with coextensive circumscriptive morphemes and their effects on Māori allomorphy. Section 4 sketches a formal implementation, and Section 5 outlines typological predictions.

2. Non-coextensive circumscriptive morphemes

There are several non-coextensive circumscriptive morphemes (NCRs) in Māori. In all cases, the NCRs are foot-sized in the output: i.e. (C)V(C)V (de Lacy 2004).[1] They are suffixal in that they merge with the rightmost set of elements in the stems they attach to: e.g. /wahine+NCR/ → [wahine].

The Māori kinship plural morpheme is an NCR. Its presence can be detected through a side effect of Māori's alignment restrictions on stems (de Lacy 2003): stem edges must align with Prosodic Word (PrWd) edges. For example, in /wahine+NCR/ → [wahine], [h] is the left edge of the NCR morpheme, and so it must align with a PrWd edge: [wa{hine}]. However, doing so strands [wa] – as root material [wa] must also inhabit its own PrWd: [{wa}{hine}]. The problem with {wa} is that it is too small – it is below Māori's minimum size requirement of one foot per PrWd (de Lacy 2004, McCarthy & Prince 1986). Consequently, /wa/ must augment, economically achieved by

[1] There are several theories of how the surface form of underlyingly contentless morphemes is determined. In templatic theories, the Māori NCRs would consist of a Ft node underlyingly (e.g. Marantz 1982; McCarthy & Prince 1986); in Generalized Template Theory, they would be marked as stems, and constraints would emergently enforce foot size (McCarthy & Prince 1999; Urbanczyk 1996).

lengthening: [{waː}{hine}].

2. *Formation of wāhine*

Input:	/wahine+NCR$_{PLURAL}$/
Merger of NCR:	[wa.hi.ne]
PrWd boundaries:	[{wa}{hi.ne}]
Subminimal word lengthening:	[{waː}{hi.ne}] <wāhine> ‘woman+PL’

So, the effect of the kinship plural NCR is to cause vowel lengthening. However, exactly *which* vowel can be lengthened is very restricted: the kinship plural NCR can only cause lengthening of the initial vowel in *trimoraic* stems /(C)V(C)V(C)V/. For bimoraic stems like /tama-NCR/ ‘son, nephew’, the NCR merges with the *entire root*: [{tama}]. Consequently, no segments are stranded, so lengthening does not occur.

3. *Kinship plural*[2]

a. No lengthening in bimoraic roots

Singular		*Plural*		*Gloss*
[ta.ma]	<tama>	[ta.ma]	<tama>	boy
[tau]	<tau>	[tau]	<tau>	husband, spouse

b. Lengthening in Trimoraic roots (Hohepa 1967: 12; Bauer 1993: 354)

[ma.tu.a]	<matua>	[maː.tu.a]	<mātua>	parent
[ta.ŋa.ta]	<tangata>	[taː.ŋa.ta]	<tāngata>	man
[tu.pu.na]	<tupuna>	[tuː.pu.na]	<tūpuna>	ancestor
[wa.hi.ne]	<wahine>	[waː.hi.ne]	<wāhine>	woman
[ɸae.a]	<'a>	[ɸaː.e.a]	<whāea>	mother
[tei.na]	<teina>	[teː.i.na]	<tēina>	younger sibling (same sex)

Notice that the NCR carves out (circumscribes) a foot disregarding the derivational base’s syllable structure: e.g. /teina/ is [tei.na] in the singular, but [teː.i.na] in the plural, where the NCR has effectively cut the base’s first syllable in two.

The NCR analysis has two subtle effects. One is that the NCR cannot affect trimoraic roots of the form /CVːCV/. Māori does not permit adjacent identical vowels within a stem. So, the plural of [taːne] <tāne> ‘husband’ is the unaugmented [taːne]; it is blocked from becoming *[taː.a.ne] because [aː.a] consists of adjacent identical vowels.

Using the NCR also means that longer roots will not have lengthening. For example, [paːpaː] <pāpā> ‘father’ merges with the NCR as [paːpaː]; the remnant [paː] already has enough material for a PrWd, so does not lengthen: [{paː}{paː}]. The same goes for [tamaːhine] ‘daughter(s)’, [hu.nao.ŋa] <hunaonga> ‘son(s)/daughter(s)-in-law’, [tuakana] <tuakana> ‘elder same-sex sibling(s), cousin(s)’, and [mokopuna] <mokopuna> ‘grandchild/grandchildren’.

Like all minor plural formation rules, exactly what belongs in the ‘kinship’ class is

[2] Unless stated otherwise, words and their glosses are from Williams (1957). If a word has several meanings, only one is provided here to save space.

somewhat capricious and fluid. For example, *whaea* ‘mother’ has lengthening in some dialects and not others (Bauer 1993: 354).

There are two problematic words for this analysis: [tuahine] <tuahine> ‘sister (of male)’ and [tuakana] <tuakana> ‘older sibling (same sex)’. Both words have more than three moras and so lengthening should not occur, yet it does: their plurals are [tuaːhine] and [tuaːkana] respectively. It is possible that these forms are suppletive (i.e. listed in the Lexicon), similar to [ta.mai.ti] ‘child’~[ta.ma.ɾi.ki] ‘children’. Certainly, other four-mora roots do not undergo lengthening, as listed above.

2.1. Minor passive formation

Apart from the kinship plural morpheme, there is also an NCR morpheme that accompanies the passive suffix in a particular class of words. Such forms involve lengthening the first mora in the passive form, but only in trimoraic words. Notice that all the words below in (4) have the same output shape in the passive: [(C)Vː(C)V(C)V]. There are no lengthened bimoraic or four-mora passive forms.

4. *Passive lengthening (Harlow 1991: 118–120; 2007: 117; Keegan 1996: 63)*

Active		*Passive*		*Gloss*
[a.ko]	<ako>	[aː.ko.-na]	<ākona>	teach
[ɸai]	<whai>	[ɸaː.i.-a]	<whāia>	chase
[ɸa.ka-ɾe.ɾe]	<whakarere>	[ɸa.ka-.ɾeː.ɾe.-a]	<whakarērea>	reject
[hu.ti]	<huti>	[huː.ti.-a]	<hūtia>	hoist
[ki.-ki]	<kiki>	[kiː.ki.-a]	<kīkia>	kick
[ku.me]	<kume>	[kuː.me.-a]	<kūmea>	pull
[ku.ti]	<kuti>	[kuː.ti.-a]	<kūtia>	cut
[mi.-mi]	<mimi>	[miː.i.-a]	<mī.ia>	urinate
[pa.ki]	<papaki>	[paː.ki.-a]	<pākia>	slap
[po.ki]	<poki>	[poː.ki.-a]	<pōkia>	cover over
[pu.-pu.hi]	<pupuhi>	[puː.hi.-a]	<pūhia>	shoot
[ɾi.ɾi]	<riri>	[riː.ɾi.-a]	<rīria>	scold
[ta.-ta.ɾi]	<tatari>	[taː.ɾi.-a]	<tāria>	wait
[ti.ki]	<tiki>	[tiː.ki.-na]	<tīkina>	fetch

Lengthening happens in the passive for the same reason as in kinship plural formation. The NCR morpheme accompanying the passive in these words carves out a foot’s worth of material from the passivized form: e.g. /poki+PASS+NCR/ → [po.ki.a+NCR] → [po.ki.a]. The remnant [po] is stranded and subminimal, and so must lengthen for the same reason as in the kinship plural forms: [{poː}{kia}].

Notice that it is the shape of the *output* that matters, not of the input. Almost all the roots above are /(C)V(C)V/ underlyingly because suffixing the passive typically adds a mora, resulting in a trimoraic form that the NCR can then circumscribe. However, some roots are other sizes underlyingly. For example, ‘urinate’ is underlyingly /mi/. In the singular, it is augmented to minimal word size by a reduplicant: [{(mi+mi)}]. However, reduplicants are typically lost in the passive (as in *tatari~tāria* above), so the NCR

morpheme ends up circumscribing the passive morpheme alone, leaving the root to augment: /mi+PASS+NCR/ → [mi.i.a+NCR] → [mi.i.a] → [{(miː)}{(i.a)}], *[mīa].

Not every trimoraic verb undergoes lengthening in the passive. In fact, the usual situation is to not lengthen: e.g. [kati]~[katia], *[kaːtia] <kati> 'shut', [numi]~[numia], *[nuːmia] 'fold' (de Lacy 2004). However, all of the verbs that lengthen have the same shape: they form [CVːCVCV] in their passivized output form. *Kiki* 'kick' is interesting as it has two passive options: [kikia] and [kiːkia] (Keegan 1996: 64). In present terms, the different passive forms for this word depend on whether the NCR morpheme is present: i.e. [ki.ki-a] without the NCR, and [{kiː}{kia}] with the NCR. The variation is due to *kiki*'s variable class membership.

Finally, the NCR also appears with the gerund of some stems: e.g. [a.ko]~[aː.ko.-ŋa] 'learner', [ɾei]~[ɾeː.i.ŋa] 'leap', [ta.ɾi]~[taː.ɾi.-ŋa] 'wait', [ɸai]~[ɸaː.i.-ŋa] 'follow'.[3] However, such stems seem to form a different class from those that take the NCR in the gerund: e.g. <hutinga> 'raising', *<hūtinga> (cf. <hūtia>); <wāhi puhinga> 'shooting target', <riringa> 'anger', <taringa> 'ear'.

2.2 Base of affixation

Māori's third NCR morpheme is found in combination with its reduplicants. The NCR serves as the reduplicants' locus of affixation. An example is provided in (5) for <maranga> 'rise up' and its reduplicated form <mārangaranga> 'rise up one by one, bob up and down'. Importantly, the reduplicative morpheme prefixes to the left edge of the NCR morpheme – [ɾ]. Consequently, the reduplicant cuts the root /maraŋa/ in two, and the remnant [ma] must lengthen, just as with the passive and kinship plural.

5.	Input	/REDUP-maraŋa-NCR/
	NCR merger	[REDUP-maɾaŋa]
	Reduplicative alignment	[ma-REDUP-ɾaŋa]
	Reduplication	[ma-ɾaŋa-ɾaŋa]
	Subminimal lengthening	[{maː}{ɾaŋa}{ɾaŋa}]

As with the passive and kinship plural, using an NCR as the locus of reduplicant prefixation makes several predictions. One is that there will be no lengthening in longer forms: e.g. [hara-ŋote-ŋote] 'piecemeal', *[haraː-ŋote-ŋote], *[haːra-ŋote-ŋote]. Accordingly, in Keegan (1996)'s list of 672 trimoraic and four-mora stems with an infixed Ft size reduplicant, all infixed reduplicated trimoraic stems undergo lengthening, but no four-mora stems do.

The NCR morpheme can also appear with Māori's shorter reduplicant, too. It has the same lengthening effect on trimoraic roots: e.g. [kanono]~[kaː-no-nono] *Coprosma*

[3] Keegan (1996: 64) lists [ɸa.ti]~[ɸa.tiː.a.ŋa] <whatīanga> 'angle'. However, the verb *whati* 'flee' is listed with the nominalized form *whatinga* in Williams (1957). *Whatīanga* may be an unrelated form.

australis; [ma.nei]~[maː-ne-nei] ‘waver’; [pakini]~[paː-ki-kini] ‘ache’.

Māori’s reduplicants can also appear without a circumscriptive morpheme. When alone, they prefix to the stem: e.g. [ha.wa-ha.wai] ‘hillocks’, [ku.i-ku.i.a] ‘elderly women’, [ma.ne-ma.ne.a] ‘satisfied’, [ma-mao.a] ‘steam’. For more on reduplication, see de Lacy (2006) and Keegan (2006).

Without the NCR, it is difficult to explain why and where the reduplicant infixes. The NCR provides a boundary internal to the root’s segments which serves as the locus of reduplicative prefixation. It also explains why infixation only applies to trimoraic or longer roots, and not bimoraic ones (e.g. there are no forms like [pata]~*[pa-ta-ta]).

Alternative methods of explaining the locus of Māori reduplicative infixation present challenges. The reduplicant clearly cannot prefix to the derivational base’s foot because the foot is word-initial in some words and non-initial in others, as shown in (6) (foot structure is from de Lacy 2004).

6. a. [ho(púː)] ‘be swollen’ [{(hóː)}{puː}{puː}] ‘blistered’
   b. [hu(tói)] ‘stunted’ [{huː}{toi}{toi}] ‘stunted’
   c. [(háe)re] ‘go’ [{haː}{e.re}{e.re}] ‘wander about’
   d. [(ŋói)o] ‘wheezing’ [{(ŋóː)}{io}{io}] ‘weak’
   e. [(má.ra)ke] ‘bald’ [{maː}{rake}{rake}] ‘bald’

In other cases of infixation, morphemes have been argued to appear inside their stems to either avoid a phonological problem (such as a syllable with a coda), or to attach to a particular prosodic element or edge (Ussishkin 2007). The fact that there are both prefixing and infixing versions of reduplication in Māori means that there can be no phonological motivation for reduplicative infixation: i.e. there is nothing phonologically ill-formed about [{(ma.ra)}-{(ma.ra)ke}].

It is clear that the Māori infixing reduplicants do not seek to prefix to a particular output prosodic constituent, either. There is certainly a prosodic similarity in the outputs of reduplicated forms (all reduplicants end up with shape […RED+{Ft}]). However, this infixation cannot be motivated by an output requirement that the reduplicant prefix to a foot because this would be more faithfully satisfied by merely prefixing: e.g. [{(ma.ra)}-{(má.ra)ke}] – notice that such a form has the added advantage of being faithful to underlying vowel length and contiguity. The Māori reduplicative infixes are classic cases of ‘circumscription’ (McCarthy & Prince 1986) – morphological operations that involve carving out prosodic constituents, hence the need for a special ‘circumscriptive’ morpheme like the one proposed here.

To conclude, Māori has three NCR morphemes.[4] One is the kinship plural, the other

[4] Keegan (1996: 61ff) lists 31 words that have variable initial syllable length (also see Harlow 1991). However, these forms are not obviously morphologically related; they seem to be variant pronunciations: e.g. <aki>~<āki> ‘to encourage’, <hou>~<hōu> ‘new, recent, fresh’. Several of the forms are not trimoraic, and where they lengthen is unpredictable: e.g. <tuara>~<tuarā> ‘back’, cf. <tauhou>~<tauhōu> ‘strange’, <tawere>~<tāwere> ‘hang’. These

accompanies a small class of words in the passive form, and the third accompanies reduplicants. Their effects are similar: they circumscribe part of the stem they attach to, forcing any remaining segments to form their own PrWd, which then causes lengthening if the remainder is subminimal. For reduplicants, the NCR is the reduplicant's base of prefixation. The NCR analysis explains why only trimoraic roots are affected in the plural, and why the form of the lengthened output in passivized forms is always [(C)V:(C)V(C)V].

3. Coextensive circumscriptive morphemes

Māori also has a *coextensive* type of circumscriptive morpheme – 'CCR'. Like the Māori NCRs, the Māori CCR is one foot in size. Unlike the NCRs, *both* of the CCR morpheme's edges must coincide with its stem's edges. So, its edge requirements mean that it only appears with foot-sized (i.e. bimoraic) roots. For example, the output of /pai-CCR/ is [pai] because the CCR is both foot-sized and aligned with both edges of the root. In contrast, the CCR cannot occur with the trimoraic root /taŋata/: in [taŋata] the CCR aligns with the root's edges but is larger than a foot, and in [taŋata] the CCR is foot-sized but fails to align with both edges.

The CCR's presence is detectable in Māori because it conditions suppletive allomorphy. For example, the tense marker *ka* has two allomorphs [ka] and [ka:] (Biggs 1969: 28; Bauer 1993: 243ff). The long [ka:] appears when the rest of the phrase contains two morae, otherwise the short allomorph [ka] appears. This allomorphy can be explained by using the CCR. *Ka* has two lexically-listed (suppletive) allomorphs: /ka:+CCR/ and /ka/ (Mascaró 1996). The allomorph [ka:] appears when it can attach to the CCR morph, and [ka] appears elsewhere. So, in /'ka' mataku/, the output [ka mataku] wins because the alternative [ka mataku] does not have a coextensive CCR morpheme. In contrast, for /'ka' pai/, the output [ka: pai] can win because [ka:] can attach to the CCR morph.

There are a number of monomoraic particles in Māori that do *not* undergo lengthening, such as the indirect object marker *ki*. In the present approach, the difference between *ka* and *ki* is that *ki* has only one allomorph in the lexicon: /ki/.

The CCR also conditions the imperative morpheme. The imperative is [e], but only if the following phrase consists of two morae: e.g. [e tu:] <e tū> IMP+stand 'Stand up!' (Bauer 1993: 30). If the phrase has more than two morae, *e* is not used: e.g. [hae.re a.tu] 'go away!', *[e hae.re a.tu]. In present terms, the imperative has two suppletive allomorphs: /e+CCR/ and /∅/. /e/ appears whenever CCR can be satisfied – i.e. with a bimoraic base.

So, CCR morphemes can be used to condition allomorphy. They are not useful for Māori alone: this same conditioning environment occurs in many English dialects for the comparative *-er*. For example, in my speech (an idiolect of New Zealand English),

---

forms clearly do not show the effect of an NCR morpheme – there is no clear morphological process here, and their length variants are unpredictable, indicating that they are suppletive.

/-ɨɻ/ <er> *comparative* only suffixes to bases that consist of a foot: e.g. [(gɻiːn)-ɨ] <greener>, [(pɜp)l-ɨ] <purpler>, [(klév)ɻ-ɨ] <cleverer>, [(hǽ.pi)-jɨ] <happier>, [(mɔ́d)n-ɨ] <moderner>. It doesn't suffix to longer bases: *[(ʔé.lɨ)gɨnt-ɨ] <eleganter>, *[dɨ(láit)fɨl-ɨ] <delightfuller>, *[(zél)ɨs-ɨ] <zealouser>, *[ɨ(mɛ́ns)-ɨ] <immenser>. As with the Māori imperative and *ka*, the comparative's allomorphs are conditioned by whether the base is bimoraic, and the CCR provides a way to account for such sensitivity.

4. Theory

It is likely that any theory that implements morphological haplology as coalescence (e.g. de Lacy 1999) and has some way of dealing with templatic morphology will permit haplologizing reduplicants.

In theories with templates (e.g. Marantz 1982), haplologizing morphemes have the same underlying shape as reduplicative templates. Instead of being filled in by spreading/copying, they merge (or, in serial terms, they reduplicate then haplologise).

In Generalized Template Theory, templatic shape is due to the emergent effect of markedness constraints on morpheme size and prosodic alignment. In serialist Optimality Theories, circumscriptive reduplicants may be implemented as involving reduplication, followed by morphological haplology. A greater challenge is found in one-level theories, like classical Optimality Theory (Prince & Smolensky 2004).

In one-level OT with Generalized Template Theory, circumscriptive morpheme shape is due to the emergent effect of markedness constraints. In Māori, stem edges must align with PrWd edges. So, in /wahine+NCR/, where NCR is a stem, a candidate such as [{(wáhi)ne}] will fatally violate stem-PrWd edge alignment constraints because the left edge of the NCR is not at the left edge of a PrWd. In contrast, both the winner [{(wáː)}{(hine)}] and loser *[{wahine}] have PrWd boundaries at both left and right edges of both stems – *wahine* and NCR. The crucial distinction between these candidates involve constraints that favour a smaller size of PrWd (for a constraint-based analysis of Māori PrWds, see de Lacy 2003).

Finally, there is probably no representational difference between coextensive and non-coextensive circumscriptive morphemes. The sensitivity of suppletive allomorphy to coextensiveness in Section 3 can be ascribed to the constraints that select for particular suppletive allomorphs. Those constraints favour faithfulness and small PrWds, so they select allomorphs with a circumscriptive morph for foot-sized stems and allomorphs without the circumscriptive morph for smaller or larger stems.

5. Typology

Cross-linguistically, circumscriptive morphemes should have the kinds of effects discussed above: serving as the locus of infixation, and conditioning edge-sensitive phonological processes.

Yu (2003) has argued that infixes attach to a variety of constituents. For example, the KiChaga intensive is formed by prefixing /n-/ to the final syllable of a stem: e.g.

/muili-n/ → [mui-n-li] 'white+intensive' (Yu 2003: 32). Such 'pivots' could instead be analysed as circumscriptive morphemes. So, the KiChaga intensive would consist of /n/ and a NCR that is σ-sized (an 'affix' in Generalized Template Theory). The NCR merges with the final syllable, and the /n-/ prefixes to it: /muili-n-NCR/ → [mui-n-li].

Circumscriptive morphemes could come in even smaller forms when their template specifies it (e.g. they are /C/) or when constraints compress them. Such cases would appear as involving infixation to initial consonants or subsyllabic constituents. For example, the Atayal animate actor focus suffix can be seen as suffixing to a circumscriptive prefix that is forced to a minimal size: e.g. / NCR-m-qul/ → [q-m-ul] 'snatch' + animate actor focus (Yu 2003: 13). The famous Tagalog *um* infix can be reanalysed as infixation to an NCR that spans the first onset: e.g. [gr-u.m-ad.wet]. In such an analysis, phonological constraints do not motivate infixation; instead, *um* suffixes to an NCR (cf. Prince & Smolensky 2004; McCarthy 2003; Yu 2003: 23).

If circumscriptive morphemes exist, it may be possible to conceive of all morph concatenation as involving attachment to morpheme edges or prosodic heads. Cases where infixes seem to attach to other prosodic constituents would really involve attachment to circumscriptive morphemes. In such an approach, there is no role for direct phonologically-driven affixation, where phonological restrictions determine the position of morphemes (according with proposals by Yu 2003; Zimmerman & Trommer 2013; Ussishkin 2007).

In terms of the phonological effects of circumscriptive morphemes, Māori shows that alignment of PrWd and circumscriptive morpheme edges can affect vowel length. It is likely that other visible effects will be side-effects of cutting stems in two, as in Māori. The visible phonological effects of circumscriptive morphemes, then, will most obviously be those restrictions and processes that apply at certain constituent edges – the PrWd in the case of larger circumscriptive morphemes.

6. Conclusions

I have argued here that there are morphemes that have no segmental material of their own, but merge with their stems both coextensively and non-coextensively. Such 'haplologizing reduplicants' can have both phonological and morphological effects, and account for Māori's lengthening and reduplicative infixation. It is possible that such 'circumscriptive' morphemes are found in many other phonological systems, and may in fact be present in all cases of infixation.

Acknowledgements

I wrote this paper in honour of Liz Pearce, whose work on the formal structure of Māori I have found both engaging and inspiring.

References

Bauer, Winifred. 2003. *Maori* (Routledge Descriptive Grammars). London: Routledge.

Biggs, Bruce. 1969. *Let's learn Maori*. Wellington, NZ: A.H. and A.W. Reed.

de Lacy, Paul. 1996. Circumscription revisited: An analysis of Maori reduplication. *Rutgers Optimality Archive* 133.

de Lacy, Paul. 1999. Haplology and correspondence. In de Lacy, Paul & Nowak, Anita (eds.) *Papers from the 25th Anniversary* (University of Massachusetts Occasional Papers, UMOP 24), 51–88. Amherst, MA: GLSA Publications.

de Lacy, Paul. 2003. Constraint universality and prosodic phrasing in Māori. In Carpenter, Angela & Coetzee, Andries & de Lacy, Paul (eds.) *Papers in Optimality Theory II* (University of Massachusetts Occasional Papers, UMOP 26), 59–79. Amherst, MA: GLSA Publications.

de Lacy, Paul. 2004. Maximal words and the Māori passive. In McCarthy, John (ed.) *Optimality Theory in phonology: A reader*, 495–513. Oxford: Blackwell.

Harlow, Ray. 1991. Consonant dissimilation in Maori. In Blust, Robert (ed.) *Currents in Pacific linguistics: Papers on Austronesian languages and ethnolinguistics in honor of George W. Grace* (Pacific Linguistics C-117), 117–128. Canberra: Australian National University.

Harlow, Ray. 2007. *Māori: A linguistic introduction.* Cambridge: Cambridge University Press.

Hohepa, Patrick W. 1967. *A profile-generative grammar of Maori* (Indiana University Publications in Anthropology and Linguistics, Memoir 20). Bloomington: University of Indiana.

Keegan, Peter. 1996. *Reduplication in Maori*. Hamilton, NZ: The University of Waikato. (M.Phil. thesis.)

Marantz, Alec. 1982. Re reduplication. *Linguistic Inquiry* 13. 435–482.

Mascaró, Joan. 1996. External allomorphy as emergence of the unmarked. In Durand, Jacques & Laks, Bernard (eds.) *Current trends in phonology: models and methods*, 473–483. Salford, Manchester: European Studies Research Institute, University of Salford.

McCarthy, John. 2003. OT constraints are categorical. *Phonology* 20. 75–138.

McCarthy, John & Prince, Alan. 1986. *Prosodic morphology* (Rutgers Technical Report TR-32). New Brunswick, NJ: Rutgers University Center for Cognitive Science.

McCarthy, John & Prince, Alan. 1999. Faithfulness and identity in Prosodic Morphology. In Kager, René & van der Hulst, Harry & Zonneveld, Wim (eds.) *The prosody-morphology interface,* 218–309. Cambridge: Cambridge University Press.

Prince, Alan & Smolensky, Paul. 2004. *Optimality Theory: Constraint interaction in generative grammar*. Oxford: Basil Blackwell.

Urbanczyk, Suzanne. 1996. *Patterns of reduplication in Lushootseed*. Amherst, MA: University of Massachusetts, Amherst. (Doctoral dissertation)

Urbanczyk, Suzanne. 2007. Reduplication. In de Lacy, Paul (ed.) *The Cambridge handbook of phonology,* 473–494. Cambridge: Cambridge University Press.

Ussishkin, Adam. 2007. Morpheme position. In de Lacy, Paul (ed.) *The Cambridge handbook of phonology,* 457–472. Cambridge: Cambridge University Press.

Williams, Herbert W. 1957. *A dictionary of the Maori language*. 6th edn. Wellington, NZ: Government Printer.

Yu, Alan. 2003. *The morphology and phonology of infixation*. Berkeley: University of California, Berkeley. (Doctoral dissertation)

Zimmerman, Eva & Trommer, Jochen. 2013. The linearization of morphological weight. In Heck, Fabian & Assmann, Anke (eds.) *Rule interaction in grammar* (Linguistische Arbeitsberichte 90). Leipzig: Universität Leipzig.

Paul de Lacy
Rutgers University
delacy@rutgers.edu

Phonetics, phonology and syntax in synchrony and diachrony

Mark Hale

1. The "big picture"

The honorand of this volume has shown how valuable Oceanic language data can be in addressing some "big picture issues" in linguistics. While I won't be able to live up to the admirable model she provides in this domain, I would like to sketch out a research program that might lead to interesting insights into some important concerns in linguistics. The particular "big picture issue" I am interested in is simply this: imagine that you, like me, believe that a substantial amount of sound change arises via misparsing of input data in acquisition ("the Ohala Hypothesis", or OH, on which see, e.g., Hale 2012, with relevant literature). Let us assume further that there is some phenomenon which we can call, for lack of a better term, and hopefully without the fuzziness often associated with the literature on this topic, "grammaticalisation". One kind of diachronic event that this label will cover is the shifting of an etymon from "open class" (lexical) to "closed class" (functional) status.

Finally, let us assume, following Selkirk since at least the days of her dissertation (Selkirk 1972), that we will find frequent cross-linguistic phenomena whereby the relationship between two open-class lexical items may show *different phonology* than the relationship between an open-class and a closed-class lexical item. We expect this type of 'phonology-syntax interaction' even when some 'functional' element x_1 is the *grammaticalised* descendant of some element *X, which was also the antecedent of some 'open class' lexeme X_2.[1] For a concrete example consider the evolution of the English indefinite article (a functional element) *an* relative to the numeral *one*. Both descend from a single Proto-Germanic etymon, the word for 'one', and thus as some point in their prehistory were phonologically identical (since they were the same entity!). They now show a number of phonological differences.

It then follows that there must have been diachronic phonological developments, which, recall, the Ohala Hypothesis holds should generally be based on the misparsing of input data, whereby the sequence *X Y split, via grammaticalisation, into x_1 Y and X_2 Y (*an apple* vs. *one apple*, both from a single Proto-Germanic proto-string). Clearly these sequences, phonetically identical at the moment before grammaticalisation, underwent (or failed to undergo) some sound change(s) differently, and thus are now phonetically different. But this divergence between the 'x_1 Y' and 'X_2 Y' cases does not seem to involve a 'random walk' around the phonological space: it seems that, when

[1] I will use lower case variables to represent functional elements, upper case variables for open class elements.

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 54–62

the two contexts come to be phonologically differentiated, the 'x_1 Y' context *always* shows *greater* phonological interaction between the phonological elements involved, the 'X_2 Y' more phonological independence. In the English case we can see this in that the alternation between *a* and *an* is dependent on the properties of Y in the case of the indefinite article, but no such dependency exists in the case of the numeral.

But since the strings started out *identically* (back when they were both *X Y), what misparse of the input data allowed their divergent diachronic treatment? If 'x_1 Y' is *always robustly* differentiated from 'X_2 Y' in its low-level phonetics, then the phonological input itself would serve as a serious block on the reanalysis of *X Y (with two open class lexemes) as 'x_1 Y' (i.e., on "grammaticalisation").[2] But if there is no reliable, robust phonetic contrast between 'x_1 Y' and 'X_2 Y', how do we account for the seemingly consistent *directionality* of sound changes involved in the diachronic evolution of the former structure?

Here's a concrete example: Sanskrit /s/ normally debuccalises in codas before following voiceless stops to [h] (written in transliteration as <ḥ>, to distinguish it from a 'voiced h', transcribed <h>).[3] We see this in an example such as (1), in which we have the ablative singular of the word for 'heaven' followed by a sequence of the so-called preverb *pári* and its verb *srava*.[4]

1. vr̥ṣṭíṃ diváḥ pári srava
   rain-ACC.SG heaven-ABL.SG PREVERB make-flow
   'make the rain flow from heaven' (Rigveda 9.8.8a)

When a postposition follows a bare-N object, however, this debuccalisation does not take place (2).

2. ā́ no yātaṃ divás pári
   PREVERB us-DAT.PL travel-IMPV-du heaven-ABL.SG from-POSTP
   'Travel hither to us from heaven.' (Rigveda 8.8.4a)

The most sensible analysis of the data above seems to be that the *s* has *resyllabified* in the latter case (so *di.vá.spá.ri*), thus avoiding the impending doom of debuccalisation in a coda (since it is no longer in a coda), but not in the former case (so *di.vás.pá.ri*). The resyllabification in (2) is a direct function of the close syntactic connection between the postposition and its complement, clearly lacking in (1), where *pári* forms a morphosyntactic constituent with *srava*, rather than *diváḥ*. That is, in cases of … *divás] [pári …* we see no resyllabification, while in cases of … [*divás pári*] … we do see resyllabification. The parallels for this type of phenomenon are well known (including, e.g., French liaison).

[2] The literature on 'grammaticalisation' spends a lot of energy on the morpho-syntactic preconditions for a reanalysis, but none on the phonetic ones.
[3] I suppress lots of fascinating detail, for which see Hale, forthcoming.
[4] Most of the 'preverbs' of Vedic Sanskrit are homophonous with prepositions in that language, as in other IE languages (think of English *forget* beside *for*, or *undergo* next to *under*).

2. The phonology-syntax interface in Marshallese

Marshallese is an Oceanic (Micronesian) Language, widely spoken on the Marshalls (and more narrowly in Arkansas). Assuming the phonemicisation of Bender (1968), the language has a 'vertical' vowel system with a 4-way height contrast, which I will write here /I/ "high", /Ẹ/ "upper mid", /E/ "lower mid", and /A/ "low".[5] These vowels are not specified along the front-back or round-unround dimensions. By contrast, the language has a rich consonant inventory, with 3 classes of consonants: palatalized (also called "light") C^{j}, velarized (also called "heavy") $C^{ɯ}$, and round C^{w}.[6] Every vowel must have a consonant on each side of it: i.e., there are no vowel-initial or vowel-final words, nor are there sequences of vowels within words.

Vowel realization along the front-back and unround-round dimensions is determined by adjacent consonants; vowels themselves are *phonetically underspecified* along these dimensions (for the concept in general, Keating 1985; for Marshallese specifically, Choi 1992). Palatalized consonants give the adjacent portion of a vowel a front quality, velarized consonants give the adjacent portion of a vowel a back unround quality, and round consonants give the adjacent portion of a vowel a back round quality.

As a result, we find 'bodily output' realizations something like those below (sequences of vowels represent nuclei which move from one articulatory position to the other).

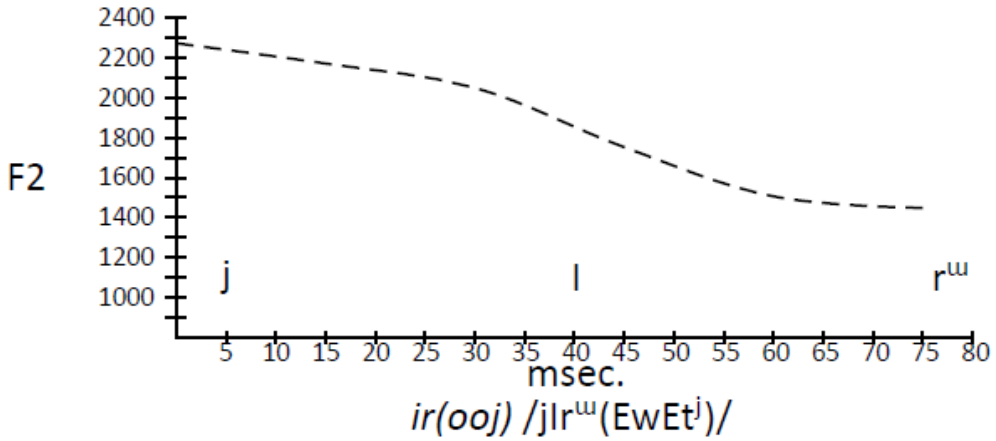
3. *nin* /$n^{j}In^{j}$/ 'pound' > $n^{j}in^{j}$
   kūk /kIk/ 'bite' > kɯk
   kuk /$k^{w}Ik^{w}$/ 'huddle' > $k^{w}uk^{w}$
   nuknuk /$n^{j}Ik^{w}n^{j}Ik^{w}$/ 'clothing' > $n^{j}iuk^{w}n^{j}iuk^{w}$
   kin /kIn^{j}/ 'bed of fronds on which the tails of porpoises are placed' > kɯin^{j}
   etc., in parallel fashion

In general, F2 correlates well with the front-back dimension, which is all we will be worrying about here. High F2 generally indicates frontness, lower F2 backness, and roundness tends to suppress F2 yet more. As Choi showed (see, e.g. 1992:49) the raising effect on F2 of 'palatalized' consonants is stronger in syllable onsets (weaker in codas), and the lowering effect of 'velarized' consonants is stronger in codas (weaker in onsets). This is more detail than we will need, I hope.

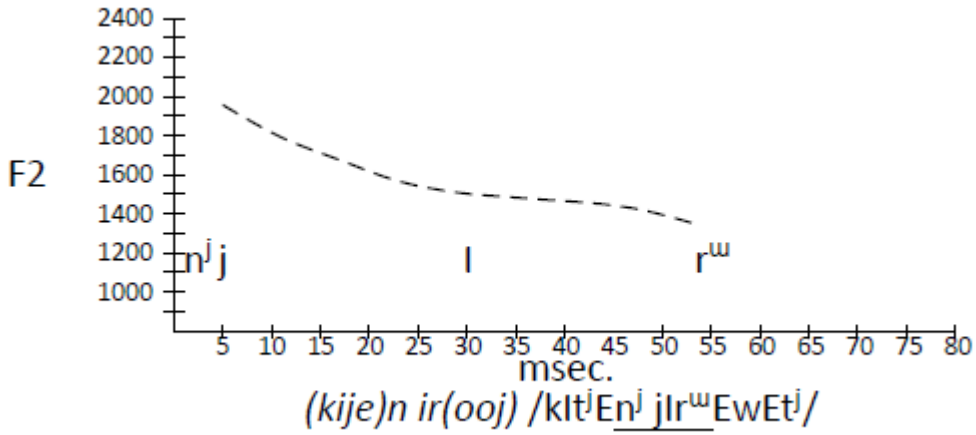
[5] It is difficult to decide how best to depict in a phonemic representation the underspecified vowels of Marshallese. Bender (1968) uses the symbols *i, &, e, a*. In my experience, when the discussion is focusing directly on matters for which the underspecification is relevant, as here, these representations quickly become confusing for the reader. It is hard to remember that sometimes an *i* is a high, front, non-round vowel and at other times it is a vowel specified only as "high", with no front or roundness specification. Here I will represent these vowels with capitals, as *I, Ẹ , E, A*, in the hopes that the capitalization will allow the reader to bear in mind the unusual nature of the vowel phonemes.

[6] Non-round velars do not traditionally get the "velarized" symbol, but are phonologically "back" consonants.

Here's what the data we are going to be looking at will look like.[7] This is the F2 trajectory for a sentence-initial utterance of the vocalic portion of the first syllable of *irooj* /jIrʷEwEtʲ/ 'chief':



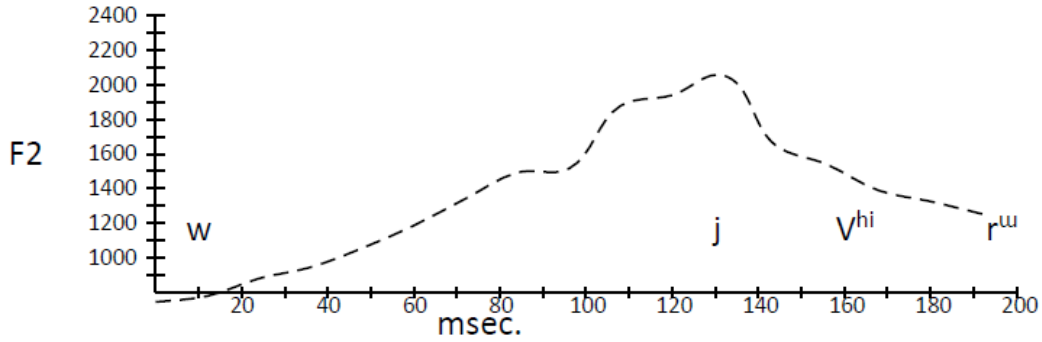
And here's the same syllable in the context *kijen irooj en* /kItʲEnʲ jIrʷEwEtʲ jEnʷ/ 'food=his chief this' 'the food of this chief'; note that the consonant preceding the glide is a palatal one:



[7] The data was recorded on Majuro, including both extensive text reading and story telling, as well as limited word-list elicitation. There were two primary informants, both male: Wilbert Alik and Hermon Lajar. I am deeply indebted to these speakers for the data provided, as well as to Ruth Abbott for help coordinating our activities.

And just to show you how much fun this can be, here's the F2 trajectory for the /w jIrʷ/ part of the string below:

4. mʲEnʲ kEw jIrʷEwEtʲ jEw
   men *ko* *irooj* *eo*
   thing DEF.PL.INANIM chief DEF.SG.ANIM
   '(to find) the things the chief (mentioned to them)'



Note that if the /j/ glide were not present, *ko irooj* would have the form /kEw IrʷEwEtʲ/, and the sequence /kEwIrʷ/ should be realized something like kɯowuɯrʷ, and the very high F2 at about 130ms in the trajectory above, indicating a very front articulation, would make no sense. The front glide posited by Bender's analysis thus appears to be necessary, supporting his analysis as a whole.

Note additionally that the sequence *ko irooj* above involves no direct syntactic relationship between the article (which goes with the preceding *men*) and the 'chief' word (which forms a constituent with the *following* definite article). That is, we are looking at, in structural terms,

5. … *men ko*] [*irooj eo* …

So far, everything is as we would have expected, given the literature to date.

3. The synchrony and diachrony of the phonology-syntax interface in Marshallese

I pointed out above that every Marshallese vowel must have a consonant flanking it on both sides. But this is not at all true of Proto-Micronesian, which happily possessed both word-final vowels and word-initial ones. How did Marshallese get the phonotactic property it now enjoys?

Descriptively, this is not a difficult problem. Words which we might expect to start or end with vowels now start or end with glides (/j/, /w/, or /ɯ/, corresponding to our

three consonant series) instead. Some of the /j/ and /w/ glides come from earlier **j*[8] and **w* glides, but not very many.[9]

But in general, a process of glide insertion is responsible: word-initial **i* and **e* developed into **ji-* and **je-* (then, ultimately, into /jI-/ and /*jE-/,[10] final[11] **i* and **e* into **-ij* and **-ej* (later /-Ij/ and /*-Ej/). Similarly, initial **u* and **o* became **wu-* and **wo-* (later /wI-/ and /*wE-/), respectively, with their final versions becoming **-uw* and **-ow* (eventually /-Iw/ and /*-Ew/). Finally, initial and final **a* acquired a /ɰ/ glide before (when initial, thus **ɰa-* > /ɰA-/) or after (when final, thus **-aɰ*) > /-Aɰ/ itself.

This is probably already clear enough, but here are a few examples. PMc **iri* 'rub' > **jirʷij* > MRS /jIrʷIj/; PMc **amʷini* 'to wash the hands' > **ɰamʷinʲ* > MRS /ɰAmʷInʲ/; **una* 'hair; scale' > **wunʲ* > MRS */wInʲ/. The glide insertion affected, as you can see, only initial and final vowels. It's worth noting that many of the glides are 'very weakly articulated', indeed, so weakly that one might say that they have been posited essentially on structural grounds.

Like all human languages I know of, Marshallese shows phonological effects based in part (ultimately) on the nature of the *syntactic* relationship between the relevant elements. Intervocalic voicing (which applies word-internally as well) applies postlexically (under appropriate conditions), e.g., as do a variety of assimilations (particularly to following nasals). There is an epenthesis process which inserts a vowel between word-final and word-initial C's of various types – this phenomenon differs from the regular epenthesis which the language employs to break up geminates (!), under well-defined phonotactic conditions. Most of these have been observed in passing, though no careful work has been done on any of them. The process I am interested in today has never been described, to my knowledge.

We saw above how word-initial /jI-/ sequences acted, phonetically, after a variety of final consonants. Here's a pair of similar examples (this time involving initial /jE-/ after a velarized consonant). The dotted line shows the F2 trajectory for the underlined portion of (6) below, the dashed line the F2 trajectory for the underlined portion of (7). Both of these involve a sequence of a "heavy" consonant at the end of one word, followed by a front glide /j/, an underspecified lower mid vowel /E/, and another "heavy" consonant.

6. jEkArʷ pʷArʷAjInʲwEtʷ jEtʷAlʲ tʲEpʷtʷIj
   ekar *barāinwōt* *etal* *Jebti* …
   III.SG.SUBJ-PAST likewise go Jebti …
   'Jebti (and his mother and his father) likewise went.'

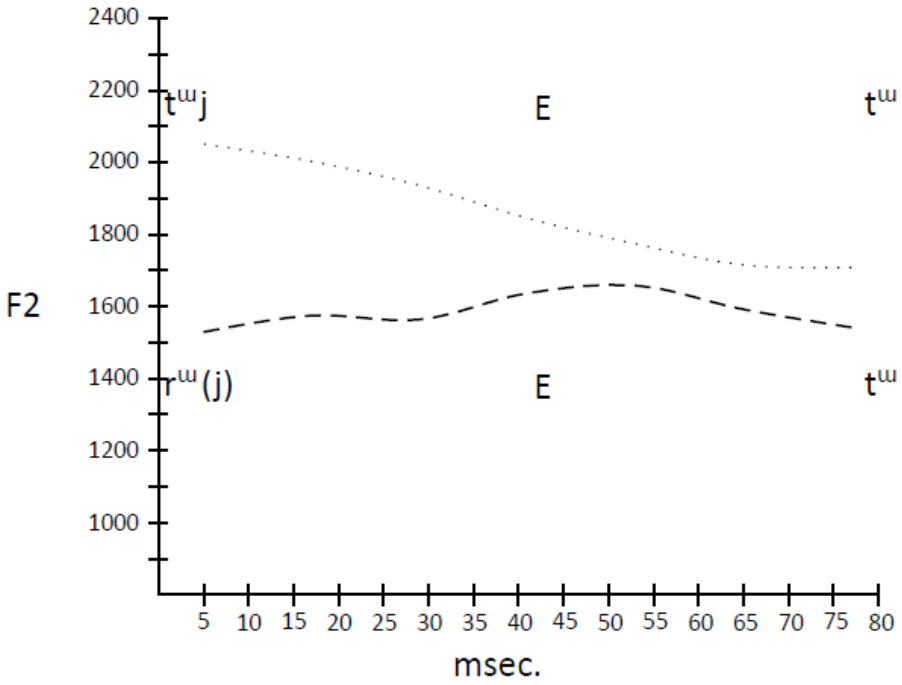
[8] Generally written **y* by Oceanists.

[9] For example, it is clear that before PMc **f* was lost in Marshallese, it, like other non-velarized labials, was palatalized (thus **fʲ*) such that when it disappeared, it generally left behind a **j* glide. There are wrinkles to this story; see Hale 2007 for a brief discussion of a few of them.

[10] This E vowel later splits into 'upper' and 'lower' mid values, so I've marked it as an intermediate reconstruction here.

[11] These are not PMc final vowels: those were lost in Marshallese. These are neo-finals.

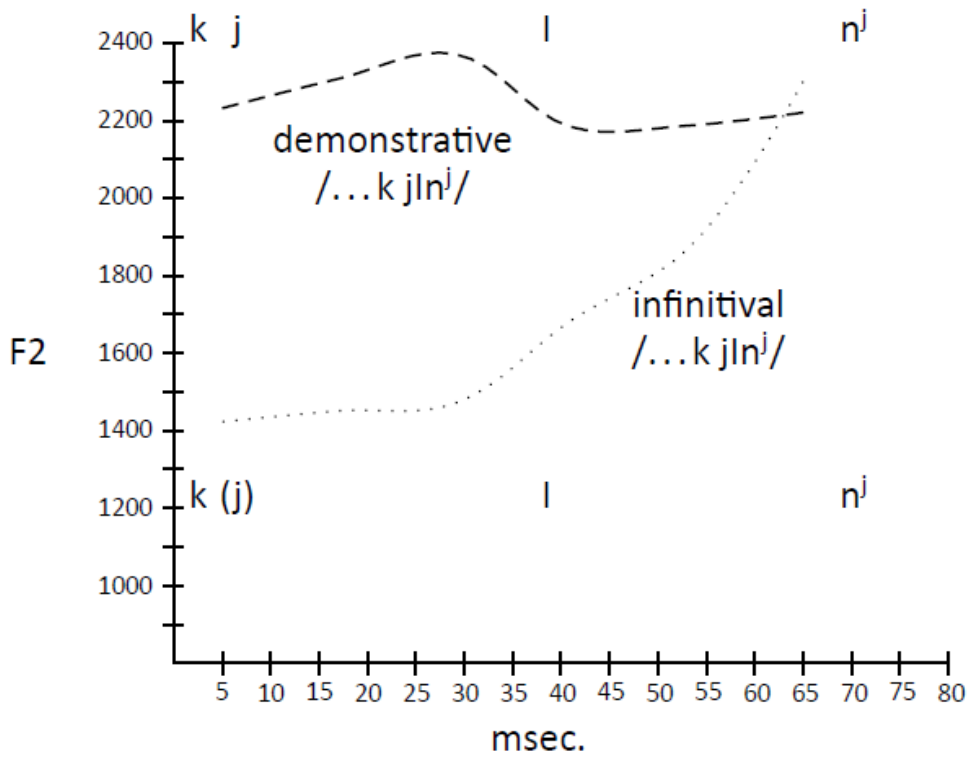
7. rʷAɯArʷ jEtʷAlʲ …
   raar *etal* …
   III.PL.SUBJ-PAST go
   'They went.'



It seems apparent that in the latter example (lower in the figure above), the vowel realization is staying in the back unround space throughout its duration (in contrast to the first, where we clearly see the effect of the initial /j-/ of *etal* 'go' in the high F2 at the beginning of the F2 tracing). It is as if the phonological input in the second string was not /…rʷjEtʷ…/, but rather /…rʷEtʷ.../, without the front glide.

It hardly needs to be pointed out that there is a marked syntactic difference between the two strings, with the 'likewise' adverbial having a less close relationship to the finite verb that does the Subject Agreement and TAM marker *raar*.

And this pattern repeats itself readily. An interesting contrast is that between the emphatic demonstrative *in* /jInʲ/ 'this (one)' and the infinitival marker *in* /jInʲ/ 'to', where the latter form under goes a kind of *gonna*, 'ready to' (in my dialect, [rɛri rə]) type close sandhi development, while the emphatic demonstrative does not. Contrast their F2 trajectories in *wiik in* `this week' (/wIjIk jInʲ/) and *pojak in* 'ready to' (/pʲEwtʲAk jInʲ) below.



The demonstrative acts like there is no preceding /k/, the effects of that consonant being essentially blocked by the intervening glide. But the infinitival marker, in sharp contrast, acts like there is *only* a /k/, no /j/; /k/, being a back consonant, induces a low F2.

So I feel confident, though we have only had a glimpse at the data, that Marshallese has an interesting phonology-syntax interface effect involving glides. Does this matter in some way?

First, I think it offers pretty strong support for Bender's phonemicisation, because it is fairly difficult for me to see how to get these dramatic effects if we are dealing with a "regular" (i.e., fully-specified) /i/ phoneme (in the two *in* cases above, for example). The phonetic data seems to point in the direction of phonetic underspecification in the case of these post-lexical processes (just as Choi argued it did for word-internal phonology).

But let me return briefly to the initial issue I raised, regarding the diachrony of the phonology-syntax interface. We know that Marshallese had an event of 'glide insertion', and there are two plausible stories one could tell about how our data relates to that event.

The first possibility is that glide-insertion was triggered only before and after a boundary of a certain strength. If at this Pre-Marshallese stage, the phonological phrasing that we see evidence for was already in place, the tight juncture created by the phrasing between, e.g., a verb 'be ready' and the infinitival marker **in* would have been such that glide insertion was *precluded*. This would mean that the diachronic

reason there is no glide in *pojak in* 'be ready to' is because the conditions on glide insertion were not met: there is no glide now, because (in a diachronic sense) there never was a glide.

Alternatively, of course, it could be that glide insertion took place as expected in the case of the infinitival marker **in*, but that the 'tight connection' which it shared with preceding governing verbs was such that, e.g., a preceding velar triggered assimilation of the glide (in backness), and, eventually, this was re-analysed simply as "glide deletion" in close sandhi.

The two stories differ in the nature of the lexical entries for the items involved. Under the first scenario, there would presumably be no reason for elements like the infinitival marker /(j)In^{j}/ to be stored with initial glides in UR at all (since the marker is presumably always in close connection with what precedes, though I have not checked this). Under the second, of course, it would have the glide in UR.

However, the behaviour of *etal* /jEt^{ɯ}Al^{j}/ 'go' certainly seems to indicate that we need a glide deletion rule, synchronically, in any event (since it showed no initial front glide in one syntactic environment above, but did show the glide in the other). So, the second story seems the better one for now: if we need to delete the glide in *etal* just when it is in 'close sandhi', and, e.g., the infinitival marker is always in close sandhi, then we can safely assume an underlying representation with a glide, taking the free ride to invariably delete it, if indeed this is what is going on.

But diachronically, of course, it is just as plausible that the glide in *etal* was only inserted when the word was in an appropriate environment for glide insertion (i.e., showed enough independence from the preceding word to form part of, let's say, a distinct phonological phrase). The synchronic state (invariable underlying glide, deletion in the relevant environments) is the reanalysis of this diachronic patterning by a learner who wanted to make a sensible grammar. The phenomenon thus patterns with well-known diachronic processes whereby (diachronically) deleted material is reanalysed as "inserted" in the environment complementary to the loss (think "linking-*r*"), or (diachronically) inserted material is reanalysed as "deleted" in the environment complementary to the insertion.

So, imagine we've reached the point where glide-deletion in close sandhi (i.e., between an "open-class" and certain types of "closed-class" lexical items) is a rule and that this rule, phonetically low-level enough to have escaped scholarly attention to date, arose because of a difference in the *phonetics* of the strings in question (in keeping with the OH). Since there was only *one string* before the assumed grammaticalisation of the elements in question (i.e., when these elements would have been "open class" lexemes), how did the listener analyse this (single) string in two different ways, as the "grammaticalisation" story requires?

This brings us back to the "big picture" issue of interest in this paper. The grammatical-isation literature has expended significant energy in identifying conditions on morpho-syntactic (or sometimes semantic) misanalysis in an attempt to constrain grammaticalisation paths. But if functional elements have a consistently different

phonological relationship to their co-constituent elements, we need there to be *both* morpho-syntactic "noise" (an unambiguous string will not be subject to misanalysis) *and* phonetic "noise" in the reanalysis context (a phonetically unambiguous strings would also not be subject to misanalysis). The listener will need to believe that the grammaticalised element can be construed as having an appropriate morphosyntactic relationship to the other elements of its constituent in the context of reanalysis, and an appropriate phonetic one in that context.

It follows from this, it seems, that there may be instances in which a string contains an element in an appropriate context for a morphosyntactic misparse, but *not* for a misparse of the phonetic relationships between elements that would be required for grammaticalisation. Grammaticalization in such cases would be blocked.

More globally, understanding that the "noise" required for misanalysis, and thus change, can result from the nature of the Language Acquisition Device parsing operations operating over data relevant to any of the modules of UG, and may, in cases such as those under discussion, require simultaneous misparsing at several levels, seems like it may play an important role in the development of a constrained model of change, and thus of a constrained model of the nature of UG.

References

Bender, Byron W. 1968. Marshallese phonology. *Oceanic Linguistics* 7(1). 16–35.

Choi, John. 1992. *Phonetic underspecification and target interpolation: an acoustic study of Marshallese vowel allophony* (UCLA Working Papers in Phonetics 82). Los Angeles: UCLA Department of Linguistics.

Hale, Mark. forthcoming. *Wackernagel's law: phonology and syntax in the Rigveda.*

Hale, Mark. 2012. Foundational concepts in the scientific study of sound change. In Solé, Maria-Josep & Recasens, Daniel (eds.) *The initiation of sound change*, 235–246. Amsterdam: Benjamins.

Hale, Mark. 2007. *Historical linguistics: theory and method.* Oxford: Wiley-Blackwell.

Keating, Patricia. 1985. Underspecification in phonetics. *Phonology* 5. 275–292.

Selkirk, Elizabeth. 1972. *The phrasal phonology of English and French*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation)

Mark Hale
Concordia University
mark.hale@concordia.ca

Indo-European linguistics meets Micronesian and Sunda-Sulawesi

Hans Henrich Hock

The common Indo-Europeanist view is that reconstructed Proto-Indo-European (PIE) cannot go back farther than the beginning of the fourth millennium BC and that its speakers must have been located somewhere in the Eurasian steppes. This perspective is based on the evidence of linguistic palaeontology – the linguistic evidence for reconstructing PIE words for 'horse' (**h₁(e)ḱwos*) and 'wheel' (**kʷekʷlos, *rotHo*),[1] combined with the archaeological evidence for the earliest horse domestication in the Eurasian steppes in the early fourth millennium BC, and the first archaeological evidence for wheels and wheeled vehicles in the mid-fourth millennium BC (see Anthony (2007) for a comprehensive discussion and Outram et al. (2009) for additional evidence of horse domestication). Significantly, conditions in other areas of Eurasia were not conducive for horse domestication, except for the Iberian peninsula (Bendrey 2012). Outside the steppes (or Iberia), therefore, words for 'horse' can only refer to already domesticated horses.

Publications by Atkinson, Gray, and colleagues (e.g. Atkinson & Gray 2006a; 2006b; Atkinson et al. 2005; Bouckaert et al. 2012; Gray & Atkinson 2003) – henceforth "Atkinson et al." – argue that, contrary to the standard Indo-Europeanist perspective, Proto-Indo-European goes back to ca. 6,500 BC. They question proposed reconstructions of the words for 'horse' and 'wheel', and instead suggest that the words were diffused through borrowing in post-PIE times. In support of the borrowing hypothesis for 'horse' and 'wheel', Atkinson and Gray (n.d.; 2006a: 294) argue that it is possible to reconstruct a Proto-Micronesian word for 'motor car', based on systematic correspondences in the daughter languages, such as Pohnpeian *sidōsa* and Woleian *sitōosa*, but that the word obviously is a recent borrowing from Japanese. Reconstructability, therefore, does not necessarily imply inheritance.

Atkinson et al.'s claims have met with a great number of responses – some online, some in print – that question both methodology and conclusions. Methodologically, Holm (2007) raises important questions about the mathematical models and statistical

[1] The fact that Anatolian does not have cognates of **kʷekʷlo /*rotHo*, but rather Hittite *ḫūrkis* without exact correspondences in the rest of Indo-European, might suggest that the Anatolian languages moved away from the other languages before the invention of the wheel (but after the domestication of horses), but other interpretations are possible. For instance, Hittite might have replaced an original **kʷekʷlo* with *ḫūrkis*. The root underlying Hittite *ḫūrkis* has counterparts in Tocharian A *wärkänt* and Tocharian B *yerkwantai* but these differ in formation. Ringe (2009) concludes that the Tocharian and Hittite words 'were derived independently' and that a common word for 'wheel' can be reconstructed only for the non-Anatolian languages.

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 63–67

approaches. Chang et al. (2015) have shown that when 'ancestry constraints' are incorporated, a phylogenetic method very similar to that of Atkinson et al. supports a Eurasian steppe origin of the Indo-European languages. Pereltsvaig & Lewis (2015) present a comprehensive (and somewhat intemperate) critique of many other aspects of the Atkinson et al. approach, with broad coverage of earlier literature, especially the evidence of linguistic palaeontology.

The primary aim of this paper is more narrow within the context of this debate, namely to address the claim that it is possible to reconstruct a Proto-Micronesian word for 'motor car', based on systematic correspondences in languages such as Pohnpeian (*sidōsa*) and Woleian (*sitōosa*), even though the words in questions are borrowings from Japanese. The significance of whether the claim of Atkinson et al. holds, however, extends beyond this debate, for if correct, it raises serious questions about the ability of historical linguists to distinguish borrowings from inherited cognates.

As it turns out, a closer look at Micronesian (and beyond) shows that the claim is problematic.

First, Satawalese, closely related to Woleian, not only offers the form *sitoosa*, which is phonologically similar to the Woleian and Pohnpeian forms and is likewise a borrowing from Japanese *jidōsha* 'motor car', but it also has the variant *stosa*, probably reflecting the Japanese allegro form *j$^{(i)}$dōsha*. Significantly, this form contains an initial consonant cluster that does not occur in native words, but is only found in borrowings (Roddy 2007: 31, 189, 190). Even without knowing that Japanese was the source for these words, proper historical methodology would lead to the conclusion that *stosa* and its variant *sitoosa*, as well as the similar forms in the other Micronesian languages are borrowings (from a non-Micronesian language) and that hence a reconstruction to Proto-Micronesian must be excluded.

To this evidence must be added that Palauan also has a borrowing, *sidosia* (Bradshaw 2007). Unlike Pohnpeian, Woleian, and Satawalese, however, Palauan does not belong to the Micronesian or even the larger Oceanic branch, but is a geographic outlier of Sunda-Sulawesi (Wouk & Ross 2002[2]). The fact that among the Sunda-Sulawesi languages, Palauan is the only one to show this word, combined with its geographical proximity to the Micronesian languages, suggests that we are dealing with a regional innovation, rather than with something that is inherited from a proto-language common to Micronesian and Palauan/Sunda-Sulawesi.

Now, there are indeed cases where linguistic forms showing regular phonological correspondences should not be reconstructed to the proto-language even though they do contain inherited elements (see e.g. the discussion in Hock 1991: §18.8). These cases involve morphologically complex forms, such as English *there-by* : German *da-bei* 'with that', where the component elements can be reconstructed, but the combinations of these elements could be (and in fact, were) created independently in the two languages. The Micronesian and Palauan words for 'motor car', however, are

[2] Dyen (1965) argued for Palauan as an independent branch of Malayo-Polynesian.

not morphologically complex (within these languages), and hence a hypothesis that they were created from inherited morphological material can be safely excluded.

The phonological similarity between the Micronesian and Palauan words contrasts markedly with the phonological differences between the earliest attested Indo-European words for 'horse' and 'wheel'. These differences (e.g. Sanskrit *aśvas* : Greek *hippos* : Latin *equus* for 'horse') show that the words cannot be recent borrowings, diffused from one Indo-European language to another. If they were, then a form like Sanskrit *aśvas*, or its Proto-Indo-Iranian ancestor **aćwas* would wind up as something like Greek *as(s)as** or Latin *aswas**; a nativisation as *hippos* or *equus* would be preposterous. Similar concerns apply if any of the other languages were chosen as the source language.

Atkinson et al. are aware of this problem and try to address it by claiming that the borrowing must have been made at a time "that most of the major Indo-European groups were just beginning to diverge" (Atkinson & Gray, n.d.; 2006b: 103). But how can this scenario be distinguished from one where the Indo-European linguistic groups were still part of a dialectally diversified Proto-Indo-European? Further, for their hypothesis to work it would have to make the gratuitous assumption that over the first ~3000 years the Indo-European languages underwent virtually no changes, but then changed extremely rapidly during the next ~2000 years.

Under the circumstances, Atkinson and Gray's (n.d.; 2006a: 294) opting for a borrowing account in post-PIE rather than inheritance from PIE is arbitrary and looks like special pleading. The phonological correspondences between the different languages can only be explained in terms of divergent historical changes from an ancestral, reconstructed form **h₁(e)ḱwos*.[3]

The fact that Atkinson et al. failed to inquire more deeply into the fuller range of Micronesian and Palauan evidence may be considered to be significant, and not just a minor slip. It is reminiscent of the problem pointed out by Donohue et al. (2012a: 519) about Atkinson et al.'s approach, namely that it is not able to distinguish between "social and spacial proximity" and "inheritance" in the lexicon-based classification of Polynesian.[4] Ultimately, the difficulty appears to be attributable to the fact that Atkinson et al. draw on corpora produced by other scholars, without themselves examining the reliability of these corpora by employing standard comparative-historical methodology to critically scrutinize the evidence that they are based on.

[3] The difference between **h₁eḱwos* and *h₁ḱwos* follows well-established PIE patterns of morphophonemic variation.

[4] This publication is part of a larger discussion between Donohue et al. (2012a; 2012b) and Greenhill & Gray (2012).

References

Anthony, David W. 2007. *The horse, the wheel, and language: How Bronze-Age riders from the Eurasian steppes shaped the modern world*. Princeton: Princeton University Press.

Atkinson, Quentin D. & Gray, Russell D. 2006a. Are accurate dates an intractable problem for historical linguistics? In Lipo, Carl P. & O'Brien, Michael J. & Collard, Mark & Shennan, Stephen J. (eds.) *Mapping our ancestors: Phylogenetic methods in anthropology and prehistory*, 269–296. Chicago: Aldine.

Atkinson, Quentin D. & Gray, Russell D. 2006b. How old is the Indo-European language family? Illumination or more moths to the flame? In Forster, Peter & Renfrew, Colin (eds.) *Phylogenetic methods and the prehistory of languages*, 91–109. Cambridge: The McDonald Institute for Archaeological Research.

Atkinson, Quentin D. & Gray, Russell D. n.d. Second response to Trask. Originally posted at http://www.psych.auckland.ac.nz/psych/research/Evolution/Response%20to%20Trask%20Take2.doc. (The link is no longer active.)

Atkinson, Quentin D. & Nicholls, Geoff & Welch, David & Gray, Russell D. 2005. From words to dates: Water into wine, mathemagic or phylogenetic interference? *Transactions of the Philological Society* 103(2). 193–219.

Bendrey, Robin. 2012. From wild horses to domesticated horses: A European perspective. *World Archaeology* 44(1). 135–157.

Bouckaert, Remco & Lemey, Philippe & Dunn, Michael & Greenhill, Simon J & Alekseyenko, Alexander V. & Drummond, Alexei J. & Gray, Russell D. & Suchard, Marc A. & Atkinson, Quentin D. 2012. Mapping the origins and expansion of the Indo-European language family. *Science* 337. 957–960.

Bradshaw Joel. 2007. Japanese and other loanwords in Palauan. (http://faroutliers.wordpress.com/2007/07/28/japanese-and-other-loanwords-in-palauan)

Chang, Will & Cathcart, Chundra & Hall, David & Garrett, Andrew. 2015. Ancestry-constrained phylogenetic analysis supports the Indo-European steppe hypothesis. *Language* 91. 194–244.

Donohue, Mark & Denham, Tim & Oppenheimer, Stephen. 2012a. New methods for historical linguistics? Calibrating a lexicon-based methodology for diffusion vs. subgrouping. *Diachronica* 29(4). 505–521.

Donohue, Mark & Denham, Tim & Oppenheimer, Stephen. 2012b. Consensus and the lexicon in historical linguistics: Rejoinder to "Basic vocabulary and Bayesian phylolinguistics". *Diachronica* 29(4). 538–546.

Dyen, Isidore. 1965. *A lexicostatistical classification of the Austronesian languages*. Baltimore: Waverly Press.

Gray, Russell D. & Atkinson, Quentin D. 2003. Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* 426. 435–439.

Greenhill, Simon J. & Gray, Russell D. 2012. Basic vocabulary and Bayesian phylolinguistics: Issues of understanding and representation. *Diachronica* 29(4). 523–537.

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

Holm, Hans J. 2007. The new arboretum of Indo-European "trees": Can new algorithms reveal the phylogeny and even prehistory of IE? *Journal of Quantitative Linguistics* 14(2–3). 167–214

Outram, Alan K. & Stear, Natalie A. & Bendrey, Robin & Olsen, Sandra & Kasparov, Alexei & Zaibert, Victor & Thorpe, Nick & Evershed, Richard P. 2009. The earliest horse harnessing and milking. *Science* 323. 1332–1335.

Pereltsvaig, Asya & Martin W. Lewis. 2015. *The Indo-European controversy: Facts and fallacies in historical linguistics.* Cambridge: Cambridge University Press.

Ringe, Don. 2009. Inheritance vs. borrowing of reconstructed vocabulary. *Language Log* January 10, 2009. (http://languagelog.ldc.upenn.edu/nll/?p=994)

Roddy, Kevin M. 2007. *A sketch grammar of Satawalese, the language of Satawal Island, Yap State, Micronesia*. Manoa, HI: University of Hawaii. (MA Thesis)

Wouk, Fay & Ross, Malcolm (eds.) 2002. *The history and typology of western Austronesian voice systems.* Canberra: Australian National University.

Hans Henrich Hock
University of Illinois at Urbana-Champaign
hhhock@illinois.edu

Nembangahu – The big stone

Leina Isno

This is a Ninde tribal story I worked on with Liz a few years ago. I have translated the story into Vanuatu Bislama, Solomon Islands Pijin, Papua New Guinea Tok Pisin, French and New Zealand Te Reo Māori, and I am currently in the process of getting it published as a children's book for the Pacific children. The two versions included here are the Ninde version of the story which will appear in the children's book, as well as a line-by-line translation into English of the version I worked on with Liz.

Nembangahu – The big stone (A tribal story by Leina Isno in the Ninde Language – Vanuatu)

Introduction

Na xonta ka bitiworx nimbitiworyene tuwa.
Beh nimbitiworyene nge kha tor ha nembangahu, sene na wul nge.
Kala nga kine hor kha rewul lele nevet did'lepe ha nembangahu-sene naha rekor pa ne.

Story

We sei, kaiwut rakh ma watarwor a niye rakh khor. Rakh khor ha nembangahu. Rakh vian nami. Rakh skeske mboilelow.
Tahai tuwa kha kaiwut tho ve ngolow nanyene. Niye tho sla nap. Niye honta rakh pe ngaha nemen ai nambalai ai titu. Niyep lo sla ras, pe thasongone mbilbil.

Lele nute tuwa eti Lorbunwoi hane kha wian kaiwut. Niye tor awut niye wian nambalai ma titu ma nemen. Niye liliworhx maqas nakha.
Kaiwut lele nembangahu, niye tor kha niye mas pe lo te pe gaha nemen ma titu ma nanyene te niye pal nge pul l'ami. We sei, niye lome. Niye lip narxveh ga niye, yel nge te ve ngolow nemen ma titu. Kaiwut lo kha wul saha tele li titu a kaiwut lorbunwoi nge tho wornetes. Yenta kha sene Kaiwut Lorbunwoi tor pa nge ya. Niye ngar narxveh nga niye, ye luworow paha titu. Titu mes mbin ye niye lipluwo titu ye yel pul ha nembangahu. Rahk ma watawor rahk pan luwo titu nge-rahk wisin ge lambh ai rahk piyor niyow nene te rahk pan ge.
Kaiwut nene nembangahu yor tartar semba mange tahait lapo. Le numhoine sei, huwo, til, ves, selme kha niye tho te pe gaha titu lorbunwoi.

We sei kha, kaiwut Lorbunwoi golow bil pa titu a niye hor we nute. Niye liwor titu kha li man wud re nono semba. Niye wusuworh tartar rome. Niye timan,
"O titu a kine re kho sene ya? sumo yayaha ba naha. Sumo wena me nemen ma titu ma nambalai ge ho! sumo ya. Sumo ge tor mene? Sumo ge lo mene pul?"

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 69–75

Tahaine re lo. We sei niye lipa man titu kha re skeske pa. Titu res tortor sikwei mintumko ai wenelis ai lahap. Niye laasaha sepme. Ngome baha ye ti mankanege,
“kaiwut tuwa d’tor ha nembangahu haindeve nati, niye kha niye pulme yanganeke kha paha titu a kine hor ma nemen ma nambalai a kine. Na pe gatep munt e rahk ha nembangahu!”

Gom baha ye lip narxve ma numbutluwo ga niye. Niye waha ge ambustikone ye wor netes. Niye humbutei me wor netes. Niye logoteh wor netes ma niye yale rome. Niye yale nambi tuwa e yale manke, (te niye honta pho nele ne titu ho)
“Nitiven betep, ku wen tor mene, rahn bangaru, rahn bangaru.”
“Kukura ku!” Nemen yohoho ha nembangahu erei.
Tahait titu ma nemen re totoho lele nalse kha kaiwut Lorbunwoi hor me naha. Niye yale lis ma wisarow lis narxve sei lis. Niye lo rome pul sasa Labo, niye yor rome mange lis; yale ma wisarow nubutluwo sei lis.
“Nitiven betep, ku wen tor mene, rahn bangaru, rahn bangaru.”
“Kukura ku!” Nemen yohoho ma totoho ha nembangahu erei.

Kaiwut Lorbunwoi yantep we neitlo kha yale lis ma wisarow lis narxve sei lis. Niye lo rome pul sasa Lambetep. Niye yale romba,
“Nitiven betep, ku wen tor mene, rahn bangaru, rahn bangaru.”
“Kukura ku!” Nemen yohoho ma totoho ha nembangahu erei.
Nemen kha re mes luwa awut naha re kukura semba lele nalse ha nembangahu erei. Kaiwut rakh ma watawor rakh kor lele nami ya rakh ha nembangahu kha rakh ho pa man norpo kha re welela pa. Nemen re ma titu kha re yoyoho ras pa lele nalse. Wiya kha rap miya baya. Raph langlang sepme nimiyathane te rakh horh paya.

Kaiwut Lorbunwoi yantep Lambetep kha yale lis ma wisarow lis narxve sei lis. Narxve dum pa lumete. Niye lo rome pul sasa yale rome nambi ge lis,
“Nitiven betep, ku wen tor mene, rahn bangaru, rahn bangaru.”
“Kukura ku!” Nemen yohoho ma totoho ha nembangahu erei.

Kaiwut Lorbunwoi wul sasa nembangahu wortes kha y’aa me nevet lepe ge. Niye luworo lis narxve ka we dum pa ha nami ha nopmo nevet erei. Lele nami erei kha kaiwut rakh ma watawor kha nimiyathene lepe spo rakh. Rakh hor ras nemela. Rakh skom sikvei lami te rap lothloth. Rakh holangre pa kaiwut Lorbunwoi kha niye wul paya tele puhus rakh. Watawor ti tele kaiwut ti mangkhe,
“nisa khati pa ge ya, kus ptho ras wei kha kut tukulwul-wiya ye kup li pa ge ya!”
Kaiwut Lorbunwoi yantep wor nami aat kha wurow meh,
“Ai kaiwut ma watawor, kamor ha khor me lami? Na goloworh titu ma nemen a kineh roh. Kamor ha lime naha sene?”
Rahk ti mangeh,
“Ai Kaiwut, kamarkh kha ha mia nap! Kamarh ha lameh nisa ge ho!” Rakh lang sepme.
Kaiwut Lorbunwoi sar netelu ye wul lami. Ghom baha yete li kaiwut mia nap. Niye kunkun sepba niye. Watawor su mia nap. Niye spo. Niye ghom baha ye wusuwor tele kaiwut,

“naho nemen tuwa titoho yankhe (yankaneke), ma titu tuwa su tortorho me yanke!?”
Kaiwut timankanenge,
“O, nalame nisa ge! Khamar ha lameh nisa ge!” Niye ti ge lele nimiyathene dlepe. Awud niye langlang sepme.
Kaiwut Lorbunwoi gom baha ye ti mangke,
“na wian titu man nemen ma nambalai ho re kor haindeveh awut re we nono pa!”

Watawor timankhe,
“kamar ha lameh nisa nuk ti nge. Kine kha na mia, ka mia, ka mia nap-nalame nisa nuk ti nge. Nalameh. Titu ma nemem re wul yakanege awut kamar has li wei naha!”

Tahait niye timan kha kaiwut lorbunwoi ngom baha ye ti man ge,
“kamor ha langlang sepme. Nemen a kine kha re khor yanke. Naholangre re kor yanke. Kamor kha lip langal ai selme. Kamor hayaha nemen langal ai selme re kor yanke te naholangre nele ne naha. Naha Re wul re kor yankanekhe!”

Niye gom baha ye lip narxve nga niye, ye luworow narxveh, hrus luwo kaiwut rah ma watawor, gom baha yete lip nemen ho, titu ho, gom baha ye yel lis naha pe Lorbunwoi. Naha re kor lis Lorbunwoi.
The end!
i sep me ya

The Nembangahu Story – A tribal story of Ninde Translation into English by Leina Isno (Vanuatu)

Introduction

Na xonta ka bitiworx story tuwa.
I would like to tell a story of a kind.

Be storian nge kha tor ha nembangahu, sene na wul.
But the story began on the big stone called Nembangahu, where I come from.

Family a kine kha wul lele nevet did’lepe ha nembangahu.
My family originated (or came from) inside the big stone, Nembangahu.

Story

We sei, kaiwut xha ma watarworhx xha khor. Xha khor ha nembangahu. Xha vian nami, xha skeske mboilelow.
One day, a man and his wife stayed. They lived on the big stone Nembangahu. They have a house but they don’t have any children.

Be tahaituwa kha kaiwut lo ve ngolow nanyene…..nanyene lele….Niye honta pe ngaha nemen ai nambalai ai titu….Niyep lo sla sla sla nap
But some days, the man would go to look for food….food in the…..He would go to get birds or hunt pigs or chickens/hens/roosters. He used to go very far from home.

Lorbunwoi (kha) vian kaiwut niye tor be niye vian nambalai ma titu ma nemen. Niye liliworhx nahax.
A place called Lorbunwoi that has a man who lived there, but he had pigs and chickens and birds. He looked after them.

Kaiwut lele nembangahu, niye mas pe lo te pe ngaha nemen ma titu ma nanyene te niye pal nge pul l'ami.
The man from the big stone Nembangahu always went to steal/take the birds and chickens and the food and he would take them back home to his place.

Alei, niye lo, lip narxveh nga niye, yel nge te ve ngolow nemen ma titu. Kaiwut lo kha wul saha tele li titu a kaiwut lele….sene niye (the farmer) tor. Niye ngar narxveh nga niye, niye luworow, lipluwo titu.
Then, one day, he went, got his bow (and arrow), took it and went to look for birds and chickens. The man went and came to the chicken that belonged to the farmer. This is where the farmer lived. Then he took his bow, shot his arrow and took away the farmer's chicken.

Beh kaiwut tuwa titor, tuwa ti'vian titu ma nambalai, niye liliwor ma'as nambalai a niye hor ma'as.
But the farmer who raised the livestock (or lived there), the one who had the chickens and the pigs has really looked after his pigs well.

Beh tahai tuwa kha niye liwor titu ma nambalai, liman,
But some days (or one day) that he looks at the chickens and the pigs, he sees that

"ooooooo, sumo lipluwo pa nge. Sumo lipluwo nemen ai sumo lipluwo nambalai ai sumo lipluwo titu."
"Oh, someone has taken it! Someone has taken the bird or someone has taken the pigs or someone has taken the chicken!"

Tahait d'lor, ve sei niye lipa man titu kha re skeske pa. Niye laa sepme sene titu re lo pe. Ngome baha ye ti mankanege,
When the days go by, one day he realised that the chickens are minimal or not many in numbers. He doesn't know where the chickens have gone to. The he said,

"kaiwut tuwa d'tor ha nembangahu nati, niye kha mestem kha niye pulme yanganeke kha paha titu a kine hor ai nemen a kine hor ai nambalai a kine."
"The man that lives on the big stone Nembangahu, I think, he must come here and take my chickens and the birds and the pigs!"

Alei, niye….nambi tuwa niye tor tele man niye, niye…….manmene ye napbiti nge ya…te niye tor te yel narxveh nga niye. Alei niye humbutei wor netes (sorry mistake). Niye logoteh wor netes be niye yale. Niye yale nambi tuwa.
Then, he (this is the song that he sang)…..Leina is trying to figure out how to say it properly here….that the farmer has now taken his bow. Then he ran along the sand beach. He walked along the sea but he kept singing. He was singing a special song.

"Nitiven betep, ku wen tor mene, rahn bangaru, rahn bangaru."
"Nitiven Betep, where did you go to, closer to Nembangahu, closer to Nembangahu"

Kaiwut tor kha hor pa wut……nemen titu yohoho (totoho) lele name ya kaiwut tuwa tor ha nembangahu.
The man at Nembangahu stayed and heard…..the bird and the chickens in the house crowed

"kukura ku!"

Alei, niye luworow lis narxve nga niye alei, logote wor netes alei yale lis,
Then, he shot another arrow from his bow, then kept walking along the seashore and sang again

"Nitiven betep, ku wen tor mene, rahn bangaru, rahn bangaru."
"Nitiven Betep, where did you go to, closer to Nembangahu, closer to Nembangahu"

"Kukura ku!"

Nemen, nemen kha mes luwa be man niye, niye kukura sebmba lele nalse.
The bird/chicken has already been dead but it just kept crowing in the earth oven.

Alei, kaiwut tor tartar rome tele hup bul ve nembangahu. Narxveh nga niye, niye luworow nge. Alei niye wulsasa (romeh) luwo (laha). Tahait niye wul pa lele narxveh nga niye wud sasa niye pno nge, narxveh, niye luworow narxveh kha narxveh we torpa ma name ya kaiwut ha nembangahu.
Then, the man kept running towards the big stone Nembangahu. His bow, he kept shooting with the arrow. He kept coming closer. When he was about to run out from his arrows, he shot one last arrow and it went inside the house on the big stone.

Alei, niye, niye wor romeh, lo lo lo we kasem name ya kaiwut lele nembangahu. Ghom baha yete lop'h l'ami we li kaiwut lami; watawor spo. Niye ghom baha ye wusuwor tele kaiwut,
Then he kept on going, walking until he got to the house on the big stone Nembangahu. Then he went inside the house to see the man; the woman his wife sat there too. Then he asked the man,

"naho nemen tuwa tortor (titoho) yankhe (yankaneke), titu tuwa tortorho yanke!?"
"I heard a bird singing in here, a chicken crowing in here!?"

Kaiwut timankanenge,
The man said,

"Oh, nalame nisa ge!"
"Oh I don't know what that is!"

Beh niye langlang sepme.
But he was only lying.

"Nalame nisa ge!"
"I don't know what that is!"

Niye ghom baha ye ti tele…..kaiwut ghom baha ye ti mankaneke,
Then he got up and said….The farmer man got up and said,

"na wian titu man nemen ma nambalai ho kor haindeveh beh re lus pa-re lus re lus yaha pa."
"I have chickens and birds and pigs which lived over there (at Lorbunwoi) but they have all gone lost. They have been lost forever."

Awut watarwor tor lele name kha ti mankhe,
But the woman inside the house said,

"kamem lameh nisa nuk ti nge. Kine kha na mia, ka mia, ka mia nap-nalame nisa nuk ti nge. Kames-nalame man….titu ai nambalai re wul yakanege….nas….kamem es li wei naha."
"We don't know what you are saying!" I am very sick, very sick, and very sick indeed – I really do not know what you are saying. We – I do not know whether the chickens or the pigs came here…..we did not see them!"

Tahait niye timan kha kaiwut lorbunwoi ngom baha ye ti man ge,
When she said that, the farmer guy then said,

"kamor ha langlang sepme. Nemen a kine kha re khor yanke. Naholangre re kor yanke ai kamu kha e¨ lip langal. Kamu e¨yaha nemen langal re kor yanke te naholangre nge. Re wul re kor yankanekhe!"
"You two are lying. My birds are here, I know they are here. The two of you took 10 of my birds which are here and I know. They came and they stayed here!"

Rah gom baha ye rah ti mankeh,
Then, they said,

"kamem eh lame¨ nisa nuk ti nge!"
"We don't know what you are saying!"

Niye gom baha ye lip narxve nga niye, gom baha ye luworow narxveh, hrus luwo kaiwut rah ma watawor, gom baha yete lip nemen ho, titu ho, gom baha ye yel lis naha pe lorbunwoi.
Then, he took his bow and he shot his arrow, killed the man and his wife, then took all his birds and chickens and returned home to Lorbunwoi.

(i sep me ya)
(That is all)

Leina Isno
leina.isno@gmail.com

The syntax of *wherewithal*

Richard S. Kayne

Abstract

This paper discusses the derivation of sentences containing *wherewithal*, arguing that its three morphemes should be treated as atoms by Merge. The proposed rich syntactic analysis suggests an account for the absence of words such as **whatwithal*, and it also suggests that *wherewithal* is not a syntactic constituent in sentences in which it occurs.

1. Introduction

English, or at least some English, allows (1) in approximately the sense of *They don't have the means/money to buy that big a house*.

1. They don't have the wherewithal to buy that big a house.

A lexicalist approach to syntax in the spirit of Chomsky (1970), although it might well recognize that *wherewithal* is composed of three morphemes, *where, with* and *al(l)*, would, since it takes words to be atomic with respect to Merge, deny that the three morphemes of *wherewithal* are put together by Merge.

On the other hand, if the morphemes that constitute a word can be put together by Merge, as they could be in an updated and generalized version of Chomsky's (1957) affix-hopping analysis of English forms such as *played,*[1] then the component morphemes of *wherewithal* could be too. Yet allowing that the three morphemes of *wherewithal* are treated as atoms by Merge underdetermines the derivation of (the relevant part of) (1) to a significant degree. We don't automatically know whether *where, with* and *al(l)* are affected solely by external merge or whether they are also subject to internal merge (movement). Nor do we automatically know in what order Merge affects them. Finally we don't automatically know whether or not there are in (1) additional relevant morphemes that are not pronounced.

In this paper, I will try to spell out what an analysis might look like that takes the component morphemes of *wherewithal* to be manipulated by Merge. In so doing, I will suggest (the beginnings of) an account of the absence in English of various words that resemble *wherewithal* but are not at all possible. For example, the following are unacceptable:

[1] Cf. Baker (1985; 1988), Pollock (1989), and Roberts (2017), among others.

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 77–86

2. *whatwithal; *whenwithal; *herewithal; *therewithal; *wherewithboth

A merge-based account of (2) is, I think, within our reach, in a way that would seem not to be the case within a lexicalist (non-merge) approach to internally complex words.

It should be noted that to call *wherewithal* a word is in effect to take it to be a constituent. But words in standard orthography, although they may well correspond to syntactic constituents in many cases, do not necessarily do so,[2] and in fact the analysis to be suggested in what follows leans toward the conclusion that *wherewithal* is not a syntactic constituent in sentences in which it occurs.

2. The subparts

The *wherewith* subpart of *wherewithal* recalls the following English words:[3]

3. whereby; thereby; whereupon; therefore; wherein

These are to one degree or another still part of contemporary English. In earlier English, words of this form were more productive (e.g. *whereof, thereof, wherefore*), in the manner (as a first approximation) of contemporary Dutch and German.[4] As in van Riemsdijk's (1978) original discussion, we can take, say, *therefore* to be very close to *for that* (or more exactly to 'for that REASON' with a silent noun corresponding to *reason*), and in the same way we can take *wherewith* to be very close to *with what*, apart from the difference in order between adposition and wh-word, and the difference between the R-pronoun *where* and the non-R-pronoun *what*.

As for the *-al* subpart of *wherewithal*, it seems virtually certain that it is simply *all*.[5] (On the spelling difference, cf. *alright, altogether, almost, also*, *already*, all with *al-* rather than *all*; also *careful, spoonful* vs. *full.*) Taking the *-al* of *wherewithal* to equal ordinary *all* might seem a bit surprising at first, insofar as *all* in English cannot normally be preceded by *the*, whereas the *-al* of *wherewithal* normally is (though not immediately preceded), as in (1).

However, the definite article in (1) recalls that found overtly with *whole* (as in *the whole day*), which has a fair amount in common with *all*. In addition, the fact that *the* in (1) co-occurs with *all* (written as *-al*) recalls even more specifically (4), in which *all* is not in its more usual DP-initial position.[6]

---

[2] Cf. Myers (1987); Koopman & Szabolcsi (2000); Julien (2002).

[3] Also *whereabouts*, which raises somewhat different questions; see Kayne (2014).

[4] Cf. Kayne (2004: Part III) on a possible link within Germanic to 'OV' vs.. 'VO'. For recent discussion of OV/VO, see Cinque (2016).

[5] Note that the final vowel of *wherewithal* (equal to that of *all*) is not reduced, as opposed to that of *refusal*, *removal*.

[6] In this example, there is very likely a silent N, too (akin to EFFORT perhaps), within the DP containing *his all*.

4. He gave it his all.

Returning to the *wherewith* component of *wherewithal*, let me take it to be an instance of a relative pronoun followed by an adposition, parallel to (5).

5. They had a plan whereby they would both leave at the same time.

From this perspective, (1) should be thought of as (6).[7]

6. ...the wherewith all to...

In contrast to (5), though, there is no visible noun in (6) between the determiner (*a, the*) and the relative pronoun *where*. Two possibilities come to mind. First, the apparently missing noun might be *all* itself. The second possibility, which is the one I will prefer here, has there being a silent noun in (6), much as there is arguably a silent noun present in headless relatives of the sort seen in (7).[8]

7. They ate what was put in front of them.

If *all* itself is not the nominal head of the relative in (6),[9] the question arises as to what the status of this *all* is. Let me take as a clue the existence in earlier English of interrogative *wherewithal*, as in (8).[10]

8. Wherewithall shall wee be clothed?

This to my more recent ear recalls McCloskey (2000: 66) on West Ulster English sentences such as (9) and (10).

9. Who all did you give tea to?

10. Where all did you move the books to?

If so, then we can think of (8) as (11), with the difference in position of the adposition reflecting the extensive use in earlier English of *wherewith, whereof* et al.

11. Wherewith all...?

---

[7] Note the parallelism with:

i. They don't have the money with which to buy that big a house.

Some speakers accept the following, recalling Collins (2014) on deleted relatives.

ii. They would like to buy that car, but they don't have the wherewithal.

iii. They don't have the wherewithal for that car.

[8] Cf. Groos and van Riemsdijk (1981).

[9] Keeping in mind that the term 'head (of a relative)' here is not the X-bar sense of that term.

[10] This example is from the King James version of the Bible, as given by the Oxford English Dictionary (*wherewithal*). Archaic *withal* is found in the OED too; in some of its uses, it may be accompanied by a silent WHERE.

If the *-al* of earlier English interrogative *wherewithal* is, as I am suggesting, related to McCloskey's *all* in (9) and (10), then it is plausible to think that current English relative *wherewithal*, as in (1)/(6), also contains an *all* that is substantially the same as that of West Ulster English. Thus, we now have additional grounding for taking the *-al* of current *wherewithal* to equal *all*.

As for the constituent question alluded to earlier, i.e. the question whether *wherewithal* in (1)/(6) is a syntactic constituent, it is logically possible that it is a constituent, but the presence of the adposition *with* between relative pronoun *where* and *al(l)* makes that unlikely. If so, then *wherewithal* is an example of a word that is not a syntactic constituent.

3. Derivation

As for more specific questions of properties and derivation, we can take *wherewith(al)* + infinitival relative in (1)/(6), repeated in (12) to have a lot in common with (13).

12. They don't have the wherewithal to buy that big a house.

13. They don't have the money with which to buy that big a house.

This is the case even though a simpler *where*-initial infinitival relative is at best marginal in English (as opposed to French).[11]

14. ?They need a place where to store their books.

Sharply impossible, on the other hand, is the result of adding *all* to (13), as below.

15. *They don't have the money with which all to buy that big a house.

It may be that the contrast between (12), which allows *...wherewith all...*, and (15), which disallows *...with which all...*, can be understood as follows. Contrary to appearances, (12) does have a pronounced relative head, namely *all*, or more likely *all + silent noun (N)*, meaning that (12) would have at an intermediate stage in its derivation.

16. all N wherewith to buy...

Wherewith would then move to the left of *all* (perhaps via remnant movement). This movement would have something in common with the movement of relative clauses from post-nominal (post-head) position to pre-nominal (pre-head) position suggested for languages like Chinese and Japanese in Kayne (1994, sect. 8.3), especially for cases like *the recently arrived letter*, in which the preposed relative ends up being preceded

[11] Note that the English restriction found in infinitival relatives, as in (i):

i. *We're trying to think of someone who to invite.

is found in both French and Italian even with finite relatives, with direct object *qui*, *cui*, cf. Kayne (1976) and Cinque (1982).

by *the*, as well as having something in common with the movement of *destruct*- past -*ion* in the derivation of derived nominals like *the destruction of the bridge*, as proposed in Collins (2006) and Kayne (2008).

The idea would then be that the specificity of *wherewith* in contemporary English is precisely that it, and it alone, is subject to movement past *all*, so that (15) is not derivable. Of course (15) is not derivable for a second reason, from this perspective, namely that it would have two heads, both *money* and *all* (or *all + N*, henceforth just *all*), as also seen in (17).

17. *They don't have the money wherewithal to buy that big a house.

Removing *(the) money* from (15) and undoing the movement of *with which* yields (18).

18. *They don't have all with which to buy that big a house.

The unacceptability of (18) is almost certainly linked to that of (19).

19. *They bought all which was on the top shelf.

Relatives headed by *all* are hardly possible with relative *which*.

In a similar vein, we can note the impossibility in contemporary English of (20).

20. *They don't have the money wherewith to buy that big a house.

In effect, *wherewith* is currently possible only in combination with *al(l)*, in a way that recalls the limitation of German relative *was* ('what') to light-headed relatives.[12]

English relatives headed by *all* are to some extent possible with *that* rather than *which*.

21. ?They bought all that was on the top shelf.

In contrast to (12) with *wherewithal*, though, (21) cannot have *the*, as shown below.

22. *They bought the all that was on the top shelf.

That the *all* of *wherewithal* can be preceded by *the* within its DP recalls French *le tout* (on the assumption that French *tout* is a good match to English *all*, which seems plausible), as in (23).[13]

23. Le tout a été envoyé à Paris.
    the all has been sent to Paris
    'the whole thing has been sent to Paris.'

It also recalls, though the determiner is different, (4), repeated below.

[12] Cf. Citko (2004).
[13] On *le tout* vs. *tout* alone, see Obenauer (1994).

24. He gave it his all.

In addition, it recalls, less directly, (25) and (26).

25. They watched his every step.

26. They ate the whole cake.

Why universal quantifier-like elements can sometimes be preceded by *the* and sometimes not is left an open question.

4. *al(l)* vs. *both*

That the *al(l)* of *wherewithal* and the *all* of (24) are closely related is supported by the fact that neither can be replaced by *both*.

27. *wherewithboth

28. *He gave it his both.

Nor can they be replaced with *all*+numeral.

29. *wherewithalthree

30. *He gave it his all three.

In addition, they cannot be replaced with *whole* or *half* or *almost all*.

31. *wherewithwhole; *wherewithhalf; *wherewithalmostal

32. *He gave it his whole/*his half/*his almost all.

Possible, on the other hand, is (33).

33. He gave it his whole effort.

But (34) is not possible, due to the fact mentioned in the discussion of (20), namely that *wherewith* is restricted to occurring in combination with bare *all* (or *all* + silent N) as relative head.

34. *wherewithwholeeffort

The restriction seen in (27) is also found in West Ulster English, in the sense that the West Ulster English examples (9) and (10), repeated In (35) and (36), have no counterparts with *both*, as seen in (37) and (38).[14]

35. Who all did you give tea to?

[14] I am grateful to Jim McCloskey (p.c.) for the judgments.

36. Where all did you move the books to?

37. *Who both did you give tea to?

38. *Where both did you move the books to?

This gives further support to the idea that the *all* of *wherewithal* is related to the *all* studied by McCloskey (2000).

We can now see clearly, I think, that the non-existence of (27) is not an accidental isolated fact about the English lexicon, but is tied to differential syntactic properties of *all* vs. *both* seen in (24) vs. (27) and in (35) - (38). This linkage is expressible, though, only if words like *wherewithal* (and by plausible extension all words containing more than one morpheme) are put together in the syntax (via Merge, both external and internal).

There remains, needless to say, the question why *all* and *both* diverge in their behaviour is these specific ways. The core of an answer may be as follows. *Both* is more complex than *all*, just as *all three* is more complex than *all*.[15] *Both* is comparable to *all two*, perhaps with a silent ALL. The appearance of *all* in (24) and in (35)/(36), and therefore in *wherewithal*, depends on its (relative) lack of complexity (i.e. on its not being accompanied by a numeral). Consequently, there can be no examples of words like (27) (and perhaps similarly for (31), at least for *almost all* and for *half*).

5. Romance

In addition to accounting for the absence of certain logically possible words in English, the perspective outlined above can account for the absence of *wherewithal* itself in any Romance language (as far as I know), in terms of the absence of *wherewith* (along with *whereby, whereof, wherefore*) in any Romance language (as far as I know).[16]

6. Other impossible words

The syntax of *(the) wherewithal* that I have proposed takes its *where* component to be a relative pronoun. If a relative clause structure is the only way to reach that sort of complex word, then we may have an account of the following, since *here* and *there* are not possible in English as relative pronouns.

---

[15] The *all* in question may have something in common with the *ever* of free relatives.

[16] As for why these are absent from Romance languages, see note 4 on the VO vs. OV question, keeping in mind that English is more 'OV' than any Romance language when it comes to compounds of the *magazine reader* or *magazine reading* sort - cf. Emonds and Faarlund (2014: 20) - though the details of this linkage remain to be worked out. Some 'OV'-ness may also be a necessary condition for 'V DP Prt' order of the *pick the book up* sort, which seems to be absent from all Romance languages (Andrea Padovan, p.c.), as opposed to 'V Prt DP', which is found in some Romance languages. More complex than the *where-X* cases are those with *there*; it may be, though, that French *là-dedans* ('there of in') is closer to *in there* than to *therein* - cf. McCawley (1988: note 12) and Rizzi (1988).

39. *They don't have the herewithal to buy that big a house

40. *They don't have the therewithal to buy that big a house.

Somewhat similarly, if *wherewith* in *wherewithal* is of the *whereby* type found more robustly in Dutch and German, then we can exclude (41) and (42) in terms of the general exclusion within Germanic of forms like (43),[17] with non-R-wh-words.

41. *They don't have the whatwithal to buy that big a house.

42. *They don't have the whenwithal to buy that big a house.

43. *whatby, *whenby

7. Challenges

In my English, *wherewithal* has no plural.

44. *We don't have the wherewithals to buy that big a house.

From the present perspective this may be related to (45).

45. *We gave it our alls.

To judge by a Google search, however, there seem to be speakers for whom sentences like (44) are possible. It may be that the silent N arguably present with *wherewithal* can be accompanied by plural *-s* for some speakers, but not for others, though it remains to be understood why I allow plural *-s* with a silent noun in (46), while disallowing it in (44).

46. the others; two four-year-olds; the extra-wides

A second challenge comes from the fact that some speakers accept (47).

47. We don't have the wherewithal with which to do it.

Such speakers must be allowing adposition doubling, recalling the existence, in some English, of cases like (48).

48. the problem to which they're alluding to

The speakers in question must also be allowing, from the present perspective, two wh-phrases (*where, which*) both of which have *all N* as antecedent, thereby having something in common with (49),[18] and perhaps also with German (50).[19]

---

[17] In all likelihood, examples like (i), which are acceptable to Bob Frank, are of a different character.

i. What about were you guys talking?

[18] Cf. Kayne (1983; to appear).

49. Prof. Mary Smith, whose students' admiration for whom is well-known, ...

50. Ich habe Blumen gebracht. Was für welche?
I have flowers brought. what for which
'I have brought flowers. What kind?'

8. Conclusion

In conclusion, then, sentences containing *wherewithal* have a richer syntax than might at first glance be apparent. This richer syntax provides a handle on the absence of various potential words such as **whatwithal, *whenwithal, *herewithal, *therewithal, *wherewithwhole, *wherewithhalf, *wherewithalmostal, *wherewithboth,* that an approach associated with a less rich syntax could not provide.

References

Baker, Mark. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16. 373–415.

Baker, Mark. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: The University of Chicago Press.

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.

Chomsky, Noam. 1970. Remarks on nominalization. In Jacobs, Roderick A. & Rosenbaum, Peter S. (eds.) *Readings in English transformational grammar*, 184–221. Waltham: Ginn.

Cinque, Guglielmo. 1982. On the theory of relative clauses and markedness. *The Linguistic Review* 1. 247–294.

Cinque, Guglielmo. 2016. A microparametric approach to the head-initial/head-final parameter. (To appear in Karimi, Simin & Piattelli Palmarini, Massimo (eds.) *Parameters: what are they? Where are they? Linguistic Analysis*, vol. 42)

Citko, Barbara. 2004. On headed, headless, and light-headed relatives. *Natural Language and Linguistic Theory* 22. 95–126.

Collins, Chris. 2014. Relative clause deletion. New York University. (Unpublished ms.)

Collins, Chris. 2006. A note on derivational morphology. New York University. (Unpublished ms.)

Durrell, Martin. 2002. *Hammer's German grammar and usage*. 4th edn. Chicago: McGraw-Hill.

Emonds, Joseph Embley & Faarlund, Jan Terje. 2014. *English: The language of the Vikings*. Olomouc: Palacký University.

Groos, Anneke & van Riemsdijk, Henk. 1981. Matching effects in free relatives: A parameter of core grammar. In Belletti, Adriana & Brandi, Luciana & Rizzi, Luigi (eds.) *Theory of markedness in generative grammar: Proceedings of the 1979 GLOW Conference*, 171–216. Pisa: Scuola Normale Superiore.

Julien, Marit. 2002. *Syntactic heads and word formation*. Oxford: Oxford University Press.

[19] Durrell (2002: 93).

Kayne, Richard S. 1976. French relative 'que'. In Hensey, Fritz & Luján, Marta (eds.) *Current studies in Romance linguistics*, 255–299. Washington D.C.: Georgetown University Press.

Kayne, Richard S. 1983. Connectedness. *Linguistic Inquiry* 14. 223–249. (Reprinted in Kayne 1984.)

Kayne, Richard S. 1984. *Connectedness and binary branching*. Dordrecht: Foris.

Kayne, Richard S. 1994. *The antisymmetry of syntax.* Cambridge, MA: The MIT Press.

Kayne, Richard S. 2004. Here and there. In Leclère, Christian & Laporte, Éric & Piot, Mireille & Silberztein, Max (eds.) *Syntax, lexis & lexicon-grammar: Papers in honour of Maurice Gross*, 253–273. Amsterdam: John Benjamins. (Reprinted in Kayne 2005.)

Kayne, Richard S. 2005. *Movement and silence*. Oxford: Oxford University Press.

Kayne, Richard S. 2008. Antisymmetry and the lexicon. *Linguistic Variation Yearbook* 8. 1–31. (Also in di Sciullo, Anne-Marie & Boeckx, Cedric (eds.) 2011. *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*, 329–353. Oxford: Oxford University Press.) (Reprinted in Kayne 2010.)

Kayne, Richard S. 2010. *Comparisons and contrasts*. Oxford: Oxford University Press.

Kayne, Richard S. 2014. *Once* and *twice*. In Contemori, Carla & Dal Pozzo, Lena (eds.) *Inquiries into linguistic theory and language acquisition: Papers offered to Adriana Belletti*. Siena: CISCL Press (Centro Interdipartimentale di Studi Cognitivi sul Linguaggio). (Also in *Studies in Chinese Linguistics* 36. 1–20.)

Kayne, Richard S. To appear. A note on some even more unusual relative clauses. In Bailey, Laura & Sheehan, Michelle (eds.) *Order and structure in syntax*. Language Science Press.

Koopman, Hilda & Szabolcsi, Anna. 2000. *Verbal complexes*. Cambridge MA: The MIT Press.

McCawley, James D. 1998. Adverbial NPs: Bare or clad in see-through garb? *Language* 64. 583–590.

McCloskey, James. 2000. Quantifier float and *wh*-Movement in an Irish English. *Linguistic Inquiry* 31. 57–84.

Myers, Scott P. 1987. *Tone and the structure of words in Shona*. Amherst: University of Massachusetts. (Doctoral dissertation)

Obenauer, Hans-Georg. 1994. *Aspects de la syntaxe A-barre: Effets d'intervention et mouvements des quantifieurs*. Paris: Université de Paris VIII. (Thèse de doctorat d'Etat)

Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20. 365–424.

Rizzi, Luigi. 1988. Il sintagma preposizionale. In Renzi, Lorenzo (ed.) *Grande grammatica italiana di consultazione Volume 1: La frase: I sintagmi nominale e preposizionale*, 507–531. Bologna: Il Mulino.

Roberts, Ian. 2017. Verb movement and cartography in English and Romance (and elsewhere). (Paper presented at Yale University)

van Riemsdijk, Henk. 1978. *A case study in syntactic markedness*. Lisse, Netherlands: The Peter de Ridder Press.

Richard Kayne
New York University
richard.kayne@nyu.edu

A note on the phonology and phonetics of CR, RC, and SC consonant clusters in Italian

Michael J. Kenstowicz

1. Introduction

Previous generative research on Italian phonology starting with Vogel (1982) and Chierchia (1986) has proposed that intervocalic consonant clusters are parsed into contrasting tauto- vs. heterosyllabic categories based on several factors: phonotactic restrictions on word-initial consonant sequences, syllable weight as reflected in the distribution of stress and the length of a preceding tonic vowel, the distribution of prenominal allomorphs of various determiners, and the application of syntactic gemination (*radoppiamento sintattico*). Based on these criteria, clusters of rising sonority (in particular stop plus liquid) fall into the tautosyllabic category while falling sonority clusters composed of a sonorant plus obstruent are heterosyllabic. Clusters composed of /s/ plus a stop display mixed behaviour but generally pattern with the heterosyllabic group. In her 2004 UCLA Ph.D. dissertation, Kristie McCrary investigated corpus-external reflexes of these cluster distinctions with a psycholinguistic test of word division and measurements of the phonetic duration of segments (both consonants and vowels). Her results support some aspects of the traditional phonological analysis but call into question others. In this squib we summarize the literature supporting the traditional distinction among these clusters and then review McCrary's results. An important finding in McCrary's study was that stops in VCV and VCRV contexts (R = a liquid) were significantly shorter than stops in VRCV contexts. She observed that these contexts align with the distribution of geminates in Italian and proposed that singleton stops are significantly shorter in the VCV and VCRV contexts in order to enhance their paradigmatic contrast with geminates. We review this finding and then explore two consequences. First, we look at the distribution of the lenition found in the Tuscan dialects known as the *gorgia toscana* and see that it targets the shorter stops. Second, we introduce data from a pilot study of Brazilian Portuguese, which has phonotactic restrictions similar to Italian but crucially lacks the systematic contrast of singleton vs. geminates. The prediction is that the duration of the stops in these contexts should not differ systematically in the way reported for Tuscan. This prediction is confirmed.

2. Background

A stressed penultimate syllable in standard Italian is canonically analysed as bimoraic (Krämer 2009 and references). In the OT framework (Prince & Smolensky 2004), this generalization reflects the activity of the Stress to Weight constraint ("If stressed then heavy"). In an open syllable, the tonic vowel is lengthened to make the syllable heavy

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 87–100

and has been shown by a number of researchers to have significantly greater duration compared to the stressed vowel of a syllable closed by the first half of a geminate consonant. For example, in a study of seven Italian speakers D'Imperio & Rosenthall (1999) found a mean duration of 177 ms. for the stressed vowel of [fa.te] vs. 126 ms. for [fat.te]. On the other hand, word-final stressed vowels are barred from lengthening by a markedness constraint against final long vowels: *V:#. Following up on previous observations by Vogel (1982) and earlier literature, Chierchia (1986) showed that the bimoraic constraint for stressed syllables is satisfied in an alternative fashion for word-final vowels by geminating the initial consonant of the following word (*radoppiamento sintattico*): *città [pp]ulita* 'clean city' and *città [ss]anta* 'holy city'.

An interesting question that these analyses raise is just which consonant clusters count as closing the penultimate syllable and thereby bleed the vowel lengthening process as well as preventing stress from receding to the antepenultimate syllable by the Latin Stress Rule (see Section 3). Cross-linguistically the syllabic parsing of intervocalic consonant clusters is often guided by what are possible word-initial and word-final consonants and consonant clusters (Pulgram 1970; Steriade 1999). This strategy is based on the assumption that the word is exhaustively parsed into syllables, entailing that a word-initial consonant (cluster) occupies the syllable onset and a word-final consonant (cluster) occupies the syllable coda. As in most languages, a single intervocalic consonant parses as an onset with the following vowel in Italian. This generalization follows automatically in an OT grammar, given the constraints of Onset and No-Coda and the absence of any higher-ranking constraint to counteract a V.CV parse. The inventory of permissible word-initial (and hence syllable onset) clusters is largely controlled by the Sonority Sequencing Generalization discovered by 19th century scholars such as Sievers (1876) and resuscitated by Steriade (1982) and Clements (1990) and invoked by many other scholars since. According to this generalization, onsets rise in sonority and codas fall in sonority. By this criterion, a stop plus sonorant (liquid or glide) cluster is a valid onset while a sonorant plus stop cluster is not and hence the latter will parse with a syllable break in word-medial position. The one situation where the Sonority Sequencing Generalization breaks down in Italian is with SC clusters, which are possible word-initially (*spago* 'string', *stato* 'state', *scala* 'stairs') but are claimed to block tonic vowel lengthening, as seen in the data of (1) taken from Morelli (1999: 166–173).

1.	CV	fá:.to	'fate'	CC	fát.to	'fact'
		pé:.lo	'hair'		mán.to	'coat'
	CR	ká:.pra	'goat'	SC	vés.pa	'wasp'
		ré:.tro	'behind'		pás.ta	'pasta'
		sá:.kro	'sacred'		mós.ka	'fly'

An SC cluster also stands out in selecting vowel-final allomorphs for various prenominal determiners, as in the paradigm in (2) from Davis (1990).

2.	ponte	proposito	specchio
	il ponte	il proposito	lo specchio
	un ponte	un proposito	uno specchio
	quel ponte	quel proposito	quello specchio
	nessun ponte	nessun proposito	nessuno specchio
	'bridge'	'purpose'	'mirror'

Lastly, word-initial SC clusters differ from TR clusters in failing to geminate in the *radoppiamento sintattico* context (Chierchia 1986): *città [s]porca*, **città [ss]porca* 'dirty city' vs. *città [tt]riste* 'sad city'. This property has been explained by postulating that the [s] in SC clusters does not have the same phonological status as the stop of CR clusters and is licensed by a separate stipulation or at a higher level of prosodic structure as a syllable appendix. As a syllabically 'stray' consonant (Steriade 1982), the [s] of *sporca* can freely associate to the coda of the preceding stressed syllable to satisfy the bimoraic Stress to Weight constraint. But the onset consonant of *triste* requires the addition of an autosegmental association to the coda of the stressed syllable to create a geminate and hence close the tonic syllable.

3. Consonantal interludes and stress

Italian inherited the Latin Stress Rule according to which the accent fell on the penultimate syllable unless it was light, in which case stress receded to the preceding antepenultimate syllable if there was one. In the development of Italian from Proto-Romance, long vowels shortened when outside of a penultimate open syllable. This results in a state of affairs in which there is a tradeoff between phonemicising stress and predicting length with a rule CV́. > CV́: or phonemicising length and predicting stress with the rule CV: > CV́:. The latter interpretation was proposed by Saltarelli (1970) while the former is the view adopted by most of the later literature (e.g. Krämer 2009). From the majority perspective, stress is primarily penultimate in Italian. It can – but need not – recede to the antepenult if the penult is an underlying light syllable. A number of factors bias stress to the paroxytone or proparoxytone categories and these factors have been shown to play a role in novel word experiments (Krämer 2009, Burani et al. 2014). How do the three cluster types (CR, RC, and SC) behave with respect to the distribution of stress when they form the hinge between the penultimate and final vowels? The data of Table 1 below show our counts of the type frequencies for the locus of stress in trisyllabic verbs and nouns, respectively, taken from the stress-marked corpora of Delmonte (1999) and Thornton et al. (1997).

Table 1: Accent placement from two stress-marked corpora

Delmonte	verbs		Thornton	nouns	
	penult	antepenult		penult	antepenult
VCV	307	252		557	289
VCRV	7	4		8	10
VRCV	59	0		75	1
VSCV	30	0		57	0

The VCV structures counted in this table exclude geminates as well as palatal consonants such as [ʎ] (orthographic *gl*) and [ɲ] (*gn*) that pattern with geminates. Even when these factors are controlled for, there is a bias to penultimate stress in VCV structures. The rising sonority CR clusters are compatible with both antepenult and penult stress, while RC clusters uniformly make the syllable heavy and hence force the word into the paroxytone class. CR clusters are much smaller in number and so the learner cannot make a reliable choice here just based on the statistics of the lexicon. See Appendix 1 for a list of these items. On the other hand, SC clusters uniformly make the syllable heavy and so there is consistency between stress and *radoppiamento sintattico* with respect to this cluster type. These generalizations are utilized by speakers in novel word experiments. For example, Krämer (2009: 185) reports that his subjects uniformly assigned penultimate stress to the nonce words *tapirco* and *grotulfo* while stimuli with all light syllables such as *frunaco*, *frudalo*, and *nalico* were more variable but displayed an overall bias to penultimate accent.

Table 2 below summarizes the various phonological factors that bear on the tauto-syllabic versus hetero-syllabic parses of the CR, RC, and SC clusters. The only inconsistency is in the unexpected licensing of the SC clusters at the beginning of the word.

Table 2: Phonological factors reflecting syllable parse

	V.CRV	VR.CV	VS.CV
SSG	✓	✓	✓
Initial cluster	✓	✓	*
Antepenult stress	✓	✓	✓
Allomorphy	✓		✓
Radoppiamento sintattico	✓		✓
Tonic length	✓	✓	✓

4. Phonetic duration reflexes

In reviews of the available data on the duration of tonic vowels in paroxytone structures, Vogel (1982) and later McCrary (2004) report that the most secure finding is a complementarity between the durations of the consonantal interlude and the tonic vowel. A geminate consonant is approximately twice as long as a singleton consonant and the tonic vowel is invariably shorter before the geminate than before the singleton, although the ratio varies quite bit. In addition, the tonic vowel is significantly longer than the following consonant in V.CV structures. These findings are summarized in Table 3 below.

Table 3: Average duration in milliseconds of tonic vowel and consonant interludes from Josselyn (1900), Parmenter & Carmen (1932), Fava & Magno Caldognetto (1976), and Vogel (1982)

	Josselyn	P & C	F & MC	Vogel
V.CV	260	200	207	130
V.***C***V	140	110		84
VC_i.C_iV	170	150	107	105
V***C***$_i$.***C***$_i$V	240	250		162

5. McCrary (2004)

McCrary utilized two different experimental methods to investigate behavioural and phonetic reflexes of the phonological contrasts in syllabification among the various types of post-tonic consonantal interludes. The first was a syllable break task in which 50 subjects were asked to divide an orally presented nonsense word into two parts. The word "syllable" was not used in the instructions and the subjects were trained on CVCV stimuli such as *Roma* and *Bari*, all of which were divided CV.CV. In the test trials, words with consonant clusters were introduced. The major findings are as follows. First, CL clusters (L = liquid) were parsed as tautosyllabic by 88% of the subjects versus only 8% for NC clusters and 5% for LC clusters—a significant difference both in terms of statistical reliability as well as magnitude. SC clusters showed much more variability and were not reliably distinguished from the baseline. These results are consistent with the phonological evidence distinguishing the CR and RC clusters (R = liquid) as tautosyllabic versus heterosyllabic. The subjects' uncertainty on how to divide SC clusters might reflect a conflict between the fact that SC appears as a well-formed word-initial cluster, suggesting a tautosyllabic parse, vis à vis its uniformly counting for weight with respect to the distribution of stress, motivating a heterosyllabic parse.

Another of McCrary's findings concerns the duration of the tonic vowel and the consonants occupying the interlude between the penultimate and ultimate vowels. The phonological evidence leads to the expectation that CR clusters should be associated with a longer tonic vowel than RC clusters, while SC clusters should occupy an intermediate position. This hypothesis is based on the assumption that syllable structure will be a primary determinant of phonetic duration — a legitimate expectation if the moraic structure of the syllable plays a major role in determining its phonetic duration, as suggested by the findings of Broselow et al. (1997). Of course, other factors such as the inherent durational properties of the individual segments may also be relevant.

To investigate this question, McCrary (2004) constructed the corpus of disyllabic nonsense words seen in Table 4 that held the vowel constant as [a] and systematically varied the consonantal interlude among the clusters of interest. Fifteen subjects residing in Pisa produced six repetitions of these items in the frame "non trovo la parola X nel dizionario".

Table 4: Nonsense words for tonic vowel duration from McCrary (2004)

Tauto C_1C_2 and CVCV stimuli		Hetero C_1C_2 stimuli	
CV́CRV	pápra, pátra, pákra	CV́RCV	párpa, párta, parka
CV́CLV	pápla, pátla, pákla	CV́LCV	pálpa, pálta, pálka
CV́CNV	pápna, pátna, pákna	CV́NCV	pánpa, pánta, pánka
CV́CV	pápa, páta, páka	CV́SCV	páspa, pásta, páska
	pása, pána, pála, pára	CV́CSV	pápsa, páksa
		CV́CTV	pápta

In calculating the results, the first and last repetitions were discarded, leaving 1,851 tokens. Segmentation was based on the appearance/disappearance of high amplitude in the second formant of the vowel. By this criterion, the VOT associated with the voiceless stops was counted as part of the consonant. The corpus contains a CV.CV baseline for the expected longest duration but no CVCCV with a geminate as baseline for the expected shortest duration. The main findings of this part of McCrary's study are summarized as follows (see Table 5). First, while interludes beginning with a stop consonant (*pápsa*, *páksa*, *pápta*) were associated with the shortest tonic vowel duration as expected, there was essentially no significant difference in the duration of the tonic vowel between CV.CRV and CVR.CV structures (R = liquid). Second, the tonic vowel in CVS.CV structures was significantly shorter than in CV.CV structures but longer than in interludes starting with a stop, a finding that would be consistent with the conflicting evidence regarding the syllabic affiliation of SC clusters mentioned earlier. Third, the duration of the tonic vowel was significantly correlated with the duration of the entire consonantal interlude regardless of whether it was phonologically tautosyllabic or heterosyllabic. Indeed, the duration of the interlude proved more predictive of the duration of the tonic vowel than the tautosyllabic vs. heterosyllabic distinction. Finally, the particular consonants composing the interlude had a significant effect on the duration of the tonic vowel. The tap [r] was associated with a longer tonic vowel than the lateral [l], which in turn was associated with a longer tonic vowel than the fricative [s]. This point held true regardless of whether the consonant occupied either the first or second position in the interlude. In sum, it seems that the differences in the duration of the tonic vowel depended more on the particular consonants occupying the interval independent of their position in the onset or coda of the phonological syllable. At least this was true for the two liquids. Nasals at the beginning of the cluster were associated with a longer vowel compared to when they terminated the cluster.

Table 5: Duration of tonic vowel in various interlude structures

Test Category	Mean Stressed Vowel Duration (ms.)	Std. Dev.	Claimed Syllable Type
CV́.CV	186	34	Open
CV́.CRV	169	28	Open
CV́R.CV	167	28	Closed
CV́.CLV	158	27	Open
CV́N.CV	158	25	Closed
CV́L.VC	156	27	Closed
CV́S.CV	145	24	Closed
CV́.CNV	135	24	Open
CV́.C.SV	127	23	Closed
CV́C.TV	125	21	Closed

When viewed from the perspective of Broselow et al. (1997), one might propose that in the case of the SC clusters the coda [s] shares the second mora of the tonic syllable with the nuclear vowel. Taking the CV.CV structure as a baseline, the duration of 145 ms for SC is approximately 3/4ths of 186 ms: 186/2 = 93 + 47 = 140. However, it is far from obvious how the behaviour of the interludes containing a liquid consonant could be explained by judicious allocations of the moraic structure of the tonic syllable.

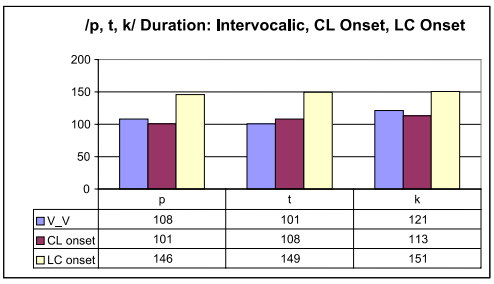
A second aspect of McCrary's study examined the duration of stop consonants in three positions: VCV, VCRV, and VRCV (R = liquid). Once again, nonsense words were utilized (Table 6). In this table, McCrary used L to stand for either liquid consonant.

Table 6: Nonsense words for measurements of stop duration (McCrary 2004:236)

	Onset – V_V	Onset - CV**C**LV	Onset - CVL**C**V
/p/	*pápa*	*pápra, pápla*	*párpa, pálpa*
/t/	*páta*	*pátra, pátla*	*párta, pálta*
/k/	*páka*	*pákra, pákla*	*párka, pálka*

Other things being equal, one might expect the stop to be shortest in the CVCLV contexts since it must share the syllable's onset with the following liquid under the CV.CLV parse. The surprising finding was that the stop in the simple VCV context was also significantly shorter in comparison to the stop in the VLCV structure, as shown in the chart of Table 7, where L = liquid.

Table 7: Duration of stops in various interludes (McCrary 2004:238)



McCrary insightfully connects this surprising distribution with another one in Italian phonology—namely, the positions where a geminate consonant is permitted. While a singleton-geminate contrast is possible in the VCV (*fato* vs. *fatto*) and VCRV (*te.atro* vs. *quattro)* contexts (R = liquid), no contrast is possible in VRCV contexts (*quarto* but **quartto*). In traditional terms, the geminate must span the syllable boundary and the overriding ban on complex codas blocks a VRC.CV parse. These requirements are active in the adaptation of word-final stops in loanwords, which are normally geminated and accompanied by an epenthetic vowel (*sud* > *sudde*); but gemination is blocked in *nord* > *norde* (Marotta 2008). McCrary sees the shortening of the stop in the VCRV (R = liquid) and especially the VCV structures of Table 7 as a manoeuvre by the phonetics to enhance the paradigmatic contrast between singleton and geminate consonants. Duration is of course the major cue to this contrast and so curtailing the duration of the singleton is an effective strategy to sharpen the difference between the two consonantal categories. Recent research allows us to connect this distribution with another asymmetry in Italian phonology—the *gorgia toscana*.

6. Gorgia Toscana

The Tuscan dialects of Italian exhibit a lenition process known traditionally as the *gorgia* in which post-vocalic stops are realized as fricatives. The dialects differ in terms of which consonants are more likely to undergo the process. One secure generalization is that the process does not affect geminates. Marotta (2008) views the *gorgia* as a weakening process that reduces the articulatory pressure of the consonantal constriction gesture as well as the segment's duration. The canonical locations for the process are either simple or complex onsets (i.e. VCV and VCRV contexts). It does not apply after a consonant or word-initially after pause. The discussion of McCrary's study at the end of Section 5 suggests that the *gorgia* targets the stops with curtailed duration that stand in opposition to the geminates and may reflect an incomplete closure of the stop's constriction under increased time pressure. McCrary does not indicate whether or not her Pisan subjects spirantised any of their stops in her experimental results. According to Marotta (2008), the gorgia is primarily found with the velar stops in Pisa while essentially all stops are lenited in the Florentine variety. From a phonological perspective, the *gorgia* enhances the geminate-singleton opposition with a difference in manner of articulation—a common cross-linguistic phenomenon found in languages such as Berber (Saib 1974). A recent study by Ulfsbjorninn (2016) on the phonological behaviour of different consonant clusters from the Government Phonology perspective reports data that corroborate our hypothesis that the *gorgia* targets the polarized short consonants. In particular, he finds that

consonants which occupy the coda of the syllable when followed by an onset consonant (his "bogus" clusters) fail to lenite even though they are postvocalic. As Ulfsbjorninn observes, the failure of lenition in these contexts cannot be attributed to their status as loanwords since the *gorgia* is readily extended to borrowings and even applies in the same word containing a bogus cluster when its conditions are met, as in [at.lɛθɛ] 'athlete'. The data in (3) illustrate these points.

3. Lenition of postvocalic stops (Marotta 2008; Ulfsbjorninn 2016)

VCV	abete	[aβeθe]	'fir'	lago	[laɣo]	'lake'
CRV	capra	[kaɸra]	'goat'	magro	[maɣro]	'thin'
	la preghiera	[laɸreɣjɛ́:ra]	'the prayer'	i crackers	[ixré:xe(r)]	'the crackers'
VRC	porta	[pɔrta]	'door'	largo	[largo]	'large'
VCC	atlete	[at.lɛθɛ]	'athlete'	ɛtna	[ɛt.na]	'Mt. Etna'

We don't have duration measures from McCrary for stops in coda position. But if the lenitions really are motivated by the reduced duration that enhances the singleton-geminate contrast, then the stops in the coda of words like *atlete* and *Etna* should be longer than the stops in the simple and complex onsets of words like *abete* and *capra*. In any case, even in the absence of such phonetic evidence, we can say that the *gorgia* targets precisely those stops that McCrary found to be polarized in duration with respect to the geminates: the intervocalic singleton VCV and the rising sonority VCRV.

7. Portuguese

The preceding discussion suggests that if a language lacks the singleton-geminate contrast then the motivation for the duration differences between stops in the VCV and VCRV vs. VRCV contexts found by McCrary for Italian will be absent. Moreover, the large differences in tonic vowel duration between VCV and VCCV structures should also be absent if the relatively short vowel duration before geminates is a cue to the geminate category. Finally, the relatively large differences between the duration of the tonic vowel and the following consonant in the simple CVCV structures seen in Table 1 should not be found if the vowel duration cues the singleton-geminate contrast in the following consonant.

In order to pursue these points, we recorded a corpus of 62 disyllabic words (see Appendix 2) from two female Brazilian Portuguese (BP) speakers from San Paolo (thanks to Suzana Fong and Karin Vivanco for sharing their language). BP lacks systematic geminates. Moreover, in its historical development from Proto-Romance, SC clusters underwent initial epenthesis in BP, as in Spanish. Thus, unlike in Italian, there is no conflict between the phonotactic distribution of these clusters and the fact that they are associated with penultimate accent. In other respects the two languages are similar in virtue of having inherited the same basic stock of vocabulary and the Latin Stress Rule (at least in nominals). The words in our corpus varied the consonantal interludes among the five types of interest (VCV, VCRV, VRCV, VSCV, and VCCV) and distributed the voiceless stops [p, t, k] evenly across the first three structures. The stressed vowel was predominantly [a]; in some cases a different vowel had to be used in order to complete the paradigm. Unlike in McCrary's study, we used existing words

(all nouns) rather than nonsense words. The speakers were recorded in a sound-proof booth reading a randomized version of the list of the target words in both an isolation form as well as in the frame "*Eu não tô encontrando a palaura X no dicionário* "I can't find the word X in the dictionary". The data were analysed in Praat (Boersma and Weenink 1992–2017) with text grids delimiting the word and the individual phonemes. Segmentation followed the procedure used by McCrary (2004) by including VOT as part of the stop consonant. The duration measures were normalized with respect to the duration of the entire word in order to accommodate any differences in speech rate between the two speakers. To analyse the results, a series of regression analyses were run with word, speaker, and trial (repetition) as random factors and the five consonantal interludes as predictor factors. Here we report only the duration measures for the words as spoken in isolation. As far as the duration of the tonic vowel is concerned, the VCV structure served as the baseline. As seen in Table 8, it was significantly longer than in each of the other structures (using t = 2.0 as the baseline level of significance). However, multiple comparisons (Tukey) among the other clusters found no significant differences except for SC vs. RC (p = 0.03).

Table 8: Tonic vowel duration (normalized milliseconds)

	mean	Std error	t
VCV	322	15	
VSCV	294	13	-2.0
VCRV	274	14	-3.4
VCCV	262	6	-3.5
VRCV	252	13	-5.1

The second regression analysed the duration of the stop consonants in the various interludes. Once again VCV was the baseline. As seen in Table 9, the duration of the stops in the clusters were all significantly shorter than the baseline. Multiple comparisons (Tukey) found that the SC context was significantly shorter than CR and RC contexts but there was no significant difference between the latter two clusters themselves. Thus, just as in McCrary's findings with respect to Italian, the duration of the stop seems to be more affected by the consonant it is paired with in the interval (in our case the rhotic) rather than its position as onset or coda in the syllable.

Table 9: Stop duration (normalized milliseconds)

	mean	Std error	t
VCV	315	15	
VSCV	178	14	-9.5
VCRV	268	15	-3.2
VCCV	198	18	-3.5
VRCV	260	14	-6.4

The final regression compared the duration of the tonic vowel to the duration of the following consonant in the VCV structures. Recall from Table 1 that in Italian the consonant was noticeably shorter in this context compared to the tonic vowel. This was attributed to two factors: minimization of the duration of the consonant to enhance its paradigmatic contrast with a geminate and maximization of the duration of

the stressed vowel in order to implement a bimoraic structure. Since BP lacks geminates, there is no particular motivation for the single intervocalic consonant to be significantly shorter than the tonic vowel. And in fact this is what our data reveal. The average duration of the tonic vowel was 323 ms and that of the consonant was 326 ms.

Table 10: Singleton consonant and tonic vowel duration

	mean	Std error	t
V***C***V	325	10	
VCV	323	4	-0.66

There was however a significant difference when the comparison was restricted to the stop consonants.

Table 11: Singleton stop consonant and tonic vowel duration

	mean	Std error	t
V***C***V	243	10	
VCV	288	3	14.00

However, the ratio between the duration of the tonic vowel and the intervocalic stop in BP is much smaller compared to what is reported in the earlier studies of Italian mentioned in Table 1 above and summarized in Table 12 below.

Table 12: Mean duration and ratio of tonic vowel to following stop

	Josselyn (1900)	P&C (1932)	Vogel (1982)	BP
V.CV	260	200	130	288
V.***C***V	140	110	84	243
ratio	1.85	1.81	1.54	1.18

We thus tentatively conclude that there is a linguistically significant difference between Italian and Brazilian Portuguese regarding the duration of an intervocalic stop in CVCV structures that can be explained by the hypothesis that the consonant is much shorter than otherwise expected in Italian in order to enhance its paradigmatic contrast with a geminate.

8. Summary and Conclusion

In this squib we reviewed the various phonological factors that converge on a tautosyllabic parse for intervocalic CR clusters versus a heterosyllabic parse for RC clusters in Standard Italian. For the most part, these factors also motivate a heterosyllabic structure for SC clusters. We then reviewed some the behavioural and phonetic experiments of McCrary (2004) which sought corpus-external confirmation for these structural differences. The results of a syllable-parsing task were largely consistent with the phonological distinctions while the duration of the tonic vowel failed to differentiate the tautosyllabic CR from the heterosyllabic RC interludes. We concluded with McCrary that the duration requirements of the individual consonants masked any differences that should have emerged from the putative syllabic

structures. We then reviewed McCrary's other surprising finding from her study of Italian—namely that the duration of the stop in VCV interludes was significantly shorter than in VRCV interludes and comparable to its duration in onset-sharing VCRV structures. McCrary insightfully connected this finding to another asymmetry in Italian phonology—the distribution of geminate consonants. Geminates are permitted in intervocalic VCCV and rising sonority VCCRV clusters but are banned from falling sonority VRCCV clusters. McCrary views the shortening as an enhancement strategy to polarize the singleton's paradigmatic contrast with geminates. We followed up on this point with the observation that the *gorgia toscana* lenition process targets precisely the stops which stand in paradigmatic opposition to the geminates. Finally, we called attention to Portuguese, a language largely comparable to Italian save for the lack of a systematic singleton-geminate contrast. A small pilot study of two Brazilian Portuguese speakers did not find the large differences in the duration ratio of the tonic vowel to the following singleton consonant reported in earlier studies of Italian. This failure would be consistent with the absence of geminate consonants in this Romance variety, which in turn would remove any motivation for dramatically shortening a single intervocalic consonant and magnifying the duration of the preceding tonic vowel. A critical task for future research is to replicate McCrary's study with actual instead of nonsense words. Also, other dialects such as the Roman variety should be studied to establish the independence of the Gorgia lenition and the paradigmatic shortening of the intervocalic singletons.

Appendix 1 – Italian words with CR clusters

verbs		nouns	
proparoxytone	paroxytone	proparoxytone	paroxytone
calibro	massacro	palpebra	ottobre
integro	allegro	tenebra	massacro
arbitro	riciclo	vertebra	mezzadra
penetro	consacro	calibro	puledra
cronometro		cattedra	mezzadro
centimetro		anatra	puledro
millimetro		arbitro	psichiatra
		scheletro	aratro
			teatro
			manovra

Appendix 2 – Brazilian Portuguese disyllabic words

vaca	‘cow’	opta	‘opts’	certo	‘right’
lassa	‘weary’ (f.)	pacto	‘pact’	porca	‘dirty’ (f.)
pato	‘duck’	rapto	‘kidnapping’	carta	‘paper’
coca	‘coca’	táxi	‘taxi’	marco	‘landmark’
passo	‘step’	lapso	‘lapse’	barca	‘ship’
data	‘data’	acre	‘acrid’	parto	‘childbirth’
passa	‘raisin’	potro	‘colt’	surto	‘outburst’
caco	‘shard’	litro	‘litre’	arco	‘arch’
gato	‘cat’	cetro	‘sceptre’	basta	‘basta’
bata	‘blouse’	metro	‘metre’	casca	‘crust’
rapa	‘scrap’	ocre	‘ochre’	tasco	‘bit’
sopa	‘soup’	macro	‘macro’	casta	‘caste’
saco	‘bag’	sopro	‘puff’	casto	‘chaste’
mapa	‘map’	atriu	‘atrium’	vasto	‘vast’
classe	‘class’	lucro	‘profit’	raspa	‘shavings’
mata	‘forest’	sacro	‘sacred’	asco	‘disgust’
capa	‘cover’	parca	‘scanty’ (f.)	rasto	‘track’
paca	‘bloody’	corpo	‘body’	lasca	‘silver’
tapa	‘slap’	sorte	‘luck’	basca	‘Basque’ (f.)
massa	‘mass’	parco	‘scanty’	pasto	‘pasture’
				casco	‘shell’

References

Boersma, Paul & Weenink, David. 1992–2017. Praat. www.praat.org.

Broselow, Ellen & Chen, Su-I & Huffman, Marie. 1997. Syllable weight: Convergence of phonology and phonetics. *Phonology* 14. 47–82.

Burani, Cristina & Paizi, Despina & Sulpizio, Simone. 2014. Stress assignment in reading Italian: Friendship outweighs dominance. *Journal of Memory and Cognition* 42. 662–75.

Chierchia, Gennaro. 1986. Length, syllabification, and the phonological cycle in Italian. *Journal of Italian Linguistics* 8. 5–34.

Clements, George N. 1990. The role of the sonority cycle in core syllabification. In Kingston, John & Beckman, Mary (eds.) *Between the grammar and physics of speech*, 283–333. Cambridge: Cambridge University Press.

Davis, Stuart. 1990. Italian onset structure and the distribution of *il* and *lo*. *Linguistics* 28. 43–55.

Delmonte, Rudolfo. 1999. SIWL data base.

D'Imperio, Mariapaola & Rosenthall, Sam. 1999. Phonetics and phonology of main stress in Italian. *Phonology* 16. 1–28.

Fava, Elisabetta & Magno Caldognetto, Emanuela. 1976. Studio sperimentale delle caratteristiche elettroacustiche delle vocali toniche e atone in bisillabi italiani. In *Studi di fonologia e fonetica*, 35–80. Rome.

Josselyn, Freeman Marshall. 1900. *Etude sur la phonétique italienne*. Paris.

Krämer, Martin. 2009. *The Phonology of Italian*. Oxford: Oxford University Press.

Marotta, Giovanna. 2008. Lenition in Tuscan Italian (Gorgia Toscana). In Brandão de Caravalho, Joaquim & Scheer, Tobias & Ségéral, Philippe (eds.) *Lenition and fortition*, 235–271. New York: Mouton de Gruyter.

McCrary, Kristie. 2004. *Reassessing the role of the syllable in Italian phonology: An experimental study of consonant cluster syllabification, definite article allomorphy and segment duration*. Los Angeles: University of California. (Doctoral dissertation)

Morelli, Frida. 1999. *The phonotactics and phonology of obstruent clusters in Optimality Theory*. College Park: University of Maryland. (Doctoral dissertation)

Parmenter, C.R. & Carmen, J.N. 1932. Some remarks on Italian quantity. *Italica* 9. 103–108.

Prince, Alan & Smolensky, Paul. 2004. *Optimality Theory*. Cambridge: MIT Press.

Pulgram, Ernst. 1970. *Syllable, word, nexus, cursus*. The Hague: Mouton.

Saib, Jilali. 1974. Gemination and spirantization in Berber: Diachrony and synchrony. *Studies in African Linguistics* 5. 1–25.

Saltarelli, Mario. 1970. *A phonology of Italian in a generative grammar*. The Hague: Mouton.

Sievers, Eduard. 1876. *Grundzüge der Lautphysiologie*. In *Bibliothek Indogermanischer Grammatiken, 1*. Leipzig: Breitkopf und Hartel.

Steriade, Donca. 1982. Greek prosodies and the nature of syllabification. Cambridge: MIT. (Doctoral dissertation)

Steriade, Donca. 1999. Alternatives to syllable-based accounts of consonantal phonotactics. In Steriade, Donca & Fujimura, Osamu & Joseph, Brian & Palek, Bohumil (eds.) *Proceedings of the 1998 Linguistics and Phonetics Conference*, 205–242. Prague: The Karolinum Press.

Thornton, Anna M. & Iacobini, Claudio & Cristina Burani. 1997. *Una base di dati sul Vocabolario di Base della lingua italiana*. Roma: Bulzoni.

Ulfsbjorninn, Shanti. 2016. Bogus clusters and lenition in Tuscan Italian: Implications for the theory of sonority. University of Lyon/University College London. (Unpublished ms.)

Vogel, Irene. 1982. *La sillaba come unità fonologica*. Bologna: Zanichelli.

Michael Kenstowicz
MIT
kenstow@mit.edu

A sensorimotor interpretation of Logical Form, and its application in a model of Māori sentences

Alistair Knott and Martin Takac

1. Introduction

In this paper we will present a model of the syntax of Māori transitive clauses. In one sense, the model will be familiar to linguists working within the Government-Binding/Minimalist tradition, but in another sense, it will be novel, because it is implemented as a neural network, within a model of nonlinguistic sensory and motor processing.

The guiding hypothesis behind the model is that the logical form (LF) of a sentence reporting some directly observable event in the world can be interpreted as a description of the sensory and/or motor processes involved in *experiencing* this event. This proposal is described in more detail elsewhere (see e.g. Knott 2012; Takac et al. 2012). We hypothesise that language is intimately connected with the sensorimotor mechanisms through which we experience the world. This idea has received a lot of attention in cognitive science, within 'embodied' models of cognition (see e.g. Feldman & Narayanan 2004; Barsalou 2008). But the idea has some interesting implications about language universals that cognitive scientists do not typically pursue. If language is connected to sensorimotor mechanisms, then we expect structural similarities between all languages, because speakers of all languages have the same sensorimotor mechanisms. If language is *strongly* connected to sensorimotor mechanisms, as many embodied linguists believe, then we should expect a substantial set of structural similarities between languages. Such similarities are clearly not visible 'on the surface', so the only way to maintain a strongly embodied model of language is to adopt some linguistic theory that posits cross-linguistic universals at some 'underlying' level of structural representation. This argument provides an interesting way of thinking about Chomskyan models of syntax. Chomskyan models take linguistic universals seriously: identifying underlying structures that obtain in many languages is at the heart of the Chomsykan research programme. From this perspective, a Chomskyan account of syntax might provide an ideal vehicle for the expression of 'strongly' embodied models of language.

This suggestion upsets Chomskyan linguists and cognitive scientists in equal measure. Cognitive scientists tend not to like Chomskyan models – they are not implemented, they provide no account of sentence processing, and they cannot represent the collocational surface structures in text that modern statistical linguistics is so good at characterising. Chomskyan linguists often see the project of looking for neural correlates of syntactic structure as peripheral to the main work to be done. Liz, it must

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 101–114

be said, has been quite supportive of the line of work we are pursuing – though she did say she was glad she's not the one doing it!

We'll begin in Section 2 by sketching a simple LF template for a transitive sentence. In Section 3 we will outline a sensorimotor interpretation for this LF structure. The basic idea is that the LF structure describes the process of 'rehearsing' a sensorimotor process – in this case, the process of perceiving an event involving a transitive action. In Section 4 we will describe a neural network mechanism that implements this sensorimotor rehearsal process. This mechanism doubles up as a sentence generator: during rehearsal, sensorimotor representations that become active can trigger output phonological representations, through a network that is trained by exposure to a particular language. The training process involves learning the meanings of individual word stems and inflections, but also involves a process akin to parameter-setting: the network has several opportunities to generate phonological signals reflecting the semantic constituents of a transitive sentence, and learns to take the opportunities that result in surface structures in the exposure language. Our hope is that when Chomskyan linguists look at this network, and screw up their eyes a bit, they can see its sensorimotor representations as encoding LF structures, and the mechanism that maps sensorimotor representations onto output phonology as a device for learning the parameters that map LF to PF in a particular language. In Section 5 we will show how the network can learn some simple Māori sentences.

2. LF structure of a transitive sentence

The LF structure we assume for a transitive sentence is shown in Figure 1.

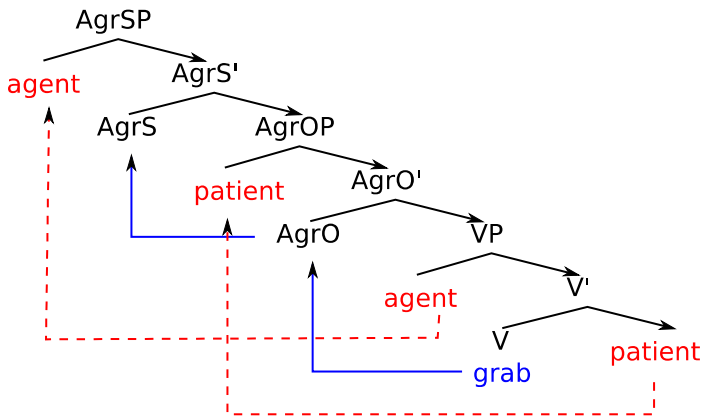


Figure 1: Schematic LF structure of a transitive clause

This corresponds roughly to the structure sketched by Chomsky (1995), summarising the GB model as it had advanced to by that time, in preparation for his initial statement of the Minimalist model. It incorporates Pollock's (1989) suggestion that the agreement features of the verb occupy a separate functional projection above VP, and Koopman & Sportiche's (1991) suggestion that the subject of a sentence is base-generated at the specifier of VP. Chomsky's (1995) analysis added the suggestion that there are two agreement projections in a transitive clause, one for the subject and one

for the object. In this model, both the subject and the object of a transitive clause are base-generated in VP: the subject in [Spec,VP] and the object in [Comp,V]. They each raise to the specifier of a higher functional projection, to be assigned Case: the subject to [Spec,AgrSP] and the object to [Spec,AgrOP] (see the red arcs in the figure). The verb raises by head-movement successively to the heads of AgrO and AgrS (see the blue arcs in the figure). The justification for the movement of the verb is easier to give in Minimalist terms. The verb is fully inflected when generated in the V head, and has to raise to these two heads to 'check' its agreement features. For now, we will omit the tense projection that featured in Chomsky's model, but we will introduce it in Section 5.1.

This model is attractive because it simplifies both Case-assignment mechanisms and theta-role-assignment mechanisms. Chomsky's positing of two agreement projections simplifies Case assignment, because Case is now uniformly assigned by a functional head to its specifier. Koopman & Sportiche's positing of a VP-internal subject simplifies theta-role assignment, because 'agent' and 'patient' roles (or 'proto-agent' and 'proto-patient', to use Dowty's 1991 terminology) can now be assigned by the verb locally, within its maximal projection, to its specifier and complement positions respectively.

The structure sketched in Figure 1 provides the basis for a simple account of Māori transitive sentences. To account for the VSO structure that is typical of these sentences, we can posit that in Māori, V raises to its high position before spell-out, while S and O raise to their Case-assigning positions after spell-out, so that at PF, V is pronounced at its 'high' position (in our sketch, at the AgrS head), while S and O are pronounced at their base positions in VP. This account of VSO languages was one of the motivations for Koopman & Sportiche's model of VP-internal subjects, and several models of Māori along these lines have been developed, among which Liz's models feature prominently (see e.g. Pearce & Waite 1997; Pearce 2000). More recent models of Māori sometimes extend or revise this scheme (e.g. Pearce 1998; 2002), but for our account we will adopt this Chomskyan model, preserved in aspic from 1995.

3. A sensorimotor interpretation of the LF structure of a transitive clause

All linguists think of syntactic structures as having cognitive significance: they portray something about how sentences are represented in the brain. How does the LF structure sketched in Figure 1 do this? One suggestion is that it somehow describes a cognitive *representation*: something stored in a pattern of activity somewhere in the brain, or in a pattern of synaptic connections. Our suggestion is that it describes a *process* that takes place in the brain, rather than a static representation. Specifically, it describes a process whereby a particular sensorimotor experience is *rehearsed*, or relived. We assume a particular model of sensorimotor processing, which emphasises the *sequential structure* of the sensory and motor operations through which we interact with the world. The basic principles of this model were introduced by Ballard et al. (1997). As set out persuasively in that paper, sensorimotor operations often have to be executed in a particular sequence: for instance, an agent cannot readily classify

an object presented visually until she has attended to it, overtly or covertly, and cannot reach for a target object until it has been both attended to and classified.

3.1. A sensorimotor model of reaching-to-grasp

Drawing on a large body of experiments in neuroscience, we have developed a model of the sequence of sensorimotor processes that an observer must execute in order to experience an event involving a simple transitive action – a reach-to-grasp action (see Knott 2012 for details). Following Ballard et al., we hypothesise that the atomic elements of this sequence are all operations of the same basic type: a sensory or motor operation is executed (which we term a **deictic operation**), which updates the observer's current physical and cognitive state (which we term the observer's **context**), generating a sensory representation as a side-effect (which we term the **reafferent signal**). The new context permits the execution of other deictic operations; thus sensorimotor processing is naturally structured into *sequences* of deictic operations. We call these sequences **deictic routines**, again following Ballard et al.

The deictic routine involved in experiencing a reach-to-grasp action is illustrated in Figure 2.

Context	Deictic operation	Reafferent signal
C1	Attend to agent	Attending-agent
C2	Attend to patient	Attending-patient
C3	Execute motor action	Reattending-agent
C4		Reattending-patient

Figure 2: The deictic routine involved in experiencing a reach-to-grasp action

It comprises three deictic operations. The first operation is an action of attention to the agent. This could be implemented in an operation like a saccade, that points the observer's fovea towards a particular external agent in the world. But it could also be a more internal action of attention that focuses the observer's attention on herself: this is what happens when the observer 'decides to act', thereby selecting herself as the agent of whatever action takes place next. In each case, the attentional action allows activation of a representation of the agent as a reafferent consequence.

The second operation is an action of attention to the target of the reach action. If the observer is the agent, this involves directing attention to an object in her peripersonal space. If the agent is some external actor, it involves following the gaze of this actor to identify the intended target of her reach action. Again, in either case, the attentional action allows activation of a representation of the target object as a reafferent consequence.

The third operation is one whereby the observer monitors a continuous motor action in real time, until it is completed. Interestingly, during this process, the observer activates the category of the action in question, but also activates a second representation *of the agent*, as a reafferent consequence of action monitoring. This time the agent is represented as a dynamic, articulated entity, rather than just as a static object of attention. At the completion of the monitored action, the observer also activates a second representation *of the target object*, again within the motor modality: roughly speaking, the location of the object is represented by the location of the agent's arm, and the shape of the object is represented by the shape of the agent's hand.

3.2. LF structure as a representation of a rehearsed deictic routine

There are many similarities between the structure of the deictic routine for transitive actions sketched in Section 3.1 and the LF structure of a transitive sentence outlined in Section 2. In each case, the structure is composed of instances of a recursively defined 'basic building block'. For LF structure, the building block is the X-bar schema, and the recursive principle is the one which allows an XP to occupy the complement of another XP. For the deictic routine, the building blocks are deictic operations. These building blocks also align well with each other: in each case, we have an element associated with the (proto-)agent, followed by an element associated with the (proto-)patient, followed by an element associated with the action. Finally, in each case, there are two representations of the agent, and two representations of the patient.

These similarities suggest an interesting cognitive interpretation of LF structure. According to this interpretation, the LF structure of a sentence reporting a transitive event represents the deictic routine through which this event was experienced. Each XP in the right-branching LF structure identifies one of the deictic operations in the routine. The head of each XP denotes a deictic operation, while its specifier denotes its reafferent consequence. The right-branching structure of XPs identifies the sequential order in which the deictic operations occur. This interpretation is illustrated in Figure 3.

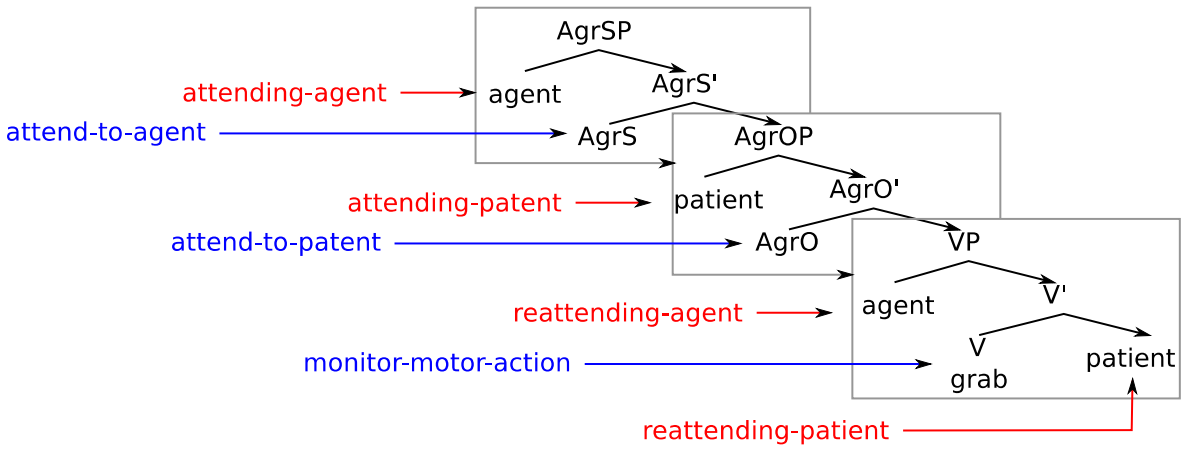


Figure 3: A sensorimotor interpretation of the LF of a transitive clause

This way of thinking about LF requires some mental adjustments. We now have to think about an LF structure not as the product of some abstract process of derivation, but as a reflection of an *actual* cognitive process, that can be directly studied. (What is more, this process is not a linguistic process *per se*, but a sensorimotor one.) Moreover, the operations that 'move' constituents from one LF position to another now have a completely different significance. These movements now have a *temporal* interpretation. For instance, when we see the subject DP raising from a VP-internal position to a higher position, we have to interpret this as implying that there are two *times* when the subject 'appears'. Similarly, the raising of the object DP tells us there are two *times* when the object 'appears'.

However, we suggest that thinking of LF structures as representing processes in this way can be extremely useful. For instance, consider the topic of DP-raising. The requirement that DPs raise 'to get Case' is ultimately motivated because it contributes to a descriptively adequate and economical model of many languages. But it would be nice to justify it in more concrete terms as well. In the sensorimotor interpretation of LF that we propose, the raising of DPs is a manifestation of a constraint *on sensorimotor processing*: an observer has to *attend* to the agent and patient of a transitive action (in that order) before she can monitor this action (both for actions she executes herself, and for actions she perceives being executed by other agents). Thinking of an LF structure as representing a sensorimotor sequence is also helpful from the perspective of linking models of LF derivation to models of sentence processing. In the interpretation we suggest, an LF structure does not just represent a speaker's declarative knowledge of language – it *directly represents a cognitive process* – and moreover, one which is plausibly involved in the actual generation of sentences. We will flesh this idea out in Section 4.

The idea that LF structure encodes a sequence is not completely foreign to linguists. Kayne's (1994) model of LF, which was another of the influences in Chomsky's (1995) model, stipulates that the specifier of an XP appears before its complement at PF. Kayne does also explicitly state that LF has hierarchical, and not 'linear' (i.e. temporal) structure. However, it has at least an *implicit* temporal structure, in the structure it imposes on PF. And in fact, he tangentially suggests that the right-branching form of LF structures in his model may have a temporal origin.

Before we introduce our network model, we need to consider what head movement means in our reinterpreted conception of LF. Head movement allows a fully inflected verb to 'raise' from the head of V, through the head of AgrO, to the head of AgrS. If the head of each XP denotes a deictic operation, and the right-branching structure of these XPs denotes the sequence in which they occur, then the mechanism of head movement allows for deictic operations to be reported '*out-of-sequence*': the motor action denoted by the V head appears 'too soon' when the inflected verb occupies its 'high' positions, and the actions of attention to agent and patient denoted by the Agr heads appear 'too late' when the inflected verb occupies its 'base' position. In our interpretation of this phenomenon, we introduce another component to the sensorimotor model. Sentences are not generated as a direct side-effect of sensorimotor experience: rather, they are produced from a representation of an experienced event held in *working memory*. Our model of sensorimotor processing

includes a model of working memory for experienced events. In this model, an agent stores an experienced event as a *prepared deictic routine* that can be *replayed*. In particular, it can be replayed in a special mode, in which activated *sensorimotor* representations can trigger output *phonological* representations. A great deal is known about how deictic routines are stored in the brain. Crucially, the working memory representation of a deictic routine holds representations of all its component operations active *in parallel*, even though they are executed sequentially (see for instance Averbeck et al. 2002). We propose that the phenomenon of head-raising arises because heads are phonological expressions of deictic operations *as represented in working memory*, rather than in the sensorimotor media where they occur transiently during actual sensorimotor experience. We propose that the LF of a sentence describes a deictic routine *replayed from working memory*, rather than one occurring in real time. During this process, there is an interesting mixture of 'sustained' and 'transient' representations. We propose that heads are read from the 'sustained' ones, and specifiers are read from the 'transient' ones.

With these preliminaries, we can now introduce our model of sentence generation.

4. A neural network model of sentence generation

Our sensorimotor conception of LF lends itself to a model of sentence processing – specifically, a model of sentence generation. As just described, we think of an LF structure as a representation of the mixture of sustained and transient sensorimotor representations that are activated when a deictic routine, encoding a recently experienced event, is replayed from working memory. We envisage that this replay operation can happen in a special mode, in which active sensorimotor representations trigger phonological representations: in this mode, a sequence of phonological representations will be produced. Thus we see the process of sentence generation as a process that maps a rehearsed deictic routine onto a phonological sequence.

In this section, we will introduce a computational model of this generation process. It is implemented as a neural network. For details about the architecture of the network and its training, see Takac et al. (2012).

The basic structure of the network is shown in Figure 4. It takes a sequence of inputs, at three successive time steps, and at each step has the opportunity to generate a phonological output. Its inputs come both from the 'sustained' representations of the complete deictic routine held in working memory, which are the same at each time step, and from the 'transient' representations of individual operations in the routine that change at each step. At each step, there is a mechanism that selects first the transient representation and then the sustained one.

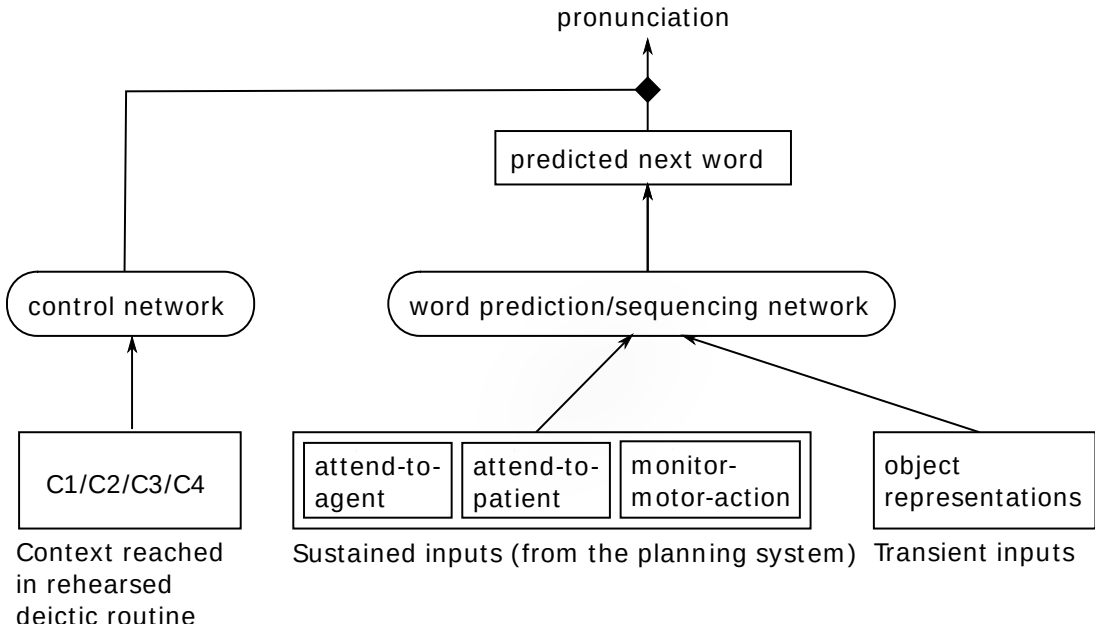


Figure 4: Architecture of the sentence generation network

There are two sub-networks. The **word production and sequencing network** (or WPSN) takes the currently selected sensorimotor input representation, and maps it onto a predicted output phonological representation (a stem and an inflection). The **control network** decides whether or not this phonological representation should be explicitly *pronounced*. It does this using information about what 'stage' of rehearsal has been reached: this is an encoding of the 'context' that is updated after each sensorimotor operation.

Both networks are trained on sentences from a given language that denote concrete events. Each sentence is paired with the deictic routine through which the associated event is experienced, according to the model outlined in Section 3.1. At the start of training, we assume the learner's ability to rehearse a deictic routine is poor, and items from the routine are paired indiscriminately with words in the associated sentence. During this time, the WPSN slowly learns a small set of word meanings – that is, associations between sensorimotor signals and output word stems – through a process called 'cross-situational learning' (Siskind 1996). At a certain point, the learner becomes able to rehearse deictic routines accurately. At this point, the control network starts to be trained in addition. Training now involves rehearsing each deictic routine in its proper sequence, to produce a sequence of output words: these words are compared to the words in the associated training sentence, beginning with the first word.

The WPSN is trained to produce the 'current word' in the training sentence. Meanwhile, the control network is trained in the meta-level task of when to overtly *pronounce* the words produced by the WPSN. As already mentioned, a rehearsed deictic routine provides several opportunities to produce the key constituents in a transitive sentence. There is an 'early' opportunity to produce a word denoting the agent in the first rehearsed deictic operation (attention to the agent), and then a 'late'

opportunity to produce such a word in the third rehearsed operation (monitoring of the motor routine, where the agent's characteristic pattern of movement is represented). Similarly, there is an 'early' opportunity to produce a word denoting the patient in the second rehearsed operation (attention to the patient), and a 'late' opportunity to pronounce such a word in the third rehearsed operation (where the patient is represented as a motor state of the agent's hand/arm). Finally, there are several opportunities to produce phonological outputs denoting the deictic operations themselves, because these outputs are generated from tonic representations of these operations in the planning system, which are active throughout the rehearsed routine. From all these opportunities to pronounce words, the control network learns to take opportunities that result in surface sentence forms resembling those of the training language.

Consider a schematic Māori training sentence, comprising a verb, a subject and an object, in that order, and the deictic routine paired with this sentence in training. The network receives each operation in the rehearsed deictic routine in turn. In the first operation, it receives first the agent (an opportunity to pronounce the subject), and then the set of planned deictic operations (an opportunity to pronounce the verb and associated inflections). Assume the WPSN correctly generates the word denoting the agent. This word is compared to the first word in the training sentence – because they are not the same, the control network will learn (incrementally) that it should *not pronounce* the agent in this early 'context'. Assume the WPSN also correctly generates a word denoting the action. This word is compared to the first word in the training sentence – this time it does match, and the control network learns (incrementally) that it *should pronounce* the action in this early 'context'. In syntactic terminology, the training sentence provides a small piece of evidence in favour of pronouncing verbs/inflections 'high' in Māori, and against pronouncing subjects 'high'.

The above example assumes that the WPSN has already learned the word forms denoting the relevant sensorimotor symbols. However, this is not always the case: the WPSN still has to learn many words. But as the control network learns the right 'opportunities' to pronounce words in the training language, it also generates improved training data for the WPSN. When the control network learns not to pronounce the subject 'high' in Māori, it refrains from training the WPSN to map the agent onto the first word in a Māori sentence. As the WPSN's learning of word meanings improves, it in turn generates cleaner training data for the control network, so the two networks 'bootstrap' each other. This simulates the effect whereby knowledge of syntax aids word learning (see e.g. Aslin et al. 1996).

There is one other important feature of the WPSN to introduce. This network takes sensorimotor signals and learns to generate word forms, as already noted. But it also maintains a record of the sequence of words *produced so far* in the sentence being generated: its decision about how to map sensorimotor signals onto word forms is conditioned on this sequence of recent words. This provides the network with a mechanism for producing sequences of words that conform to surface regularities in the exposure language. It has the ability to learn 'idiomatic' or collocational structures in language – an ability that is hard to model within a traditional GB or Minimalist paradigm (as notoriously discussed by Jackendoff (2002) in his criticism of the

paradigm). We see this ability to learn idioms as one of the key benefits of our proposed reinterpretation of LF structures. The model outlined in this section is a model of how LF structures participate in actual sentence processing – that is, in routines that generate surface sentence forms. Within this model, we can introduce machinery that learns regularities in surface sentence structure, *in addition to* the machinery that learns traditional GB/Minimalist-style parameter settings. We hope that a Chomskyan linguist will be able to look at our model (with eyes screwed up) and see LF structures, head movement, DP-raising, and parameter-setting. But an empiricist linguist should also be able to look at our network and see standard neural network mechanisms for learning surface collocations and idioms.

5. Training the network on a corpus of Māori sentences

5.1. Adding tense and causative actions to the deictic routine

One of the distinctive features of Māori (along with other Polynesian languages) is its use of tense markers. In Chomsky's (1995) model, tense information is contributed by a separate functional projection, 'high' in the LF structure: we will assume the tense projection TP is the highest projection, above AgrSP and AgrOP. Importantly, we have to extend our sensorimotor model of transitive events to incorporate a deictic operation that occurs before 'attention to the agent', which plausibly contributes tense information. Our suggestion is that this operation is one which determines whether the observer attends to the perceptual here-and-now as a source of incoming events, or to his own episodic memory. The operation of attending to the here-and-now corresponds to a 'present-tense' head; the operation of 'engaging episodic memory' corresponds to a 'past-tense' head. We assume a particular neural network model of episodic memory, presented elsewhere (Takac & Knott 2016a; 2016b), in which events are stored in and retrieved from long-term memory in the form of deictic routines, with the same structure as those generated during experience. In this model, events are retrieved from episodic memory *into working memory*, from where they can be rehearsed like events that have just been perceived.

Another distinctive feature of Māori is its productive use of the causative prefix *whaka* on verbs. Again, we can extend the LF structure of the clause to model this. Our model of causatives, like many others, is based on Larson's (1988) concept of VP shells: we assume an 'outer' VP headed by 'cause' that introduces an 'inner' VP denoting the caused event. Again, we must extend the sensorimotor model, to provide a plausible sensorimotor correlate for the outer VP, and its relationship to the inner VP. We have developed a neural network model of causative actions, again presented elsewhere (Lee-Hand & Knott 2015) in which there are correlates both of the causative action and of the caused event.

With these preliminaries, we can introduce an experiment in which our sentence generation network was trained on a corpus of Māori sentences.

5.2. A training corpus of Māori sentences

Our training corpus consisted of 160,000 sentences (16 training epochs each with 10,000 sentences) generated at random from a simple grammar, out of which approximately 60% were transitive sentences, 27% causative sentences and 13% intransitive sentences. Intransitive sentences lacked the AgrO projection, and had no object. Each sentence had a tense/aspect marker, which was either *i* (for past tense) or *kei te* (for present continuous) or *e (…) ana* (for continuous aspect, with tense unspecified). The parentheses in *e (…) ana* indicate the position of the verb. We did not include a separate projection for aspect (this is a topic for ongoing work). Instead, our network was required to learn to produce *e (…) ana* from the same deictic routines as *i*. We used a range of DPs that were semantically suited to the argument roles of verbs. DPs could include pronouns (first/second/third-person and singular/dual/plural); dual and plural first-person pronouns could be exclusive or inclusive. We also included reflexive pronouns. Third-person DPs using common nouns could use definite or indefinite determiners. Common nouns could be singular or plural; we included some irregular plural nouns (*tīpuna* 'ancestors, grandparents'*, wāhine* 'women'*, tamariki* 'children'*).* We used 31 open-class verbs in our example sentences, and 42 open-class nouns. Intransitive verbs could participate in causative constructions; in that case, for technical reasons, the causative prefix *whaka* appeared as a separate word. Object DPs were introduced with the particle *i*: again, this particle had to be learned as an idiom. Finally, we included some additional continuous idioms in the training sentences (*kai moana* 'seafood, shellfish'*, pene rākau* 'pencil'*, tipuna whaea* 'great grandmother'*, tipuna matua* 'great grandfather'*, taonga tākaro* 'traditional games'). In addition, since we had not provided a dedicated sensorimotor operation for *ana* in the *e (…) ana* construction, this construction also functions as an 'idiom' for our network: in this case, a discontinuous one. And since our deictic routines do not model the internal structure of DPs, determiner-noun constructs also function as (continuous) idioms for our network – as they apparently do for infants at a certain developmental stage (see e.g. Pine and Lieven 1997).

Some examples of the training sentences in our corpus are given below.

1. E whaka hoki ana kōrua i te parāoa.
   CONT CAUS return CONT 2DU OBJ DET.SG bread
   'You [dual] are returning the bread.'

2. I mātakitaki tāua i ngā taonga tākaro.
   PAST watch 1DU.INCL OBJ DET.PL games
   'We [dual, inclusive] watched the games.'

3. Kei te horoi tātou i a kōrua.
   PRES wash 1PL.INCL OBJ PERS 2DU
   'We [plural, inclusive] wash you [dual].'

4. Kei te whaka ngaro ahau i a māua.
   PRES CAUS be.hidden 1SG OBJ PERS 1DU.EXCL
   'I hide us [dual, exclusive].'

5. E whaka makere ana he tamaiti i ngā kau.
   CONT CAUS fall CONT INDEF child OBJ DET.PL cow
   'A child drops the cows.'

5.3. Results

In each training run, we trained our network on 10,000 sentences of the kind described above, each paired with its associated deictic routine. To assess the network's performance, we tested it by presenting it with the deictic routines associated with each of the 10,000 training sentences and the deictic routines associated with an additional 1,000 sentences unseen during training, and asking it to generate a sentence from each. We compared the generated sentence to the sentence paired with the deictic routine during training. The sentence was judged to be 'correctly' generated if it matched the paired sentence in every respect, modulo synonyms. The model was able to correctly generate 99.2% of training sentences and 98.5% of unseen ones.

4. Discussion

In this paper we have described a neural network that can learn a fragment of a natural language grammar, when trained on sentences from a given language, paired with semantic representations. The semantic representations we use are distinctive, in that they derive directly from a model of sensorimotor processing, rather than being expressed in an artificial logical language. But they are also distinctive in having a direct correspondence with Chomskyan LF structures. In our model, constraints on the structure of sensorimotor routines, and on the working memory mechanisms that store and replay them, are reflected in the space of possible surface languages: our network makes use of these constraints to learn the grammar of its exposure language. In this model, the innate 'knowledge of language' that is captured by LF structure is (at least partly) due to the structure of the sensorimotor system. In this sense our model is an 'embodied' model of language. But by the same token, it is also a 'nativist' model, of an interesting new kind. At the same time, our sentence generation network can also learn idiomatic surface structures in the exposure language: it thus implements a mixture of nativist and empiricist models.

Note that if the network is exposed to training corpora from other languages, it will learn different parameter settings that choose different positions for verb heads and their arguments. The network has also been trained on SVO languages (English and Slovak) and SOV languages (Japanese), and performs at a similar level. It can learn to express tense in verb inflections as well as in stand-alone particles. It can learn to produce subject and object agreement inflections on verbs, or to omit these. It can learn to produce pronouns as clitics adjoined to verb heads (as in Slovak) or in regular argument positions (as in English). It can learn to realise causative actions with an explicit prefix (as in Māori), or without (as in English). For details of these experiments, see Takac et al. (2012) (plus papers in preparation on Slovak and Japanese). The Māori experiments reported here are particularly useful in demonstrating an ability to learn a

rich pronoun paradigm, stand-alone tense markers, and morphology realising the 'cause' concept in causative sentences.

Of course, we are just scratching the surface of the complexity of Māori sentence structure. We do not have a model of passive sentences, which often provide the most natural way of rendering events in Māori. We have not begun to model the distinctive topicalising projections in the left periphery, or the internal structure of Māori DPs – or indeed any forms of predicative or stative sentence – all topics that Liz has studied in great depth. Our grammar development methodology is rather slow compared to that of a theoretical linguist: every LF structure has to be justified not only by its role in a wider model of grammar, but also as a plausible deictic routine, motivated by research in neuroscience, and tested in a neural network model. However, we are not deterred by this slower pace of progress: we think it is helpful to use deictic routines to 'triangulate' on LF structures in this way. One difficulty with theoretical linguistics in general is that there are often many plausible theoretical analyses of a given phenomenon: the data frequently underdetermine the space of possible theories. If models of sensorimotor processing can provide additional constraints on the process of building syntactic models, that could be a good thing in the long run.

Acknowledgements

The authors would like to acknowledge Marsden grant 13-UOO-048 for supporting this work.

References

Aslin, Richard N. & Woodward, Julide Z. & LaMendola, Nicholas P. & Bever, Thomas G. 1996. Models of word segmentation in fluent speech to infants. In Morgan, James L. & Demuth, Katherine (eds.) *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, 117–134. Mahwah, NJ: Erlbaum.

Averbeck, Bruno B. & Chafee, Matthew V. & Crowe, David A. & Georgopoulos, Apostolos P. 2002. Parallel processing of serial movements in pre-frontal cortex. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* 99(20). 13172–13177.

Ballard, Dana H. & Hayhoe, Mary M. & Pook, Polly K. & Rao, Rajesh P.N. 1997. Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences* 20(4). 723–767.

Barsalou, Lawrence W. 2008. Grounded cognition. *Annual Review of Psychology* 59. 617–645.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619.

Feldman, Jerome & Narayanan, Srinivas. 2004. Embodied meaning in a neural theory of language. *Brain and Language* 89(2). 385–392.

Jackendoff, Ray. 2002. *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Knott, Alistair. 2012. *Sensorimotor cognition and natural language syntax*. Cambridge, MA: MIT Press.

Koopman, Hilda & Sportiche, Dominique. 1991. The position of subjects. *Lingua* 85. 211–258.

Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19(3). 335–391.

Lee-Hand, Jeremy & Knott, Alistair. 2015. A neural network model of causative actions. *Frontiers in Neurorobotics* 9. Article 4.

Pearce, Elizabeth. 1998. Incorporation in Māori syntax. In *Proceedings of AFLA V*, University of Hawai'i.

Pearce, Elizabeth. 2000. Argument positions and anaphora in the Māori clause. In Fisher, Steven Roger & Sperlich, Wolfgang (eds.) *Leo Pasifika: Proceedings of the Fourth International Conference on Oceanic Linguistics*, vol. 2, 313–325. Auckland: The Institute of Polynesian Languages and Literatures.

Pearce, Elizabeth. 2002. VP versus V raising in Māori. In Rackowski, Andrea & Richards, Norvin (eds.) *Proceedings of AFLA VIII: The Eighth Meeting of the Austronesian Formal Linguistics Association*, 225–240. Cambridge, MA: MIT Working Papers in Linguistics.

Pearce, Elizabeth & Waite, Jeffrey. 1997. Kia and ki te complementation in Māori: An unaccusative analysis. *Te Reo* 40. 45–75.

Pine, Julian M. & Lieven, Elena V.M. 1997. Slot and frame patterns in the development of the determiner category. *Applied Psycholinguistics* 18. 123–138.

Pollock, Jean-Yves. 1989. Verb movement, universal grammar and the structure of IP. *Linguistic Inquiry* 20(3). 365–424.

Siskind, Jeffrey Mark. 1996. A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition* 61(1–2). 39–91.

Takac, Martin & Knott, Alistair. 2016a. Working memory encoding of events and their participants: A neural network model with applications in sensorimotor processing and sentence generation. In *Proceedings of the 38th Annual Meeting of the Cognitive Science Society* (CogSci), 2345–2350.

Takac, Martin & Knott, Alistair. 2016b. Mechanisms for storing and accessing event representations in episodic memory, and their expression in language: a neural network model. In *Proceedings of the 38th Annual Meeting of the Cognitive Science Society* (CogSci), 532–537.

Takac, Martin & Benuskova, L & Knott, Alistair. 2012. Mapping sensorimotor sequences to word sequences: A connectionist model of language acquisition and sentence generation. *Cognition* 125. 288–308.

Alistair Knott
University of Otago
alik@cs.otago.ac.nz

Martin Takac
University of Otago
takac@cs.otago.ac.nz

Some issues on verbal reciprocals in Malagasy[1]

Paul Law

Abstract

This paper claims that there is but one morphological realization of reciprocals in Malagasy. It clarifies Keenan & Razafimamonji's (2004) arguments against the syntactic analysis of the reciprocal morpheme *if-*. It shows that most of them are in fact quite compatible with the syntactic treatment of *if-*. It nevertheless argues that certain facts about comparative ellipsis cannot be accounted for by the syntactic account, but can be explained by their view that *if-* is derived from an affixation rule. Split plural argument, however, remains problematic for both views.

1. Introduction

In their detailed study of reciprocals in Malagasy, Keenan & Razafimamonjy (2004), henceforth K&R, suggested that reciprocal verbs are formed by prefixing *if-* to a derived verb, i.e., verbs with a Focus affix, and that *if-* has two allomorphs *ifamp-* and *ifanka-* for certain verbs. More significantly, they concluded that *if-* does not originate in argument position and incorporate into the verb, but is a verbal affix. The combination *if-*V (or *ifamp-*V and *ifanka-*V) is an intransitive verb requiring the subject to be plural. They called the first account the *if-*=Anaphor view, and the second the *if-*=Affix view. I will use these terms for ease of reference.

In this paper I argue that their suggestion of *if-* having two allomorphs misses a morphological generalization that the appearance of *ifamp-* or *ifanka-* on V is possible just in case the form *amp-*V or *anka-*V is independently possible. I agree with their general conclusion that the reciprocal morpheme is a verbal affix and does not have its origin in argument position. However, most of their arguments against the *if-*=Anaphor analysis do not stand. The facts they adduced for their arguments are not necessarily problematic for the *if-*=Anaphor account, given that the problems that arise are not specific to the analysis of the reciprocals. Their general conclusion nevertheless holds, for it is clearly supported by the interpretive property of the reciprocals in comparative ellipsis. Facts concerning split plural arguments nevertheless remain a problem for both views.

[1] I gratefully acknowledge the sabbatical leave granted to me by the City University of Hong Kong. It allowed me the time to think about the various issues discussed here. I also would like thank Ed Keenan for spending many hours discussing Malagasy with me.

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 115–127

2. Morphological realization of the reciprocal

Reciprocal *if-* combines directly with Actor Focus (AF) verbs, whence the grammatical difference between (1b,c) (K&R 2004: 172, 181):

1. a. m+aN+enjika (Manenjika) an-dRabe Rakoto.
   PRE+ACT+chase ACC-Rabe Rakoto
   'Rakoto is chasing Rabe.'

   b. m+if+aN+enjika (Mifanenjika) Rabe sy Rakoto.
   PRE+REC+ACT+chase Rabe and Rakoto
   'Rabe and Rakoto are chasing each other.'

   c. * if+enjika+ina+Rabe sy Rakoto (ifenjehin-dRabe sy Rakoto).
   REC+CHASE+THM+Rabe and Rakoto
   'Rabe and Rakoto are chasing each other.'

Importantly, it combines only with verbs with the AF affix *aN-*, not those with the AF affix *i-* or the causative/potentiality prefix *aha*, as can be seen in the contrast in (2) and (3) (K&R 2004: 182, 183):

2. a. m+i+jery azy aho.
   PRES+ACT+look.at 3SG 1SG
   'I am looking at him.'

   b.* m+if+i+jery isika.
   PRES+REC+ACT+look.at 1PL.INCL
   'We (you and I) are looking at each other.'

3. a. m+aha+lala an-dRabe Rakoto.
   PRES+CAUSE/POT+know ACC-Rabe Rakoto
   'Rakoto knows Rabe.'

   b.* m+if+aha+lala Rabe sy Rakoto.
   PRES+REC+CAUSE/POT+know Rabe and Rakoto
   'Rakoto and Rabe know each other.'

The examples in (2b) and (3b) would be grammatical, if *ifamp-* and *ifanka-*, allomorphs of *if-* according to K&R, appear instead:

4. a. m+ifamp+i+jery isika.
   PRES+REC+ACT+look at 1PL.INCL
   'We (you and I) are looking at each other.'

   b. m+ifank+aha+lala Rabe sy Rakoto.
   PRES+REC+CAUSE/POT+know Rabe and Rakoto
   'Rakoto and Rabe know each other.'

They gave the affixation rule in (5) for the formation of reciprocal verbs (^ is concatenation), and the rule in (6) for interpreting them (K&R 2004: 183, 178–179):[2]

5. *if*^pref^root
   Rec(pref^root) = { if pref=*aN-*, *ana-*, *amp-* or *anka-*
   ifamp^pref^root if pref=*i-* or Ø
   ifank^pref^root if pref=*a-* or *aha-*

6. verb: [NP_{ACC}, NP_{NOM}] ⇒ **Rec**(verb) [$NP_{NOM.PL}$] where θ > θ'in (7)
   θ θ' θ'

7. Agent > Experiencer > Theme/Patient > Other

Rule (6) relates a reciprocal verb to a transitive verb, effectively deriving an intransitive verb from a transitive one. The argument of the derived verb is required to be plural. The truth condition for **Rec**(verb) is as in (8), where IF stands for *if-*, *ifamp-* and *ifanka-*:

8. IF(p)(A) = True if and only if for all distinct x, y in A, p(y)(x) = True.

A reason for taking reciprocal *ifamp-* to be an allomorph of *if-* and not decomposable into *if-* and causative *amp-* is that it lacks the causative meaning regularly associated with *amp-* (see also Keenan & Polinsky 1998: 606). For example, the verb *mifampatoky*, built from *matoky* (*m+a+toky*) 'trusts, has confidence in', is semantically ambiguous. If *ifamp-* is analysed as *if+amp*, then it is not clear how the two readings are to be accounted for (K&R 2004: 182, note 7):

9. a. m+a+toky azy aho.
      PRES+ACT+trust 3SG 1SG
      'I trust him.'

   b. m+ifamp+a+toky isika.
      PRES+REC+ACT+trust 1PL.INCL
      'We trust each other.'

   c. m+if+amp+a+toky azy isika.
      PRES+REC+ACT+trust 3SG 1PL.INCL
      'We inspire each other to trust him.'

Similarly, *mifankahita* 'see each other' cannot be analyzed as *m+if+anka+hita*, nor can *mifankatia* 'like/love each other' be decomposed as *m+if+anka+tia*, for neither ***mankahita* nor ***mankatia* exist. They gave several cases of this kind, e.g., *mifampidera* (<*dera*) 'praise each other', *mifampilaza* (<*laza*) 'say to each other' and *mifampitantara* (<*tantara*) 'narrate to each other.'

The point about the interpretive difference between (9b,c) and the non-existence of forms from which the reciprocal verbs can be said to be derived is well-made. Their

[2] The example in (9b) below shows that *ifamp-* may be followed by the prefix *a-*. It thus seems appropriate to place *a-* in the second line in (5) as well.

suggestion that *ifamp-* and *ifanka-* are morphologically whole allomorphs of *if-* is nevertheless problematic. Conceptually, there is no reason for the allomorphs to have a part that looks exactly like causative *amp-* or *anka-*. They might well be of any other form. Empirically, there is apparently regularity with verbs with *ifamp-*. In many cases *ifamp-* attaches to the same derived verbs with the AF affix *i-* and *a-* to which *amp-* attaches, without reciprocal *if-*. The same holds for (bare) verb roots *ino* 'believe', *aka* 'take' and *hay* 'know' (Abinal & Malzac 1888):

10. a. m+amp+i+dera 'cause to praise', m+amp+i+jery 'cause to look'
m+amp+i+laza 'cause to say', m+amp+i+tantara 'cause to narrate'

    b. m+amp+ino 'cause to believe', m+amp+aka 'cause to take'
m+amp+a+hay 'cause to know'

The forms ***mankahita*, ***mankatia* and apparently **mankahalala* as well, indeed do not exist, but the alternatives with *amp-* do. Abinal & Malzac (1888) listed *mampahita*, *mampitia* and *mampahalala*. The existence of the latter probably alleviates the need for the former, especially when causative *anka-* is not very productive. As we might expect, reciprocal *if-* may attach to the causative verbs to yield *mifampahita* and *mifampahalala* (Abinal & Malzac 1888), although *mifampitia* does not seem to exist.

There are some cases with causative *amp-* and *anka-* in which the interpretation is not predictable, the interpretive difference between (9b,c) is therefore not totally surprising. The causative of the verb root *vangy* 'visit' does not always have causative meaning (Andrianierenana 1996: 72):

11. a. m+aN+vangy (>mamangy) azy aho.
PRES+ACT+visit 3SG 1SG
'I visit him.'

    b. m+amp+aN+vangy (>mampamangy) azy aho.
PRES+CAUSE+ACT+visit 3SG 1SG
'I send him compliments.'

    c. m+amp+aN+vangy (>mampamangy) an-dRavelo an-dRabe aho.
PRES+CAUSE+ACT+visit ACC-Ravelo ACC-Rabe 1SG
'I have Rabe visit Ravelo.'

    d. m+if+amp+aN+vangy (>mifampamangy) Ranaivo sy Rakoto.
PRES+REC+CAUSE+ACT+visit Ravelo and Rakoto
'Ranaivo and Rakoto send each other compliments.'

    e. m+if+amp+aN+vangy (>mifampamangy) an-dRabe Ranaivo sy Rakoto.
PRES+CAUSE+ACT+visit ACC-Rabe Ravelo and Rakoto
'Each of Ranaivo and Rakoto has Rabe send compliments to the other.'
*'Ranaivo and Rakoto make each other visit Rabe.'

The verbs in (11b,d) have causative morphology but no causative meaning.

Similarly, derived verbs with the least productive causative *anka*- sometimes do not have the causative meaning (Andrianierenana 1996: 62). Examples (13b,c) clearly do not have causative meaning:

12. a. marary Rabe.
sick Rabe
'Rabe is sick.'

b. n+anka+marary an-dRabe ny paoma manta.
PAST+CAUSE+sick ACC-Rabe the apple green
'The green apple made Rabe sick.'

13. a. fy ny nahandro.
delicious the dish
'The dishes (prepared food) are delicious.'

b. m+anka+fy ny nataon'ny havako aho.
PRES+CAUSE+delicious the relative.1SG 1SG
'I cherish my relatives.'

c. m+anka+fy ny nahandro aho.
PRES+CAUSE+delicious the dish 1SG
'I find the dishes delicious.'

Given the non-causative interpretations of examples (11b,c) and (13b,c), it is perhaps not too surprising that example (9b) has no causative meaning. As well, it is worth noting that a derived verb with causative *anka*- may have an unpredictable interpretation when prefixed with reciprocal *if*-. The causative of the verb root *hay* 'capable' is *mankahay* 'make capable' (Abinal & Malzac 1888), but *mifankahay* means 'agree, get along' (K&R 2004: 183), without the causative meaning.

In sum, there seems to be little reason for taking *ifamp*- and *ifanka*- to be morphologically whole allomorphs of reciprocal *if*-. The idiosyncratic properties are likely to be accidental, for they arise elsewhere as well.

3. *if*-=Anaphor vs. *if*-=Affix

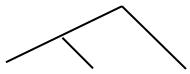
K&R (2004: 202–205) gave seven arguments against the *if*-=Anaphor analysis. I will first consider them in turn, in a slightly different order from the original presentation, and turn to certain facts concerning comparative ellipsis. These can be accounted for in the *if*-=Affix view, but not by the *if*-=Anaphor view. Lastly, I will present examples of split plural argument that are problematic for both views.

First, according to them, on the *if*-=Anaphor view, the reciprocal verbs are not semantically interpreted. The nonreciprocal verb and its object are interpreted compositionally. On this view it is unnatural to handle noncompositional cases. For example, *mifampody* is the reciprocal of the causative of *mody* (<*fody*) 'return home' and literally means 'make each other go home'. But in fact, it means 'to reconcile, said of husband and wife'. If *if*- is an argument of the verb, then it is unclear how the actual, largely unpredictable reading arises from compositional interpretation. The point is well-made, but to the extent that phrasal idioms are possible, e.g., *kick the bucket*

meaning ‘to die’, the idiomatic reading of *mifampody* can be given a phrasal analysis in which *if-* is an anaphor incorporated from argument position. The *if-*=Affix view, too, cannot explain how the idiomatic reading comes about, insofar as it is not predicted by the rule in (6).

Second, they argued that if *if-* incorporates into V, then it is likely that the empty category [e] will fail to be property bound, in violation of the Empty Category Principle (Chomsky 1981):

14.



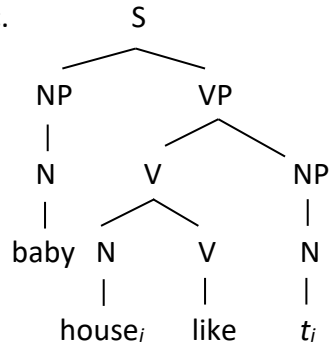
[if+anenjika] [e]

In (14), the moved *if-* does not c-command [e]. The point is well-taken, but is by no means unusual. Syntactic incorporation of this sort is attested in a wide variety of languages, including Mohawk (Postal 1962), cited in Baker (1988: 82, 83):

15. a. yao-wir-a?a ye-nuhwe?-s ne ka-nuhs-a?
PRE-baby-SUF 3FS/3N-like-ASP the PRE-house-SUF
‘The baby likes the house.’

b. yao-wir-a?a ye-**nuhs**-nuhwe?-s.
PRE-baby-SUF 3FS/3N-house-like-ASP
‘The baby likes the house.’

c.



In (15c), *house* does not c-command the position *t* from which it incorporates into V. The structural relationship between the incorporated noun and the empty category it leaves behind is the same as that between *if-* and [e] in (14). A way out is to assume that the condition for syntactic binding of an empty category is m-command (Aoun & Sportiche 1983), where *a* m-commands *b* iff the first maximal projection dominating *a* dominates *b*. In (15c), VP is the first maximal projection dominating *house* dominates the empty category. The same can be said of the structure in (14), taking the reciprocal verb as V and its combination with [e] as VP.

Third, they contended that specifying the class of verbs that *if-* combines with is more natural on the *if-*=Affix view, because it is given as a derivational operation that takes predicates as arguments. It is thus natural that it selects its predicates: AF ones. So *if-* forms part of a natural derivational class in Malagasy, much like causative *amp-* and the agent nominalizer *mp-*. All these prefix to AF verbs. On the *if-*=Anaphor view, the

class of verbs that the object can prefix to would have to be specified. Moreover, no other nominal expressions incorporate as prefixes. This makes incorporation of *if-* to V as a prefix look unmotivated. Nevertheless, these conceptual problems with the *if-*=Anaphor view arise in the *if-*=Affix analysis as well. There seems to be no good reason for why *if-* cannot select predicates with Theme Focus (TF). In fact, the *if-*=Anaphor view readily explains why *if-* cannot appear on TF verbs. Sentences with these verbs require that the Theme argument be externalized to a position outside VP (Keenan 1976; Pearson 2005). But *if-*, a verbal affix, cannot be externalized to a position for phrases.

Certain facts concerning *mp-* nominalization indicate that *if-* may originate in argument position. The examples in (16) show that an argument of the verb root may occur in agent nominalization with the prefix *mp-* (Keenan & Polinsky 1998: 615):

16. a. ny mpihaino azy.
the listener 3.ACC
'The people listening to it/him.'

b. ny mpampianatra azy.
the teacher 3.ACC
'the teacher of him (= 'his teacher')

This sort of nominalization presents a bracketing paradox. Morphologically the verb root forms a unit with the agentive nominalizer *mp-*, but semantically it should form a unit with the argument it is related to:

17. a. morphology: [mp+amp+i+anatra] azy.
b. semantics: mp [amp+i+anatra azy]

It remains unclear how the paradox can be resolved. But what is of interest is that if the bracketing in (17a) is possible, then it is conceivable that agent nominals with the prefix *mp-* and reciprocal *if-* of the sort in (18) (K&R 2004: 198) have the structures in (19) where *if-* originates in argument position, much like the accusative pronoun *azy* in (17a), and incorporates into the verb:

18. mifankahala 'detest e.o.' mpifankahala 'individuals who detest e.o'
mifanampy 'help e.o.' mpifanampy 'individuals who help e.o.'

19. mp [anka+hala] if ⇒ mp [if+anka+hala] [e]
mp [an+ampy] if ⇒ mp [if+an+ampy] [e]

It is true that no other nominal expressions incorporate as prefixes. But the same issue arises in the *if-*=Affix view as well, if it cannot be explained why *if-* is introduced by the affixation rule in (5) as a prefix.

Fourth, according to K&R, the *if-*=Anaphor view provides a less satisfactory account for why we do not get multiple reciprocals in Malagasy (*protect each other from each other*, etc.). They claimed that the absence of iterated *if-* is a theorem, as is the fact that a single *if-* does not license more than one gap. No *if-* can be introduced by rule

(5) to a verb already with *if-*, for the structural condition is not met. *If-* can only concatenate with an AF affix. A single *if-* does not license more than one gap, since the reciprocal verb is related by rule (6) to a transitive verb. Licensing more than one gap would require a rule relating a ditransitive verb to an intransitive verb.

But the lack of iterated *if-* and the failure of *if-* licensing more than one gap fall under the *if-*=Anaphor analysis as well. If *if-* originates as the Theme argument, and a verb has at most one Theme argument, then the lack of multiple *if*s follows directly. In addition, *if-* does not license more than one gap, for *if-* can only originate in one position.

Fifth, coordination raises an interesting issue for the analysis of reciprocal *if-*. In coordinated V (or VP), *if-* occurs in both conjuncts (K&R 2004: 204):

20. a. n+ifank+a+hita sy n+ifamp+i+arahaba izahay.
PAST+REC+ACT+see and PAST+REC+ACT+greet we.EXCL
'We saw each other and greeted each other.'

b.* n+ifank+a+hita sy niarahaba izahay.
PAST+REC+ACT+see and PAST+REC+ACT+greet we.EXCL

The problem here is that if it is possible for two Vs to conjoin and take one syntactic argument as in (21a), then there is no reason why reciprocal *if-* cannot incorporate into the conjoined verb as in (20b):

21. a. n+a+hita sy n+i+arahaba azy ireo aho.
PAST+ACT+see and PAST+ACT+greet 3ACC DEM+PL 1.SG.NOM
'I saw and greeted them.'

b. * [n+if+anka+hita sy n+i+arahaba] [e] izahay.

The account for the ungrammatical example in (21b) is quite straightforward for the *if-*=Affix view. The verbs here are of different arities, one is intransitive and the other is transitive. They can therefore not coordinate, if coordination requires that the conjuncts be of the same category.[3] However, on the *if-*=Anaphor view, the ungrammaticality of the example can also be explained by appealing to the plausible ban on movement into a coordinate structure (cf. Ross 1967).

The structure in (22a) is in principle possible on the *if-*=Anaphor view (K&R 2004: 205). Here, the Theme argument is a coordinate Theme argument, with the reciprocal allomorph *ifamp-* in the left conjunct and a full NP in the right conjunct:

[3] This is in fact not clear. The examples in (i) are perfectly fine:

i. a. John broke the glass and felt bad.
b. They put the books on the table and left.

22. a. [n+i+arahaba [ifamp- sy ny vadiny]] Rabe sy Rakoto.
PAST+ACT+greet each other and the spouse+3GEN Rabe and Rakoto
'Rabe and Rakoto greeted each other and their spouses.'

b.* n+ifamp+i+arahaba [[e] sy ny vadiny] Rabe sy Rakoto.

K&R argued that it is unclear how to block movement out of the left conjunct. By contrast, on the *if*-=Affix view, the structure in (22a) is impossible, because the derived predicate takes just one (plural) argument, *Rabe sy Rakoto*, whence nothing licenses the additional presence of *sy ny vadiny*. While their own account surely can explain the ungrammaticality of the example, the *if*-=Anaphor view can explain it too, by appealing to the general constraint on movement, namely, nothing may move out of a coordinate structure (Ross 1967). This is plausibly the same constraint that excludes the movement in (21b).

Sixth, K&R (2004: 205) argued that on the *if*-=Anaphor view, it is necessary to stipulate the attachment site for *if*-. For example, if *if*- may occur as a possessor, raise to object and incorporate into the verb (Keenan & Ralalaoherivony 2000) as in (23a), then why couldn't it just prefix to the head of the possessive *volo* 'hair'? Similarly, if causatives are treated as biclausal, then 'Rabe and Rakoto make each other dance' would, on the *if*-=Anaphor view, have the structure in (23b):

23. a. m+if+aN+intona (mifanintona) volo [e] [i Vao sy i Velo]
PRES+REC+ACT+pull hair ART Vao and ART Velo
'Vao and Velo are pulling each other's hair.'

b. [amp+ [[an+dihy] if]] [Rabe sy Rakoto]
CAUSE+ACT+dance REC Rabe and Rakoto
'Rabe and Rakoto make each other dance.'

They raised the question of what prevents *if*- from prefixing to the immediately preceding verb, yielding the ungrammatical and senseless **mifandihy*.

The question that K&R raised for the example in (23a) is likely to be related to the independent fact that possessive pronouns are in the genitive case (Keenan & Polinsky 1998) suffixed to the head noun, e.g., *volo-ny* 'his/her/their hair'. Reciprocal *if*- is clearly not a genitive pronoun; it is therefore expected that it may not suffix to the possessed noun.[4]

The problem in (23b) is understandable from the English perspective. As the structure in (24) shows, the reciprocal is the syntactic argument of the embedded verb:

[4] It remains unclear how the possessor raising construction discussed by Keenan & Ralalaoherivony (2000) is to be analyzed syntactically. While the possessive relation between the raised possessor and the head noun can be maintained if the possessor literally raises from inside the NP, it must be explained how it is possible for an argument in the genitive case to become accusative, whether syntactic raising is motivated elsewhere, what the motivation for it is, etc. Before these issues are resolved, the *if*-=Anaphor account need not assume that reciprocal *if*- may originate inside a possessive NP.

24. They make [each other dance].

The structural relation between reciprocal *if-* and the verb *andihy* 'dance' in (23b) is the same as that between *each other* and *dance* in (24). Reciprocal *if-* in (23b) cannot incorporate into the verb, since it would involve downward movement. **mifandihy* is therefore impossible.

The example in (25a) is of some interest, with the causative prefix *amp-* first combining with the verb root to which reciprocal *if-* is prefixed:

25. a. m+if+amp+aN+dihy (>mifampandihy) Rabe sy Rakoto.
PRES+REC+CAUSE+ACT+dance Rabe and Rakoto
'Rabe and Rakoto make each other dance.'

b. if [amp+aN+dihy] [e] Rabe sy Rakoto.

The structure in (25b) is consistent with the *if-*=Anaphor view with reciprocal *if-* incorporating into the verb from argument position. It is also consistent with the *if-*=Affix view, for *if-* prefixes to a causative verb, a transitive verb.

The facts considered above are largely compatible with both the *if-*=Affix view and *if-*=Anaphor view. The most compelling support for the former taking *if-* to be an affix on the verb comes from the semantic ambiguity in comparative ellipsis (see also Dalrymple, Mchombo & Peters 1994 for similar facts in Chichewa). I consider this case in detail below.

Comparative ellipsis may sometimes give rise to ambiguous interpretation of what is elided. As shown in (26a), the argument in the *noho*-phrase, the equivalent of English *than*-phrase, can be understood to be the object or subject of the comparative predicate (K&R 2004: 203):

26. a. m+aN+haja (manaja) an-dRabe kokoa Rasoa noho Ranaivo.
PRES+ACT+respect ACC-Rabe more Rasoa than Ranaivo
'Rasoa respects Rabe more than (she respects) Ranaivo.'
OR 'Rasoa respects Rabe more than Ranaivo (does).'

b. m+if+aN+haja (mifanaja) kokoa izy ireo noho Ravelo sy Ravao.
PRES+REC+ACT+respect more 3 DEM+PL than Ravelo and Ravao
'They respect each other more than Ravelo and Ravao respect each other.'
*'They respect each other more than (they respect) Ravelo and Ravao.'

However, when reciprocal *if-* occurs on the verb, the argument in the *noho*-phrase can only be understood to be the subject. K&R argued that the interpretive difference between (26a) and (26b) can be accounted for on the *if-*=Affix view, but not on the *if-*=Anaphor view. They did not elaborate on the argument, but it is not too difficult to see the point.

In the *if-*=Affix analysis, the reason why the second reading in (26b) is unavailable is because there is no object to compare. Recall that on this view, the reciprocal verb is intransitive and requires a plural subject. In the *if-*=Anaphor account, the second

reading should be available, because there the transitive verb *manaja* 'respects' has an object. I believe the conclusion is correct, but we need to consider a few facts to appreciate it.

Potsdam (2011) gave several arguments against deriving comparative ellipsis from deleting portions of a clausal structure containing the *noho*-phrase, and proposed a direct analysis in which the *noho*-phrase is simply a PP. As the lack of the second reading in (26b) is what distinguishes the *if*-=Affix view and the *if*-=Anaphor view,[5] we only need to look at examples in which the argument in the *noho*-phrase is understood to be the object of the comparative predicate. The example in (26a) has the alternative order in (27a), where the *noho*-phrase is related to the accusative argument; the order in (26a) is derived from movement of the *noho*-phrase to the right:

27. a. m+aN+haja (manaja) an-dRabe kokoa noho Ranaivo Rasoa.
PRES+ACT+respect ACC-Rabe more than Ranaivo Rasoa
'Rasoa respects Rabe more than (she respects) Ranaivo.'

b. m+aN+haja (mifanaja) if kokoa noho Ravelo sy Ravao izy ireo.
PRES+ACT+respect REC more than Ravelo and Ravao 3 DEM+PL
'They respect each other more than (they respect) Ravelo and Ravao.'

c. m+if+aN+haja (mifanaja) [e] kokoa izy ireo noho Ravelo sy Ravao.

It is therefore conceivable that, on the *if*-=Anaphor view, reciprocal *if*- originates in the same position as the accusative argument in (27a), as in (27b), and then incorporates into the verb, as in (27c). The *noho*-phrase subsequently moves to the right, yielding the surface word in (26a). There seems to be no reason why such derivation is impossible. If this is correct, then the second reading in (26b) should be available, contrary to fact.

On the *if*-=Affix view, reciprocal *if*- does not originate in argument position. The *noho*-phrase cannot be related to the object of the verb since there is no object, the verb *mifanaja* 'respect each other' being intransitive. It then follows that the argument in the *noho*-phrase cannot be understood to be the object of the verb.

Split plural argument seems to be problematic for both the *if*-=Anaphor view and the *if*-=Affix view. The problem can be seen in the examples in (28a-c) (K&R2004: 183; Rajaona 1972: 586):

28. a. Mifanenjika amin-dRabe Rakoto.
PRES+REC+ACT+chase with-Rabe Rakoto
'Rakoto is engaged in mutual chasing with Rabe.'

b. m+if+anka+hita (>mifankahita) t+amin-dRabe Rakoto.
PRES+REC+ACT+see PAST+with-Rabe Rakoto
'Rakoto and Rabe saw each other.'

[5] Potsdam did not account for the semantic ambiguity of the argument in *noho*-phrase in (26a). However, this has no effect in distinguishing the *if*-=Anaphor view and the *if*-=Affix view.

c. m+if+aN+tolotra (>mifanolotra) boky amin-dRabe Rakoto.
PRES+REC+ACT+offer book with-Rabe Rakoto
'Rakoto and Rabe offer books to each other.'

d. Rakoto no mifanenjika amin-dRabe.
Rakoto FOC PRES+REC+ACT+chase with-Rabe
'It's Rakoto that is engaged in mutual chasing with Rabe.'

The interpretations of examples (28a–c) indicate that there is a semantically plural argument. Example (28d) clearly shows that the argument *Rakoto* does not form a constituent with the *amin*-phrase. This effectively excludes the possibility that there is a syntactic plural argument. These facts are as problematic for the *if*-=Anaphor view as it is for the *if*-=Affix view. For the former, there is no argument that can be said to bind the reciprocal *-if*. For the latter, there is no plural subject for the reciprocal verb, contrary to rule (6).

4. Conclusions

If the foregoing discussions are correct, then there is some reason to suppose that there is only one reciprocal morpheme *if-* and that most of K&R's (2004) arguments against the *if*-=Anaphor view in fact do not stand. However, their *if*-=Affix view is decisively supported by facts concerning comparative ellipsis. Split plural argument with the *amin*-phrase is nevertheless problematic for both views. It would be of special interest to see how general the distribution of split plural argument, e.g., whether it can be related to accusative object, genitive complement or object of a preposition.

References

Abinal, Antoine & Malzac, Victorin. 1888. *Dictionnaire Malgache-français*. Tananarive: Imprimerie mission catholique.

Andrianierenana, Clément Luc. 1996. Morphological causatives in Malagasy. In Pearson, Matthew & Paul, Ileana (eds.) *The structure of Malagasy*, vol. I (UCLA Occasional Papers in Linguistics 17), 58–75. Los Angeles: Department of Linguistics, UCLA.

Aoun, Joseph & Sportiche, Dominique. 1983. On the formal theory of government. *Linguistic Review* 2. 211–236.

Baker, Mark. 1988. *Incorporation: A theory of grammatical function changing*. London and Chicago: The University of Chicago Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Cambridge, MA: MIT Press.

Dalrymple, Mary, & Mchombo, Sam & Peters, Stanley. 1994. Semantic similarities and syntactic contrasts between Chichewa and English reciprocals. *Linguistic Inquiry* 25. 145–163.

Keenan, Edward. 1976. Remarkable subjects in Malagasy. In Li, Charles & Thompson, Sandra (eds.) *Subject and topic*, 247–301. Oxford: Blackwell.

Keenan, Edward & Polinsky, Maria. 1998. Malagasy. In Spencer, Andrew & Zwicky, Arnold (eds.) *The handbook of morphology*, 563–624. Oxford: Blackwell.

Keenan, Edward & Ralalaoherivony, Baholisoa Simone. 2000. Raising from NP in Malagasy. *Lingvisticae Investigationes* 23. 1–44.

Keenan, Edward & Razafimamonji, Jean-Paulin. 2004. Reciprocals in Malagasy. *Oceanic Linguistics* 43. 177–207.

Pearson, Matthew. 2005. The Malagasy subject-topic as an A' element. *Natural Language and Linguistic Theory* 23(2). 381–457.

Postal, Paul. 1962. *Some syntactic rules of Mohawk*. New Haven: Yale University. (Doctoral dissertation)

Potsdam, Eric. 2011. A direct analysis of Malagasy phrasal comparatives. In Clemens, Lauren Eby & Scontras, Gregory & Polinsky, Maria (eds.) *The proceedings of the 18th Annual Meeting of the Austronesian Formal Linguistics Association*, 140–155. Cambridge, MA: Harvard University.

Rajaona, Siméon. 1972. *Structure du malgache*. Fianarantsoa: Ambozontany.

Ross, John Robert. 1967. *Constraints on variables in syntax*. Cambridge, MA: MIT. (Doctoral dissertation)

Paul Law
City University of Hong Kong
paul.law@cityu.edu.hk

Why did Erromangan wind names turn 90 degrees?

John Lynch

Abstract

The non-Polynesian languages of Southern Vanuatu have borrowed much of their maritime terminology from the neighbouring Polynesian Outlier Futuna-Aniwa. Included in this terminology are terms for wind directions. Although Anejom̃ and the Tanna languages fairly faithfully reflect the semantics of the Futuna terms, the Erromangan languages have made a shift such that most terms are 90 degrees further clockwise from the Futuna original (thus original 'north wind' means 'east wind', etc.). I attempt to explain this as due to imperfect learning.

1. Introduction[1]

The non-Polynesian languages of Southern Vanuatu have always had the reputation of being a bit "odd". More than half a century ago, Arthur Capell said of them that "the languages of Eromanga [sic], Tanna, and Aneityum diverge most of all from the rest of the New Hebrides [now Vanuatu], while not agreeing among themselves" (1962: 383). The titles of two papers by Terry Crowley (2000b; 2001) – "How did Erromangan verbs get so messy?" and "What happened to Erromangan possessive morphology?" – also evoke the notion of "some funny things going on".

In this short paper, I look at one other "funny" thing: the names of winds in the languages of Erromango. As in the languages of Tanna and Aneityum, these have been borrowed from the nearby Polynesian Outlier Futuna-Aniwa (Lynch 2001: 190–193). Unlike the languages of neighbouring islands, however, the Erromangan languages have made significant changes to the semantics of many of these wind terms, changing the direction of most of them 90 degrees.

2. Polynesian borrowings in Southern Vanuatu languages

The Southern Vanuatu (SV) subgroup of Southern Oceanic consists of (i) Anejom̃, the sole language of Aneityum; (ii) the five languages of Tanna; and (iii) all the languages of Erromango – Sye (also known as Erromangan), the moribund Ura, and probably three now extinct languages. Located not far to the east of Tanna are the islands of Futuna and Aniwa, home to a Polynesian Outlier language.

[1] It is a pleasure to dedicate this paper to Liz Pearce, who has made significant contributions to the study of Vanuatu's languages. I am grateful to Stuart Bedford, Piet Lincoln, and Jeff Marck for comments on an earlier version of this paper.

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 129–137

Quite a significant number of borrowings have taken place from a Polynesian language – almost certainly Futuna-Aniwa – into the languages of Southern Vanuatu. Lynch (1994) documents a number of these, especially in the semantic domains of maritime terminology, terms relating to kava-drinking and -preparation, and names of various other cultural items and activities (see also Lynch 2001: 190–195).

Identification of Polynesian borrowings is generally quite simple. In their development from Proto-Oceanic and Proto-Southern Oceanic,[2] SV languages have regularly deleted vowels in various environments, mainly word-finally, and word-medially when unstressed. Certain consonants, too, show some unusual developments, like *l,*r > Lenakel *i* / _ *a,*u (see *bulut in 1), or *p > Anejom̃ *h-, -h-, -Ø* (see *lipon-gu in 1). Forms like those in (1) are regular inheritances, showing regular phonological developments; a slash separates non-cognate material (accretions and the like).[3]

1.	**Protoform**	**Sye**	**Lenakel**	**Anejom̃**
	POc *lipon-gu 'my tooth'	*ne/lve-ŋ*	*ne/lu-k*	*ne/jhe-k*
	POc *bulut 'sticky'	*a/mplet*	*a/p^{w}iit*	*a/p^{w}ol*
	POc *kasupe 'rat'	*na/kih*	*kahau*	*in/ɣeʊ̯o*
	POc *ta-m^{w}aqane 'man'	*na/tman*	*ie/ram^{w}aan*	*na/tam^{w}añ*

Now compare forms like those in (1) with the ones in (2), which are Polynesian loans, where the last column contains predictions as to what these forms would be if they were directly inherited:

2.	**Protoform**	**Futuna**	**S. Vanuatu**	**[Expected if inherited]**
	POc *malino '(sea) calm'	*marino*	LEN *a/melinu*	*[ə/mlin* or *a/mlin]*
	POc *kiajo 'outrigger float'	*kiato*	ANJ *na/kiato*	*[na/ɣet* or *na/ɣyet]*
	POc *panaq 'bow'	*fana*	SYE *ne/vane*	*[ne/ven]*
	PPN *mako 'dance'	*mako*	ANJ *na/mako* (n.)	*[in/maɣ]*

In comparing these sets of forms, note that those in (2) show no vowel loss: the root undergoes no structural change. Other regular changes – like *k > ANJ *ɣ* – have also not taken place. Polynesian loans in these languages, then, are usually pretty obvious.

To give some idea of the scope of borrowing in the area of maritime terminology, in addition to the wind names that form the subject of this paper, the following are some reconstructed Polynesian terms that have been borrowed through Futuna into one or

[2] Southern Vanuatu is a subgroup of Southern Oceanic, whose other members are the non-Polynesian languages of (i) the remainder of Vanuatu and (ii) New Caledonia. There is no published body of Southern Oceanic reconstructions, but Proto-North-Central Vanuatu reconstructions (Clark 2009) would be identical or very similar.

[3] I do not use the standard orthography but rather a phonemic orthography. The following language abbreviations need explanation. (1) Modern languages: ANJ, Anejom̃; FUT, (West) Futuna-Aniwa; KWM, Kwamera; LEN, Lenakel; NTN, North Tanna; SWT, Southwest Tanna; SYE, Sye (also Erromangan); URA, Ura; WSN, Whitesands. (2) Protolanguages: PCEPN, Proto-Central-Eastern Polynesian; PNCV, Proto-North-Central Vanuatu; POC, Proto-Oceanic; PPN, Proto-Polynesian; PSO, Proto-Samoic-Outlier. Data sources are listed in the appendix.

more SV languages (Lynch 1994):[4] *faqi-ava 'bay', *malino '(sea) calm', *peau 'a wave', *tafo-laqa 'whale', *sasawe 'flying fish', *moa-moa 'boxfish', *palaŋi 'surgeonfish', *malau 'soldierfish, squirrelfish', *tapatuu 'barracuda', *pusi 'moray eel', *tila 'mast', *kiato 'outrigger-boom', *mataqu 'fish-hook'.

3. Wind terms

POc terms for winds were associated with the two major seasons, *raki for the dry season with southeast trades, and *apaRat for the wet season with the northwest monsoon; and these terms may have referred in a portmanteau-like fashion to the season, the nature of the wind, and cardinal directions (Ross 2003: 127). Ross goes on to say that "whereas the POc terms evidently referred prototypically to seasonal winds, the central meanings of the PPn terms seem to have been winds from a certain portion – apparently a quadrant – of the compass" (2003: 131). Nevertheless, there has been shifting of the referents of wind terms in different Polynesian languages, as will be mentioned below.

Almost all the terms for winds in all SV languages are obvious Polynesian loans, and bear a remarkable similarity to the names of the winds in Futuna-Aniwa. The term for 'wind', or a certain kind of wind, derives clearly from PPn *mataŋi and not from POc *aŋin or *jaŋi or from PNCV (or Proto-Southern Oceanic) *laŋi:

3. PPn *mataŋi 'wind' > Sye *ne/metaŋi* 'cyclone', NTn *metaŋ*, Wsn *nə/mətaŋi*, Len Swt *nə/mataaŋ*, Kwm *nə/mataŋi*, Anj *ne/mtañ/jap*^{w} 'wind from a particular direction' (*injap*^{w} = 'sea')

None of the following PNCV or POc terms for winds from particular directions are reflected in SV languages:

4.	**PNCV**	**POc**
'northwest'		*aparat
'south or east'	*dualiu	*timu(R)
'southeast'		*raki, *karak(a), *marau, *aqura
'southerly'		*yawana
'southwest'	*kadua	

The only high-level term that appears to be retained in these languages is POc *tokalau(r), PNCV *tokalau/*tokolau 'northerly wind'. However, it is clear that the Erromangan and Anejom̃ terms at least – and thus by implication the Tanna forms as well – are not directly inherited but borrowed from Futuna. POc *l and *r merge in Tanna (Kwamera *r*, others *l*), but remain distinct in other SV languages: *l > Anejom̃, Sye *l* (Anejom̃ *j* in some environments), *r > Anejom̃, Sye *r* (Anejom̃ final Ø). The fact that reflexes of POc *tokalau(r) – Anejom̃ *na/tokorau*, Sye *na/toɣrau*—have a medial *r*

[4] Some of these terms differ formally from those give in Lynch (1994), due to updates in POLLEX.

rather than *l* suggests that they are loans from Futuna, where POC, PPN *l regularly became *r* (see reflexes of *luatuqu and *paa-lapu in Table 1).

Erromangan wind names are given in Table 1, which does not include all wind names in all languages, but only all recorded Erromangan wind names, their sources in Futuna-Aniwa, and forms in other SV languages that have been borrowed from the same source; these forms are *formally* related but not necessarily *semantically* identical to one another. Table 1 also includes the corresponding forms in Proto-Oceanic, Proto-Polynesian, or some lower-level Polynesian protolanguage. Terms are listed clockwise from north based on the Futuna system. Names are followed by directions in parentheses.

Table 1: Erromangan wind terms

Protoforms	**Futuna**	**Anejom̃**	**Tanna**	**Erromango**
PSO *luatuqu 'wind direction (northeast?)'	*ruetu* (N)	*narutu* (due N)	LEN *luatu* (N) SWT *luatu* (NE) KWM *ruatu* (N)	SYE URA *norwotu* (E)
	retuamlai (NE)	*narutu-amlai* (NE)	LEN *luatuamlaai* (NE) SWT *luatuamlaai* (N) KWM *ruatu amrai* (NE)	SYE *norwotamlai* (ESE)
PPN *toŋa 'southeasterly quadrant, southeast wind'	*etoŋa* (SE)	*natooŋa a-nwai* (S of due E)†	LEN SWT KWM *natoŋa* (E)	SYE URA *natuŋa* (S)
PCEPN *uru 'wind from a westerly quarter'	*uritoŋa / ruitoŋa / ritoŋa* (due S)	*nauritooŋa* (due S)	LEN *uritoŋa* (SE) KWM *uritoŋa* (S) KWM *uritoŋa natoŋa* (SE)	SYE URA *nourituŋo* (W)
POC *tokalau(r) 'northerly wind' PPN *tokelau/ *tokolau 'northwesterly quadrant, northwest wind'	*tokorau* (just S of W)	*natokorau* (WNW)	LEN *tokolau*(S) SWT *tokolau* (SE) KWM *tak^{w}arau* (SSE)	SYE *natoɣrau* (SE)
PSO *paa-lapu 'westerly wind'	*parapu* (W)	—‡	LEN *nəp^{w}elaap^{w}* (S) SWT *nəpelaap* (W) KWM *nəparapu* (W)	SYE *nomporavu* (N) URA *noboravu* (N)

† Anejom̃ *a-nwai* = LOC-water/river.

‡ Anejom̃ has two terms for westerly winds that are not loans: *netoranm^{w}al* (almost W) and *neðinaej* (W by SW).

Figure 1 presents these wind names in rough schematic form. The top of each cell is headed by the Futuna-Aniwa term, and I will use that term in referring to the set of borrowings in other languages; thus, when I say something like "in the case of *tokorau* …", I mean "in the case of Futuna *tokorau* and all forms borrowed from it in SV languages". In each cell, arrows show the directions of winds marked by the term in Futuna and in a Polynesian protolanguage, as well as the directions of winds marked by terms derived from this Futuna term in all languages in the table. (Note that 'Erro' is

used to represent 'Erromangan languages' and 'Tanna' to represent 'all or most Tanna languages'.)

There have been some changes of direction in Futuna, in comparison with earlier protolanguages. Two terms have unchanged meanings – *toŋa > *etoŋa* (SE) and *paa-lapu > *parapu* (W) – while a third, *luatuqu (NE?) > *ruetu* (N) shows only a minor semantic shift. The case of PCEPN *uru 'wind from a westerly quarter' is a bit more tricky. This does not occur by itself in Futuna: rather, it compounds with *toŋa as *uritoŋa* ~ *ruitoŋa* (due S), possibly originally as a three-term compound *uru + qi (LOC) + toŋa ('west wind in/from the south'?), with subsequent loss of the second *u in *uru. But consider the following compounds (not in Table 1) that have a westerly suggestion to them: *ruitoga tane* (SSW), where *tane* means 'male', and *ruitoga fine* (WSW), where *fine* means 'female'.

The final term is POC *tokalau(r) 'northerly wind' > PPN *tokelau/*tokolau 'northwesterly quadrant, northwest wind' > Futuna *tokorau* (just S of W). This is a major shift; but as can be seen in Figure 1, there are other major shifts associated with this term, which has the most messy semantics of all. Its meaning shifted from northerly in POC to southwesterly in Futuna, and there were further changes in the SV languages: west-northwest in Anejom̃, south / south-southeast / southeast in Tanna, and southeast in Erromango.

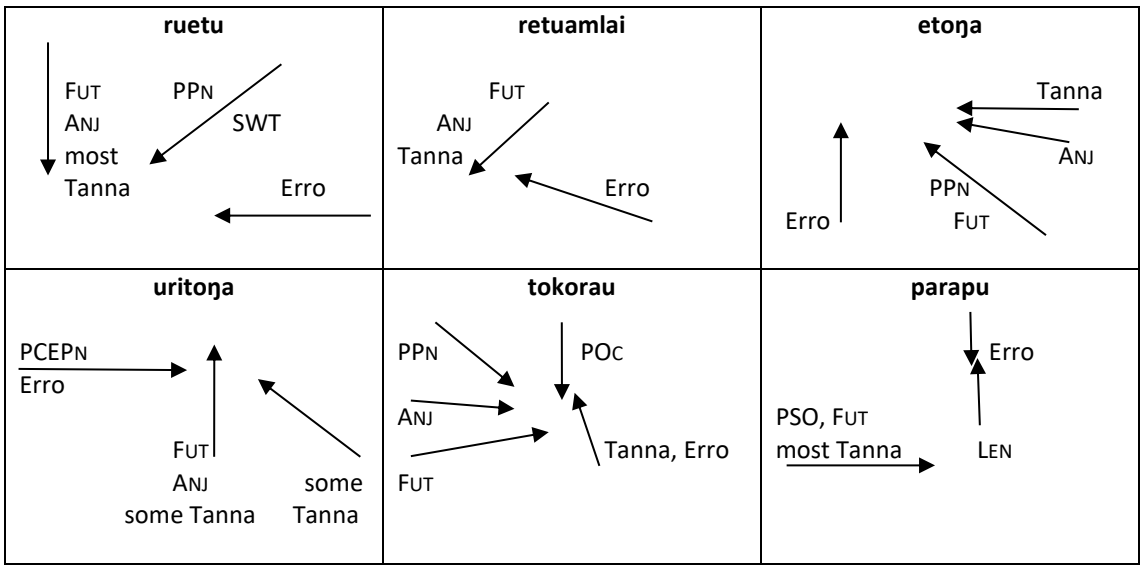


Figure 1: Schematic diagram of wind directions

There are a number of features to note in relation to Erromangan wind terms:

1. With the exception of *etoŋa*, each Erromangan wind is at 90 degrees to the Futuna wind; in the case of *etoŋa*, it is more like 45 degrees. But in no case do a Futuna term and the Erromangan term borrowed from it share the same direction.
2. In all six cases, it is as if the Erromangan wind has turned 90° (or 45° in the case of *etoŋa*) *clockwise* to the Futuna wind.
3. Except in the case of *tokorau*, no other SV language has made the same change.
4. Except in the case of *tokorau*, the direction implied by each Erromangan wind term is also significantly different – i.e., more than 45°, and often 90° or more – from that implied by the same borrowed term in Anejom̃ or the Tanna languages. On the other hand, in most cases, the wind directions of the Futuna, Tanna and Anejom̃ terms are the same, or are similar enough – within 45° of each other. The only exception is *parapu*, where Lenakel is at 90° to the others.

In discussing in my 1994 paper the pattern of borrowing of Polynesian terms into Southern Vanuatu languages, I had the following to say:

> One possible explanation for the number of Polynesian borrowings in this area of maritime terminology is that, having arrived by sea, speakers of Proto Southern Vanuatu turned their attention to the reef and the land, and did not exploit the deep sea to nearly the same extent as they did in more recent times. That is, it is possible that they treated the sea as if it were a river – fishing from the beach, exploiting the reef, but not venturing much beyond.
> …
> It would appear, therefore, that having pretty much abandoned the use of canoes and the exploitation of maritime resources beyond the reef for some considerable time, the people of Southern Vanuatu were subsequently reintroduced to this technology by Polynesians from neighbouring Futuna and Aniwa, from whom they derived much of their modern-day maritime vocabulary (Lynch 1994: 298–299).

The ancestors of modern Futunese and Aniwans presumably taught their non-Polynesian neighbours how to use the open sea (again), and passed on to them the relevant terminology, including the names of the wind directions. If we ignore *tokorau*, where all sorts of changes have taken place, and if we also ignore one or two minor aberrations – like Futuna *parapu* 'west wind' being borrowed into most Tanna languages with the same meaning but into Lenakel with the meaning 'south wind' – then it is fair to say that the Anejom̃ and Tanna terms have basically the same meaning as the Futuna terms from which they derive: at least, they refer to winds in the same quadrant. The teaching and learning process was successful, suggesting prolonged and/or sustained contact over some period of time, in which Tannese and

Aneityumese learned sailing and fishing skills from Futunese sailors, and with these skills the terms for anything new to them.[5]

The Erromangan terms, however, *consistently* refer to winds from the next quadrant (moving in a clockwise direction). This would suggest that, while the transmission of this aspect of maritime knowledge from Futunese to Tannese and Aneityumese was direct and accurate, something else happened with Erromangans. When the Futuna wind system was learned by the ancestors of today's Erromangans, it was not just a matter of them mis-hearing, or mis-identifying, one term in the system – something that can easily happen in borrowing situations. It was, rather, a matter of getting the whole system wrong, by 90 degrees.

4. So what happened?

Let us first deal with the phonology of the borrowed wind terms. With the exception of some forms deriving from *parapu*, wind names in Tanna and Aneityum basically reflect Futuna forms unchanged (though some take the fused article, which derives from POc *na, and there is also the *r > *l* change in some Tanna languages discussed above). In Erromango, however, there has been quite a bit of phonological change. Compare the Sye and Lenakel forms below, with unnecessary changes underlined.[6]

5.

Futuna	**Sye**	**Lenakel**
ruetu	*no/rwotu*	*luatu*
toŋa	*na/tuŋa*	*na/toŋa*
uritoŋa	*no/rituŋo*	*uritoŋa*
tokorau	*na/toɣ_rau*	*tokolau*
parapu	*no/mporavu*	*nə/p^{w}elaap^{w}_*

A possible explanation of the Sye and Ura forms is imperfect learning. All of the Futuna terms in (5) are possible Erromangan words, with no phonological changes necessary. They have just been mispronounced. In comparison with the situation in which the Tannese and Aneityumese learned the wind system, it suggests that the Erromangans' learning of it was either not very thorough, or took place through some third party who didn't really speak Futunese well, or was flawed in some other respect. And in getting the pronunciation wrong, they also got the semantics wrong.

Stuart Bedford (p.c.) makes an interesting point in this regard: "In terms of island layout, Erromango is an outlier when it comes to Futuna and Aniwa." All of the other islands were also part of an interaction sphere requiring sailing with the Loyalties,[7] but

[5] There are currently two Futunese villages on Tanna and one on Aneityum. My understanding is that they are relatively recent settlements—within the last hundred years or so. But there may well have been earlier such settlements, which would have facilitated this thorough transfer of knowledge and terminology.

[6] An underlined blank space indicates loss of a vowel.

[7] However, Fagauvea, the Polynesian Outlier spoken in the Loyalties, does not seem to have been the source of wind terms in the SV languages. The Fagauvea terms *luetuu* 'north (wind)'

Erromango not so much. Erromango, he suggests, simply didn't undergo the same influences as Tanna and Aneityum, or at least underwent only insipid versions of them. It would follow from this suggestion that, while the transfer of skills, and by implication vocabulary, from Polynesians to Melanesians in Tanna and Aneityum may have been reasonably intensive, it was much less so when Erromangans were involved, due to much less intense contact. This would help to explain (i) why Erromangan languages show so many phonological shifts in these terms, and (ii) why the referents are incorrect.

Exactly *why* these referents were consistently wrong by 90 degrees is not easy to explain, however. In the modern world, I can envisage a teacher sketching the wind directions on a piece of paper, and the learner accidentally turning the paper the wrong way and getting everything twisted 90 degrees. That is unlikely to have happened hundreds of years ago. But it is possible that, once one term was mis-learned, the whole system collapsed like dominoes. Once the early Erromangans mis-learned, say, *ruetu* as 'east wind' rather than 'north wind', for whatever reason, all the other errors followed.

Appendix: Data sources

Reconstructed forms:	Clark (2009); Greenhill & Clark (Ongoing); Ross (2003)
Futuna:	Dougherty (1983)
Fagauvea:	Hollyman (1987)
SV languages generally:	Lynch (1994; 2001)
Anejom̃:	Lynch & Tepahae (2001)
Tanna languages:	Lindstrom (1986); Lynch (1977; 1982)
Erromangan languages:	Crowley (1999; 2000a)
Polynesian languages generally	Pyrek & Feinberg (2016)

References

Capell, Arthur. 1962. Oceanic linguistics today. *Current Anthropology* 3. 371–428.

Clark, Ross. 2009. **Leo tuai: A comparative study of North and Central Vanuatu languages*. Canberra: Pacific Linguistics.

Crowley, Terry. 1999. *Ura: A disappearing language of southern Vanuatu*. Canberra: Pacific Linguistics.

Crowley, Terry. 2000a. *An Erromangan (Sye) dictionary*. Canberra: Pacific Linguistics.

Crowley, Terry. 2000b. How did Erromangan verbs get so messy? In Palmer, Bill, & Geraghty, Paul (eds.) *SICOL: Proceedings of the Second International Conference on Oceanic Linguistics*, vol. 2: *Historical and descriptive studies,* 185–199. Canberra: Pacific Linguistics.

---

and *toŋa* 'east (wind)' are of Polynesian origin, but *hwadrœ* and *gikoo* 'east (wind)', *hwaiœu* 'south (wind)', and *imuli* 'west (wind)' are almost certainly not.

Crowley, Terry. 2001. What happened to Erromangan possessive morphology? In Bradshaw, Joel, & Rehg, Kenneth L. (eds.) *Issues in Austronesian morphology: A focusschrift for Byron W. Bender*, 87–104. Canberra: Pacific Linguistics.

Dougherty, Janet W. D. 1983. *West Futuna-Aniwa: An introduction to a Polynesian Outlier language*. Berkeley: University of California Press.

Greenhill, Simon J., & Clark, Ross. Ongoing. POLLEX-Online: The Polynesian lexicon project online. (https://pollex.shh.mpg.de/)

Hollyman, K. J. 1987. *De muna fagauvea I: Dictionnaire fagauvea–français.* Auckland: Linguistic Society of New Zealand.

Lindstrom, Lamont. 1986. *Kwamera dictionary / N̶i̶kukua sai nagkiariien N̶i̶ni̶n̶i̶fe*. Canberra: Pacific Linguistics.

Lynch, John. 1977. *Lenakel dictionary*. Canberra: Pacific Linguistics.

Lynch, John. 1982. South-west Tanna grammar outline and vocabulary. *Pacific Linguistics*, A-64, 1–91.

Lynch, John. 1994. Melanesian sailors on a Polynesian sea: Maritime vocabulary in Southern Vanuatu. In Pawley, Andrew K. & Ross, Malcolm D. (eds.) *Austronesian terminologies: Continuity and change*, 289–300. Canberra: Pacific Linguistics.

Lynch, John. 2001. *The linguistic history of Southern Vanuatu.* Canberra: Pacific Linguistics.

Lynch, John, & Tepahae, Philip. 2001. *Anejom̃ dictionary / Diksonari blong Anejom̃ / Nitasviitai a nijitas antas Anejom̃.* Canberra: Pacific Linguistics.

Pyrek, Cathleen Conboy, & Feinberg, Richard. 2016. The Vaeakau-Taumako wind compass as part of a "navigational toolkit". *Structure and Dynamics* 9(1). 41–69.

Ross, Malcolm. 2003. Meteorological phenomena. In Ross, Malcolm & Pawley, Andrew & Osmond, Meredith (eds.) *The lexicon of Proto Oceanic*, vol. 2: *The physical environment*, 115–147. Canberra: Pacific Linguistics.

John Lynch
University of the South Pacific
johnlynch123@gmail.com

The curious case of preverbal *ko* in Niuean[1]

Lisa Matthewson[2], Heidi Quinn[3], Diane Massam[4], and Lynsey Talagi

1. Introduction

Many Oceanic languages have cognates of an element *ko,* which according to Pollex Online (http://pollex.org.nz, Greenhill & Clark 2011) may descend from two possible sources. **Ko.1* is reconstructed to the Fijic group as a preposition which marked 'topic, nominal predicate and other functions; specifier', and **ko.2* is reconstructed to Samoic-Outlier Polynesian as a marker of progressive aspect. Pollex Online lists 26 languages with modern reflexes of **ko.1*, while **ko.2* is listed as appearing only in Samoan and Pukapuka.

Niuean (Tongic) has an element *ko* which seems to perform both an information-structuring function (perhaps similar to Pollex Online's **ko.1*), and a temporal or aspectual function (perhaps similar to **ko.2*). However, Niuean *ko* poses both syntactic and semantic puzzles. Massam et al. (2006) document no fewer than ten different uses of *ko*, and although they provide a uniform analysis of it (as an expletive case-marking preposition), it is still unclear whether it might be possible (or desirable) to provide a unified (non-expletive) semantic analysis for *ko*. Examples of some of the major uses of *ko* are given in (1–6).[5]

---

[1] Liz Pearce has contributed greatly to our work on Niuean and to much else besides. In many ways, it is thanks to Liz that we started our collaborative research on Niuean, because Liz and her PhD student Laura Dimock provided valuable assistance for reviving the LING407 honours field methods course at the University of Canterbury, which eventually brought all of us together.

[2] Liz was one of my earliest and best linguistics teachers; she taught me generative grammar and syntactic argumentation, and has given me unfailing support and friendship ever since.

[3] Liz has been a wonderful mentor and friend since I first met her at the Australian Linguistics Institute in 1998. Her passion for syntactic analysis and Pacific languages has been a constant source of inspiration for my own research and teaching, and I am so very grateful for all her encouragement and support throughout the years.

[4] Liz showed overwhelming hospitality to me and my family on my first trip to New Zealand, making it possible for me to start my fieldwork on Vagahau Niue, and she has been a great friend and colleague ever since.

[5] Abbreviations used in this paper are: ABS: absolutive; C: common; DEM: demonstrative; DIR: direction; DU: dual; EMPH: emphatic; ERG: ergative; EX: exclusive; GEN: genitive; INCL: inclusive; NEG: negation; NFT: nonfuture; NSP: nonspecific; PL: plural; PRES: present; PRO: locative/temporal resumptive pronoun; PST: past; SBJ: subjunctive; SG: singular. We have provided uniform glosses, rather than following those of the data sources. Note we gloss *ko* simply as KO, since its meaning is the subject of this paper. Note also that case markers and prepositions vary

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 139–157

1. **Ko** Pita ne fano ki Niu Silani.[6]
   KO Pita NFT go to New Zealand
   'It's Pita who went to New Zealand.' (Seiter 1980: 99) CLEFT

2. **Ko** hai ne fifili a koe ke vagahau?
   KO who NFT choose ABS you SBJ speak
   'Who chose you to speak?' (Seiter 1980: 109) WH-QUESTION

3. **Ko** Pule e faiaoga.[7]
   KO Pule ABS teacher
   'The teacher is Pule.' (Seiter 1980: 54) EQUATIONAL

4. **Ko** e[8] kamuta a au.
   KO C carpenter ABS 1.SG
   'I am a carpenter.' (Seiter 1980: 53) NOMINAL PREDICATE

5. **Ko** e matua fifine haana, mate tuai.
   KO C parent female GEN.3.SG die PERFECT
   'As for his mother, she's dead.' (Seiter 1980: 116) TOPIC

6. **Ko** e kai a mautolu he talo.
   KO C eat ABS 1.PL.EX at taro
   'We are (now) eating taro.' (Seiter 1980: 6) PREVERBAL

In this squib, we provide a preliminary investigation of the last usage of *ko* in (6), in which *ko* appears immediately preceding a verbal predicate. The predicate is obligatorily introduced by the common article *e,* as shown by the unacceptability of (6').[9]

---

depending on whether the following noun is proper or common, but we do not include this in the glosses, for simplicity.

[6] Compare (1) with the non-clefted version in (i):

(i) Ne fano a Pita ki Niu Silani.
PST go ABS Pita to NewZealand
'Pita went to New Zealand.'

[7] Pearce (1998: 254f) gives cases parallel to (3) in Māori. Interestingly, Māori does not seem to use *ko* for cases parallel to (4) (Bauer 1997; Pearce 1998).

[8] We gloss the *e* following *ko* here as C: 'common'. Since it appears before verbs as well as before common nouns, it is more properly considered 'non-proper', but its full function and meaning remain open to debate. It is arguably not absolutive in this context, as it does not contrast here with the proper absolutive marker *a*. Before proper nouns, *ko* appears bare, as in (1). See Massam (2016) for further discussion.

[9] As mentioned in Note 8, in its nominal uses, *ko* takes *e* only when the noun is non-proper, as seen for example in the contrast between (1) above and (i).

(i) **Ko** e tama haau ne kumi ai a maua.
KO ABS child GEN.2.SG NFT search PRO ABS 1.DU.EX
'It's [your child]$_{FOC}$ we're looking for.' (Seiter 1980: 101)

6'. ***Ko*** kai a mautolu he talo.
KO eat ABS 1.PL.EX at taro
Intended: 'We are (now) eating taro.'

This preverbal usage of *ko* has received little attention in prior literature. Our modest goal here is to present some empirical generalizations about its distribution and function, and to speculate about potential lines of analysis.

The data and generalizations given in this paper are based primarily on original fieldwork with the fourth author, who is a native speaker of Niuean from the village of Lakepa. All unreferenced data represent the fourth author's judgments.

2. Background on *ko*

2.1 Syntax

In its preverbal use, the sequence *ko e* appears – at least at first glance – to be patterning like a TAM [Tense-Aspect-Modal] marker. TAM markers in Niuean appear clause-initially (Seiter 1980: 2; Massam 2009: 1), as illustrated in (7–8). Further, it is generally the case that *ko e* is in complementary distribution with (other) TAM markers, as for example in (9).

7. **Ne** nofo a au i Makefu.
PST live ABS 1.SG in Makefu
'I lived in Makefu.' (Massam 2009: 1, citing Seiter 1980: 4)

8. **To** nākai liu feleveia foki a taua.
FUT not return meet also ABS 1.DU.INCL
'We will never again meet.' (Massam 2009: 1, citing Seiter 1980: 16)

9. ***Ne ko e** huva a John.
PST KO C vacuum ABS John
Intended: 'John was vacuuming.'

However, it has been pointed out that preverbal *ko* differs from other TAM marking in Niuean in that it can follow negation. Thus, *ko e* in (10) contrasts with the future marker *to* in (8) in its position with respect to negation (note that there are two forms of negation in Niuean: *nākai* and *ai;* see Seiter 1980: 14f).[10]

10. **Ai ko e** onoono (a) au ke he ha mena.
NEG KO C look (ABS) 1.SG at NSP thing
'I'm not looking at anything.' (Seiter 1980: 82; checked with our fourth author)

We return to the issue of *ko*'s interaction with negation in Section 6.1 below.

[10] Seiter writes that (10) is good for 'a couple of [his] youngest consultants'.

Massam et al. (2006) provide a uniform analysis of *ko* across all its uses as a case-marking preposition used with non-argument nominals. Semantically, they consider it to be an expletive. However, the pre-verbal use of *ko* is problematic for their analysis, since it leads to the claim that predicates following *ko* are actually nominalized (Clark 1976), and are combined with a (phonologically null) light verb, although there are no independent arguments for this analysis. This opens up the possibility that *ko* is not an expletive case marker, but that semantically, it serves some other purpose, for example, focus. We turn now to some semantic analyses that have been proposed for *ko*, and we return to syntactic issues in Section 6.

2.2 Semantics

It has been observed for many Oceanic languages that cognates of *ko* can convey **focus**. Under this umbrella we can certainly fit the uses in (1) and (2) above, and probably those in (3) and (4) as well. For discussion of the focus semantics of *ko* and its cognates, see Bauer (1991) and Pearce (1999) on *ko* in Māori; Hohaus & Howell (2015) on *ʻo* in Samoan; and Brown & Koch (2013) on Polynesian generally.[11]

Ko is also frequently used to mark **topics,** as in (5) above, and as discussed for Māori by Bauer (1991) and Pearce (1999). Both Brown & Koch (2013) and Hohaus & Howell (2015) propose that the topic uses of *ko*-cognates can be semantically assimilated to the focus uses, in that they both rely on sets of alternatives (Rooth 1985; 1992; Roberts 1996; Büring 1997; 2003; Krifka 2007).

The preverbal use of *ko* which is our concern in this squib has not so far received dedicated attention in the formal literature. According to Seiter (1980), the use of *ko e* before verbs conveys ‘actual present’. Further examples are given in (11–12). We use Seiter’s gloss for *ko e* in (11–14), to emphasize his analysis of the complex *ko e* as a present tense marker.

11. **Ko e** tohitohi a au mogonei aki e pene fōu.
    PRES write ABS 1.SG now with ABS pen new
    I'm writing at the moment with a new pen.’ (Seiter 1980: 5)

12. **Ko e** kumi agaia au he tama haau.
    PRES seek still 1.SG at child GEN.2.SG
    ‘I'm still looking for your child.’ (Seiter 1980: 5)

Seiter observes that this use of *ko e* is ‘not explicitly progressive, since it may introduce stative verbs’ (1980: 5):

[11] According to Pollex Online, reflexes of **ko.1* in Kapingamarangi, Nukuoro and West Uvea take the form *go*. Interestingly, we also find a focus marker of the form *go* in the rather more distantly related Central Vanuatu language Unua (cf. Pearce 2015). Like Niuean *ko*, Unua *go* can mark verbs as well as nominal elements, but Unua *go* follows rather than precedes the elements it marks (cf. Pearce 2015: 350-353). It would be interesting to explore whether the semantic properties of postverbal *go* in Unua resemble the semantics of preverbal *ko* in Niuean.

13. **Ko e** tokoluga koa kia e tagata ne kitia e au he hola.
    PRES tall EMPH Q ABS man NFT see ERG 1.SG at run
    'The man I saw running away is really tall.' (Seiter 1980: 5)

He further observes that '[u]nlike sentences which bear no tense/aspect marker, ones marked with *ko e* cannot express a habitual or characteristic situation.' The example in (14), repeated from (6), is thus said not to be able to mean 'We eat taro':

14. **Ko e** kai a mautolu he talo.
    PRES eat ABS 1.PL.EX at taro
    'We are (now) eating taro.' (Seiter 1980: 6)

Seiter's proposal that preverbal *ko e* conveys present tense semantics is challenged by Massam et al. (2006), who point out that the use of *ko* with verbal predicates does not always give rise to a present tense interpretation. They provide data such as in (15–16) in support of this.[12]

15. ka fakatatai atu pehē ka ha, **ko** e hau au he tau [19XX].
    but similar DIR2 like that=is, **KO** C come 1.SG in year [19XX].
    'say...like..., I came in 19XX here...' (Massam et al. 2006: 14, citing data from the Languages of Manukau Project provided by Donna Starks)

16. Ko e eke hā a mua neafi?
    KO C do what ABS 2.DU yesterday
    'What did you do yesterday?'

    Ko e ta kiilikiki a maua.[13]
    KO C play cricket ABS 1.DU.EX
    'We played cricket.' (Massam et al. 2006: 14, citing Kaulima & Beaumont 2002: 42)

Massam et al. conclude that *ko* 'cannot be indicative of actual present or present progressive.' Since their purpose is not to provide a semantic analysis of this usage of *ko*, they do not investigate further, noting merely that *ko* 'presumably provides some sort of tense or aspectual meaning to the clause, with a possible focus component as well.'

Our working hypothesis here is that an extension of the focus (or more broadly, alternatives-based) analysis to these uses may be possible, so that preverbal *ko* marks **predicate focus** (Zimmermann 2016, among others). In this we follow a suggestion by Massam (2009), who writes that 'Since [*ko e* V] behaves exactly like [*ko e* N], where NP is focused, I assume [*ko e* V] is a predicate focusing device and is not a member of the

[12] Our Niuean speaker/fourth author comments that she initially started out thinking that preverbal *ko* was only for the present tense, but then realized that it can be used in past tense contexts.

[13] Our fourth author corrects this to 'Ko e tā kilikiki a maua.' There is some variation between Niuean speakers with respect to vowel length.

TAM paradigm.' However, as readers will see, our investigation is at a preliminary stage and the evidence for a predicate focus analysis is not yet conclusive.

3. Evidence for and against preverbal *ko* as a TAM marker

In this section, we address whether preverbal *ko* conveys temporal and/or aspectual semantics. We show that in the speech of our fourth author, *ko* marks neither present tense nor progressive aspect. However, we confirm that *ko* is incompatible with habitual interpretations, which suggests that it may have an aspectual component. We also show that there is a preference for *ko* to avoid past/completive perfective contexts.

3.1. Preverbal ko *is not present tense*

Massam et al.'s (2006) claim that *ko e* is not a present tense is confirmed in our data. (17–19) involve past events and using *ko* in preverbal position is fine:

17. Context: I call you up and you don't answer. The next day, I ask you 'What were you doing yesterday when I rang?'

    Ko e kai a au he talo.
    KO C eat ABS 1.SG at taro
    'I was eating taro.'

18. **Ko** e kai a au he talo he magaaho ne hoko mai
    KO C eat ABS 1.SG ABS taro at time NFT arrive DIR.1
    e matua fifine haaku.
    ABS parent female GEN.1.SG
    'I was eating taro when my mother arrived.'

19. **Ko** e tāmate e koe e kapitiga haaku.
    KO C kill ERG 2.SG ABS friend GEN.1.SG
    'You killed my friend.'

3.2. Preverbal ko *is not a progressive*

Seiter's observation that preverbal *ko* is compatible with stative predicates is also confirmed, as in (20–22). This suggests that *ko e* is not a progressive aspect.

20. **Ko** e eto/mule e uasi haau.
    KO C be.slow/be.slow ABS watch GEN.2.SG.
    'Your watch is slow.'[14]

21. **Ko** e loa koa kia e tagata ne kitia e au he hola.
    KO C tall EMPH EMPH ABS man NFT see ERG 1.SG at run
    'The man I saw running away is really tall.'

22. Context: Aren't you excited?

    Ē – **ko** e mategūgū[15] mo e hoge au.
    yes **KO** C be.tired and C hunger 1.SG
    'Yeah – but I'm tired and hungry.' (*Haia*)

Further suggestive evidence against a progressive analysis of *ko e* is given in (23), where the sequence *ko e* is optional but acceptable under the interpretation where I start to eat after my mother arrives. (Note that this is another past-tense case of preverbal *ko*.)

23. Context: I started eating taro once my mother arrived.

    (Ko e) kai (a) au he talo he magaaho ne hoko mai/age
    (**KO** c) eat (ABS) 1.SG at taro at time NFT arrive DIR.1/DIR.3
    e matua fifine haaku.
    ABS parent female GEN.1.SG
    'I ate taro when my mother arrived.'

(15) above is another case where a progressive would be unlikely to appear, yet *ko* is fine.

Despite this evidence that preverbal *ko* is not progressive, there does seem to be a preference for it to be used with ongoing events rather than with inceptive ones; this is shown in the contrast between the speaker's volunteered versions in (24a) vs. (24b).

[14] This example is adapted from Seiter's version, which is given in (i), with his glosses. Our fourth author translates (i) as 'You are delaying the proceedings/event,' which suggests that it may not actually be stative: and notably the predicate here contains the causative prefix *faka-*.

(i) **Ko e** fakatuai e uasi haau, …
    PRES slow ABS watch your
    'Your watch is slow, … ' (Seiter 1980: 5, citing FVTI13: Verne, Jules)

[15] This word has been corrected to conform to our fourth author's spelling, which matches how it is spelled in Sperlich (1997) and elsewhere in *Haia*.

24. a. Context: I started laughing when my sister arrived (e.g., because she looked funny).

Ne kata au he kitia e mahakitaga haaku.
PST laugh.SG 1.SG when see ABS man's.sister GEN.1.SG
'I laughed when I saw my sister.'

b. Context: I was laughing when my sister arrived (e.g., because I was reading a funny magazine).

Ko e kata au he magaaho ne hoko age/mai
KO C laugh.SG 1.SG at time NFT arrive DIR3/DIR1
e mahakitaga haaku.
ABS man's.sister GEN.2.SG
'I was laughing when my sister arrived.'

3.3. Preverbal ko *is episodic*

Seiter's claim that preverbal *ko* is incompatible with habitual events is confirmed by our speaker, as illustrated in (25–26). This restriction suggests that *ko* may convey some aspectual content, such as requiring episodic interpretations.

25. Context: I am cooking taro and am wondering whether you eat it or not (generally). I ask 'Do you eat taro?' and you answer:

Ē, (**#ko** e) kai (a) au he talo.
yes (**KO** C) eat (ABS) 1.SG at taro
'Yes, I eat taro.'

26. Context: A first date, trying to get to know the other person and their hobbies.

Her: (**#Ko** e) koukou tahi nakai a koe?
(**KO** C) bathe sea Q ABS 2.SG
'Do you swim?'

Him: Ē, (**#ko** e) koukou tahi au.
yes (**KO** C) bathe sea 1.SG
'Yes, I swim.'

3.4. Preverbal ko *disprefers completed/perfective contexts*

A further suggestive connection of *ko* to aspectual semantics is that our fourth author tends to reject it in contexts of past, completed events.

27. Context: I ask you 'What did you do yesterday?' You say:

(#Ko e) ala tuai (a) au. (**#Ko** e) matike hake (a) au, (**#Ko** e) kai (a) au
KO C wake early ABS 1.SG **KO** C arise up ABS 1.SG **KO** C eat ABS 1.SG
he talo, mo e fano ke feleveia mo e nena haaku.
at taro, and C go SBJ visit with C grandmother GEN.1.SG
'I woke up early. I got up, I ate taro, and I went to visit my grandmother.'

28. Context: We talked on the phone yesterday. The next day:

Me: Ko e heigoa haau ne taute he oti e tutala a taua
KO C what GEN.2.SG NFT do when finish ABS talk ABS 1.DU.INCL
he telefoni?
at telephone
'What did you do after we got off the phone?'

You: (**#Ko** e) fano au gahua.[16]
KO C go 1.SG work
'I went to work.'

3.5. Summary

We have shown that while preverbal *ko* is neither straightforwardly a present tense nor a progressive aspect, it seems to convey some aspectual information: it is incompatible with habituals, prefers ongoing readings, and disprefers past completed contexts. In the next section we investigate the evidence for our working hypothesis that preverbal *ko* encodes predicate focus, and in Section 5 we try to tie the aspectual effects in with this hypothesis.

4. Exploring the idea that preverbal *ko* is a predicate focus marker

Following Zimmermann (2016), we adopt a fairly broad definition of predicate focus, which includes focus on the verb, verb phrase, or TAM marking. Commonly, focus is diagnosed by a preceding wh-question or a corrective context, as illustrated in (29–30) (for VP focus), from Zimmermann (2016: 314).

29. A: What did Peter do?
B: Peter [petted the CAT]$_{FOC}$.

30. A: Peter stayed away from the animals.
B: No, Peter [petted the CAT]$_{FOC}$.

In line with the predictions of the predicate focus hypothesis, question-answer conversations such as in (31) are acceptable in Niuean. (31B) plausibly involves focus on the predicate *huva* 'vacuum.'

[16] Our fourth author has the intuition that *ko* is not good here because preverbal *ko* is 'more for present tense'. Although as we have seen above, *ko* is not in fact restricted to present tense, we interpret this intuition as reflecting some temporal or aspectual contribution of *ko.*

31. A: Ko e eke hā a Sione?
KO C do Q ABS Sione
'What is John doing?'

B: **Ko** e huva a Sione.
KO C vacuum ABS Sione
'John is vacuuming.' (adapted from *Haia*; confirmed by our fourth author)

In predicate correction contexts, our fourth author also volunteers preverbal *ko*, as in (32–33).

32. Context: Your mother calls you and asks what everyone in the family is doing right now.

You: Ko e koukou tahi e tehina haaku.
KO C bathe sea ABS sister GEN.1.SG
'My sister is swimming.'

Mother: Ai koukou tahi e tehina haau, **ko** e gahua e patu.
NEG bathe sea ABS sister GEN.2.SG **KO** C work ABS person
'Your sister isn't swimming, she's working.'

33. A: (Ko e) lātau ka[17] e tau tagata kō?
KO C fight Q ABS PL person DEM
'Are those people fighting?

B: Nākai, ai lātau (ka e) **ko** e koli.
no NEG fight but C **KO** C dance
'No, they're not fighting, they're dancing.'

Although preverbal *ko* is *possible* in predicate focus environments, it is not obligatory, as shown in (34–35):

34. A: Ko e eke hā a mua (i) neafi?
KO c do what ABS 2.DU (on) yesterday
'What did you (two) do yesterday?'

B1: **(Ko** e) tā kilikiki a maua.
KO c play cricket ABS 2.DU
'We played cricket.'

B2: (**Ko** e) gahua au.
KO c work 1.SG
'I worked.'

[17] Our fourth author observes that this word is said with a long vowel, but not written with a macron.

35. A: Ko e heigoa e mena ne tupu?
KO C what ABS thing NFT happen
'What happened?'

B: (**Ko** e) malona e fale.
KO C collapse ABS house
'The house collapsed.'

This correlates with Zimmermann's observation that cross-linguistically, predicate focus marking is often optional (to a much greater extent than focus on argument DPs). Zimmermann suggests (2016: 332) that since verbal predicates 'constitute the default focus of an utterance', they 'need not be marked as such in order to be properly identified as focus'.

The optionality of predicate focus marking means that it is challenging to formulate strong predictions of the predicate focus analysis. We seem to predict a weak one-way correlation, whereby preverbal *ko* should be disallowed when an argument nominal is in focus. This prediction receives some support from (36). Here, there is contrastive nominal focus. The mother's first response (volunteered by the speaker) overtly focuses the nominals; her second response involves no *ko*-marking in the first clause and is acceptable. However, preverbal *ko* is disallowed in this context, as shown in the mother's third and fourth responses:

36. Context: Your mother calls you and asks what everyone in the family is doing right now.

You: Ko e koukou tahi e tehina haaku.
KO C bathe sea ABS sister GEN.1.SG
'My sister is swimming.'

Mother1: **Ko** e tugaane haau ne koukou tahi. Ai **ko** e tehina haau.
KO C brother GEN.2.SG NFT bathe sea NEG **KO** C sister GEN.2.SG
'It's your brother who's swimming. Not your sister.'

Mother2: Ai koukou tahi e tehina haau, ko e tugaane haau.
NEG bathe sea ABS sister GEN.2.SG KO C brother GEN.2.SG
'Your sister isn't swimming, your brother is.[18]

Mother3:* **Ko** e **ai** koukou tahi e tehina haau, ko e tugaane haau.
KO C **NEG** bathe sea ABS sister GEN.2.SG KO C brother GEN.2.SG

Mother4:? **Ai** **ko** e koukou tahi e tehina haau, ko e tugaane haau.
NEG **KO** C bathe sea ABS sister GEN.2.SG KO C brother GEN.2.SG

Further suggestive support for the predicate focus idea comes from (37). This is not a predicate focus environment, and our fourth author volunteers the answer with nominal focus in (B1). While she does accept the version with preverbal *ko* in (B2), her initial reaction suggests that she would not *expect* predicate focus in this discourse

[18] *Tehina* and *tugaane* are used here if addressee is female and has a younger sister and a brother.

context. (B2) is a more appropriate answer to a predicate-focus question, as was shown in (31) above.

37. A: Ko hai ka taute e tau fekau?
KO who Q do ABS PL errand
'Who is doing the chores?'

B1: Ko Sione.
KO Sione
'It's John.'

B2: **Ko** e huva a Sione.
KO C vacuum ABS Sione
'John is vacuuming.'
Consultant's initial reaction to B2: "John is a vacuum cleaner!"

The unexpectedness of *ko* in (B2) here contrasts with the apparent full acceptability of the same string in a contrastive predicate context as in (38), again providing suggestive support for the predicate focus analysis.

38. A: Ko hai ka taute e tau fekau?
KO who Q do ABS PL errand
'Who is doing the chores?'

B: **Ko** e huva a John, **ko** e fakameā a Moka he motoka.
KO C vacuum ABS John **KO** C CLEAN ABS Moka at car
'John is vacuuming, Moka is cleaning the car.'

In spite of these suggestive pieces of evidence for the predicate focus analysis, we have to admit that when attempting to use the question-answer diagnostic with simple two-sentence conversations, there are results which are challenging for the hypothesis. This is true in both directions: we find both that a predicate-centred question allows nominal focus in the answer, as in (39), and that a nominal-centred question allows (hypothesized) predicate focus in the answer, as in (40). However, it is also possible that *ko* functions as a topic marker in (39B) and (40B). We would need more context to confidently rule out the possibility that *ko Sione* in (39B) and *ko e huva* in (40B) are contrastive topics. So there may well be a flaw in the methodology of using the simple question-answer diagnostic and asking for an acceptability judgment; further investigation is required.[19]

39. A: Ko e eke hā a Sione?
KO c do Q ABS Sione
'What is John doing?'

B: **Ko** Sione ne huva.
KO Sione NFT vacuum
'John is vacuuming.'

[19] Also possibly relevant is that the answer in (40B) is not preferred by our fourth author. Her volunteered answer is *Ko Sione* 'John is.'

40. A: Ko hai ne huva?
KO who NFT vacuum
'Who is vacuuming?'

B: **Ko** e huva a Sione.
KO C vacuum ABS Sione
'John is vacuuming.'

In addition to these prima facie counter-examples, another challenge is that in many contexts, the *ko* form appears to be interchangeable with a *ko*-less form, without an obvious semantic or pragmatic effect. Our fourth author cannot detect a meaning difference between the versions of (41) with and without *ko,* for example. She feels that the form with *ko* may be the 'proper' or 'full' way to say it. ((41) is our fourth author's version of Seiter's example in (11) above.)

41. (**Ko** e) tohitohi a au mogonei aki/mo e pene foou.[20]
KO C writing ABS 1.SG now with/with ABS pen new
'I'm writing at the moment with a new pen.'

Similarly, in (42), there is no perceived semantic difference between the two versions, but the version with *ko* is judged as the proper/full way to say it:

42. Context: I call you up and ask 'what are you doing right now?' You answer:

(**K**o e) kai (a) au he talo.
KO C eat ABS 1.SG at taro
'I am eating taro.'

The lack of a clear meaning difference between the versions with and without *ko* in (41–42) may simply reflect the pitfalls of eliciting judgments without rich enough discourse contexts.

5. First steps towards tying things together

We have seen that preverbal *ko* displays (a) some strong aspectual effects (disallowing habitual interpretations), (b) some temporal/aspectual tendencies (preferring present ongoing events, dispreferring past/completed events), and (c) at least a tendency to match the predictions of a predicate focus analysis. The challenge now is whether, and how, these observations can all be accounted for. One important question is how far it is possible, or desirable, to unify *ko* across all its uses. This question arises both at the level of preverbal *ko* itself – is there just one? – and with respect to trying to unify preverbal *ko* with its other uses in (1–5) above.

With respect to the question of how many *ko*s there are, Seiter (1980: 86) suggests a historical, but not necessarily a synchronic, unification; he comments that 'the predicate marker *ko* plus absolutive *e* is homophonous with the actual present marker

[20] Generally speaking, *aki* is instrumental 'with', and *mo* is comitative 'with'.

ko e. The tense marker *ko e* is a Niuean innovation, and it is reasonable to suggest that it is historically related to predicate nominal marking.'

The working hypothesis we adopted about preverbal *ko* – that it is a predicate focus marker – was obviously driven by the idea that it would be conceptually appealing if all the uses of *ko* were related. We offer here a few remarks about the connection between the aspectual and information-structure effects of preverbal *ko.*

It turns out that such a connection is well-attested in other languages. For example, De Kind et al. (2015) argue that in languages of the Kikongo Language Cluster (Bantu), both the 'fronted-infinitive' and the 'locative-infinitive' constructions mark predicate focus as well as progressive aspect. (See also Hyman & Watters 1984; Hyman 1999; Güldemann 2003 on the predicate focus/progressive connection in Bantu.) De Kind et al. argue that the unifying concept is 'event-centrality': 'The utterance is centred around the event expressed by the verb' (2015: 122). Güldemann (2003) argues that 'the continuous, ongoing nature of an event is that information which is viewed by the speaker to be the most relevant', and even goes so far as to say that 'the progressive is a verb category with inherent focus' (Güldemann 2003: 350, cited in De Kind et al. 2015: 138). The connection is further supported by the existence of languages where progressive aspect is incompatible with focus on an argumental phrase (De Kind et al. 2015: 139, citing Güldemann 2003; Zimmermann 2016: 317, citing Hyman 1999).

The connection between predicate focus and progressivity is intriguing and bodes well for an eventual unification of Niuean *ko*. The tendency for predicate focus markers to prefer present-tense contexts also seems to have cross-linguistic validity; see Güldemann (2003: 350–351; De Kind et al. 2015: 146).

Obviously, much future empirical and analytical work remains to be done. With respect to the source of the predicate focus/progressive connection, Zimmermann (2016) is sceptical of the idea that progressives are inherently focused. He suggests that 'A more promising possibility – to be investigated in future research – would be that the difficulties with realising term focus in progressive sentences follow from conceptual problems with the backgrounding of temporally unbounded situations, which form a characteristic part of the progressive meaning' (Zimmermann 2016: 317).

In the following section we briefly return to the syntax of preverbal *ko.* We will see that when the interaction of *ko* with negation is investigated in more detail, an extra piece of support for the predicate focus analysis emerges.

6. The syntax of preverbal *ko*

6.1. Ko e *is not a TAM marker*

Massam et al. (2012: 244) provide evidence for the following order of elements in the Niuean clause:

TAM – **NEG** – MODAL – PREDICATE – PARTICLES/ADVERBIALS – ARGUMENTS

As mentioned in Section 2.1, preverbal *ko* differs from Niuean TAM markers in that it can follow the negative marker *ai*, as in example (10), repeated here as (43).

43. **Ai** **ko** e onoono (a) au ke he ha mena.
NEG **KO** C look (ABS) 1SG at NSP thing
'I'm not looking at anything.' (Seiter 1980: 82)

While our fourth author accepted Seiter's example, the forms she volunteered in this context both lack *ko* (44–45).

44. Ai onoono (a) au ke he ha mena. (volunteered)
NEG look (ABS) 1.SG at NSP thing
'I'm not looking at things.'

45. Ai onoono mena au. (volunteered)
NEG look thing 1.SG
'I wasn't looking.'

This suggests that for our fourth author, preverbal *ko* is disfavoured with simple sentence negation.

In contrastive contexts, the occurrence of preverbal *ko* and the relative position of *ai* and *ko* appears to depend on the nature of the contrast. In the predicate correction context in (46), *koukou tahi* 'swimming' is contrasted with *gahua* 'working' in a structure involving the coordinating conjunction *ka* 'but'. Both predicates are marked with *ko* here, and the negative marker *ai* necessarily appears in initial position.

46. Context: Your mother calls you and asks what everyone in the family is doing right now.

You: Ko e koukou tahi e {tehina/taokete/mahakitaga} haaku.[21]
KO C bathe sea ABS sister GEN.1.SG
'My sister is swimming.'

Mother: **Ai** **ko** e koukou tahi ka e ko e gahua (a ia). (accepted)
NEG **KO** C bathe sea but C KO C work (ABS 3.SG)
'She isn't swimming but working.'

Mother: ***Ko** e **ai** koukou tahi …'
KO C **NEG** bathe sea ...

However, when the negative phrase 'she is not swimming' is contrasted with the modal construction 'she is probably swimming', as in (47), *ko* (plus *e*) appears before the negative marker.[22]

[21] *Tehina* is the appropriate term if an older sister is talking; *taokete* if a younger sister is talking; and *mahakitaga* if a brother is talking.
[22] Our fourth author volunteered the two versions of the mother's response and also accepted the *ko e ai* ordering when it was re-checked in a later elicitation.

47. Context: Your mother asks you what your sister is doing right now.

You: Liga haia ne koukou tahi. (volunteered)
likely right NFT bathe sea
'She's probably swimming.'

Mother: **Ko** e **ai** koukou tahi he gagao.
KO C **NEG** bathe sea because be.sick
'She's not swimming because she is sick.' (volunteered)

Ko e **ai** fano ke koukou tahi he gagao.
KO C **NEG** go.SG SBJ bathe sea because be.sick
She's not going swimming because she is sick.' (volunteered)

When the subject is contrasted as in (48), our fourth author disfavours *ko* in preverbal position as well as before *ai*.

48. Your mother calls you and asks what everyone in the family is doing right now.

You: Ko e koukou tahi e {tehina/taokete} haaku.
KO C bathe sea ABS sister GEN.1.SG
'My sister is swimming.'

Mother: Ai koukou tahi e {tehina/taokete} haau
NEG bathe sea ABS sister GEN.2.SG
ka ko e tugaane haau
but KO C brother GEN.2.SG
'Your sister isn't swimming, but your brother is.'[23]

Mother: **?Ai** **ko** e koukou tahi …
NEG **KO** C bathe sea ...
'Your sister isn't swimming, but your brother is.'

Mother: ***Ko** e **ai** koukou tahi …
KO C **NEG** bathe sea ...
'Your sister isn't swimming, but your brother is.'

If the sequence *ko e* was a TAM marker in (47), we would expect it to be equally acceptable before *ai* in (48), since the conversation revolves around an ongoing event in both examples. The fact that our fourth author volunteered *ko* before *ai* in (47) but rejected it in (48) indicates that the use of preverbal (pre-NEG) *ko* is focus related.

This raises the question why preverbal *ko* seems to be incompatible with TAM marking. In the following section, we will argue that this is due to the internal properties of the *ko e* construction.

[23] This sentence is used when the addressee is female and has a brother and a sister. *Tehina* indicates that the sister is younger, whereas *taokete* indicates that the sister is older.

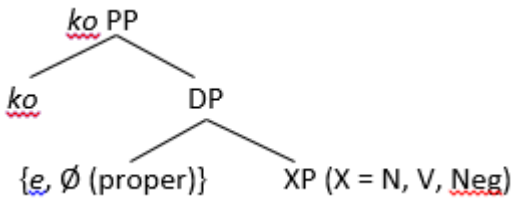
6.2. The internal structure of ko *constructions*

A particularly intriguing feature of preverbal *ko* constructions is the appearance of *e*, which Massam (2016: 10) treats as an overt realisation of D in the context of case markers/prepositions such as comitative *mo*, benefactive *ma*, and *ko*.

According to Massam et al (2012: 4), when *mo* is used as a coordinator, it is a P that selects a DP. This DP is headed by *e* when the complement of D is not a proper noun. Massam et al. (2012: 4f) emphasise that the complement of the determiner *e* can be not just an NP, but also an AP, VP, or CP/IP.

The data presented above suggest that *ko* resembles *mo* in that it selects a DP that is headed by *e* when its complement is not a proper noun. However, while *mo* can be followed by a nominalised CP that contains a TAM marker (cf. Massam et al. 2012: 2, example (8)), the absence of TAM markers in preverbal *ko* constructions indicates that *ko* is only compatible with nominalisations up to NegP-level (49).

49.



6.3. The position of ko *constructions in the clause*

We hypothesise that the alternative semantics that unifies topic and focus uses of *ko* is associated with the internal properties of the *ko* PP. What distinguishes topic uses from focus uses is the overall position of the *ko* PP in the clause. Drawing on Pearce's (1999) analysis of *ko* in Māori, we assume that *ko*-marked topics appear in SpecTopP, whereas *ko*-marked focus constituents form the predicate of a cleft-like construction, which arguably sits in SpecTP (cf. Massam et al. 2012). It remains to be determined if the differences between other uses of *ko* can also be analysed as involving a similar basic semantic reading, with different syntactic positions yielding somewhat different effects.

7. Conclusion

In this preliminary investigation of Niuean preverbal *ko* we have shown that it is not a present tense marker, nor a progressive aspect. *Ko* seems to convey *some* aspectual information, in that it is incompatible with habitual interpretations, prefers ongoing readings, and disprefers past completed contexts. We have pursued the tentative hypothesis that preverbal *ko* can be unified with prenominal *ko,* by analysing it as a device for marking predicate focus. There is broad support for the hypothesis, but also some apparent counter-evidence which requires further investigation. The predicate focus analysis has the potential to account for the aspectual effects, following similar correlations in other languages.

We have also investigated the syntax of preverbal *ko,* and argued that it is not in a TAM position. We suggested that the obligatory presence of *e* can be accounted for by assuming that *ko*'s complement is a nominalised constituent (not larger than NegP).

We propose that it may be fruitful to study the use of preverbal *ko* in wider discourse contexts and explore the syntax/semantics interface as a route to explaining all the differing functions of *ko.*

Acknowledgements

We are very grateful to Rebecca Tollan for her help with data collection.

References

Bauer, Winifred. 1991. Māori *ko* again. *Te Reo* 34. 3–14.

Bauer, Winifred. 1997. *The Reed reference grammar of Māori*. Auckland: Reed Publishing.

Büring, Daniel. 1997. *The meaning of topic and focus*. London: Routledge.

Büring, Daniel. 2003. On D-trees, beans, and B-accents. *Linguistics & Philosophy* 26(5). 511–545.

Brown, Jason & Koch, Karsten. 2013. Focus, Polynesian *ko, and language change (How morphosyntax is, and is not, like rats). *Proceedings of the 2013 annual conference of the Canadian Linguistic Association*. (http://homes.chass.utoronto.ca/~cla-acl/actes2013/Brown_and_Koch-2013.pdf)

Clark, Ross. 1976. *Aspects of Proto-Polynesian syntax*. Auckland: Linguistic Society of New Zealand.

De Kind, Jasper & Dom, Sebastian & de Schryver, Gilles-Maurice & Bostoen, Koen. 2015. Event-centrality and the pragmatics-semantics interface in Kikongo: From predication focus to progressive aspect and vice versa. *Folia Linguistica Historica* 36. 113–163.

Greenhill, Simon J. & Clark, Ross. 2011. POLLEX-Online: The Polynesian Lexicon Project Online. *Oceanic Linguistics* 50(2). 551–559

Güldemann, Tom. 2003. Present progressive vis-à-vis predication focus in Bantu: A verbal category between semantics and pragmatics. *Studies in Language* 27. 323–360.

Haia! An introduction to Vagahau Niue. 2010. Wellington, New Zealand: Ministry of Education. (http://pasifika.tki.org.nz/Media/Files/Haia!-an-introduction-to-Vagahau-Niue)

Hohaus, Vera & Howell, Anna. 2015. Alternative semantics for focus and questions: Evidence from Sāmoan. In Camp, Amber & Otsuka, Yuko & Stabile, Claire & Tanaka, Nozomi (eds.) *Proceedings of the 21st Meeting of the Austronesian Formal Linguistics Association*, Asia Pacific Linguistics, 69–86.

Hyman, Larry M. 1999. The interaction between focus and tone in Bantu. In Rebuschi, Georges & Tuller, Laurice (eds.) *The grammar of focus*, 151-178. Amsterdam: John Benjamins.

Hyman, Larry M. & Watters, John Robert. 1984. Auxiliary focus. *Studies in African Linguistics* 15. 233–273.

Kaulima, Aiao & Beaumont, Clive H. 2002. *Learning Niuean: Combined edition. – Tohi ako Vagahaau Niue*. Auckland: C.H. and D.J.M. Beaumont and A. Kaulima.

Krifka, Manfred. 2007. Basic notions of information structure. In Féry, Caroline & Fanselow, Gisbert & Krifka, Manfred (eds.) *The notions of information structure* (Interdisciplinary studies on information structure 6), 13-56. Potsdam: Universitätsverlag Potsdam.

Massam, Diane. 2009. The morpho-syntax of tense particles in Niuean. *Proceedings of the 2009 annual conference of the Canadian Linguistic Association*. (http://homes.chass.utoronto.ca/ ~cla-acl/actes2009/CLA2009_Massam.pdf)

Massam, Diane. 2016. Rising from the ashes: D in Niuean. (Paper presented at the Dimensions of D workshop, Rochester University, 16–18 September 2016)

Massam, Diane & Chen, Jingwei & Lokay, Hedvig. 2012. *Mo (e)*: Putting (Niuean) things together. *Toronto Working Papers in Linguistics (TWPL)* 35. 1–9.

Massam, Diane & Lee, Josephine & Rolle, Nicholas. 2006. Still a preposition: The category of *ko*. *Te Reo* 49. 3–37.

Pearce, Elizabeth. 1999. Topic and focus in a head-initial language: Māori. *Proceedings of the sixth meeting of the Austronesian Formal Linguistics Association (AFLA VI)* (*Toronto Working Papers in Linguistics* 16), 249–263.

Pearce, Elizabeth. 2015. *A grammar of Unua* (Pacific Linguistics 647). Berlin: De Gruyter Mouton.

Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. In Yoon, Jae-Hak & Kathol, Andreas (eds.) *OSU Working Papers in Linguistics 49: Papers in semantics*, 91–136. Columbus: The Ohio State University.

Rooth, Mats. 1985. *Association with focus*. Amherst: University of Massachusetts. (Doctoral dissertation)

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1(1). 75–116.

Seiter, William John. 1980. *Studies in Niuean syntax*. New York: Garland Press.

Sperlich, Wolfgang B. 1997. *Tohi vagahau Niue: Niue dictionary*. Honolulu: Government of Niue and University of Hawai'i Press.

Verne, Jules. trans. by Tongakilo. *Koe Fenoga ne Fano Viko Takai ke he Valugofulu ni e tau Aho* (Around the World in Eighty Days). (as cited in Seiter 1980).

Zimmermann, Malte. 2016. Predicate focus. In Féry, Caroline & Ishihara, Shinichiro (eds.) *Handbook of information structure.* Oxford: Oxford University Press. 314–335.

Lisa Matthewson
University of British Columbia
lisa.matthewson@ubc.ca

Heidi Quinn
University of Canterbury
heidi.quinn@canterbury.ac.nz

Diane Massam
University of Toronto
diane.massam@utoronto.ca

Lynsey Talagi
Ministry of Infrastructure, Government of Niue
lynsey-talagi@mail.gov.nu

Phrase-level stem alternations in Sumatran Malayic[1]

Timothy Mckinnon, Peter Cole, Yanti, and Gabriella Hermon

1. Introduction

Stem ablaut is viewed by many linguists as uncommon in most Sumatran Malayic varieties other than those of Kerinci; however, it is in fact found in many Malayic varieties of the region. In this short paper we will discuss phrase-level alternations in the form of stems. As we shall demonstrate, phrase-level phonological alternations are found in several Malayic varieties in Sumatra discussed in this paper. Such alternations are geographically widespread, but we shall argue that they developed independently in different dialects in the region. The fact that their development is independent in each variety is evident because they exhibit distinct phonological manifestations in different dialects; for example, in some cases they affect final consonants, while in other cases they affect final vowels. These differing manifestations suggest that they developed independently in each variety, but the fact that they occur in many varieties in the same environment suggests that they developed as differing responses to shared non-optimal properties of the parent language. More concretely, we shall suggest that the phrase-level alternations are the effects of phrasal prominence or the lack thereof on consonants and vowels. Our general point in this paper will be to show that phrasally conditioned changes are, in fact, not unusual in this group of languages, but that such changes were common throughout the region. What varies from one variety to another is the nature of the change that occurred and not the environment for the change.

We begin with a discussion of alternations manifested in the reflexes of historical word-final nasal stops. Following this, we describe other attested types of phrasal alternations. Through developing a typology of phrasal alternations in the region, we support our general claim that such alternations are well attested, and, thus, languages like Jangkat, which exhibit both phrasal-level alternations, are not exceptional varieties, but rather are characteristic varieties of Sumatran Malayic.

2. Alternation of final nasals

As we have mentioned, phrase-level alternations are well attested in the Malayic varieties of Sumatra. One type of phrasal alternation that is found in several varieties across the region involves word-final nasal sounds. In this section, we will illustrate this type of phrasal alternation in three varieties, Tanjung Raden, Sarolangun, and Jangkat.

[1] The research reported on here was supported by the Max Planck Institute for Evolutionary Anthropology and the National Science Foundation (grant #1126149).

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 159–167

These varieties represent different 'stages' in the development of the alternation. Following our discussion of alternating final nasals, we will discuss other types of phrasally conditioned alternations attested in the region. The prevalence and diverse manifestation of phrasal alternations provides support for our claim that such alternations are characteristic of the region.

The first variety which we look at is Tanjung Raden, a Malay variety spoken in the traditional village of the same name located on the banks of the Batanghari River, directly across from Jambi City (Jambi Province). Yanti (2010) provides a detailed description of the phonology, morphology, and syntax of this variety. Tanjung Raden offers a very simple example of a phrase-level alternation, which we shall see below. As Yanti notes, word-final nasal sounds exhibit slight oral occlusion, that is the reflexes of **m#*, **n#*, and **ŋ#* are realized as ^{b}m, ^{d}n, and $^{g}ŋ$.

1. Preoccluded final nasals in Jambi Seberang (Tanjung Raden)

mala^{b}m	'night'
aza^{d}n	'prayer call' (Arabic)
lawa^{g}ŋ	'door'

Pre-occlusion tends not to occur in certain phrase- and word-medial environments. Although Yanti does not discuss the specific environments in which nasals which are otherwise occluded occur in unoccluded form, she does provide the following example of full reduplication, in which the medial form of the reduplicated base *malam* 'night' surfaces with a final unoccluded nasal stop *m*.

2. Reduplication and loss of final occlusion

(final)		(medial-final)
mala^{b}m	>	malam-mala^{b}m

This form reflects a broad phonetic tendency in Tanjung Raden, whereby final nasals are more likely to occur with occlusion in phrase final positions, especially when the syllable containing the occluded nasal shows greater relative prominence. This is a gradient phonetic tendency rather than an allophonic alternation between two discrete allophones of a phoneme.

Jernih, Sarolangun: Jernih also exhibits a phrase-level alternation affecting final nasals which appears superficially to be very similar to the alternation in Tanjung Raden. Upon careful examination, however, the alternation in Jernih can be seen to be a case of regular phonological allophony, whereby the choice between allophones is conditioned by specific environments, rather than a case of gradient phonetic variation, as was seen in Tanjung Raden.

Much like Tanjung Raden, the reflexes of Proto Malay root-final nasal stops in Jernih are pre-occluded in citation form.[2]

[2] As in Tanjung Raden, occlusion is blocked in Jernih in forms in which the final nasal is locally preceded by a nasal stop.

3. Occluded reflexes of Proto Malayic nasal stops

Malay	**Jernih (citation form)**	**English**
makan	makat^{n}	‘eat’
jalan	jalɒt^{n}	‘road’
terbang	tə:bɒk^{ŋ}	‘fly’

Also, like in Tanjung Raden, these historical final nasals are realized as plain nasals in certain phrase medial positions. When the word *makat^{n}* ‘eat’ is followed by a direct object, it surfaces as *makan*.

4. Phrase final and phrase medial forms (Jernih, Sarolangun): *makat^{n}* / *makan*

Phrase final	Phrase medial
makat^{n} ‘eat’	makan nasi^{y} ‘eat rice’

However, unlike in Tanjung Raden, the occluded portion of the final nasal in Jernih has greater duration and is more perceptually salient in Jernih than in Tanjung Raden. In fact, in elicitation form (the form given when eliciting this form in isolation), the nasal articulation can be omitted, whereas the oral articulation is obligatory.

5. Variability in realization phrase-final forms

Phrase final		Phrase medial
makat^{n} ~ makat	vs.	makan

Moreover, whereas the alternation in Tanjung Raden is probabalistic and affected by speech rate, the alternation in Jernih is categorial. For most syntactic environments, speakers have categorial judgements about whether an alternating root must surface in the medial (plain nasal stop final) form or final (‘post-nasalized’ oral stop final) form.

With regard to the syntactic distribution of the phrase-final and phrase-medial forms in Jernih, we observe a pattern which is nearly identical to the distributional pattern of phrase-medial and phrase-final forms in Jangkat. Within the noun phrase, the noun appears in the phrase-medial form with a possessor, attributive adjective, demonstrative, and (optionally) with a relative clause modifier.

6. Environments where N appears in phrase medial form in Jernih

   a. N+Possessor

   ɲo nɪŋoʔ kʊčɪŋ/*kʊčɪk aliʸ
   3 N.see cat Ali
   ‘I saw Ali’s cat.’

   b. N+Attributive Adjective

   awaʔ mᵇəlɪ kʊčɪŋ/*kʊčɪk lɪaʁ
   1SG N.buy cat wild
   ‘I bought a stray cat.’

c. N+DEM

awaʔ mᵇəlɪ kʊčɪŋ/*kʊčɪk tuʷ
1SG N.buy cat that
'I bought that cat.'

d. N+REL

ɲo ɲaʁi kʊčɪŋ/kʊčɪk nəŋ laʁi
3 N.search cat REL run
'He is looking for the cat that ran away.'

In other environments (i.e. when the noun is followed by a PP, a numeral/classifier, or when it appears at the right edge of DP), the noun appears in the phrase-final form.

7. Environments where N appears in phrase-final form in Jernih

a. Noun at phrase edge

ɲo mᵇəlɪ kʊčɪk/*kʊčɪŋ
3 N.buy cat
'He bought a cat.'

b. Noun + [Numeral + Classifier]

awaʔ kəteʔ kʊčɪk/*ʔkʊčɪŋ tɪgʊ ɪkʊʔ
1SG possess cat three CLASS
'I have three cats.'

Moreover, active transitive verbs must surface in the medial form when followed by a direct object. In other environments (e.g. with an adverbial modifier, PP phrase complement/adjunct, in utterance final position) the verb must appear in its phrase-final form.

8. Environments where V appears in phrase-medial form in Jernih

Verb + object

ɲo padeʔ maleŋ/*malek dʊɪnʔ
3 often N.steal money
'He often steals money.'

9. Environments where V appears in phrase-final form in Jernih

a. Verb (in final position)

ɲo padeʔ malek/*maleŋ
3 often N.steal
'He often steals.'

b. Verb + PP

ɲo padeʔ malek/*maleŋ dɪ pasaʁ
3 often N.steal in market
'He often steals at the market.'

c. Verb + ADV

ɲo malek/*maleŋ tɪamʔ ariy
3 N.steal each day
'He steals every day.'

Jernih speakers are sensitive to the difference between phrase-final and phrase-medial forms, and have very clear judgments regarding the environments in which medial vs. final forms can appear. In this respect, the phrasal alternation in Jernih is very similar to the phrasal alternation in Jangkat, and unlike the phrasal alternation in Tanjung Raden (where speakers are largely unaware of the phrasal alternation between plain and pre-occluded nasals).

The following diagram in (10) illustrates the phrase-medial and phrase-final forms in Tanjung Raden, Jernih, and Jangkat. In Tanjung Raden, final nasals are realized with slight oral occlusion which is most salient in phrase-final positions (i.e. those positions which are most prosodically prominent). This occlusion tends to disappear or be less salient in phrase-medial (non-prosodically prominent) positions. In Jernih, the distinction between the medial form (i.e. the form exhibiting a final stop nasal) and the final form (i.e. the form exhibiting a final oral stop, optionally followed by a nasal articulation) has become categorical, and speakers exhibit firm judgements regarding which forms may occur in which syntactic positions. Jangkat also exhibits a categorical distinction between medial forms (i.e. forms ending with a plain nasal) and phrase-final forms (i.e. forms ending with plain oral stop). Dotted arrows in the table indicate gradient phonetic variability (i.e. free variation) between two forms, whereas 'vs.' indicates categorial contrast.

10. Variability in realization phrase-final forms

	Phrase final (higher prominence)				Phrase medial (less prominence)
Tanjung Raden	*makadn*	◄- - - - - - - ►	*maka^{d}n*	◄- - - - - - - ►	*makan*
Jernih	*makat^{n}*	◄- - ►	*makat*	*vs.*	*makan*
Jangkat	*makat*		*vs.*		*makan*

Comparison of these forms suggests that the Jangkat and Jernih alternations both developed from an earlier phonetic alternation like that observed in Tanjung Raden, in which final nasals (not preceded locally by a nasal stop) exhibited a phonetic oral occlusion. Phonetic occlusion of nasals in final position is well attested throughout the region (cf. Anderbeck 2008). We hypothesize that the relative salience of this occlusion is greater in stressed or phrase-final syllables, since the rimes of such syllables exhibited greater relative duration. Jangkat and Jernih represent varieties in which phonetically variable final nasals were reinterpreted as discrete allophones, and the gradient phonetic variability between these two types of sound was reinterpreted as resulting from a discrete phonological rule conditioned by the phonological phrase edge.

The fact that the conditioning environment (i.e. the position in the phrase) for the alternations in both Jangkat and Jernih happens to be the nearly identical is not

surprising, if we look at the alternation as being related to sentential prosody. In general, phrasal stress in the Malayic languages of this region has the effect that syllables situated in the phrase-final position manifest greater relative duration. It is therefore not surprising that occluded final nasal segments could develop into oral stops in phrase-final positions, those in which the oral occlusion is most phonetically salient, whereas they would remain plain nasals in phrase medial positions, those in which their oral occlusion is absent or relatively non-salient.

3. Other phrasal alternation in the region

All of the phrase-level alternations we have described so far have involved historical root-final nasal sounds. Now we would like to show that phrasal-level alternations are not limited to alternations between preoccluded nasals and plain nasals, but in fact have diverse phonological manifestations. Moreover, we would also like to point out that, despite their diverse phonological manifestations, phrase-medial and phrase-final forms occur in nearly the exact same environments in many varieties where these alternations have phonologised. This commonality can be attributed to similar prosodic phrasing across the Malayic languages of the region.

In the Malayic varieties spoken in Dusun Baru (Sarolangun) and Tapan (West Sumatra), roots which ended historically with the high vowels **i* and **u* exhibit an 'excrescent' (i.e. historically inserted) nasal stop in final position. In Tapan, the nasal in these forms exhibits a velar place of articulation; whereas in Dusun Baru, the excrescent nasal stop exhibits an alveolar place of articulation following **i#* and a labial place of articulation following **u#* (n.b. final excrescent nasals Dusun Baru, as well as other final nasals are often occluded).

11. Excrescent nasals in Tapan following **i#* and **u#*

Tapan	**Indonesian**	**English**
kakiŋ	kaki	'leg'
taliŋ	tali	'rope'
atuŋ	hantu	'ghost'
kayuŋ	kayu	'wood'

12. Excrescent nasals in Dusun Baru following **i#* and **u#*

Dusun Baru	**Indonesian**	**English**
kakin	kaki	'leg'
talin	tali	'rope'
tamum	tamu	'guest'
abum	abu	'ash'

These final nasal sounds disappear in certain medial environments. For example, in Dusun Baru, when the cognate of the Malay root *baju* 'clothing' appears in utterance final position, it is pronounced with a final nasal; however, when it appears in phrase-medial position (e.g. when modified by an attributive adjective) speakers prefer that the same root be pronounced without a nasal coda.

13. ɲo mli beju^{p}m
    3 ACT.buy clothes
    'He/she bought clothes.'

14. ɲo mli beju biʁu^{p}m
    3 ACT.buy clothes blue
    'He/she bought blue clothes.'

Moreover, both Tapan and Dusun Baru happen to be weakly alternating varieties, insofar as they exhibit a 3rd person pronoun which has been incorporated into the root. In the secondary forms of roots the excrescent nasal in these forms disappears.

15. Basic and Ablaut forms with **i#* and **u#* in Tapan

Basic	**Ablaut**	
kakiŋ	kaki^{ə}	'leg'
taliŋ	tali^{ə}	'rope'
atuŋ	atu^{ə}	'ghost'
kayuŋ	kayu^{ə}	'wood'

16. Basic and Secondary forms with **i#* and **u#* in Dusun Baru

Basic	**Secondary**	**Malay**	**Gloss**
talin	tali:^{a}h	tali-nya	'his/her/the rope'
kakin	kaki:^{a}h	kaki-nya	'his/her/the leg'
tamum	tamo^{a}h	tamu-nya	'his/her/the guest'
abum	kabo-abo:^{a}h	keabu-abuan	'greyish'

Phrasal alternations are not just realised in word-final coda positions. A phrasally conditioned alternation affecting a final vowel can be seen in Jernih Sarolangun. In our discussion above, we showed that the reflexes of Proto Malay word-final nasal stops in Jernih exhibit a regular, phrasally-conditioned alternation. In the same variety, reflexes of the high vowels **i#* and **u#* are realized as *i*^{ɣ} and *u*^{w} in phrase-final position; whereas in phrase-medial position, these sounds are realized as monophthongs with a slightly lower place of articulation (perhaps best transcribed as *ɪ* and *ʊ*, respectively; our native speaker consultant stated that the *ʊ* in the medial form of a word like *kʊtʊ/kʊtu*^{w} 'louse' is identical to the final sound in roots such as *matʊ* 'eye', roots that contain the Proto Malayic sound **a# (matʊ* 'eye' <**mata*).

17. Phrasal alternation of high vowels in Jernih

	Phrase Final		**Phrase Medial**
a.	kʊtu^{w}	vs.	kutʊ kəpalʊ
	louse		louse head
	'louse'		'head louse'
b.	api^{ɣ}	vs.	apɪ ʊnggʊt
	fire		fire woodstack
	'fire'		'campfire'

c. bɒbiʸ vs. bɒbɪ səsamʔ
pig pig underbrush
'pig' 'wild boar'

The phrasal distribution of these medial and final forms is nearly identical to the distribution of alternating nasal-final forms described earlier in this section. Nouns exhibiting the vocalic alternation appear in medial forms with attributive adjectives, possessors, demonstratives, and (variably) with relative clause modifiers. In other environments, nouns surface in the phrase-final form. Additionally, active verbs exhibiting this alternation appear in the medial form when occurring with a NP direct object, whereas they surface in phrase-final form in other environments.

18. Phrasal alternation of high vowels in Jernih

Phrase Final		**Phrase Medial**
mᵇəliʸ (dɒʁɪ gʊndo)	vs.	mᵇəlɪ hape
ACT.buy (from Gundo)		ACT.buy cell.phone
'buy (from Gundo)'		'buy a cell phone'
nʊŋguʷ (halamʊ tɪgə aʁiʸ)	vs.	nʊŋgʊ nɪneʔ
wait (one.long three day)		wait grandmother
'wait (for three days)'		'wait for grandmother'

These data demonstrate that the phrasal alternation of word-final high vowels and the phrasal alternation of word-final nasals in Jernih is abstractly the same phenomenon, since the same conditioning environment is observed in both cases.

In summary, we have seen that phrasal alternations occur in diverse Malayic varieties spoken across Sumatra. These alternations exhibit a variety of phonological manifestations, including phonetic/phonological changes to both word-final vowels and codas. The alternations we describe broadly fit into two categories: *(i) Gradient alternations*: These are alternations which affect the phonetic realization of a single segment, are gradient in their realization, and are affected by factors such as speech rate, and so forth. Examples of this type of alternation include final nasals in Tanjung Raden and inglided vowels in Padang Minangkabau. *(ii) Allophonic alternations*: These are alternations between two distinct/discrete allophones. Examples of this sort of alternation include allophony of final nasals and high vowels in Jernih, and allophony of final nasals in Jangkat. For both of these varieties, speakers have clear judgements regarding which forms can occur in which environments.[3] In both Jernih and Jangkat, the phrasal distribution of forms reflects the prosodic organization of sentences.

Putting aside the differences between the gradient and phonological alternations, we have seen that both types of alternation have similar phonetic effects (e.g. increased oral occlusion in phrase-final nasal stops). We can infer that phonological alternations

[3] The fact that speakers are aware of this allophonic alternation between nasal stops and oral stops is not surprising, since both of these varieties have distinctive nasal and oral stop phonemes.

originated as phonetic alternations, which became reanalysed by speakers as phonological rules conditioned by phrase-level phonological structure.

Although the phrase-level alternations in the varieties we have described exhibit clear similarities in their grammatical distribution, we have also seen that the formal properties of the alternation differ to a great extent across varieties (marking of phrase final words involves insertion of *-ŋ* (Tapan), changes in vowel height/length (Sarolangun, Minangkabau), occlusion of final nasal stops (Jambi Seberang, Dusun Baru, Jernih Sarolangun), [+nasal, -cont.] → [-nasal] (Jangkat), etc.). In light of these differences, it is improbable that one could reconstruct a plausible proto-alternation from which all of these phrase-level alternations derive historically. We must, therefore, conclude that the phrase-level alternation developed independently in several varieties. Its development should not be seen as happenstance, however, since all of the phrasal alternations we have described share a general characteristic: the phrase final form exhibits more phonological content than its phrase medial counterpart. We believe that the development of this type of alternation is a consequence of a more general prosodic characteristic of the languages in the region: namely, that phrase final positions have high prosodic prominence, and are, therefore, more likely to develop more phonological content over time (e.g. via processes such as lengthening, diphthongization, excrescence, etc.).

References

Anderbeck, Karl. 2008. *Malay dialects of the Batanghari river basin (Jambi, Sumatra)* (SIL e–Books 6). Dallas: SIL International. (http://www.sil.org/resources/publications/entry/9245)

Yanti. 2010. *A reference grammar of Jambi Malay*. University of Delaware. (Doctoral dissertation)

Timothy Mckinnon
University of Maryland
mckinnon@udel.edu

Peter Cole
University of Delaware
pcole@udel.edu

Yanti
Atma Jaya Catholic University of Indonesia
regina.yanti@atmajaya.ac.id

Gabriella Hermon
University of Delaware
gaby@udel.edu

Possession marking in Nkep (East Santo, Vanuatu)

Miriam Meyerhoff

Abstract

This paper attempts to: (i) describe the structure of possessive NPs in Nkep (spoken in East Santo, Vanuatu); (ii) revisit the long-standing discussion about the relational versus classifier nature of Oceanic possessives; and (iii) offer some structural observations on a more recent debate over the underlying syntactic head of indirect possessives in Oceanic languages. It argues that the distribution of possessive marking in spontaneously produced narratives and oral histories suggests the Nkep indirect possessive markers function as semantic classifiers. It notes that the structure of possessed NPs is compatible with an analysis of the possessed N as the head (with the surface order of constituents realised by successive movement) or the possessive marker as a functional head of the phrase. It concludes by exploring how this discussion of synchronic variation in Nkep possessives might inform our understanding of the historical development of possessive marking systems in Oceanic over time.

1. Introduction

There is considerable debate over (mainly) the syntactic nature of possessive marking in Oceanic languages. The descriptive facts are clear enough – Oceanic languages all have more than one syntactic strategy for marking possession relations, though the complexity and richness of the systems differs across the region. Polynesian languages, for instance, have a much simplified system – in the case of Hawaiian, Baker's (2012) comprehensive analysis of the syntactic, semantic and pragmatic constraints on the *a/o* alternation argues that the *o* construction functions as the default in all regards. It is well-known, however, that some of the Melanesian languages have rich and productive systems, e.g. in Iaai there are numerous possessive classifiers. Ozanne-Riviere (1976) gives 23 forms for Iaai, though Dotte (2013) shows that some of these can no longer be elicited with picture stimuli and that younger speakers today use only a small subset of the possible forms (Figure 1). Whether this is a developmental fact, a reflex of cultural change or a result of language endangerment/shift is the concern of her detailed study and subsequent work.

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 169–179

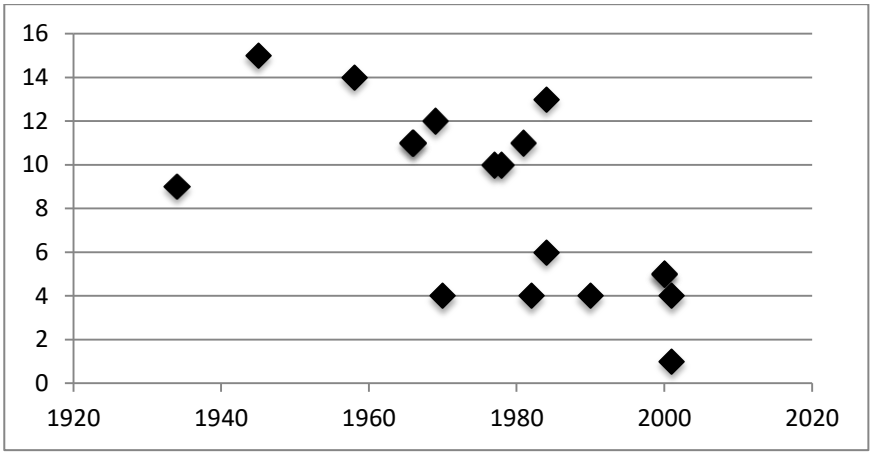


Figure 1: Number of different possessive classifiers produced by speakers of Iaai by year of birth. (Pearson correlation shows significant decrease in classifiers produced as year of birth increases, r^2 = -0.66.) Source, Dotte (2013: 278).

In Vanuatu, languages have fewer possessive categories than in Iaai, though some esoteric systems have been reported (Guy 1974 reports distinct possessive markers for edible Ns, potable Ns, 'shadows and vomit', smells, and general Ns in Sakao). In this paper, I outline the system for marking possessives in Nkep, a language closely related to Sakao, spoken by about 1000 people in and around the village of Hog Harbour, East Santo (Central Oceanic, Vanuatu).

Like most of the languages of Vanuatu, possessives in Nkep can be expressed through two main structures: the first, sometimes referred to as *direct possession,* expresses the relationship of possessum-possessee through the affixation of a suffix expressing person and number features directly on the possessed noun; the second, sometimes referred to as *indirect possession,* uses a periphrastic construction composed of the possessed NP followed by an independent possessive morpheme (expressing language-specific semantic features) and this morpheme bears the inflectional marking for person and number of the possessor. I draw on spontaneously produced data occurring in conversations and narratives recorded between 2011 and 2014 in order to:

(i) describe possessive marking in Nkep and
(ii) offer some observations on the syntax and semantics of the Nkep system.

1.1. Classifiers or heads

Most of the heat and light on the subject of possessives in this region has been expended on the question of (a) what the fundamental organising principles are that underlie the distinction between direct and indirect possession, and (b) whether the modifiers used in indirect attributive possessive constructions should be considered classifiers modifying different noun classes, or are, in fact, themselves heads of the possessive phrase. This second question problematises the internal structure (head-dependency relationship) of possessive constructions.

Pawley & Sayaba (1990) note widespread consensus since the 1970s on the fluidity of the Oceanic direct/indirect possession distinction, arguing that whether a given NP will be marked with direct or indirect possession is an empirical question that can only be resolved in context, i.e. it is pragmatically, not semantically, determined. Nevertheless, even Pawley & Sabaya's discussion shows that some nouns fall very firmly into one category rather than another, and the inventory of genuinely 'fluid' nouns is perhaps small. In this paper, I will draw on how possessives are actually used in natural speech to argue that while speakers have the capacity to play with the semantics of some nouns, in practice, virtually all nouns in Nkep can be unproblematically assigned to either the class of directly or indirectly possessed Ns.

Palmer & Brown (2007) and Lichtenberk (2009) summarise the essential points against and in favour of treating the possessive constituent in indirect constructions as a classifier or the head of the phrase. Palmer & Brown argue for Kokota (Solomon Islands) that indirect possessives do not have the cross-linguistic properties of classifiers and syntactically are N-N compounds. Under their analysis, the second N expresses semantic information about the first N and the second N bears the same possessor marking found with directly possessed Ns (thereby, as Lichtenberk 2009 points out, obviating the formal distinction between the two types of possession). Brown (2015) amends the Palmer & Brown analysis for Nasioi (Papuan, Solomon Islands), arguing that the head of the Nasioi possessive phrase is a determiner (or other functional constituent, not a noun). Lichtenberk's (2009) response to Palmer & Brown defends his earlier analysis of the constructions as head nouns modified by a classifier, and accurately (I think) identifies the lack of agreement among linguists about what constitutes a 'head' as the principal point of divergence. He concludes by quoting Matthews's (2007: 61) mordant observation: "Arguments [about heads] are often at cross purposes; and, if one view [of what constitutes a head] is indeed right and the other is indeed wrong, the facts alone seem scarcely quite sufficient to decide which."

This paper can hardly hope to resolve all aspects of the debate. However, following Pearce (2010), I find it helpful to situate the analysis of possessives in Nkep in relation to the overall structure of the NP, before providing any kind of structural account of the possessives per se. Hence, in the next section, I review the relevant descriptive facts.

2. Possessives in the context of noun phrases

Nkep is an SVO language. NPs are left-headed with an internal structure as in (1).

1. N – Adj – $Dem^{(n)}$ – Num

2. nwakër kikri momur thël[1]
   female small DEM[2] three
   'those three girls' [elicited]

Only *tha* 'some' functions as a canonical article[3] and does not combine with numerals:

3. nwakër kikri tha
   female small some
   'some girls'

4. nwakër kikri lön
   female small five
   'five girls'

5. *nwakër kikri lön tha/*nwakër kikri tha lön [elicited]

Possessive suffixes that index person and number are affixed to the head noun in direct possession constructions. The head nouns allowing direct suffixation fall into the class of what is generally considered 'inalienable' possessions, as in (6–8).

6. nam-warho-p nwase-ngur
   1.S-tell-PERF name-3.P.POSS
   'I'm telling you their names.' (Janet N)

7. n-iav-tvaac-p vooc hür neire-c m-cer wesi
   1.S-NEG-call-PERF too because blood-1.S.POSS 3.S-run strong
   'I couldn't call out because I had lost so much blood.' (Fred I)

8. walthiü-c aal ru
   child-1.S.POSS only two
   'just my two children' (Anathiel K)

The full paradigm of possessive suffixes is shown in Table 1.

[1] Examples are either from my recordings (and identified by speaker) or elicited with Shirley Warput (identified as [elicited] in the text). I use the orthography being adopted by the community for vernacular education: <th> for /ð/, <ng> for /ŋ/, <c> for /ɣ/, <oo> for /ɔ/, <aa> for /ɑ/, <ü, ë, ö> for front rounded vowels.

[2] Nkep has a rich and semantically complicated system of demonstratives which I do not yet fully understand. The glosses will not attempt to differentiate among them for the purposes of this paper.

[3] *(i)tei* 'one, a' seems to have idiosyncratic properties. For instance, it can occur before a demonstrative in a complex NP, cf. *nwakër kri tei mom* 'that one small girl' vs. **nwakër kikri thël momur* 'those three girls'.

Table 1: Possessive suffixes in Nkep

		SINGULAR	DUAL	TRIAL	PLURAL
1	INCL	—	-ro, -ru		-rthël
	EXCL	-c	-camru		-cam, -camthël
2		-m	-caru		-cei
3		-n	-ngaru	-ngarthël	-ngur

Indirect possession distinguishes between entities that are consumable possessions (9–11) and those that are general possessions (12–13).

9. cam-iel-p nov na-ngur
   3.P-strike-PERF pig food-3.P.POSS
   'They killed their pig.' i.e. 'They killed a pig for them [to eat].' (George N)

10. nithel na-ngur
    banana food-3.P.POSS
    'their bananas' (Etta S)

11. nac-el-ke-lam netnaar he-m
    1.S.IRR-IMPERF-take-come food GEN-2.S.POSS
    'I will bring your food' (Christina N)

12. nkar he-ngur
    fire GEN-3.P.POSS
    'their fire' (Etta S)

13. lom ha-n
    house GEN-3.S.POSS
    'his house' (Janet N)

Kinship terms with complex internal structure (N-N compounds as in (14) or DPs with complex internal structure as in (15)) and lexical items borrowed from Bislama (including kinship terms equivalent to those which normally would be directly possessed) occur only in indirect possessive constructions like (14–15), or in a third periphrastic construction: $N_{possessum}$ *hen* $N_{possessee}$ shown in (16).

14. a. nvaat-wari he-camthël ... cam-el-ro
       elder-male GEN.POSS-1.P.EXCL 3.P-IMPERF-stay
       'our elders/fathers were waiting' (Fred I)

    b. netnar nera ho-c
       food pig GEN-1.S
       'my pig's food/my pig food' (Freddy W)

15. lakren secren ha-n
    tomorrow morning GEN-3.S
    'the next day' (John T)

16. papa hen nathön mom
    father of friend DEM
    'the father of that friend/that friend's father' (Leci W)

The last construction is not restricted to borrowed nouns, as shown in (17):

17. nathngön am ni m-el-cen nra hen wam
    person DEM 3.S 3.S-IMPERF-eat pig of him
    'the man who was eating that guy's pig['s food]' (Freddie W)

There are numerous similarities between the Unua and Nkep possessives (even though Nkep and other NE Santo languages stand conspicuously outside the dialect chains linking SE Santo, Malo, N Malakula and Ambae, Tryon 1996). For Unua, Pearce (2010) talks about indirect possession as involving an N followed by 'a linker' analogous to the *na-* and *he-* morphemes shown in (9–13) for Nkep. Like Unua, Nkep possessives allow the speaker to 'wrap up' both the possessed N, a modifying adjective and a demonstrative within the scope of the possessive (18). Also like Unua, Nkep seems to occasionally allow the possessive marker to be interleaved with demonstratives and numerals (19–20) (cf. examples 31, 37b in Pearce 2010).

18. nwakër kri mom nio ho-c m-el-ele
    female little DEM DEM GEN.POSS-1.S 3.S-IMPERF-swim
    'My daughter was bathing.' (Anathiel K)

19. m-hün-p aal nason ha-n nio
    3.S-pull-PERF only arrow GEN-3.S.POSS DEM
    'He just pulled his/its arrow.' (Etta S)

20. ce brat ho-c ru
    and brother GEN-1.S two
    'and my two brothers' (Janet N)

It is unclear whether the same kinds of pragmatic/information structure constraints operate on examples like (19) and (20) that Pearce (2010) discusses for similar examples in Unua. Such tokens are rare in my corpus.

3. The relational or categorical quality of Oceanic possession

As noted in §1, linguists have long noted that "there is a considerable element of choice in possessive-marking for some nouns" (Pawley & Sayaba 1990: 165). This 'element of choice' is, of course, a function of the noun's semantics. Thus, linguists can elicit contrasts such as 'my coconut (for drinking)', 'my coconut (for eating)' and 'my coconut (for sale)' with edible, drinkable and general indirect possession constructions. Or contrasts such as 'my tooth (in my head' (direct possessive construction) and 'my tooth (I wear as a necklace)' (indirect possessive construction). We may concede that this does show an element of choice, without conceding that that it is considerable. While it is possible that speakers might co-opt the direct-indirect possession distinction to mark discourse-pragmatic relations (as Baker 2012 argues the *a/o* distinction is used

to signal foreground and background information in Hawaiian), this has not been claimed for the direct/indirect alternation that is possible in Melanesian languages. Rather, arguments rest on the capacity to elicit such alternations from speakers. While elicitation is useful as a tool for probing the limits of grammar, it does not tell us what the everyday grammar of speakers actually is.

3.1. In defence of classification: A quantitative perspective

For Nkep speakers, the optionality that has been central to discussions about whether the distinction between direct and indirect possession is a (semantic) property inherent to the noun, or a characteristic that emerges from the pragmatics of use is negligible. For all practical purposes, nouns seem to fall into one and only one of two classes.

I extracted all possessive constructions from my Nkep corpus (consisting of narratives, oral histories and some process texts – there were next to no possessives in the conversational speech recorded). For the purposes of this study, I focussed on the direct and indirect constructions (i.e. excluding 52 tokens of possessive constructions using forms such as (16)). There were 87 different possessed noun types, in 319 possessive constructions. Eighty-five of the noun types occur in only one construction. The only exceptions were 'child' and 'year'. Nkep can use the lexeme 'child' to metaphorically refer to an arrow (that is, an 'arrow' can be 'the child' of a bow) and 2/6 tokens of such uses of *walth-* 'child' were indirect possessives referring to arrows. The remaining four were in direct possessives as would be expected for 'child'. In reference to a human child, *walth-* occurred once (out of a total of 26 tokens) in an indirect construction, and this was in a case of double marking, shown in (21):

21. i mav-cöth walthi-c ho-c
    2.S NEG-see child-1.S.POSS GEN-1.S.POSS
    'Have you not seen my child?' (Shirleyana W)

Shirleyana was 8 years old when she produced this sentence and it is the only such doubly marked possessive in my corpus (cf. Barth, Meyerhoff & Schnell in prep. which reports a number of such constructions in the severely endangered Matukar Panau – there, we argue that the double marking in Matukar Panau is evidence of language attrition and language shift). It is unclear therefore how much we should read into it.

The lexeme *hie* 'year' occurs in two different indirect possessive constructions. Etta S (a woman in her 50s) produced it twice as *hie-c* 'my age' and Janet N (a woman in her 30s) produced it three times in an indirect construction (e.g. *nhie ning ieth* 'year food.1.S.POSS four', i.e. 'I was four'). This treatment of years as something to be consumed is idiosyncratic to Janet N in my corpus, but one can see how it might make sense – time eating away at us as we grow older. Alternatively, it might be that years are seen as only existing in the telling of them. This would link the treatment of *nhie* 'year' with another lexeme that is often possessed with the consumable possessor, namely *nesaru* 'talk, speech'. This may suggest that what I have (following tradition) been calling a 'consumable' possessor is, in fact, more accurately conceived of as [+oral].

This distribution of possessive constructions in natural, spontaneous speech strongly suggests that nouns (in practice) fall into clear semantic classes. Hence, while acknowledging the potential for some nouns to be modified in different ways (i.e. for possession to be relationally defined as most of the Oceanic linguistic tradition would agree), we might want to consider how much our grammars should reflect actual usage, as well as potentialities. The discourse semantics of possession in Nkep suggest that the general/consumable morphemes function as rather canonical classifiers, especially if we embrace "a more function-based approach to the study of classifiers" as Grinevald (2000: 86) encourages us. Elicited data is useful for telling us what speakers *can* do; usage data tells us what speakers *do* do. Since we know that grammaticalisation (and, in general, the seeds of language change) lie in the synchronic variation which characterises speakers' use of a language, the usage data affords us a different perspective on this long-standing exchange in the literature. I will return to this point in the final section.

While the syntactic arguments about the locus of agreement in Palmer & Brown (2007) and Brown (2015) do also apply to Nkep, it is not clear to me that their analysis of the possessive marker as the head of the phrase (either construed as a nominal, in Palmer & Brown, or as a functional head of a DP, in Brown) necessarily follows. Agreement marking on the possessive marker could indicate that in Nkep indirect possessives, the general/consumable morpheme is the structural head of the possessive phrase, but it could also be consistent with a wrapping analysis of sequential head movement posited by Pearce (2010) for Unua. I have no examples of possessed NPs in Nkep that consist only of the possessive marker + inflection as in the Solomon Island languages reviewed by Palmer & Brown (2007) and Brown (2015) – there is always some kind of other potential head, e.g. a demonstrative if not a lexical N.

Palmer & Brown's (2007) argument against the possessed N as the head of the phrase rests heavily on the failure of Oceanic possessive markers to satisfy a range of criteria. They derive these from their principal source on classifiers, Grinevald (2000). However, their criterial properties do not seem to me to do full justice to Grinevald's nuanced cross-linguistic review of different types of classifier systems. Specifically, the criteria that they derive from Grinevald (2000) for identifying classifiers (Palmer & Brown 2007: 203) combines properties that Grinevald (2000) identifies as being associated with several different types of classifier systems. It is hardly surprising that Oceanic possessive markers fail to demonstrate the characteristics of all the different types of classifier systems Grinevald reviews – the only relevant criteria should be those associated with what Grinevald calls 'genitive classifiers'. Here, Palmer & Brown place considerable weight on Grinevald's statement that genitive classifiers "[select] a limited set of nouns of the language for classification" (2000: 66) and argue that Oceanic possessives fail to meet this standard because "[i]n these languages, every common noun may be possessed." This is a very literal reading of one sentence in Grinevald's exploration of classifier typologies – my reading of the whole chapter suggests that when she observes that the genitive classifiers are associated with a limited set of nouns she is clearly talking about the semantically specific genitive classifiers (e.g. edible, potable, wearable), and she is not precluding the possibility that a language may have a general or default classifier which is much less particular in its lexical associations. Indeed, Grinevald (2000: 84) suggests that such general classifiers

may be a good diagnostic of how old a classifier system is. This leads me to the final observations I will make about the possible historical significance of this preliminary investigation of Nkep possessives.

5. Broader significance of these findings

It seems clear that Oceanic languages progressively lost the direct possession constructions. Hence, Polynesian languages have only an analogue of the indirect possessive system that has been reconstructed for Proto-Oceanic. Blust's (2009) work suggests that the shift from the Proto-Oceanic direct/indirect system to the (Polynesian) indirect-only system unfolded over a period of approximately 7000 years. As we know, evidence of such change only emerges through a comparison of related languages and the signal is measured over extended periods of time. Yet the variationist sociolinguistic tradition has shown us repeatedly that it is possible to discern the early stages of diachronic change in synchronic patterns of language use. For this reason, it is valuable to explore synchronic data on how speakers use the potential that their grammar affords them. By adding sociolinguistic data to theoretical (e.g. Palmer & Brown 2007; Lichtenberk 2009; Pearce 2010) and (quasi-)experimental (Dotte 2013) discussions of the nature of possession marking in Oceanic languages, we can derive some sense of what historical processes might have given rise to the patterns found across the language family as a whole today. Indeed, by further triangulating this perspective with a synchronic cross-linguistic comparison of languages that differ in their relative vitality, we may gain an even more informed sense of how this particular aspect of the diversification of Oceanic languages played out in speech communities (Barth, Meyerhoff & Schnell in prep.).

Persistence (of form and meaning), being a quality of grammaticalisation, means that speakers may well be able to produce forms showing pragmatic fluidity *for a few nouns* even when, in practice, the variability is not productive. This means that we should be cautious about building crucial aspects of the grammar based on highly infrequent, and possibly obsolescing, usage patterns. This preliminary survey of possessive marking occurring in naturally occurring speech events in Nkep suggests that it would be perfectly reasonable to talk of the indirect possession constructions as being classifiers.

The question of what the actual head of the possessive phrase is remains for further examination. I have noted that the data presented here seems compatible with either an analysis of the possessive phrase in which the head N successively raises or an analysis in which the possessive is the functional head of the phrase (with the possessed N an external argument). Clearly, there is a great deal more work to be done to reconcile functional and formal arguments about possessives in Nkep and in Oceanic languages more generally. By adding data on vernacular use (from a variationist sociolinguist's perspective), this paper has added to the ongoing debate and been able to reframe some of the discussion in ways that may prove to be more productive and more focussed in the future.

Acknowledgements

My thanks to Sarah Truesdale for her help in analysing Nkep possessives. Also to numerous people in Hog Harbour, especially the family Warput. It has always been a pleasure to discuss the syntax of Vanuatu languages with Liz Pearce, and I've been grateful for her challenging and trenchant questions over the years, in particular for her interest in and encouragement to write up my possessives data. Some of this data was presented at New Ways of Analysing Variation—Asia-Pacific 4 and at the New Zealand Linguistic Society conference; I thank audiences there for their helpful comments and questions. The inevitable faults and deficiencies in this paper are mine alone.

References

Baker, C.M. Kaliko. 2012. A-*class subject effect: A pragmatic and discourse grammar approach to* a- *and* o-*class genitive subject selection in Hawaiian.* Mānoa: University of Hawai'i. (Doctoral dissertation)

Barth, Danielle & Meyerhoff, Miriam & Schnell, Stefan. In prep. The emergence of indirect possession in Oceanic languages. Australian National University, Victoria University of Wellington and University of Melbourne. (Unpublished ms.)

Blust, Robert. 2009. *The Autronesian languages*. Canberra: Pacific Linguistics.

Brown, Jason. 2015. Headedness in Nasioi possessives. University of Auckland. (Unpublished ms.)

Dotte, Anne-Laure. 2013. *Le Iaai aujourd'hui: Évolutions sociolinguistiques et linguistiques d'une langue kanak de Nouvelle-Calédonie (Ouvéa, Îles Loyauté).* Lyon: Université Lumière Lyon 2. (Doctoral dissertation)

Grinevald, Colette. 2000. A morphosyntactic typology of classifiers. In Senft, Gunter (ed.) *Systems of nominal classification,* 50–92. Cambridge: Cambridge University Press.

Guy, Jacques B.M. 1974. *A grammar of the northern dialect of Sakao* (Pacific Linguistics B-33). Canberra: Pacific Linguistics.

Lichtenberk, Frantisek. 2009. Oceanic possessive classifiers. *Oceanic Linguistics* 48(2). 379–402.

Matthews, Peter H. 2007. *Syntactic relations: A critical survey*. Cambridge: Cambridge University Press.

Ozanne-Rivierre, Françoise. 1976. *Le Iaai: Langue mélanésienne d'Ouvéa, Nouvelle-Calédonie: Phonologie, morphologie, esquisse syntaxique*. (Langues Et Civilisations à Tradition Orale 20). Paris: Société d'études linguistiques et anthropologiques de France.

Palmer, Bill & Brown, Dunstan. 2007. Heads in Oceanic indirect possession. *Oceanic Linguistics* 46(1). 199–209.

Pawley, Andrew & Sayaba, Timoci. 1990. Possessive marking in Wayan, a Western Fijian language: Noun class or relational system? In Davidson, Jeremy H.C.S. (ed.) *Pacific island languages: Essays in honour of G.B. Milner,* 147–171. London/Honolulu: SOAS, University of London and University of Hawai'i Press.

Pearce, Elizabeth. 2010. Possession syntax in Unua DPs. In Mercado, Raphael & Potsdam, Eric & Travis, Lisa deMena (eds.) *Austronesian and theoretical linguistics*, 141–162. Amsterdam/Philadelphia: John Benjamins.

Tryon, Darrell. 1996. Dialect chaining and the use of geographical space. In Bonnemaison, Joël & Huffman, Kirk & Kaufmann, Christian & Tryon, Darrell (eds.) *Arts of Vanuatu,* 170–181. Honolulu: University of Hawai'i Press.

Miriam Meyerhoff
Victoria University of Wellington
miriam.meyerhoff@vuw.ac.nz

On Cia and C-final bases in Polynesian

Yuko Otsuka

1. Introduction

One of the most oft-discussed topics in Polynesian linguistics is the so-called "passive" suffix Cia.[1] Any learner of Māori, for example, is confronted with irregular "passive" forms, such as *inu-mia* 'to be drunk', *tiro-hia* 'to be seen', *hopu-kia* 'to be caught', and so on. All these "passive" forms end in /a/ or /ia/, the latter of which is preceded by a seemingly randomly selected consonant. In fact, Biggs 1969 lists twelve forms as "passive" suffixes in Māori: *-a, -ia, -hia, -ina, -kia, -mia, -na, -nga, -ngia, -ria, -tia,* and *-whia*. In the Polynesian linguistic literature, these suffixes are collectively referred to as Cia, with C representing a variable consonant. Based on their extensive distribution across Polynesian, these variants of Cia are regarded as reflexes of a Proto Polynesian (PPN) suffix *Cia. In this squib, I revisit this claim by examining the distribution of various forms and functions of Cia in three Polynesian languages, Hawaiian, Māori, and Tongan, and show that the suffix *Cia as such cannot be reconstructed for PPN.

2. Previous analyses of Cia

While the presence of Cia cognates across Polynesian has long been recognised by linguists, the historical development of PPN *Cia and its reflexes is still somewhat an open question. Some possible sources have been suggested for each of the components of *Cia (the initial consonant, /i/, and /a/), but it is not entirely clear how the relevant items were put together to form a single morphological unit in PPN.

2.1. Source of C in Cia

It is widely accepted that the initial C of Cia is historically the final consonant of the base. As Polynesian languages do not permit closed syllables, reflexes of consonant-final forms of Proto Oceanic (POC), an ancestral language of PPN, are vowel-final in Polynesian languages, e.g., Tongan *inu* 'to drink' vs. POC *inum, *tanu* 'to bury' vs. POC *tanum, and *piki* 'to adhere' vs. POC *pikit. Rather nicely, the initial C of Cia corresponds to the final C of the POC form of the base, as in Tongan *inu-mia* and *tanu-mia*, and *piki-sia* ([t] becomes [s] before [i] in Tongan). At some point in history, systematic loss of word-final consonants led to a reanalysis of suffixed forms in which the final consonants were retained; namely, historical base-final consonants were

[1] I use the label "passive" in double quotes, as there are some reasons to believe what is called "passive" in Polynesian languages may not be best analyzed as such (Otsuka 2011; 2012).

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 181–189

reanalysed as the initial consonant of the suffix, e.g. *inum-i became *inu-mi. Loss of final consonants is a recurrent phenomenon across Austronesian and is assumed to have taken place independently in multiple languages (Blevins 2004). With regard to Polynesian languages, final consonants were believed to have been lost in Proto Eastern Oceanic (PEO), a language ancestral to all Polynesian languages as well as the two other members of Central Pacific subgroup, Fijian and Rotuman (Clark 1976; Pawley 2001). Thus, separation of C from the base is presumed to have occurred in PEO.

2.2. Source of /i/ in Cia

Polynesianists also agree that the source of /i/ of Cia can be traced as far back as to a POC transitive suffix -*i. Clark (1976) reconstructs a transitive suffix *Ci for PPN and proposes that it is the direct source of /i/ (more precisely, of Ci) in PPN suffix *Cia. Pawley (2001) agrees with this analysis. Its closer ancestor is PEO transitive *Ci, also reconstructed with the initial C due to the presumed reanalysis of base-final consonants. An intriguing side note is that PEO transitive *Ci was used when the verb had a specific object (Clark 1976: 77). When the object is third person singular, a third person singular object clitic pronoun *a attaches to the verb, yielding a sequence Ci-a, as seen in (1) and (2).

1. Tasiriki (Clark 1976: 77)
   tani-si-a tamana
   weep-Ci-him his.father
   'mourned for his father'

2. Bauan (Clark 1976: 77)
   gunu-va(<vi-a)na yaqona
   drink-Ci-it the kava
   'drink the kava'

2.3. Two possible sources of /a/ in Cia

Clark (1976) reconstructs three suffixes for PPN: *Ci 'transitive', -*a (function unknown), and *Cia. He proposes that *Cia was formed by combining *Ci and *-a, but remains agnostic about the source of *-a as well as how *Cia acquired distinctive semantic and syntactic properties. Although POC third person singular object pronominal clitic *a is hinted at as a likely candidate, as seen in the above examples, Clark does not make any commitment to this hypothesis.

On the other hand, Pawley (2001) argues that the source of /a/ in Cia is not the POC third person singular object pronominal clitic (as suggested by Churchward 1928; 1951), but the PEO stativiser *-a. Citing data from southeast Solomonic languages, Pawley shows that the PEO transitiviser *Ci and stativiser *-a were productively used to derive a "past participle", an intransitive verb with patient subject. Table 1 provides relevant data from Arosi (Pawley 2001: 200–201).

Table 1: Arosi Ci-a forms

Base	Base-Ci	[Base-Ci]-a
age	*age-ri*	*ageri-a*
'thatch (v)'	'thatch something'	'thatched'
angi	*angi-si*	*angisi-a*
'cry'	'cry for something'	'cried for'
awanga	*awanga-hi*	*awangahi-a*
'open up, expose'	'uncover something'	'uncovered'
huna	*huna-ʻi*	*hunaʻi-a*
'tie, bind'	'tie or bind something'	'bound'

Pawley proposes that these past participles that were retained in PPN and the two PEO suffixes were reanalysed as a single unit, *Cia, as a marker of "derived passive": PEO *inu-mi-a > PPN *inu-mia.

3. What really happened to Cia?

Both Clark and Pawley reconstruct a suffix *Cia for PPN, assuming that base-final consonants were reanalysed in PEO as the initial C of the suffix. Contrary to their analysis, I argue that the relevant reanalysis took place at a later stage in individual languages, rather than in PEO or even PPN. I agree with Clark's (1976) and Pawley's (2001) reconstruction of the two independent suffixes, transitiviser -*i (-*Ci for them) and stativiser -*a, which creates a predicate denoting some affected state. The sequence Cia arose when a C-final base was simultaneously suffixed by both -*i and -*a. I argue, however, that this sequence did not constitute a morphological unit as a single suffix in PPN. Rather, I propose (a) that PPN had C-final bases; and (b) that PPN had two productive suffixes, transitiviser -*i and stativiser -*a. Below I discuss evidence to support this analysis from Māori, Hawaiian, and Tongan.

3.1. C-final stems in Māori

Māori exhibits multiple reflexes of *Cia. Collectively, Cia in Māori is quite productive in that (a) all verbs have a corresponding Cia form and (b) the meaning of the suffixed form is relatively transparent, having the semantic effect of passive. If we reconstruct a suffix *Cia for PPN, various realizations of Cia in Māori must be understood as an instance of lexically conditioned allomorphy (Hale 1968; Clark 1976). From a perspective of language acquisition, a more efficient system would be one that has a single suffix *-ia* and two allomorphs (C-final and C-less forms) for some bases.

Postulating C-final bases for Māori may seem problematic, because closed syllables were not permitted in Māori as in other Polynesian languages; at least all lexical items, including particles, end in a vowel and contain no consonant clusters. However, it should be noted that even with a C-final base, the relevant phonotactic constraint is observed at the word level. As suffixes are always vowel-initial (and vowel-final), suffixed forms would never contain a consonant cluster or end in a consonant.

Base allomorphs are easily acquired in Māori since the alternation is observable with

another highly frequent suffix, the nominaliser *-anga* (conventionally called -C*anga* in Māori linguistic literature). C-final allomorphs are thus reinforced in multiple morphological contexts. Again, if we assume C is part of the suffix (as is conventionally assumed), we would have to postulate another large set of lexically conditioned allomorphs. The Māori facts can be explained in a much simpler manner if we do not assume that C was part of the suffix *Cia (cf. de Lacy 2003). I therefore propose that PPN had allomorphs for some bases (i.e., reflexes of POC C-final bases) and two suffixes, transitive -*i and stative -*a. That is, PPN did not have a suffix of the form *Cia; the phonological sequence Cia arose when the two suffixes simultaneously attached to a C-final base, e.g., *inum-i-a. If this analysis is correct, (historical) base-final consonants have not yet been reanalysed as part of the suffix in Māori. Accordingly, the presumed reanalysis could not have occurred in PPN.

3.2. Radical reanalysis of Cia in Hawaiian

The situation is drastically different in Hawaiian, in which "passive" is formed by means of a particle *ʻia*. It is a particle rather than suffix, as some other items such as an adverb may intervene between *ʻia* and the verb (Elbert & Pukuʻi 1979: 83).

3. Hawaiian (Elbert & Pukuʻi 1979: 83, my gloss)

   a. Ua hānai maikaʻi ʻia.
      PFV feed well Cia
      '(He) was well fed.'

   b. Ua ʻike ʻole ʻia ke keiki.
      PFV see NEG Cia the child
      'The child was not seen.'

This particle was likely developed in the following manner: first, the C of the PPN Cia forms was reanalysed as the initial consonant of the suffix (e.g., *inum-i-a > *inu-mi-a); second, the two suffixes were reanalysed as a single suffix (e.g., *inu-mi-a > *inu-mia); third, various forms of Cia suffixes were reduced to a single form -*ʻia; and finally, the suffix -*ʻia was reanalysed as a particle, gaining a higher degree of lexical independence. The Hawaiian reflex of the PPN nominaliser is also a single, fixed form *ʻana* with a glottal stop and functions as a particle. This suggests that the separation of base-final C applied across the board, as the relevant suffix gained a higher degree of productivity.

Other Cia forms exist in Hawaiian as suffixes (*-hia*, *-kia*, *-lia*, *-mia*, *-nia*, and *-a*), but they are generally rare and often have unpredictable meanings (Elbert & Pukuʻi 1979). These rare occurrences of Cia suffixes can be regarded as vestiges of the PPN sequence of base-final C + *i + *a. Elbert & Pukuʻi (1979: 86) note that these Cia forms are "often redundantly followed by [the particle] *ʻia*". This phenomenon of Cia-doubling can be regarded as evidence for the declining function of the suffix Cia, which is gradually taken over by the particle *ʻia*.

3.3. Cia forms in Tongan

Cia forms in Tongan present a more complex picture. Tongan exhibits clear reflexes of PPN -*i and -*a. The Tongan suffix -a is highly productive. It attaches to a nominal base to derive an adjective meaning 'full of BASE': *niu-a* 'abounding in coconut' from *niu* 'coconut', *efu-a* 'dusty' from *efu* 'dust', *ika-a* 'abounding in fish' from *ika* 'fish', *lango-a* 'infested with flies' from *lango* 'fly', and *ongoongo-a* 'famous' from *ongoongo* 'news'. It also attaches to a verbal base to derive an adjective denoting an affected state as in *fiekai-a* 'hungry' (from *fie 'want'* and *kai* 'eat'), *fieinu-a* 'thirsty' (from *fie* 'want' and *inu* 'drink'), *fiemohe-a* 'sleepy' (from *fie* 'want' and *mohe* 'sleep'), as well as *hōnge-a* 'to suffer from starvation' (from *hōnge* 'famine-stricken'), *to'o-a* 'to be mentally carried away' (from *to'o* 'take'), and *'ilo-a* 'to be well-known' (from *'ilo* 'know').

Reflexes of PPN -*i are found in many fossilised forms in Tongan and preceded by various consonants, e.g., *huufi* from *huu* 'to enter', *kaiha'asi* from *kaiha'a* 'to steal', and *'anuhi* from *'a'anu* 'to spit'. Consonants in these forms are clearly reflexes of POC base-final C. PPN -*i is also reflected in a productive transitive suffix, *-'i*, preceded by a glottal stop. The presence of a consonant before /i/ in this suffix suggests that historical base-final C must have been reanalysed as part of the suffix in Tongan.

Tongan also has an independent suffix *-'ia* (also with a glottal stop) with some degree of productivity. This suffix attaches to an adjectival base and creates a verb meaning 'consider something to be BASE': *ifo* 'tasty' > *ifo-'ia* 'to find tasty', *faka'ofo'ofa* 'beautiful' > *faka'ofo'ofa-'ia* 'to regard as beautiful', *vale* 'foolish' > *vale'ia* 'to regard as foolish', *sai* 'nice, good' > *sai-'ia* 'to like', *hela* 'tired' > *hela-'ia* 'to be exhausted (with)', *fakalielia* 'disgusting' > *fakalielia-'ia* 'to loathe', and so on. Churchward (1953) treats this suffix as a variant of the suffix *-a*, as *-'ia* has a similar semantic effect when attached to a nominal base, e.g., *kakai* 'people' > *kakai-'ia* 'populated', *kumaa* 'rat' > *kumaa-'ia* 'infested with rats', *koloa* 'goods, wealth' > *koloa-'ia* 'wealthy', and *mamahi* 'painful' > *mamahi-'ia* 'in pain'. In addition, *-'ia* and *-a* can be used interchangeably with some nominal bases, e.g., *ika-'ia* 'abounding in fish', *niu-'ia* 'abounding in coconuts', *lango-'ia* 'infested with flies', etc. (Churchward 1959). Based on the meaning, the source of this suffix *-'ia* is PPN -*i and -*a. While it is not clear if the Tongan -'ia is a single suffix or still bi-morphemic, the existence of this and other productive suffixes, which incidentally all begin with a glottal stop, suggests that the original base-final Cs have been reanalysed as the initial C of the suffix in Tongan.

Tongan also exhibits a bunch of other Cia forms, all of which appear to have been lexicalised. These Cia forms fall into two classes based on their semantics: affected state (Table 2) and transitive (Table 3). This is intriguing in the light of Pawley's (2001) discussion on the competing sources of /a/ in Cia: PEO stativiser -*a and POC third person singular object pronominal clitic *a. Tongan Cia forms suggest that there were indeed two sources of /a/ in Cia.

Table 2: Tongan "affected state" Cia forms

Tongan		POC	
tonu-hia	'innocent'	*tonuq	'straight'
ʻau-hia	'carried by a current'	*qaRus	'current'
piki-sia	'be stuck'	*bikit	'sticking to'
tonu-mia	'to be covered accidentally with earth, buried by a fall of earth or lava'	*tanum	'bury'
mafa-sia	'to be weighed down, burdened'	*mapat	'heavy'

Table 3: Tongan transitive Cia forms

Tongan		POC
ʻanu-hia	'to mess up by spitting on'	*qanus 'spit'
huu-fia	'to enter surreptitiously (of a bird or a thief)'	*huru 'to enter'
inu-mia	'to consume by evaporation (of the sun) or absorption (of the earth)'	*inum 'to drink'
tengi-hia	'to weep for'	*tangis 'to weep'
hoko-sia	'to reach, meet with'	*sogot 'to join'

For both classes, there is some evidence to suggest that these Cia forms are lexicalised. First, the base *tanu* 'to bury' undergoes phonological change in the corresponding Cia form as in *tonu-mia*, though the alternative form *tanu-mia* is also available. Similarly, *tangi* 'to cry' becomes *tengi* in its Cia counterpart. Second, some Cia forms are not entirely semantically predictable, but rather have a specific meaning: *huu-fia* 'to enter surreptitiously (of a bird or a thief)', *inu-mia* 'to consume by evaporation (of the sun) or absorption (of the earth); to undergo, endure or experience (pain or suffering)', and *ʻanu-hia* 'to mess up by spitting on'.

Note also that some of these Cia forms involve a bound root, i.e., a base that only occurs in suffixed (derived) forms such as *mafa* in *mafasia* 'weighed down, burdened'. While this base is clearly a reflex of POC *mapat 'heavy' and occurs in other forms, e.g., *mamafa* 'heavy' and *mafataki* '(of rain) to be heavy, (of work) felt or regarded as heavy or difficult, (of body) feel heavy or lacking energy', it never occurs as a free morpheme. Another example of this is *ʻanu* in *ʻanuhia* 'to mess up by spitting on'. It is a reflex of POC *qanus 'spit, spittle', but only occurs in derived forms such as *ʻaʻanu* 'to spit' and *ʻanuhi* 'to spit on, spit out'. This suggests that these suffixed forms were likely formed in PPN with a C-final base and later fossilised in Tongan.

Tongan transitive Cia verbs are likely to have been the retention of POC forms containing the transitive suffix and the third person singular object clitic. Since Tongan, or any other Polynesian language, does not have a third person singular object pronominal clitic (or suffix) *-a*, the relevant clitic cannot be reconstructed for PPN. Therefore, it is unlikely for these verbs to be a retention of PPN forms with the same meaning. On the other hand, stative Cia verbs are likely to be reflexes of PPN forms, assuming that the PPN stativiser -*a was productive. Either way, variable Cia forms in Tongan are best understood to be part of fossilised PPN and POC forms rather than the allomorphs of a productive suffix Cia.

It should be noted, however, that Tongan exhibits some dubious Cia forms. For example, there are Cia forms with a base that is not reconstructible beyond PPN, e.g., *mokosia* '(subjectively) cold' with PPN **moko* 'cold'. How the consonant preceding /ia/ in such examples arises is a mystery, as it has no historical basis; that is, corresponding C-final POC forms do not exist. Nevertheless, the existence of these Cia forms at least suggests that Cia had been reanalyzed as a suffix prior to their formation. Intriguingly, cognates of *mokosia* have /m/ instead of /s/ in Niue (*moko-mia*) and /t/ in Tuvalu (*moko-tia*).[2]

In some other mysterious Cia forms, the consonant preceding /ia/ is different from the base-final C of the corresponding proto form: *afaa-ngia* 'devastated by a cyclone' vs. POC *apaRat 'northwest wind' (PEO *yavaRat 'storm'); *lango-mia* 'pressed down' vs. POC *langon 'rollers'; *ongo-sia* 'to be tired, exhausted' from *ongo* 'to sound, to be felt' vs. POC *dongor 'hear'; and *efu-hia* 'to become dusty, to be made dusty by something' vs. POC *qapu 'dust'. Based on the POC forms, these Tongan Cia forms should be *afaa-sia*, *lango-nia*, *ongo-ia*, and *efu-ia*, respectively (due to t > s before [i] and PPN *r > Ø in Tongan). These forms suggest that C had been detached from the base by the time these forms were formed, and that the original consonant was replaced with a new one for some reason. It is, however, beyond the scope of this squib to determine the exact reason why it was replaced and how a particular consonant was selected as replacement.

The distribution of various forms and the semantic nature of Cia-words in Māori, Hawaiian, and Tongan can be better accounted for if we assume that PPN had C-final bases instead of the suffix *Cia. Of the three languages examined in this study, only Māori retains PPN C-final base allomorphs. Tongan and Hawaiian retain the historical final consonants in some fossilised forms, but the distribution of variable Cia is not so regular as to warrant the base allomorphy analysis. Instead, the base-final C has been reanalysed as the initial C of the suffix in Tongan and Hawaiian. In Tongan, this reanalysis was followed by the development of several (semi-)productive C-initial suffixes such as transitive *-ʻi* and affected stative *-ʻia*. Hawaiian also developed a single Cia form, *ʻia*, which subsequently gained a remarkable degree of productivity as well as more lexical independence so that it now functions as a particle.

It should be noted that Māori shows a possible sign of reanalysis as well in three contexts. First, when the "passive" morpheme attaches to a nominal base, it is regularly realised as *-tia*. Second, when it attaches outside of the nominaliser *-anga*, the "passive" morpheme takes the form *-hia* instead (Boyce 2006; Yamada 2014). Third, when the base is a trimoraic or longer loanword, the form of the "passive" morpheme is *-tia* (Hale 1968; Blevins 1994: 41). In de Lacy's (2003) phonological analysis of the Māori "passive" suffix, *-tia* is treated as one of the three variants along

[2] Tongan *-sia* may be underlyingly *-tia*, as [t] regularly becomes [s] before [i] in Tongan except in English loan words. Other mysterious forms of this kind also have /s/ as the initial C of the suffix: *kaihaʻa-si* from *kaihaʻa* 'to steal' (> PPN *kaihaqa) and *ʻapa-sia* from *ʻapa* 'to show respect' (> PPN *apa). This choice of C is intriguing in the light of de Lacy's (2003) analysis of the Māori "passive" suffix. He postulates three allomorphs, *-ia*, *-a*, and *-tia,* and claims that /t/ in *-tia* is epenthetic.

with *-a* and *-ia*. The initial /t/ of *-tia* is shown to be an epenthesis rather than an instance of historical final C. However, systematic use of a particular form of Cia may lead to the reanalysis of base-final consonants and subsequently to the reduction of allomorphic variation. Whatever path Māori Cia takes in the future, it may shed some new light on our understanding of the historical development of Cia in other Polynesian languages.

References

Biggs, Bruce. 1969. *Let's learn Māori*. Wellington: A.H. and W.A. Reed.

Blevins, Juliette. 1994. A phonological and morphological reanalysis of the Maori passive. *Te Reo* 37. 29–53.

Blevins, Juliette. 2004. The mystery of Austronesian final consonant loss. *Oceanic Linguistics* 43. 208–213.

Blust, Robert & Trussel, Stephen. 2010. *Austronesian comparative dictionary*. Web edition. Last revised in March, 2017. Available at www.trussel2.com/ACD.

Boyce, Mary. 2006. *A corpus of modern spoken Māori*. Wellington: Victoria University of Wellington. (Doctoral dissertation)

Churchward, C. Maxwell. 1953. *Tongan grammar*. London: Oxford University Press.

Churchward, C. Maxwell. 1959. *Tongan dictionary*. London: Oxford University Press.

Churchward, Spencer. 1928. On the origins of the Polynesian passive. *Journal of the Polynesian Society* 37. 300–305.

Churchward, Spencer. 1951. *A Samoan grammar*. 2nd edn. Melbourne: Spectator.

Clark, Ross. 1976. *Aspects of Proto-Polynesian syntax*. Auckland: Linguistic Society of New Zealand.

de Lacy, Paul. 2003. Maximal words and the Māori passive. In Rackowski, Andrea & Richards, Norvin (eds.) *Proceedings of AFLA VIII: The eighth meeting of the Austronesian formal linguistics association.* MIT Working Papers in Linguistics 44. 20–39. Cambridge, MA: MIT Linguistics Department.

Elbert, Samuel and Mary Kawena Puku'i. 1979. *Hawaiian grammar*. Honolulu: University of Hawai'i Press.

Hale, Kenneth. 1968. Review of Hohepa (1967). *Journal of the Polynesian Society* 77. 83–99.

Otsuka, Yuko. 2011. Neither accusative not ergative: An alternative analysis of case in Eastern Polynesian. In Moyse-Faurie, Claire & Sabel, Joachim (eds.) *Topics in Oceanic morphosyntax*, 289–318. Berlin: De Gruyter Mouton.

Otsuka, Yuko. 2012. Passive morphemes in a passive-less language? In Graf, Thomas & Paperno, Denis & Szabolcsi, Anna & Tellings, Jos (eds.) *Theories of everything: In honor of Ed Keenan* (UCLA Working Papers in Linguistics 167), 316–324.

Pawley, Andrew. 2001. Proto Polynesian -CIA. In Bradshaw, Joel & Rehg, Kenneth L. (eds.) *Issues in Austronesian morphology: A focusschrift for Byron W. Bender*, 193–216. Canberra: Pacific Linguistics.

Yamada, Fumiko. 2014. *Māori as a phrase-based language*. Honolulu: University of Hawai'i. (Doctoral dissertation)

Yuko Otsuka
Sophia University
yotsuka@sophia.ac.jp

Pronouns and the DP in Hoava

Bill Palmer

1. Introduction

Following Abney (1989), the notion of the Determiner Phrase (DP) is established in syntactic theory. However, it has not been widely adopted by scholars working within descriptive and typological frameworks, and is not a feature of the Oceanic descriptive tradition, despite Lynch et al.'s (2002: 38) curious statement that articles in Oceanic languages "precede a noun phrase". Nominal phrases in Oceanic languages have rarely been analysed in terms of a DP, exceptions being largely confined to syntactic literature on Polynesian and other Central Pacific languages. In this regard, Pearce has been a leader in the field. Her early works on aspects of the DP in Maori (1997a,b; 1998a; 1998b; 2003) are among the earliest applications of the DP to an Oceanic language, preceded only by her colleague Waite (1994). Her application of the DP to the Southern Oceanic languages Iaai (Pearce 2000; 2001, etc.), Unua (Pearce 2007; 2010; 2011; 2012, etc.) and Ninde (Pearce 2012) led the way in applying the DP to Oceanic languages outside Central Pacific. This paper is intended as a tribute to Pearce's pioneering work in the syntax of Oceanic languages in general, and of the DP in Oceanic in particular.

Hoava is a member of the New Georgia subgroup of the Northwest Solomonic branch of Western Oceanic, with somewhere upward of 2,360 speakers in the 1990s (Davis 2003: 1). It has been the subject of a reference grammar (Davis 2003), couched in the Oceanic descriptive tradition (e.g. Lynch et al 2002), and a small corpus of texts (Davis 1991).

Davis does not employ the DP in her analysis of Hoava. A number of issues in the analysis of pronouns and of the NP in the language give rise to inconsistencies in the existing analysis, including the variable analysis of some forms as at times articles and at times pronouns, inconsistencies in the treatment of first and second person pronouns on the one hand and third person pronouns on the other, and the unaccounted for impossibility of pronouns to occur with articles or for NPs accompanying pronouns to occur with articles. Palmer (2017) examines these issues from the perspective of categorial flexibility between pronouns and articles in the language, proposing a DP analysis that accounts for this apparent flexibility. The present article examines the syntax of the Hoava DP in more detail to account for issues arising from Davis's analysis.

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 191–203

2. Status of *eri*

Davis analyses the form *eri* variably as a plural personal article (2003: 62–63),[1] a plural definite article with restricted distribution (2003: 55), and a pronominal form (2003: 46–47).

The singular personal article *e* occurs with personal names and the interrogative form *se* (1). *Eri* also occurs in these environments (2). Davis treats *eri* in this context as monomorphic. However, the suffix *-ri* marking plural occurs elsewhere in the language, such as with demonstratives (see e.g. (4a), (34b) etc.). This leads Palmer (2017: 417) to conclude that *eri* is formed from the personal article with a plural suffix.

1. [**E** se] podo=a [**e** Saku]?
   PERS[SG] who? give.birth=3SG.ACC PERS[SG] Saku
   'Who gave birth to Saku?'

2. Q: [**E-ri** se] mae? A: [**E-ri** Vezi ri Napi ri Jopa].
   PERS-PL who? come PERS-PL Vezi and Napi and Jopa
   'Who (PL) came?' 'Vezi and Napi and Jopa.'

Eri occurs in one other context analysed by Davis as an article – with NPs containing a numeral and a noun with a human referent (3). Enumerated NPs with non-human referents (4) cannot occur with *eri*. Instead they occur with what Davis analyses as a definite article, *ria* (see Section 3 below). This leads Davis to analyse *eri* in this context as a definite article occurring with enumerated NPs.

3. a. **eri** ka-ru koburu
   eri NUM-two child
   'the two children'

   b. **eri** ka-lima maneke
   eri NUM-five mother.and.child
   'the five mother with children'

4. a. **ria/*eri** ka-ru hore sa-ri
   ria/*eri* NUM-two canoe DIST-PL
   'those two canoes'

   b. **ria /*eri** ka-lima maŋara
   ria/*eri* NUM-five stone
   'the five stones'

However, the personal article across Oceanic occurs with a range of different types of nominals. In other words, membership of the class of personal nouns varies from language to language, extending to varying extents down an animacy hierarchy (Palmer & Smith 2016). In Standard Fijian, the personal article occurs with a personal class containing names of persons or places, the human interrogative pronoun, independent pronouns, demonstratives, and optionally with bound kin terms (Schütz 2014; Palmer 2017: 414). In Teop (Northwest Solomonic, Bougainville), the personal article occurs with a personal class containing names of persons, the human interrogative pronoun, independent pronouns, bound kin terms, terms for socially

[1] Many Oceanic languages display articles distinguishing between "common" and what is termed personal (Lynch et al 2002: 38), reflecting a distinction present in Proto Oceanic (Lynch et al 2002: 69–72). The personal article occurs with personal names, and a range of other nominals varying from language to language.

important humans ('chief', 'old woman' etc.), and terms for domestic animals ('pig' etc.) (Mosel & Spriggs 2000; Palmer 2017: 423–424). In Kubokota (Northwest Solomonic, Solomon Islands) it is confined to personal names and the human interrogative pronoun only (Chambers 2009: 88–89).

The distribution of *eri* leads Palmer (2017: 424) to conclude that *eri* is a plural personal article occurring with a personal class that in Hoava includes personal names, the human interrogative, and enumerated humans.

Davis also analyses *eri* as a pronominal form occurring in dual, trial and quadral pronouns (2003: 46–47), in effect proposing the following paradigm, in Table 1.

Table 1: Davis's (2003:46–47) analysis of Hoava pronouns

	1EXCL	1INCL	2	3
SG	*rao*	-	*ɣoe*	*sa*
DU	*ɣami karu*	*ɣita karu*	*ɣamu karu*	*eri karu*
TR	*ɣami kahike*	*ɣita kahike*	*ɣamu kahike*	*eri kahike*
QD	*ɣami kamade*	*ɣita kamade*	*ɣamu kamade*	*eri kamade*
PL	*ɣami*	*ɣita*	*ɣamu*	*ria*

Davis notes that the dual, trial and quadral forms are constructed by following the pronoun with a numeral. The numeral forms involved are the standard numerals in the language (5).

5. a. Te se [**ka-ru** hore sa-ri]?
   POSS who? NUM-two canoe DIST-PL
   'Whose are those two canoes?'

   b. [**ka-made** dia tomoko]
   NUM-four 3PL.PSSR war.canoe
   'their four war canoes'

Davis also notes that larger numbers can be used with pronouns (6).

6. a. Toka ria minate sa gurupu [ɣami **ka-lima**, **onomo**]?
   set.off 3PL people 3SG group 1EXCL.PL NUM-five six
   'The people set off, the group, we five or six.'

   b. [ɣami **hike** **ɣoɣoto**]
   1EXCL.PL three hundred
   'we three hundred'

Together, these facts suggest that the language has only singular and plural pronouns. These may be accompanied by a numeral to specify the number of individuals referred to. However, this is not confined to numerals on their own. Pronouns may occur with enumerated NPs (7). The difference between (7) and the putative trial and quadral pronouns in (8) is simply whether or not the numeral is accompanied by a lexical noun, leading to the conclusion that the putative dual, trial and quadral pronouns are simply plural pronouns dominating an enumerated NP, whether or not that NP contains a

lexical noun in addition to the numeral. In a Quantifier Phrase (QP) or Number Phrase (NUMP) analysis, the pronouns dominate QP or NUMP, which itself may or may not contain an NP.

7. a. ɣita [**ka-hike maneke**]
1INCL.PL NUM-three family
'we family of three'

   b. ɣami [**ka-made Lena ri Ladi ri Niubiri**]
1EXCL.PL NUM-four Lena and Ladi and Niubiri
'we four [me and] Lena and Ladi and Niubiri'

8. a. ɣita [**ka-hike**]
1INCL.PL NUM-three
'we three'

   b. ɣami [**ka-made**]
1EXCL.PL NUM-four
'we four'

Like the pronouns in (8), *eri* occurs with numerals in its putative pronominal function (9). However, like the pronouns in (7), it also occurs with enumerated NPs (10) as the plural personal article. A consistent analysis of *eri* in (9) and (10) on the one hand, and SAP pronouns in (7) and (8) on the other treats *eri* as a functionally unitary form occurring with a numeral, whether or not that numeral is itself accompanied by a lexical noun. It was argued above that in contexts such as (10) *eri* is a personal article. The possibility that it is instead a pronoun in all contexts can be rejected on the grounds that it cannot occur as a phrase on its own, without at least an accompanying numeral (11). On this basis *eri* is best analysed as the personal article in all contexts. In data such as (9) the article occurs with an enumerated phrase with a human referent, just one lacking a lexical noun.

9. a. Pule [eri **ka-ru**].
return *eri* NUM-two
'The two returned.'

   b. Dae tiono [eri **ka-lima**].
finish bake *eri* NUM-five
'The five of them finished baking.'

10. a. Pato [eri **ka-ru** koburu].
speak *eri* NUM-two child
'The two children spoke.'

   b. Ko dia [eri **ka-lima** maneke].
stay IPFV.3PL.SBJ *eri* NUM-five mother.and.child
'Four sons and their mother were living [there].'

11. *Pule [**eri**]. Intended: 'They returned.'

3. Status of *sa* and *ria*

Davis (2003) analyses *sa* and *ria* variably as 3SG and 3PL pronouns respectively on the one hand (2003:46), and as singular and plural definite articles on the other (2003: 55). She does not provide an explicit basis for this analysis. However, the tacit rationale is clear, and is dependent on a traditional non-DP analysis in which articles occur inside the NP. *Sa* and *ria* each encode definiteness, with *sa* also encoding singular number,

and *ria* plural. In (12), the NP containing *sa* or *ria* contains only those forms. NPs required a head, effectively in N. As *sa* and *ria* are the only forms in their respective phrases, they must be the head. Nominal heads expressing definiteness and number and lacking an external referent are pronouns. As the head of a phrase, in a traditional NP analysis *sa* and *ria* in (12) must be pronouns. In (13), however, *sa* and *ria* occur with a lexical noun. The lexical noun must be the head of the NP, so *sa* and *ria* cannot be head. As such they must be modifiers. Encoding as they do definiteness and number, they must therefore be articles. The crucial factor is whether or not *sa* and *ria* are a head in structures like those in (13). In a traditional NP analysis they cannot be.

12. a. Pule mae [**sa**] pa Raro.
return come 3SG LOC Raro
'He came back to Raro.'

    b. ɣore mae [**ria**].
go.down come 3PL
'They came back down.'

13. a. Tase-na [**sa nikana Tonga he-ni**] Samusoni Mapi.
name-3SG.PSSR 3SG man Tonga PROX-SG Samusoni Mapi
'This Tongan man's name [was] Samson Mapi.'

    b. [**Sa nikana tarai**] saɣi Boazi Suŋa.
sa man preach TOP Boazi Suŋa
'The preacher was Boazi Sunga.'

    c. Kobi vaɣi mae [**ria nikana Merika**] ta-mi ɣami.
always take come *ria* man Amerika POSS-1EXCL.PL.PSSR 1EXCL.PL
'The Americans always brought [things] for us.'

    d. Tige la ŋani paho=e sa
then go kill be.empty=ACC[3PL.ACC] 3SG
[**ria doluru nikana pa Hoeze**].
3PL all man LOC Hoeze
'Then he killed all the people of Hoeze.'

The distributional difference between *sa* and *ria* in (12) and (13) lies in whether or not they are accompanied by a lexical noun. Pronouns, encoding both definiteness and person, may occur with or without an accompanying noun (*we linguists* and *we* are both acceptable in English). The distribution of *sa* and *ria* in (12) and (13) is therefore consistent with the status of pronoun in both contexts. However, articles, encoding only definiteness and not person, must occur with a noun (*the linguists* is acceptable, but **the* alone is not). The distribution of *sa* and *ria* in (13) is consistent with status as articles, but that in (12) is not. They may have the status of pronouns in both contexts, or of pronouns in (12) and articles in (13). In the absence of other distributional differences, a parsimonious analysis would treat *sa* and *ria* as pronouns in both contexts.

An analysis in which *sa* and *ria* are pronouns in (12) but articles in (13) is further undermined by their parallel distribution with unambiguous first and second person

(S[peech] A[ct] P[articipant]) pronouns. In (14) SAP pronouns occur as phrases on their own, paralleling (12). In (15) they occur with a lexical noun, paralleling (13). Davis analyses these SAP forms as pronouns in all contexts, but inconsistently analyses *sa* and *ria* as articles in contexts identical to (14). SAP forms are pronouns in (14) and (15), the difference being only the presence or absence of an accompanying lexical noun. A consistent analysis will treat *sa* and *ria* in the same way.

14. a. Tata mate [**ɣami**].
almost be.dead 1EXCL.PL
'We were nearly dead.'

b. Koni vasi mate paho [**ɣamu**].
FUT really be.dead INTNS 2PL
'You will really die!'

15. a. Kipu mae [ria nikana Japani]
NEG come *ria* man Japan
'The Japanese men did not come

de [**ɣami nikana hupa**] mae ŋani=i [ɣami].
PURP 1EXCL.PL man black come kill=ACC 1EXCL.PL
to kill us we black men.'

b. Pa ɣuɣusu ri ta-mu [**ɣamu** Masu ri Tokopae].
LOC village PL POSS-2PL.PSSR 2PL Masu and Tokopae
'At the village of you Masu and Tokopae.'

4. The syntactic distribution of Hoava pronouns

4.1 Pronouns and articles

In many Oceanic languages, pronouns occur with articles, typically the personal article, as in, for example, Standard Fijian (16c), where pronouns occur with an article *o* also occurring with personal names (16b), in contrast with a common article *na* occurring with lexical nouns (16a).

16. a. na vale levu
DEF house be.big
'the big house'

b. o Timoci
PERS Timoci
'Timothy'

c. o ira [Fijian]
PERS 3PL
'they'

In Hoava, however, pronouns cannot occur with any article. While common nouns may occur with the specific article *na* (17), and personal names occur with the personal article *e* in its singular or plural form (18). Neither article is permissible with pronouns (19).

17. ***na*** *nikana*
SPCF man
'a/some [specific] man/men'

18. a. **e** Saku
PERS Saku
'Saku'

b. **e-ri** Vezi ri Napi
PERS-PL Vezi and Napi
'Vezi and Napi'

19. a. ***e/*na** ɣoe
PERS/SPCF 2SG
'youSG'

b. ***e-ri/*na** ɣita
PERS-PL/SPCF 1INCL.PL
'weINCL'

Instead, the distribution of pronouns parallels that of articles, as in (20) and (21), leading to the conclusion that Hoava pronouns are in the same syntactic position as articles.

20. a. **ɣita** [nikana hupa]
1INCL.PL man be.black
'we black men'

b. **na** [nikana pa soloso]
SPCF man LOC bush
'men from the bush'

21. a. **ɣamu** [Masu ri Tokopae]
2PL Masu and Tokopae
'you Masu and Tokopae'

b. **e-ri** [Vezi ri Napi]
PERS-PL Vezi and Napi
'Vezi and Napi'

4.2 Pronouns and nouns

The analysis that Hoava pronouns are located in N is further undermined by differences in the syntax of phrases with a noun head and those with a pronominal head, beyond the potential co-occurrence of articles. Pronouns do not display same lexical dependent modifier possibilities as nouns. While nouns may be modified by an AP, pronouns are modified by a full NP (or QP).

As in many Oceanic languages, Hoava nouns may be modified by a stative verb stem (22) or a noun stem (23) functioning adnominally. Passivized verbs may also modify nouns (22b). Davis (2003: 74) argues that Hoava lacks a lexical category of adjective defined on morphological criteria. However, while there is no evidence of underived adjectival roots, nouns (23c) and psych verbs (22c) typically reduplicate when functioning adjectivally.

22. a. ria ka-rua hogi [**kisi**] ra
3PL NUM-two tooth be.small DIST.PL
'those two small teeth'

b. iɣana [**ta-ŋani**]
fish PASS-eat
'edible fish'

c. sa p<in>ato [**edo~edo**]
3SG<NMLZ>speak RED~be.happy
'the happy speech'

Underived nouns typically reduplicate when functioning adjectivally (23c), but in some instances do not (23a–b). Verbs other than psych verbs typically function adnominally as nominalizations involving possessor-indexing (23d–e). Active verbs may also be

reduplication (23f). Passivized verbs may also function adnominally as nominalizations with possessor-indexing (23g).

23. a. mi vinetuŋu [**vaka**]
2SG.PSSR fish.hook ship
'your European fish hooks'

b. ria nikana [**Japani**]
3PL man Japan
'the Japanese men'

c. guhe [**ma~makariva**]
beetle RED~boy
'male beetle'

d. doluru gato [**koleo-na**]
all tree be.good-3SG.PSSR
'all good trees'

e. ria nikana [**mae-di**]
3PL man come-3PL.PSSR
'the arrivals'

f. ria nikana [**haga~haga-di**]
3PL man RED~run-3PL.PSSR
'the running men'

g. sa koburu [**ta-hapu-na**]
3SG child PASS-cover-3SG.PSSR
'the covered child'

These adnominal stems do not form VPs or NPs as they are highly restricted in their structure. They comprise a single verb stem (22) or noun stem (23). This may be accompanied by one of two intensifiers, both also functioning as verbs: *holapa* 'pass' and *hokara* 'be true', as in (24) and (25). No other modification of adnominal verb or noun stems occurs. Given these restrictions, adnominal noun and verb stems are treated here as the head of an AP.

24. a. k-eke pade [**kisi** **holapa**]
NUM-one floor be.small pass
'one very small floor'

b. *na beta* *[**kisi** **hokara**]*
SPCF betelnut be.small be.true
'a truly small betelnut'

25. a. na iɣana [**heleana** **hokara**
SPCF fish river be.true
'a true freshwater fish'

b. *sa bekoto* *[**v<in>a-da~dae** **hokara**]*
3SG thatching.stick CAUS<NMLZ>-RED~finish be.true
'the very final thatching stick'

These facts do not apply to dependents of pronouns. Unlike nouns, pronouns may not be modified by a verb stem. Also unlike nouns, they may be accompanied by an ordinary NP (26–27), or in an analysis recognising the Quantifier Phrase, a QP, as (27), repeating (7), shows. However, this NP/QP may not carry an article (28).

26. a. ɣita [**nikana** **hupa**]
1INCL.PL man be.black
'we black men'

    b. ɣamu [**Masu** **ri** **Tokopae**]
2PL Masu and Tokopae
'you Masu and Tokopae'

27. a. ɣita [**ka-hike** **maneke**]
1INCL.PL NUM-three family
'we family of three'

    b. ɣami [**ka-made** **Lena** **ri** **Ladi** **ri** **Niubiri**]
1EXCL.PL NUM-four Lena and Ladi and Niubiri
'we [me and] Lena and Ladi and Niubiri'

    c. ɣami [**hike** **ɣoɣoto**]
1EXCL.PL three hundred
'we three hundred'

28. a. ɣita [***na/*eri/*ria** nikana hupa]

    b. ɣita [***na/*eri/*ria** ka-hike maneke]

The position of quantifiers in relation to pronouns also differs to their position in relation to nouns. Numerals precede nouns but follow pronouns (29). This is consistent with an analysis in which pronouns are located above quantifiers, while nouns are located below quantifiers, itself consistent with a QP analysis.

29. a. La pita muziki [**ɣami** [**ka-made**]].
go look.for bait 1EXCL.PL NUM-four
'We four looked for bait.'

    b. Taloa [e-ri [**ka-hike** [**koburu** makariva]]].
leave PERS-PL NUM-three child boy
'The three boys left.'

4.3 Hoava pronoun summary

In summary, Hoava pronouns share the syntactic characteristics of articles, and not those of nouns. They are therefore located in the same syntactic position as articles.

Table 2: The syntactic characteristics of Hoava pronouns compared with articles and nouns

	articles	**pronouns**	**nouns**
own article	no	no	yes
verb as dependent	no	no	yes (AP)
nominal dependent	NP	NP	AP
numerals	follow	follow	precede

5. The DP in Hoava

In a traditional analysis of nominal phrasal structure, determiners are treated as located within the NP, as in (30).

30. $_{NP}$[***the*** *linguists*]

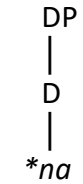
Following Abney (1987) the notion of a Determiner Phrase has been widely (though not universally) adopted in syntactic theory. In this structure, the determiner is the head of a phrase. The rest of the erstwhile NP forms an NP complement to D, as shown in (32). Articles require a complement NP (31a), while demonstratives may occur with a complement NP (32b), but need not (31b). In addition to articles and demonstratives, pronouns have been argued to be located in D (after Abney 1987: 281–284; Longobardi 1994, 2001 etc.), following an earlier pre-DP proposal by Postal (1966) that the location of pronouns in English corresponds to that of articles and demonstratives. Like demonstratives, they may (32c) but need not (31c) occur with a complement NP.

31. a. DP – D – **the* b. DP – D – *those* c. DP – D – *we*

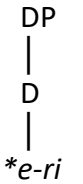
32. a. DP (D *the*, NP *linguists*) b. DP (D *those*, NP *linguists*) c. DP (D *we*, NP *linguists*)

More recent work (Cowper & Hall 2009; Déchaine & Wiltschko 2002; Ritter 1995) has argued that pronouns are not necessarily located in D and can be located in N or in various intermediate levels in the structure. However, as Table 2 shows, the syntactic distribution of pronouns in Hoava resembles that of articles. Hoava pronouns therefore appear to be located in D. In Déchaine and Wiltschko's (2002) terms, they are pro-DP.

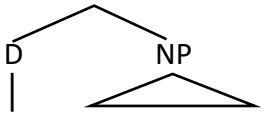
The Hoava specific and personal articles occur in D (34), and being articles, require a complement NP (33). Hoava pronouns occur in D, and being pronouns, may occur with a complement NP (36), but need not (35). The unitary syntactic status of *sa* and *ria* as pro-DP is revealed by their occurrence with (36b) or without (35b) a complement NP in the same way as non-3rd pronouns.

33. a.


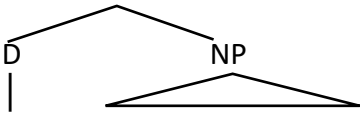
*'the'

b.


*'the' (personal)

34. a.


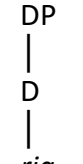
na nikana pa soloso
SPCF man LOC bush
'men from the bush'

b.


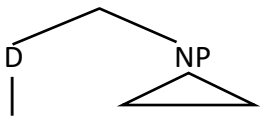
e-ri ka-ru nikana he-ri
PERS-PL NUM-two man PROX-PL
'these two men'

35. a.

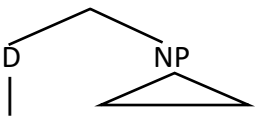

'we(excl)'

b.


'they'

36. a.


ɣami nikana hupa
1INCL.PL man be.black
'we black men'

b.


ria nikana Japani
3PL man Japan
'the Japanese men'

6. Conclusion

By employing the DP, the absence of articles with pronouns and absence of articles in NPs modifying pronouns falls automatically out of the analysis. Pronouns cannot occur with articles because the D position the article would occur in is already occupied by the pronoun. Dependent NPs also cannot occur with articles because the D head position in the DP in which they occur is already occupied by the pronoun. This DP analysis accounts for the distributional similarities between articles and pronouns shown in Table 2, allows a consistent analysis of *sa* and *ria* in their apparent article and

apparent pronoun contexts, and allows a consistent analysis of *sa* and *ria* on the one hand and first and second person pronouns on the other.

References

Abney, Steven P. 1987. *The English noun phrase in its sentential aspect.* Cambridge MA: MIT. (Doctoral dissertation)

Chambers, Mary R. 2009. *Which way is up? Motion verbs and paths of motion in Kubokota, an Austronesian language of the Solomon Islands.* London: School of Oriental and African Studies. (Doctoral dissertation)

Cowper, Elizabeth & Hall, Daniel Currie. 2009. Argumenthood, pronouns, and nominal feature geometry. In Ghomeshi, Jila & Paul, Ileana & Wiltschko, Martina (eds.) *Determiners: Universals and variation*, 97–121. London: Benjamins.

Davis, Karen. 2003. *A grammar of the Hoava language, western Solomons.* Canberra: Pacific Linguistics.

Davis, Karen (ed.) 1991. *Vivinei ruruhu pa Hoava: Custom stories from Hoava.* Gizo: Western Province Government.

Déchaine, Rose-Marie & Wiltschko, Martina. 2002. Decomposing pronouns. *Linguistic Inquiry* 33(3). 409–442.

Longobardi, Giuseppe. 2001. The structure of DPs: Some principles, parameters, and problems. In Mark Baltin & Collins, Chris (eds.) *The handbook of contemporary syntactic theory,* 562–603. Oxford: Blackwell.

Longobardi, Giuseppe. 1994. Reference and proper names: A theory of N-movement in syntax and logical form. *Linguistic Inquiry* 25(4). 609–665

Lynch, John, & Ross, Malcolm D. & Crowley, Terry. 2002. *The Oceanic languages.* London: Routledge.

Mosel, Ulrike & Spriggs, Ruth. 2000. Gender in Teop. In Unterbeck, Barbara & Rissanen, Matti (eds.) *Gender in grammar and cognition,* part 1*,* 321–349. Berlin: Mouton de Gruyter.

Palmer, Bill. 2017. Categorial flexibility as an artefact of the analysis: pronouns, articles and the DP in Hoava and Standard Fijian. *Studies in Language* 41(2). 408–444

Palmer, Bill & Smith, Ellen. 2016. *Differential noun class membership in North Bougainville.* (Paper presented at 8th Conference on Austronesian and Papuan Languages and Linguistics, School of Oriental and African Studies, University of London)

Pearce, Elizabeth. 1997a. DP licensing and SPEC roles in Maori. *Wellington Working Papers in Linguistics* 9. 17–30.

Pearce, Elizabeth. 1997b. Genitive case in the Maori DP. *Wellington Working Papers in Linguistics* 9. 31–55.

Pearce, Elizabeth. 1998a. DP licensing and SPEC roles in Maori. In Pearson, Matthew (ed.) *Recent papers in Austronesian linguistics: Proceedings of the third and fourth meetings of the Austronesian Formal Linguistics Association (AFLA).* 245–258. Los Angeles: UCLA.

Pearce, Elizabeth. 1998b. The syntax of genitives in the Maori DP. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 43. 411–434.

Pearce, Elizabeth. 2000. Transitivity, incorporation and animacy in Iaai. *Wellington Working Papers in Linguistics* 12. 21–40.

Pearce, Elizabeth. 2001. Object agreement and incorporation in Iaai. In Allan, Keith & Henderson, John (eds.) *Proceedings of ALS2k, the 2000 conference of the Australian Linguistic Society*. http://www.als.asn.au

Pearce, Elizabeth. 2003. Phrasal movement within the Māori DP. *Working Papers in Linguistics.* 34(2). 41–42. University of Hawai'i at Mānoa.

Pearce, Elizabeth. 2007. Constituent ordering within the Unua noun phrase. (Paper presented at 7th International Conference on Oceanic Linguistics, Noumea)

Pearce, Elizabeth. 2010. Possession syntax in Unua DPs. In Mercado, Raphael & Potsdam, Eric & Travis, deMena Lisa (eds.) *Austronesian and theoretical linguistics*, 141–162. Amsterdam: Benjamins.

Pearce, Elizabeth. 2011. Two or three things in Unua noun phrases. In Moyse-Faurie, Claire & Sabel, Joachim (eds.) *Morphosyntactic aspects of Oceanic languages*, 139–178. Berlin: Mouton de Gruyter.

Pearce, Elizabeth. 2012. Number within the DP: A view from Oceanic. In Brugè, Laura & Cardinaletti, Anna & Giusti, Giuliana & Monaro, Nicola & Poletto, Cecilia (eds.) *Functional heads: The cartography of syntactic structures*, vol. 7, 81–91. Oxford: Oxford University Press.

Postal Paul H. 1966. On so-called 'pronouns' in English. In Dinneen, Francis P. (ed.) *Report of the Seventeenth Annual Round Table Meeting on Linguistics and Language Studies*, 177–206. Washington: Georgetown University Press.

Ritter, Elizabeth. 1995. On the syntactic category of pronouns and agreement. *Natural Language and Linguistic Theory* 13(3). 405–443

Schütz, Albert J. 2014. *Fijian reference grammar.* Honolulu: Pacific Voices Press.

Waite, Jeffrey. 1994. Determiner phrases in Maori. *Te Reo* 37. 55–70

Bill Palmer
University of Newcastle
bill.palmer@newcastle.edu.au

A note on the typology of topic and focus markers

Luigi Rizzi

1. Introduction

Cartographic research on the left periphery led to the conception of the complementizer system as a sequence of functional heads with distinct syntactic and interface functions (Rizzi 1997 and much subsequent work). The Force-Finiteness system expresses properties of illocutionary force, clause-typing, and other global properties of the clause, such as its finite or non-finite character. This system also delimits the complementizer zone, hosting other heads, organized in subsequence, and expressing scope-discourse properties. Such heads, Top, Foc, Q, Rel, Excl and the like, have a dual function: in syntax, they attract phrases endowed with the appropriate features of topicality, focus, etc. to their Specs; at the interfaces with sound and meaning, such heads guide the assignment of the appropriate prosodic contour (Bocci 2013) and of the interpretation, also expressing conditions for felicitous use in discourse (Belletti & Rizzi 2017). These analytic guidelines define the so-called "criterial approach" to scope-discourse semantics, a program which aims at "syntacticizing" this important interpretive domain (Cinque & Rizzi 2010). Fundamental scope-discourse properties are transparently encoded in syntactic representations by a sequence of criterial heads; each head generates the familiar Specifier – Head – Complement configuration, which constitutes the configurational skeleton for expressing the fundamental scope-discourse articulations: topic – comment, focus – presupposition, operator – scope domain.

2. Criterial head is overt?

This structural approach to scope-discourse semantics is made immediately plausible by the fact that some languages overtly express the system of criterial heads through special morphemes. For instance, in the Kwa language Gungbe, topics and foci are accompanied by special morphemes which Aboh (2004) plausibly argues to lexicalize the Top and Foc heads, respectively.

1. a. [dan lo ***yà*** [Kofi hu ì]] (Gungbe)
   snake the **TOP** Kofi killed it
   'The snake, Kofi killed it.'

   b. [dan lo ***wè*** [Kofi hu __]]
   snake the **FOC** Kofi killed
   'THE SNAKE Kofi killed.' (Aboh 2004)

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 205–210

In contrast, the corresponding Italian (and English) examples do not have overt morphological markers.[1]

2. a. Il serpente, Gianni lo ha ucciso (Italian)
   the snake Gianni it has killed
   'The snake, Gianni killed it.'

   b. IL SERPENTE Gianni ha ucciso, non il topo
   the snake Gianni has killed, not the mouse
   'THE SNAKE Gianni killed, not the mouse.'

Under plausible guidelines of cross-linguistic uniformity, a reasonable initial hypothesis is that the underlying syntax is the same in the two types of languages, with criterial heads Top and Foc which have an analogous role (they attract a constituent in syntax and guide interpretation at the interface systems), but are not pronounced in Italian (and English). The minimal difference can be expressed by the following spell-out parameter:

3. Criterial head H is overt?

3. Criterial head moves?

Pearce (1999), in the first explicitly cartographic study on an Austronesian language, argues that in Māori, topics and foci occur in the left periphery in a fixed order Top > Foc, much as in Gungbe. In Italian the situation is more complex, as certain types of topics (familiarity topics, according to the typology in Frascarelli & Hinterhoelzl 2007; Bianchi & Frascarelli 2010) can follow the left-peripheral focus position, but other kinds of topics (aboutness-shift and contrastive topics, according to the same typology) necessarily precede the focus position. The following example illustrates a topicalized object and a focalized subject (thus yielding the OSV order from the unmarked VSO order) in Māori:

4. ***Ko*** te hipi ***nā*** Pita I fihore (Māori)
   KO the sheep ***NĀ*** Pita T/A fleece
   'As for the sheep, it was Pita who fleeced it.' (Pearce 1999, ex. (5))

Ko and *nā* appear to be directly linked to topic and focus. In this respect, the language looks like Gungbe w.r.t. parameter (3). On the other hand, Māori differs from Gungbe in that the markers precede topic and focus, rather than following them, so that in the pronounced string these functional elements cannot be in a Spec-head configuration with topic and focus, respectively. In the case of *ko*, Pearce explicitly considers the hypothesis that the unexpected order may follow from movement from the criterial

[1] Topic and focus constructions are sharply differentiated by intonational contour and pragmatic conditions for felicitous use in both Italian and English; moreover, at the syntactic level, in Italian the (object) topic is obligatorily resumed by a clitic, whereas a focalized element is not. The focal example illustrates corrective focus, in terms of the typology introduced in Bianchi, Bocci, & Cruschina (2015).

position to a higher head. Such a movement option has also been explicitly assumed for another case in which a scope-discourse marker precedes the corresponding criterial element, the focus marker *a* in Jamaican Creole:

5. A di bami Piita nyam (… nutn muor) (Jamaican Creole)
   A the bammy Peter eats (... nothing more)
   'It's the bammy that Peter ate (nothing else).' (Durrleman 2008: 74)

Durrleman assumes that *a* moves from the head of the FocP to a higher head position. So, there appears to be the following parameter:

6. Criterial head H moves?

Here, "moves" actually means "moves to the next higher head", under familiar assumptions of structure preservation and locality on movement. Pearce conjectures that the host head of *ko* movement may be Force because topic-marking *ko* is the initial element in the Māori clause, and moreover it can be dropped, recalling C-drop phenomena in other languages;[2] the landing site of *a* movement could not be Force in Jamaican creole, Durrleman argues, because it can be preceded by a topic and co-occurs with an overt complementizer in embedded clauses. So, the landing site may well be different, but both languages appear to share the positive setting of parameter (6) (for Top and Foc, respectively), as opposed to the negative setting of Gungbe-type languages, which overtly preserve in the surface string the criterial Spec-head configuration.

4. Criteral head movement and freezing

Much work has been devoted to the freezing effects that arise in criterial positions, a simple manifestation of which is the fact that a wh-phrase cannot be further moved from the C-system of an embedded question (see Rizzi 2006; 2014; Boskovic 2008, a.o.).

7. a. Bill wonders [what book Q [John should read __]]
   b. *What book does Bill wonder [__ Q [John should read __]]

The freezing effect also holds if the wh-phrase carries an independent criterial feature, e.g., a corrective focus feature on the lexical restriction. In such cases, the corrective focus feature is unable to pied-pipe the wh-phrase to the main complementizer system, as in (8)b, because the whole phrase is frozen in the embedded criterial position.

8. a. Bill wonders [what BOOK Q [John should read __]] (not what paper…)

[2] An alternative analysis that Pearce (1999) considers is that *ko* may be the realization of Force when the CP zone contains a TopP, a selectional property between two heads, rather than a movement relation. Both alternatives are consistent with the mechanism for leaving the edge of the clause unpronounced developed in De Lisser et al. (2016).

b. *What BOOK Bill wonders [__ Q [John should read __]] (not what paper…)

Is the further movement of the criterial head assumed by Pearce and Durrleman consistent with Criterial Freezing? It should be noticed that all the evidence discussed in the freezing literature concerns the unmovability of the specifier in the criterial configuration, whereas nothing is said about the criterial head. So, the freezing effect does not (necessarily) involve the whole criterial configuration. Only the criterial Spec is unmovable, not the head, an asymmetry between the two elements entering into the criterial configuration which is directly expressed in the statement of the effect in Rizzi (2014: 22) in terms of the probe-goal terminology.[3]

9. In a criterial configuration, the Criterial Goal is frozen in place

Moreover, the derivation of the freezing effect in Boskovic (2008) from the inactivation condition, and the derivation of the effect from labeling and maximality in Rizzi (2016) are both consistent with the possibility of a further movement step of the criterial head. So I will assume (6) to be a valid parametric option.

5. Criterial feature expressed in DP?

Pearce (1999) argues that *nā* appearing in the so-called Actor-Emphatic construction in Māori may be analyzed as a focus marker, in a broad sense, with the phrase construed with *nā* appearing in in the left-peripheral Spec-Foc. Nevertheless, the analysis of *nā* is different from the analysis of *ko*. *Nā* appears to have a more direct connection with the Case system, which leads Pearce to analyze *nā Pita* in (4) as a KP (Case phrase) sitting in Spec-Foc. *Nā* would then be a Case-like, DP internal, element marking focus. We thus seem to have the following options:

10. The morphological marker of a criterial property may be:
    a. DP external, a criterial head;
    b. DP internal, a Case-like element.

The idea that certain topic or focus markers may be Case-like DP internal elements has been proposed independently elsewhere. A case in point is, again, Jamaican Creole. Durrleman (2008) proposes that the locative element *de* (connected to English *there*) can function as a topic marker, occurring in DP internal position, in such structures as

11. [da bwaai de], mi laik im (Jamaican Creole)
    that boy LOC I like him
    'As for that boy, I like him.' (Durrleman 2008: 67)

So, both (10)a and (10)b seem to be independently attested. Notice that (10), differently from (3) and (6), should not be thought of as a UG parameter, but simply as the expression of two analytic options (head-marking and dependent marking, in a

[3] The criterial head (Q in the case of questions) enters into a probe-goal relation with the criterial goal (the wh-phrase), which is then attracted to the Spec of the criterial head in wh-movement languages. Then the criterial goal is frozen in the criterial position under (9).

classical terminology) for the status of morphological markers of scope-discourse properties. See Rizzi (2013) for the discussion of possible diagnostics to determine if a given language opts for (10a) or (10b) for a particular scope-discourse construction.

A genuine parameter, though, is the following, expressed in the microparametric format adopted for (3) and (6):

12. Criterial feature F is overtly expressed in the DP?

Under Durrleman's analysis, (12) has a positive value in Jamaican creole for F=Top, and under Pearce's analysis (12) has a positive value for F=Foc in Māori.

A priori, the two options of (3) and (12) are not mutually exclusive, i.e., a language could have overt morphological markers for topicality or focus both within the nominal expression and on the criterial head. I do not know of any such case of double marking for topic or focus. Nevertheless, a close enough analogue is the co-occurrence of the wh-marked operator and an overtly expressed Q head, a case excluded in many languages (the so-called "doubly filled COMP effect"), but possible in other languages, e.g. the Dutch varieties admitting *wie of* ("who if") sequences in embedded questions. Further research is needed to determine whether such cases of double overt morphological marking extend to other cases of criterial configurations, beyond the case of wh-constructions.

Acknowledgment

This research was supported in part by the ERC advanced grant n. 340297 "SynCart".

References

Aboh, Enoch Oladé. 2004. *The morphosyntax of complement-head sequences: Clause structure and word order patterns in Kwa*. Oxford: Oxford University Press.

Belletti, Adriana & Rizzi, Luigi. 2017. On the syntax and pragmatics of some clause-peripheral positions. In Blochowiak, Joanna & Grisot, Cristina & Durrleman, Stephanie & Laenzlinger, Cristopher (eds.) *Formal models in the study of language*, 33–48. Cham: Springer.

Bianchi, Valentina & Frascarelli, Mara. 2010. Is Topic a root phenomenon? *Iberia* 2. 43–48.

Bianchi, Valentina & Bocci, Giuliano & Cruschina, Silvio. 2015. Focus and its implicatures. In Aboh, Enoch Oladé & Schaeffer, Jeannette C. & Sleeman, Petra (eds.) *Romance languages and linguistic theory: Selected papers from Going Romance* 2013. Amsterdam: John Benjamins Publishing.

Bocci, Giuliano. 2013. *The Syntax–prosody interface: A cartographic perspective with evidence from Italian*. Amsterdam: John Benjamins Publishing.

Bošković, Željko. 2008. On the operator freezing effect. *Natural Language & Linguistic Theory* 26. 249–287.

Cinque, Guglielmo & Rizzi, Luigi. 2010. The cartography of syntactic structures. In Heine, Bernd & Narrog, Heiko (eds.) *The Oxford handbook of linguistic analysis*. Oxford: Oxford University Press.

De Lisser, Tamirand Nnena & Durrleman, Stephanie & Rizzi, Luigi & Shlonsky, Ur. 2016. The acquisition of Jamaican Creole: The null subject phenomenon. *Language Acquisition* 23. 261–292.

Durrleman, Stephanie. 2008. *The syntax of Jamaican Creole*. Amsterdam: John Benjamins Publishing.

Frascarelli, Mara, & Hinterhölzl, Roland. 2007. Types of topics in German and Italian. In Winkler, Susanne & Schwabes Kerstin (eds.) *On information structure, meaning and form: Generalizations across languages*. 87–116. Amsterdam: John Benjamins Publishing.

Haegeman, Liliane. 2012. *Adverbial clauses, main clause phenomena, and composition of the left periphery: The cartography of syntactic structures,* vol. 8. Oxford: Oxford University Press.

Pearce, Elizabeth. 1999. Topic and focus in a head-initial language: Māori. *Proceedings of the sixth meeting of the Austronesian Formal Linguistics Association (AFLA)*, *Toronto Working Papers in Linguistics* 16. 249–263.

Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. In Haegeman, Liliane (ed.) *Elements of Grammar: A Handbook of Generative Syntax*. 281–337. Dordrecht: Kluwer.

Rizzi, Luigi. 2013. Notes on cartography and further explanation. *Probus* 25. 1613–4079.

Rizzi, Luigi. 2014. Some consequences of criterial freezing. In Svenionius, Peter (ed.) *Functional structure from top to toe: The cartography of syntactic structures*, vol. 9. 19–54. Oxford: Oxford University Press.

Rizzi, Luigi & Cinque, Guglielmo. 2016. Functional categories and syntactic theory. *Annual Review of Linguistics* 2. 139–163.

Rizzi, Luigi. 2016. Labeling, maximality, and the head–phrase distinction. *The Linguistic Review* 33. 103–127.

Luigi Rizzi
University of Geneva, University of Siena
luigi.rizzi@unige.ch

Unable to say too much about *kano* in Nafsan (South Efate)[1]

Nick Thieberger

Abstract

There is a puzzle about the word *kano* 'to be unable' in Nafsan (Efate, Vanuatu). It only occurs in Nafsan and the nearby language Lelepa, but neither *kano* nor any cognate appears in any of the neighbouring languages and so is an innovation that needs to be explained. It seems likely that the term was introduced by Scottish missionaries, which leads to the question of how easily terms can be adopted by a speech community when introduced by a single outsider who is promoting a foreign religion. This squib relates to similar work I am undertaking examining the introduction by missionaries of the conditional into Nafsan that will be reported elsewhere (Thieberger 2017).

1. It looks like 'cannot'

In Nafsan (South Efate, Vanuatu) *kano* 'to be unable' can function either as a verb, as in (1), or as an auxiliary, as in (2) (for a discussion of the status of the auxiliary in Nafsan see Thieberger 2006).

1. Rui=lel-es panpan me ru=**kano**-wes
   3PL.PS=look_for-3SG.OBL until and 3PL.RS=unable-3SG.OBL
   'They looked and looked but they couldn't find them there.' (098:024)

Kano acts as an auxiliary verb in (2) where it precedes the verb complex *trau daerek pan leka* 'just go straight and look at him' containing the main verb *lek* 'to look'.

2. Boyfren neu a=**kano** trau daerek pan lek-a-ø.
   boyfriend mine 1SG.RS=unable just straight go look-TS-3SG.O
   'My boyfriend, I can't just go straight and look at him.' (066:90)

Of the neighbouring languages, only Lelepa has *kano* (see Table 2) which is not found in other languages around Nafsan. There is a form *kanokaanoa* with a similar meaning, which occurs in the Polynesian neighbour, Atare Imere (Clark 1998: 20). According to Claire Moyse-Faurie (p.c.) it is likely that *kanokaanoa* has Nafsan as a source.

[1] I'm very happy to be able to contribute to this volume celebrating Liz Pearce's career in linguistics. Note that example references are to text and sentence numbers in the corpus. Abbreviations (other than standard Leipzig forms) are as follows: OBL – Oblique Pronoun; PS – Perfect Subject; RS – Realis Subject; TS – Transitivising Suffix;

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 211–215

The Nafsan corpus on which this analysis is based includes scanned and OCRed versions of the earliest sources in the language, dating to the first Christian translations in 1864, in addition to texts written from recordings made during my fieldwork over the past 20 years[2]. It is thus possible to track the use of *kano* over time within this corpus. While we know that the early translations are not perfect representations of the language as it was spoken, they are the only sources we have for this time period. As can be seen in Table 1, the two earliest publications do not include *kano*. These are two quite short books with fewer than 1000 words each, so it may be that there is no possible context for the concept 'unable' in either of them. However, in the 1864 *Nadus Iskei nig Fat* the text *nad i tap nrogo mou* 'the person did not hear it' occurs at verse 14, which could conceivably have been rendered as 'he can't hear it', thus suggesting that the form *kano* was not yet in use in the language at this time.

From 1868 onwards *kano* appears in all published work in the language. From my experience of the language it is the normal way to express inability, occurring many times in my collection of over 100 texts and as shown in example 2 above and in the 2014 entry in Table 1 from a recent Facebook posting.

Table 1: Texts in Nafsan by year and use of *kano*

Year / Publication	**Use of kano in the text**
1864 *Nalag nig Efat*	No use of *kano*
1864 *Nadus Iskei nig Fat*	No use of *kano*
1868 *Nalag nig Efat* (Hymnbook)	One use of *kano* (out of 1700 words): I bakelag berkati tok; A **kano** bakor wis. 'He is really up on high; I can't come to him.'
1874 *Kenesis (Genesis)*	25 uses of *kano* (out of 2,800 words): Ken ki nrik Yeof kin, nanrogtesan nigneu i tob, kineu a **kano** selati. 'And Cain said unto the Lord, My punishment is greater than I can bear.' (4:13) Nawisien a nin i bi te nag Yeof i brig i; Komam ra **kano** nrik kik o ki tesa, ko te wi. 'The thing proceedeth from the LORD; we cannot speak unto thee bad or good.' (24:50)
2014 *Erakor Village Facebook page*	lakn teläp ruk4 **kano** pakot nafet MinicipalityTAX 'Because many are unable to pay the Municipal tax'

Where it has been possible to locate equivalents for 'unable' in languages neighbouring Nafsan they are all of a different form to *kano*, except for Lelepa and Atara Imere as shown in Table 2. The strong resemblance between the forms found in Ngunese and Namakir suggests that a cognate form may have been available in Nafsan, but we have no trace of it in any of the records. There is no form similar to *marisaa* or *marisa* in any text in the corpus.

[2] For a description of the Nafsan corpus see https://rebrand.ly/NafsanMaterial

It is important to understand that Nafsan was the lingua franca used by missionaries in the late 1800s and early 1900s (Miller 1991) and so *kano* is likely to have spread with the same missionary effort with which it seems to have entered Nafsan originally.

Table 2: Expression of 'to be unable' in neighbouring languages[3]

Namakir (north of Efate)	*Mar'isa* 'cannot, should not, unable to' Ni **mar'isa** ni loioh na-bitiren na-tonenio bitin' 'I cannot/shouldn't swim because my leg is sore' (Sperlich 1991: 307)
Nguna (north of Efate)	*Marisaa* E **marisaa** euaaua naga na-gorai ni na-toko-ana pota e ga laki pa-ki na-toko-ana rota 'It wouldn't allow a girl from one village to marry [someone from] another village.' (Schütz 1969: 6)
Atara Imere (Efate)	*kanokaanoa* (vi) 'be unable' (Clark 1998: 20) (Claire Moyse-Faurie (p.c.) suggests this is innovated and probably borrowed from Nafsan)
Lelepa (west of Efate)	*kano* 'be unable' (Lacrampe 2014: 280)
Ura (Erromango)	*davawi* 'to be unable' Yau **davawi** nivan yerema woreci imo ni ra dahmas. 'I cannot climb this tree because it is big.' (Crowley 1999: 36)

2. The language of the missionaries

Since it is likely that *kano* was introduced by missionaries, what do we know of the language they spoke? They were Scottish from Nova Scotia and there is one example in the correspondence of James Cosh writing to his sister Maggie that shows that he spoke Scots English: "If ye get married, gang somewhere else an seek your hame - but don't come here to pine away in solitude and sickness" (Denne 1991: 20). This suggests that he, and perhaps his predecessor Morrison and successor Mackenzie would have used the form *canna* 'can't'[4] and that it could have been a source for Nafsan *kano*, despite being what Clark (1982) would call an 'unneccessary borrowing'.

3. Introducing new forms?

How can a couple of missionaries effect a change of this kind, introducing a new form that is then taken up by the whole speech community and continues to have currency today? We have to consider that the current population of some 5,000 speakers does not reflect the size of the villages of Erakor or Pango in the late 1800s. McArthur (1981: 22) estimates the population of the whole of Efate in 1874 as only 2000. In 1853 there were 250 Pango people in church (Steel 1880: 223), and Miller (1991) says the population of Erakor in 1945 was 200, indicating a much smaller number of Nafsan speakers in the late 1800s. This smaller population makes it more plausible that *kano*

[3] Where information is available. I have not been able to find examples from Eton or Sye.

[4] See for example the online Scots English dictionary (http://scots-online.org/dictionary/english_scots.asp)

was introduced and taken up by the whole population within a period of a decade or so, especially if the Christian mission's teachings were seen as prestigious by the locals (that is not the current topic, but see Monberg (1967) for more on the conversion of small communities).

4. Conclusion

In trying to determine the origin of the unusual form *kano* I have shown that it is not related to other words with similar meanings in neighbouring languages, except for Lelepa and Atara Imere, both of which can plausibly have borrowed it from Nafsan. A small corpus of historical texts available for Nafsan shows that *kano* has been present in the language since 1868. Missionary work began in the early 1860s and the Scottish missionaries would have used a form like *canna* in their English for 'cannot', most likely providing the source for Nafsan *kano.*

References

Anon. 1864. *Nalag nig Efat.* Aneityum: Mission Press.

Anon. 1868. *Nalag nig Efat.* Sydney: Mason, Firt, nigar asler (Mason, Firth and Co).

Bible. 1864. *Nadus Iskei nig Fat*. Aneityum: Mission Press.

Bible. 1874. *Kenesis Natus a Bei Nag Moses Ki Mtir I*. Translated by James Cosh. Sydney: British and Foreign Bible Society.

Clark, R. 1982. 'Necessary' and 'unnecessary' borrowing. In Halim, Amran & Carrington, Lois & Wurm, Stephen Adolphe (eds.) *Papers from the Third International Conference on Austronesian Linguistics,* vol.3: *Accent on variety* (Pacific Linguistics Series), 137–143. Canberra: Department of Linguistics, Research School of Pacific Studies, The Australian National University.

Clark, Ross. 1998. *A dictionary of the Mele language (Atara Imere), Vanuatu.* Canberra: Pacific Linguistics.

Crowley, Terry. 1999. *Ura: A disappearing language of southern Vanuatu.* Canberra: Pacific Linguistics.

Denne, Janet. 1991. *Rev. James Cosh M.A., D.D., 1838–1900: Minister, missionary and academic*. Killara, N.S.W.: Denne Design.

Lacrampe, Sebastien. 2014. *Lelepa: Topics in the grammar of a Vanuatu language*. Canberra: ANU. (Doctoral dissertation) https://openresearch-repository.anu.edu.au/handle/10440/1026.

McArthur, N. 1981. "New Hebrides Population 1840–1967: A reinterpretation." Noumea: South Pacific Commission. (Occasional paper no.18)

Miller, Graham. 1991. Brief grammar & vocab of Erakor dialect of Efatese. Pacific Manuscripts Bureau MF 1028. (Unpublished Ms.)

Monberg, Torben. 1967. An island changes its religion: Some social implications of the conversion to Christianity on Bellona Island. In Highland, Genevieve & Howard, Alan & Sinoto, Yosihiko (eds.) *Polynesian culture history: Essays in honor of Kenneth P. Emory*. 565–589. Honolulu: Bishop Museum, Honolulu.

Sperlich, Wolfgang. 1991. *Namakir: A description of a Central Vanuatu language*. Auckland: University of Auckland. (Doctoral dissertation)

Schütz, Albert J. 1969. *Nguna grammar* (Oceanic Linguistics Special Publication). Honolulu: University of Hawai'i Press.

Steel, Robert. 1880*. The New Hebrides and Christian missions with a sketch of the labour traffic and notes of a cruise through the group in the mission vessel.* London: James Nisbet & Co.

Thieberger, Nicholas. 2006. The benefactive construction in South Efate. *Oceanic Linguistics* 45(2). 297–310.

Thieberger, Nick. 2017. Missionary-induced language change, on the trail of the conditional in Nafsan, central Vanuatu. *History and Philosophy of the Language Sciences.* (https://hiphilangsci.net/2017/05/30/missionary-induced-language-change/)

Nick Thieberger
University of Melbourne
thien@unimelb.edu.au

Adjacency and DP licensing

Lisa deMena Travis

1. Introduction

In this paper, I first present a proposal made by Levin (2015) and then give some data that may present problems for the proposal. The proposal basically says that, as an alternative to licensing a DP by case, a language may have a mechanism of adjunction (PF merger) of a nominal head to a verb that serves the same purpose as case licensing. By the time the DP reaches the PF interface, it will be licensed and the structure will be able to be interpreted.[1]

2. Levin (2015)

Levin, in a proposal similar to Lamontagne & Travis (1987) and Travis & Lamontagne (1992), correlates the Case Filter of Chomsky (1981) with the requirement that all DPs have KPs with a valued K. Within this context, he explores some interesting situations, however, where the DP must be adjacent to the potential case licenser. Even more interesting is that not only does the relevant DP have to be adjacent to, in these cases, the V, but the elements within the DP must appear in a particular order. The overall generalization is, he claims, that the highest phonetically realized head of the DP has to be string adjacent to the V.

[1] This paper owes much to Liz Pearce. Perhaps the most obvious contribution is her important theoretical work on the syntax of Austronesian languages. Her work on the structure of DP using claims of Kayne (1994) and Cinque (2005) led the way for many other researchers working on the internal structure of Austronesian DPs. Her carefully worked through theoretical accounts have opened doors for many. But beyond this, through her incisive theoretical questions, she has amassed a body of data that will be useful to generations of linguists. Finally, she embodies all that is good in a scholar. As well as being a careful conscientious researcher, she is a warm and generous colleague and mentor.

The work presented in this paper was supported by the research grants SSHRC 410-2011-0977 (Ileana Paul, PI), 435-2012- 0882 and 435-2016-133 (Lisa Travis, PI). I appreciate input from colleagues and students in Canada (especially Ileana Paul and the McGill Ergativity/Fieldwork Group), colleagues and students in Madagascar (especially Baholisoa Ralalaoherivony, Jeannot Fils Ranaivoson and the students on the SSHRC Dialect Project), and consultants Vololona Razafimbelo, Rina Rajaharison, and Rita Hanitramalala. I also benefited from discussions with Jessica Coon, who has collected Chuj data that is also problematic for Levin's account. All errors are my own.

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 217–228

2.1. The data

We will see two examples of this adjacency requirement below – one where the highest phonetically realized head in the object DP has to be adjacent to the V (similar to more traditional accusative case), and one where the highest phonetically realized head in the non-subject Agent DP has to be adjacent to the V (similar to ergative case).[2]

2.1.1. Internal order within adjacent objects

Comparing (1a) and (1b) below, we see a case of Pseudo-Noun Incorporation (PNI) in Tongan (Ball 2004; 2009). In (1a), the word order is VSO, and in (1b) it is VOS. PNI is discussed at length in Massam (2001) but for the purposes of Levin (2015) and the present paper, the main thing to note is that in the PNI construction in (1b), the object must be adjacent to the verb. For Levin, this means that this DP does not have a KP and therefore PF merger is required for proper licensing of the DP.[3,4]

1. Tongan Pseudo-Noun Incorporation (PNI)

   a. Na'e <u>tō</u> 'e Sione **'ene** **manioke** **ki'i**.
      PST plant ERG S. his cassava small
      'Sione planted his small amount of cassava.'

   b. Na'e <u>tō</u> **manioke** **ki'i** 'e Sione
      PST plant cassava small ERG S.
      'Sione planted his small amount of cassava.'

The examples that we have just seen show an order within the object DP of NA (*manioke ki'i* 'cassava' > 'small'). In the examples in (2), the order has been changed to AN (*ki'i manioke* 'small' > 'cassava'). We can see that this order is possible when there is no PNI as in (2a) but not in the PNI construction in (2b).

[2] In some of the languages under discussion such as Indonesian and Malagasy, the typological status of the language is controversial. In an ergative analysis of these languages, what I will be calling the subject would be labeled Absolutive. What I will be calling the non-subject Agent would be labeled Ergative. See, e.g. Aldridge (2011) for a more detailed ergative analysis of these languages.

[3] Licensing verbs will be <u>double underlined</u> and highest nominal heads will be boxed in order to help the reader track whether they are adjacent or not. Adjacency will only be required of KP-less nominals.

[4] ABBREVIATIONS: ACC - Accusative; ART - Article; DEF - Definite; DET - Determiner; ERG - Ergative; GEN - Genitive; IV - Instrumental Voice; NOM - Nominative; NUM - Number; OV - Object Voice; PASS - Passive; PST - Past; SG - Singular; SV - Subject Voice.

2. Tongan PNI disallows pre-nominal modifiers

   a. Na'e tō 'e Sione 'ene **ki'i** **manioke**
      PST plant ERG S. his small cassava
      'Sione planted his small amount of cassava.'

   b. *Na'e tō **ki'i** **manioke** 'e Sione
      PST plant small cassava ERG S.
      'Sione planted his small amount of cassava.'

Levin's claim is that when a nominal projection is being licensed not by having a KP but by PF head adjunction to the verb, as in the PNI construction, the highest phonologically realized head of the nominal projection must be string adjacent to the licensing verb. In (2b) above, the highest phonetically realized head is the nominal itself, *manioke* 'cassava', but the adjective intervenes between the head and the verb making the PF head adjunction impossible, causing the derivation to crash at the PF interface.

2.2. Internal order within adjacent Agents

Levin gives examples from Balinese to show that this sort of external (the entire DP) and internal (the highest phonetically realized head) adjacency requirement is found in non-subject Agents.[5] We start first with a baseline example from Balinese where the form of the verb (Subject Voice) determines that the Agent is the subject of the clause.

3. Tiang laka numbas bawi-ne punika
   I will SV.buy pig-DEF that
   'I will buy that pig.' SUBJECT VOICE

When the Theme is the subject of the clause, the form of the verb must be Object Voice, and now the Agent must be immediately post-verbal. While the Theme may appear sentence finally (4b), neither the Theme (4c) nor a prepositional phrase (4d) may intervene between the verb and the Agent.

4. a. Siap-e uber cicing ke jalan-e
      chicken-DEF OV.chase dog into street-DEF
      'A dog chased the chicken into the street.' OBJECT VOICE

   b. uber cicing ke jalan-e siap-e
      OV.chase dog into street- DEF chicken- DEF
      'It was a chicken that the dog chased into the street.'

   c. *uber siap-e cicing ke jalan-e

   d. *siap-e uber ke jalan-e cicing

[5] As mentioned in fn 2, in some analyses these Agents might be interpreted as being ergative.

Levin also points out that, as with the Tongan examples above, the order of elements within the nominal projection is restricted. We see that *liu* 'many' may appear on either side of the head when within the Agent or the Theme of a Subject Voice construction shown in (5a) and (5b) respectively, and within the Theme of an Object Voice construction shown in (5c). In (5d), however, where Levin's account claims the licensing of the DP is dependent on PF head adjunction, we see that the quantifier may not intervene between the verb and the nominal head of the projection.[6]

5. a. (liu) cicing (liu) ngugut Nyomain
   (many) dog (many) SV.bite. N.
   'Many dogs bit Nyoman.' AGENT SUBJECT

   b. Cicing-e ngugut (liu) anak cerik (liu)
   dog-DEF SV.bite (many) person small (many)
   'The dog bit many children. *v*P INTERNAL THEME

   c. (Liu) anak cerik (liu) gugut cicing
   (many) person small (many) OV.bite dog
   'A dog bit many children.' THEME SUBJECT

   d. Nyoman gugut (*liu) cicing (liu)
   N. OV.bite (*many) dog (many)
   'Many dogs bit Nyoman.' *v*P INTERNAL AGENT

Another data point from Balinese that supports Levin's observation is that fact that non-subject Agents cannot be definite. First he shows that the definite marker, while realized as a suffix, is generated in a high head position. It may appear on the nominal head as in (6a) but will appear after any other material that might follow the head as in (6b) where it is suffixed to an adjectival modifier and in (6c) where it is suffixed to the object of a preposition found in a PP that modifies the head.[7]

6. Definite marker: (right edge) (therefore in a head position)

   a. dagang-e
   trader-DEF
   'the trader'

[6] Arka (2003: 48ff) discusses the variation in the position of quantifiers. We can see that adjectives generally do not have this variable order (*luung* 'good' *umah*; *umah luung*, **luung umah*). According to Arka, the variability discussed by Levin, then, would be linked to quantifier float. Arka restricts quantifier float to definite terms but we see in (5a-5c) that the quantifier may appear (float) pronominally. I leave a more detailed look at this for future research.

[7] Since Balinese appears otherwise to be head-initial, a roll-up account of the Balinese DP along the lines of Pearce (2003; 2005) and Kahnemuyipour and Massam (2006) would be appropriate. Such an account, however, would not create problems for Levin's proposal. I will return to DP internal roll-up in Section 3.3, however, where a problem is arguably created in a similar structure.

b. dagang cekeng- e
trader pig- DEF
'the pig trader'

c. dagang ceken uli Badung- e
trader pig from Badung-def
'the pig trader from Badung'

Now placing these constructions within the non-subject Agent position, we see that the definite marker cannot appear. This is explained by the same requirement of adjacency on the highest realized head in the nominal projection. With the bare N, adjacency of the N is sufficient. But when there is also a determiner in D, if that determiner is not adjacent to the V, PF adjunction cannot occur, the DP will not be licensed, and the derivation will crash.

7. I Wayan gugut cicing / *cicing-e (ento)
ART W. ov.bite dog/*dog-DEF (that)
'A/*that dog bit Wayan.'

Further data shows that the restriction is not against having definite non-subject Agents as both pronouns (8a) and proper names (8b) can appear within this position.[8] This follows from Levin's proposal as the highest nominal head will be adjacent to the licensing verb.

8. a. be-e daar ida
fish-DEF OV.eat 3
'(S)he ate the fish.'

b. be-e daar Nyoman
fish-DEF ov.eat N.
'Nyoman ate the fish.'

In the next section I will present two data patterns that might be problematic for Levin's proposal. At this point, I simply raise the issues without providing an alternative account, leaving that for future work.

3. Problematic data patterns

In this section I begin by presenting data from Bahasa Indonesia, a language related to Balinese. Bahasa Indonesia also shows adjacency effects in a slightly different construction from the one presented for Balinese in (5) above but does allow head-final determiners, calling into question the cross-linguistic validity of the adjacency effect. Then I present data from Malagasy, another Malayo-Polynesian language that is

[8] Given Arka's claim that quantifiers can only float to the DP initial position within a definite DP, for him there would only be one relevant datapoint in Balinese (i.e. that the only definite non-subject Agents are pronouns and proper names). The ungrammaticality of (5d) would then follow from the ungrammaticality of (7).

shown by Levin to provide support for his observations. I argue that while part of the paradigm in Malagasy does show the adjacency facts needed for PF adjunction, other parts of the paradigm create a problem.

3.1. Bahasa Indonesia

Bahasa Indonesia has two different constructions that have the effect of making the Theme the subject. One is similar to the OV construction in Balinese. While there are interesting characteristics of this construction in the context of Levin's proposal, I concentrate on the other one here – one that I will call passive.[9] In the passive, there is extra morphology on the verb (the prefix *di-*) and the Agent may appear in a post-verbal by-phrase (*oleh*-phrase). As the data below show, the *oleh*-phrase may either appear adjacent to the verb, as in (9a), or separated from the verb, as in (9b). However, as we can see in (9a), when the Agent is adjacent to the verb, the preposition is optional, but when not adjacent to the verb, as in (9b), the preposition is no longer optional (data adapted from Guilfoyle, Hung & Travis 1992).

9. Bahasa Indonesia: passive
   a. Uang itu di-kirim (oleh) Ali kepada Tomo
      money the PASS-send by Ali to Tomo
      'The money was sent to Tomo by Ali.'
   b. Uang itu di-kirim kepada Tomo *(oleh) Ali
      money the PASS-send to Tomo by Ali
      'The money was sent to Tomo by Ali.'

Thinking in terms of Levin's proposal, we could say that the Agent is either within a KP valued by the P or it is a DP that gets licensed by V through PF adjunction. When it is licensed by the V, however, it must be adjacent to the V, making it look similar to the Balinese data we saw in (4). If this is the case, we would expect the order within the nominal projection to also be restricted with the highest realized head having to be string adjacent to the verb. But the following example shows that this is not what happens. Here the demonstrative, *itu*, arguably the highest realized head, appears head-finally, as it does in Balinese. Therefore, it is not adjacent to the verb.

10. Buku itu dibaca [DP lelaki itu]
    book DET PASS-read man DET
    'The man read the book.'

There may be another mechanism at play here to explain the licensing of the non-subject Agent that can explain (a) the requirement that the bare DP must be adjacent to the V, and (b) the lack of the adjacency requirement of the highest realized head, here the D. For example, perhaps the determiner is not in a head in Bahasa Indonesia whereas it is in Balinese. Or perhaps the DP in (10) is within a KP where the K is valued

[9] I do not do these constructions justice here. See e.g. Chung (1976) and Arka & Manning (2008) for more details.

in some way. If either of these were to be the case, however, the correlations would start to become quite subtle and require independent evidence.

We turn now to Malagasy, which also provides data where an adjacency requirement of a bare DP does not correlate with word order within the DP.

3.2 Malagasy

Levin uses data from Malagasy (a VOS language spoken in Madagascar) to provide nice confirmation for his proposal. Non-subject Agents in Malagasy have to be adjacent to the verb as in Balinese, however, unlike Balinese they can be definite. As Levin points out, this actually follows from the fact that determiners are head-initial in Malagasy. Further, these head-initial determiners appear to undergo observable adjunction to the verb (see Travis 2006b; 2006a; 2008 for a different account for this). This morphological adjunction is labeled N-bonding by Keenan (2000). In the Actor Topic form ((11a) glossed SV for Subject Voice) there is no N-bonding because the Agent is in the subject position. In the other two constructions, where either the Theme is the subject ((11b) glossed OV for Object Voice) or the Instrument is the subject ((11c) glossed IV for Instrumental Voice), we have N-bonding between the verb and the non-subject Agent.[10]

11. Malagasy: N-bonding (Keenan 2000)

    a. n-andrakotra azy t-amin'ny bodofotsy ny reniny
       PST-SV-cover 3ACC PST-with. GEN'DET blanket DET mother.3GEN
       'His mother wrapped him with a blanket.' ACTOR TOPIC

    b. no-rakofan'**ny** **reniny** t-amin'ny bodofotsy izy
       PST-OV.cover.GEN'DET mother.3GEN PST-with.GEN'DET blanket 3NOM
       'He was wrapped with a blanket by his mother.' THEME TOPIC

    c. n-andrakofan'**ny** **reniny** azy ny bodofotsy
       PST-IV-cover.GEN'DET mother.3GEN 3ACC DET blanket
       'He was wrapped with a blanket by his mother.' CIRCUMSTANTIAL TOPIC

We see in these forms that there appears to be visible adjunction between the highest phonetically realized head in the nominal structure, in this case the determiner *ny*, and the verb. The examples below show that there are phonological effects on the verb, and we can also see that pronouns and proper names seem to be affixed onto or compounded with the verb.

12. a. Sasana ny lamba
       OV.wash DET clothes
       'The clothes were washed.' NO AGENT

[10] The labelling of these structures and verb forms vary depending on the researcher. I have glossed the verb forms as Subject Voice, Object Voice and Instrumental Voice to parallel the glosses for Tongan and Balinese. The more traditional labels of Actor Topic, Theme Topic and Circumstantial Topic are given in the margin beside each construction.

b. Sasako ny lamba
OV.wash-1SG DET clothes
'The clothes were washed by me.' PRONOMINAL AGENT

c. Sasan-dRabe ny lamba
OV.wash.GEN-Rabe DET clothes
'The clothes were washed by Rabe.' PROPER NAME AGENT

We can see in Malagasy, then, that Levin's account can predict that definite non-subject Agents will be possible, unlike in Balinese, since the phrase-initial determiner will be adjacent to the licensing verb. Further, phonological effects of this adjacency between the V and the DP lend credence to his proposal for PF adjunction. We will see below, however, that other constructions in Malagasy may create problems for his account.

3.3 Malagasy measure phrases

I start the discussion of another possible counter-example to Levin's account with an introduction to measure phrases in Malagasy as these nominal constructions will become relevant. We can first note that cardinal numbers sometimes appear after the head noun as shown in (13a) and sometimes before it as shown in (13b).

13. a. ankizy roa
child two
'two children'

b. roa taona
two year
'two years'

I explain the difference in order using a particular view of DP internal roll-up movement (see Pearce 2003; 2005; and Kahnemuyipour & Massam 2006 for related work on related languages) and the structure of measure phrases. We see below what happens when we have full measure phrases, i.e. nominal extended projections that include a measure word, *tavoahangy* 'bottle', and the measured material, *divay* 'wine'.

14. divay roa tavoahangy
wine two bottle
'two bottles of wine'

The cardinal follows the material and precedes the measure. This fuller structure helps us understand the data in (13). The cardinal generally follows the head nominal unless it is a measure phrase, such as *taona* 'year' (also *metatra* 'meter', etc.). When it is clear what material the measure phrase is measuring (*taona* 'year' measures time), then the material part of the structure may be left unspecified. When the unit of measure is obvious as is the unmarked case with children, then the unit of measure may be left unspecified.

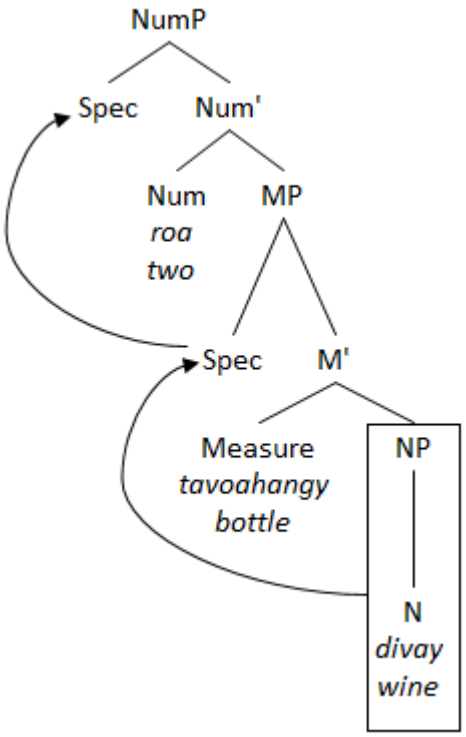
As shown below, this explains why the cardinal appears to be generated in two different positions. In fact, it is always in the same position in the structure. What varies is what part of the structure is left unrealized.

15. a. ankizy roa <measure>
child two
'two (units of) children'

b. <material> roa taona
two year
'two years (of time)'

Now we turn to the structure itself. As we have seen above, Malagasy is a VOS language. Many researchers working on the syntax of Malagasy assume an account where the VOS word order is derived from an SVO order through fronting of the VP (see e.g. Pearson 2005 for details). I propose a similar derivation for measure DPs as shown in the tree below. The NP which represents the material being measured as the semantic head of the projection is the lowest node and therefore furthest to the right. But this NP then moves to the highest Spec position.[11]

16. The structure



[11] This Spec-to-Spec movement is allowed for in Cinque's work (2005) but it does not represent the roll-up that might be expected from the work on Austronesian DPs in Pearce (2003; 2005), and Kahnemuyipour & Massam (2006), and the work on Malagasy VPs in Pearson (2000; 2005). I have shown elsewhere (2015) that dialects other than the main Merina dialect do show the word order of roll-up. I leave the questions surrounding this variation for future work.

Note that the order of the cardinal with respect to the nominal changes with the function of that nominal. For example, if *tavoahangy* 'bottle' is used as a measure, the cardinal precedes nominal. If *tavoahangy* 'bottle' is used as a concrete object, the cardinal follows the nominal.

17. a. nividy tavoahangy roa aho
PST-SV.buy bottle two 1SG
'I bought two bottles.'

b. nividy divay roa tavoahangy aho
PST-SV.buy wine two bottle 1SG
'I bought two bottles of wine.'

Turning now to the matter at hand, it is important to note that the leftmost item in the nominal phrase is not the highest phonetically realized head. The highest phonetically realized head would be the cardinal.[12]

Given Levin's proposal for Malagasy, we would then expect these full measure phrase constructions not to be able to appear in the non-subject Agent position. We can see below, however, that this is not the case. The highest head *roa* 'two' (or alternatively *tavoahangy* 'bottle' if the cardinal is assumed to be in a specifier position) is not adjacent to the verb.

18. a. lenan-divay roa tavoahangy ny lambako
wet-wine two bottle DET clothes-1SG
'My clothes were (made) wet by two bottles of wine.'

b. hitan-olona roa ankolafiny ny fantara
see-people two group DET falling.star
'The falling star was seen by two groups of people.'

In sum, Malagasy at first blush presents a nice confirmation for Levin's proposal. It allows definite non-subject Agents as expected because it is a determiner first language, placing the determiner adjacent to the verb. Further, there are morpho-phonological effects seen in the adjacency supporting the proposal for a PF adjunction process. However, going beyond the garden variety nominals and using measure phrases, we see that complications arise. In these structure, the leftmost item in the projection is not the highest head.

4. Conclusion

Austronesian languages show interesting adjacency effects between Vs and argument DPs – effects that appear to be related to some mechanism of argument licensing. We see this with objects in Pseudo-Noun Incorporation, and we see it with the licensing of

[12] If we wanted to say that the cardinal number was in a specifier position and thus would not be the highest head, the problem would still be the same because the highest head would then be the measure, in this case *tavoahangy* 'bottle'.

non-subject Agents. In this paper, I have looked particularly at the instances of non-subject Agents in Tongan, Balinese, Indonesian, and Malagasy. Levin (2015), using data from Tongan, Balinese, and Malagasy, argues that the adjacency effect is the result of a requirement that the highest phonetically realized head within the nominal must PF adjoin to the verb. This requirement explains not only the adjacency requirement on the DP but the word order within the DP. While this account does explain interesting facts of word order restrictions, I have given two examples – one from Bahasa Indonesia and one from Malagasy – that call into question whether the nominal internal word order requirement is the appropriate one. Levin's work has brought these data into the spotlight and his account has captured a promising hypothesis, but more such patterns need to be studied in order to determine the appropriate cross-linguistic facts and to understand the phenomenon more fully.

References

Aldridge, Edith. 2011. Antipassive in Austronesian alignment change. In Jonas, Dianne & Whitman, John & Garrett, Andrew (eds.) *Grammatical change: Origins, nature, outcomes*, 331–345. Oxford: Oxford University Press.

Arka, I. Wayan. 2003. *Balinese morphosyntax: A lexical-functional approach*. Canberra: Pacific Linguistics (Australian National University).

Arka, I. Wayan & Manning, Christopher D. 2008. Voice and grammatical relations in Indonesian: A new perspective. In Austin, Peter K. & Musgrave, Simon (eds.) *Voice and grammatical relations in Austronesian language*s, 45–69. Stanford University: CSLI Publications.

Ball, Douglas. 2004. Tongan noun incorporation: Lexical sharing or argument inheritance. In Müller, Stefan (ed.) *Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar*. Stanford University: CSLI Publications.

Ball, Douglas. 2009. *Clause structure and argument realization in Tongan*. Stanford University. (Doctoral dissertation)

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris Publications.

Chung, Sandra. 1976. An object-creating rule in Bahasa Indonesia. *Linguistic Inquiry* 7(1):41–87.

Cinque, Guglielmo. 2005. Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry* 36(3). 315–332.

Guilfoyle, Eithne & Hung, Henrietta & Travis, Lisa. 1992. Spec of IP and Spec of VP: Two subjects in Austronesian languages. *Natural Language and Linguistic Theory* 10. 375-414.

Kahnemuyipour, Arsalan & Massam, Diane. 2006. Patterns of phrasal movement: The Niuean DP. In Gärtner, Hans-Martin & Law, Paul, & Sabel, Joachim (eds.) *Clause structure and adjuncts in Austronesian languages*, 125–150. Berlin: Mouton de Gruyter.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Keenan, Edward. 2000. Morphology is structure: A Malagasy test case. In Paul, Ileana & Phillips, Vivianne & Travis, Lisa (eds.) *Formal issues in Austronesian linguistics,* 27–48. Dordrecht: Kluwer Academic Publishers.

Lamontagne, Greg & Travis, Lisa. 1987. The syntax of adjacency. In Crowhurst, Megan. J. (ed.) *WCCFL 6: Proceedings of the West Coast Conference on Formal Linguistics*, 173–186. Stanford University: CSLI.

Levin, Theodore. 2015. *Licensing without Case*. Cambridge MA: MIT. (Doctoral dissertation)

Massam, Diane. 2001. Pseudo noun incorporation in Niuean. *Natural Language and Linguistic Theory* 19(1). 153–197.

Pearce, Elizabeth. 2003. Phrasal movement within the Māori DP. In *Digests of selected papers presented at AFLA X* (University of Hawai'i at Mānoa Working Papers in Linguistics 34), 41–42. University of Hawai'i at Mānoa Linguistics.

Pearce, Elizabeth. 2005. Phrasal movement within the Māori DP. Wellington: Victoria University of Wellington. (Unpublished ms.)

Pearson, Matt. 2000. Two types of VO languages. In Svenonius, Peter (ed.) *The derivation of VO and OV*, 327–363. Amsterdam: John Benjamins Publishing Company.

Pearson, Matt. 2005. The Malagasy subject/topic as an A'-element. *Natural Language and Linguistic Theory* 23(2). 381–457.

Travis, Lisa. 2006a. Voice morphology in Malagasy as clitic left dislocation or Malagasy in Wonderland: Through the looking glass. In Gärtner, Hans-Martin & Law, Paul & Sabel, Joachim (eds.) *Clause structure and adjuncts in Austronesian languages*, 125–150. Berlin: Mouton de Gruyter.

Travis, Lisa. 2006b. VP-, D-movement languages. In Zanuttini, Raffaella & Campos, Hector & Herburger, Elena & Portner, Paul (eds.) *Crosslinguistic research in syntax and semantics: Negation, tense and clausal architecture*, 127–150. Georgetown: Georgetown University Press.

Travis, Lisa. 2008. Bahasa Indonesia: A window on typology. *Lingua* 118(10). 1583–1602.

Travis, Lisa 2015. Word order in measure phrases in Malagasy dialects. *Western Papers in Linguistics* 1(1). Article 20.

Travis, Lisa & Lamontagne, Greg. 1992. The Case Filter and licensing of empty K. *Canadian Journal of Linguistics* 37(2). 157–174.

Lisa deMena Travis
McGill University
lisa.travis@mcgill.ca

Surselvan 1s /-əl/, or: *Jeu anflel quei buca curios*

Dieter Wanner

1. Introduction

In Surselvan, the Raetoromance[1] variety spoken in the Anterior Rhine Valley (Surselva) in the Swiss canton of Grisons, the 1 singular present indicative and imperfect indicative ending for regular verbs is represented by /-əl/. This is unexpected in the context of standard Romance person marking on the verb. Morphs containing /l/in the verbal conjugation, /Vl/, /lV/, /VlV/, are typically associated with 3rd person as a reflex of Latin ILL- 'distal deictic' in one form or another. The typical 1s marker would rather show a reflex of /–o:/ (e.g. CANTO: present indicative), /Vm/ (e.g., CANTEM present subjunctive, CANTA:BAM imperfect indicative), or an agglutinated reduction of EGO: 'I'. The solution to the Surselvan outcome has long been presented as an extension of an originally phonological development in a rather special context (Meyer-Lübke 1890–1902(2): §133) – an analysis corroborated forcefully by Stimm (1980) (see below for discussion). With this note I wish to explore further the way in which a rather marginal form can come to occupy center stage in the regular verbal inflection. It is a study of analogical pathways, and it provides one example in support of a systematic reappraisal of analogical forces shaping language (Wanner 2006). The interest in this specific development lies in the surprising increase in range and regularity beyond expectations of frequency, conventional analogy, and systemic harmony of /-əl/ in its original setting.

Surselvan is documented (sparsely) since the 17th century. It has developed a (semi-) standardized written format and it possesses a small corpus of religious, literary and other texts (Liver 2010: 116). The current use of Surselvan is regional, but practiced in essentially all styles and media. However, the speakers are by now almost completely bilingual with the local spoken dialect of Swiss German of the Grisons, where Standard German functions in parallel with Surselvan for written and other more formal communications in this diglossic situation. While bilingualism had always had a presence in the Anterior Rhine region, it increased in the 19th century and dramatically in the 20th century. Yet, Surselvan is carried by local pride, conscious practice, and state support. The language community is nevertheless diminishing, since internal migration draws many speakers mainly to the German-speaking parts of Switzerland, foremost the metropolitan area of Zurich. Inversely, there is also considerable tourist and vacation-home migration of (Swiss) German-speaking people. Above all, the language

[1] *Bündnerromanisch* is the better term for this group of languages and dialects located in the Canton of Grisons (i.e., *Graubünden* or *Bünden* in German) and not necessarily genetically related to the Romance varieties of Friulian and Ladin (Dolomitic area) earlier classified as Raetoromance lato sensu. – Gloss for the title: 'I do not find this strange'.

Linguistic travels in time and space: Festschrift for Liz Pearce
Wellington Working Papers in Linguistics 23 (2017) 229–241

has a strictly regional application for communication. The more recent introduction of a common written standard language for Bündnerromanisch, *Rumantsch grischun*, has not been able to change this situation markedly (Schmid 1982).

2. Surselvan verbal inflections

The regular verbal inflection in Surselvan shows a solid system of person markers for the three regular conjugations (clearest in present indicative, also imperfect indicative, but less so in the present and past subjunctive). These markers are in principle phonologically derived from the corresponding Latin endings, except for 1S /-əl/. In addition, subject reference is expressed with a non-null subject strategy involving the regular exposure of a full, even though unstressed, subject pronoun. The paradigm for the /a/ conjugation class, here *cantár* 'to sing' in (1), is also representative for the endings in the /e/ and /i/ classes, e.g., *vender* 'to sell', *dromir* 'to sleep' (Grünert 2003: 85).

1.	a.	PRES IND	1s	*jeu cónt-el*	1p	*nus cant-éin*
			2s	*ti cónt-as*	2p	*vus cant-éis*
			3s	*ela cónt-aØ*	3p	*elas cónt-an*
	b.	IMPF IND	1s	*jeu cantáv-el*	1p	*nus cantáv-an*
			2s	*ti cantáv-as*	2p	*vus cantáv-as*
			3s	*ela cantáv-aØ*	3p	*elas cantáv-an*
	c.	PRES SUBJ	1s	*che jeu cónt-iØ*	1p	*che nus cant-éien / cónt-ien*
			2s	*che ti cónt-ies*	2p	*che vus cant-éies / cónt-ies*
			3s	*che ela cónt-iØ*	3p	*che elas cónt-ien*
	d.	IMPF SUBJ	1s	*che jeu cantáv-iØ*	1p	*che nus cantáv-ien*
			2s	*che ti cantáv-ies*	2p	*che vus cantáv-ies*
			3s	*che ela cantáv-iØ*	3p	*che elas cantáv-ien*

The essential Latin endings serving as the morphophonological base for the relevant 1S, 2S and 3S functions are as in (2).

2.	a.	1S	-O: > Ø	regular loss of post-tonic final *-o* ≠ /-əl/
	b.	2S	-S > *-s*	regular preservation of final *-s*
	c.	3S	-T > Ø	regular loss of final *-t*

Modern Surselvan shows the non-etymological marker /-əl/ in all regular conjugations for 1S.PRES.IND and 1S.IMPF.IND, e.g. *ieu láv-el* [jɛw 'lavəl] 'I wash' (Stimm 1980: 633; Lutz & Strehle 1988: xxv–xxvi).[2] The shape *–el* [əl] for 1S has generated a number of

[2] Irregular conjugations show varying exposure of this morpheme: descendants of *credere* 'to believe', *debere* 'must', *esse(re)* 'to be', **potere* 'can', *stare* 'to stand', **stopere* (<< *est opus*) 'to be necessary', **volere* 'to want' may not frequently show this extension even in local varieties (but still attested *súndel* 'I am', *póssel* 'I can', *víel* 'I see'); other irregulars have a variable presence of /-əl/ (e.g. *sapere* 'to know', *videre* 'to see', *venire* 'to come', *donare* 'to give', *facere* 'to do, make', *laxare* 'to let', *dicere* 'to say'); and yet others present a high

hypotheses to explain its origin (see Stimm 1980 for a complete overview). It is quite evident that this unstressed /-əl/ does not derive from the Latin or Proto-Romance morphological inventory known for the 1s function, be it the salient ending vowel /o:/ (2a) or any of its less distinctive alternates for present subjunctive /-m/ (‘CANTEM, ‘VENDAM), imperfect indicative /-‘V:bam/ (CANTA:BAM, VENDE:BAM), both with regular loss of final /-m/ into Romance, or perfect indicative /-i:/ CANTA:VI:, VE:NI:), with outcomes /i/, /e/, or Ø.

3. Pronominal agglutination

A potentially promising hypothesis for /+əl/ involves the agglutination of a pronominal element. This type of extension of verbal inflection is frequent in the area, specifically in the Lombard dialects of Italian with multiple bona fide pronominal copies integrated into the verbal morphology (3a,b). It is also found in some Raetoromance varieties, e.g. Upper Engadinian (3c), and is most prevalent in 1s and 2s forms.

3. a. *mi* *ə* *m* *lav-i* < **mi eo me lav-eo*
1s.SUBJ 1s.CL 1s.OBJ.CL wash-1s.FLEX ‘I wash myself’ (Lombard)
clitics *ə* < *eo* < EGO: *m* < ME:
tonic subject *mi* < MI:
(Jaberg and Jud 1928–1940(8): map 1683, e.g. pts. 41, 44)

b. *té* *te* *z* *láe-t* < **te te se lava te*
2s.SUBJ 2s.CL REFL.CL wash-2s.FLEX ‘you wash yourself’
clitics *te, t* < TE: *z* < SE: generalized reflexive marker
(Jaberg and Jud 1928–1940(8): map 1684: e.g. pt. 227)

c. [jaw ´tʃɔnt] [ty ´tʃɔntəʃt] [ɛl ´tʃɔntə] ‘1s, 2s, 3s, to sing’
< ´CANTO:, ´CANTA:S, ´CANTAT
(Jaberg and Jud 1928–1940(8): map 1684, pt. 47)

Following this lead, Surselvan /əl/ would be phonologically based on Latin ILL- ‘distal deictic’, adding an empty pronominal reference to the verb form. The model is widespread in the area with forms such as *l’e* for *e* ‘3s is’ where the /l/ portion is an original clitic subject pronoun later absorbed into the verbal morphology. However, this is not convincing for the Surselvan 1s development due to reference: ‘I wash it’ ≠ ‘I wash’, so there is no motivation for the [əl] part to represent a direct object (4a). In addition, the phonological correspondences are at least loose due to a stress differential between the attested outcome and the agglutinated string (4b).

4. a. /ˈlav + əl/: */+əl/ 3s.M.OBJ < ILLU

b. regular ‘ILLU > *el* [‘el], not *[əl])

---

incidence of [-əl] (e.g. *sedere* ‘to sit’, *ridere* ‘to laugh’, *tragere* ‘to pull, carry’, *cocere* ‘to cook’, *jacere* ‘to lie (down)’ (see Decurtins 1958: passim; Spescha 1989: 460). Even in the written language, the suffix is variable or preferably absent with inversion of the subject pronoun: *damónd(el) jeu* ‘I ask’ as quotative expression with V > SUBJ order (a V2 property of Surselvan) vs. regular *jeu damóndel* XP ‘I ask XP’ (Stimm 1980: 633).

Postulating an enclitically agglutinated, reduced subject pronoun instead of the object pronoun in (4a) fares even worse, since now the descendant of ILL- would have subject function (5a). Here, subject reference would be contradictory between 1S and 3S *'I_{SUBJ} wash he_{SUBJ}' (pace Lausberg 1963–1965(2): §798). The adjunction of a 1S subject pronoun would be much more expected on the pattern of the Upper Engadinian and Lombard solutions in (3). One might thus attempt a more complex phonetic development of EGO: in Surselvan 1S, leading to its eventual agglutination. However, deriving [-əl] from a reduction and agglutination of EGO: is phonetically difficult, since the 1S subject pronoun in Surselvan is ['jɛw]. Its agglutination would presuppose a stressless evolution of EGO: in postverbal, later enclitic position and a word-final lateralization of [w] to [l] (5b). However, lateralization of coda-/w/ does not belong to this language, since coda-/l/ is not even velarized with any regularity.

5. a. /'lav + əl/: */+əl/ 3S.M.SUBJ < ILLU
   b. [əl] < [əw] < [ɛw] < /eo/ < EGO: 1S.NOM

Agglutination of a subject or object pronoun simply does not obtain in Surselvan 1S (see Stimm 1980: 634–639 for a forceful refutation of such pronominal agglutination; reaffirmed in Stimm & Linder 1989: 770). The problem is even broader due to the systematic absence in Surselvan of this type of pronominal accumulation (Prader-Schucany 1970: 164; Liver 2010: 165). The Surselvan 2S form with its strictly etymological ending /-s/ compared to the Engadinian forms in /-s+t/ (3b) further discourages the quest for a solution via pronominal agglutination for 1S. Note that Surselvan did not develop subject clitic pronouns at any stage of its known development[3] and that even the object clitics, otherwise a standard feature of Romance idioms, were only weakly present in the language of the 17th century and thereafter disappeared completely, being replaced by unstressed, but full forms derived from ILL-, **mi, *ti*. By the time the /-əl/ element starts to spread in 1S, the pronominal clitics are definitely absent from usage, making the option of an agglutinative solution very unlikely in contrast to the Engadinian and Lombard varieties with active pronominal clitic paradigms. The agglutination solution of a postponed subject pronoun is not in itself inappropriate, it is rather not applicable in Surselvan for reasons of phonetics, reference, and typology.

4. Analogical extension

We are rather dealing with the (perhaps surprising) analogical extension of a specific phonetic outcome. The ending [-əl] originated in those verbs of the *a*-conjugation class that had a stem-final consonant cluster ending in a lateral (e.g. Meyer-Lübke 1890–1902(2): §133; Stimm 1980, including a complete review of the literature). Due to the regular phonetic loss of the ending -O: > /-o/ > Ø , the potentially word-final group */Cl#/ was resyllabified as [Cəl#] (6a). While the earliest texts (17th c.) do not yet

[3] There are morphologised remnants of postverbal subject clitics stemming from the V2 inverted inflection *nus mein ussa* 'we go now' ≈ *ussa mein nus* 'now we go' > **mein-nsV* postverbal atonic clitic reduction of *nus* > *meinsa* morphologised agglutination > *nus meinsa, meinsa nus* reanalysis of *(n)sa* as 1PL person marker (Hacks and Kaiser 2013: 151).

contain any non-phonetic /-əl/ forms for 1S, in the early 18th century this now morphological innovation starts appearing forcefully outside its phonological habitat of (6a). Since then, the morpheme has reached all regular conjugation classes for 1S.PRES.IND (6b), extending also to 1S.IMPF.IND (6c). The exact degree of lexical diffusion across the verbal morphology depends on the local dialect (see Jaberg & Jud 1928–1940(8): maps 1683–8, pts. 1, 11, 16).[4]

6. a. *afflar*: ‘AFFLO > **afl >* *áffel* ‘to find’
maklar: ‘MACCULO > **makl >* *mákel* ‘to stain’
schivlar: ‘SIBILO > **schivl >* *schível* ‘to whistle’

b. *contár* ‘to sing’: *jeu cóntel* 1S.PRES.IND vs. *ti cóntas* 2S.PRES.IND
temér ‘to fear’: *jeu témel* *ti témes*
romír ‘to sleep’: *jeu drómel* *ti dromes*

c. *lavár* ‘to wash’: *jeu lavável* 1S.IMPF.IND *ti lavávas* 2S.IMPF.IND

There is no immediately apparent need for either of the two extensions shown in (6b) and (6c), since Surselvan is a non-null subject language and thus exhibits clear person marking even if in 1S.PRES.IND the person marker were zero and thus identical to 3S (see (2) above).

The creation of /-əl/ as a person marker rests on the availability of an appropriately salient surface form in a specific local context of 1S.PRES.IND, equating a special phonetic string to a marker and reanalyzing it as a 1S (allo)morph. Its extension shows a morphologically determined pathway for the marked form toward the status of a morph(eme) within the function class of 1S.PRES.IND rather than staying tied to its phonologically determined origin. From there on, the new morphological element spreads by verb-class association to cover the entire available range, i.e. potentially all regular and also some irregular forms of 1S.PRES.IND. The analogical process reaches an intermediate completion with the elevation of [-əl] to /+əl/ as 1S.PRES.IND and progresses in its farthest reach further to indicate the function of 1S.PRES.IND.

Note that the original forms of type *jeu affel* in (6a), which served as point of departure for the development, also adapted to the new 1s ending and began to surface since the 18th century as *jeu affl-əl* = /affl + əl/, etc., thus obviating the original resyllabification and conferring morpheme status to the novel suffix. Since the phonological conditioning was already being supplanted in the 18th century, one might wonder whether the absence of observable extension of /-əl/ in the 17th century texts is a true reflection of the language or rather a consequence of the limited textual variety preserved in this earliest corpus (religious texts of high formality). At any rate,

[4] The extensive tabulations in Decurtins (1958) document that a given local dialect cannot be expected to exhibit a systematic application of the affix across items or even different instances of the same verb. There is no dialect with a completely regularized pattern all in /-əl/ for 1 PRES/IMPF IND, even though the Surselvan standard language tends to extend the presence of the affix to all regular conjugations (see also note 1 above). Completion of the change cannot be expected in the spoken practice of this language.

the secondary formations of type /… Cl-əl#/$_{1S.PRES.IND}$ can only be expected after the creation of a morphological variant /+əl/ at for 1S.PRES.IND.

5. 1S present indicative /-a/

A weaker second neo-formation in /-a/ for the same 1S.PRES.IND function (e.g. *jeu affl-a* 'I find', *jeu rog-a* 'I ask') co-occurred with the initial spread of /-əl/. Both developments have also been claimed to rest on a trend toward isosyllabic formation of the present indicative and imperfect indicative singular paradigms, a trend that is related to the leveling of the conjugation classes in the present subjunctive where the /a/-class extended its endings to the /e/ and /i/ classes (Stimm 1980: 653). In the /a/ class, the effect of this new ending is a formal identity between 1S and 3S: *jeu roga*, *el roga*, parallel to *che jeu/ella venda*. This is already the etymological outcome in the present subjunctive (see again (2)) and in the imperfect subjunctive (before the expansion of /-əl/ also in the imperfect indicative), as in much of Romance morphology: a rather unremarkable situation. Here an analogical extension may have been one motivation (7a). In addition, a more involved solution of reanalysis in the inversion construction with a subject pronoun has been supported by Stimm (1980: 649), e.g., *rog'el* = *roga el*, note: [rɔ´gɛl], not *[´rɔgəl]) ≈ *el roga* '3S asks' compared to *rog jeu* ≈ *jeu rog* '1S ask', where *rog jeu* was reinterpreted as involving elision of a vowel (a normal event before *jeu* as well as *el*). A corresponding analogy for this purpose might be as in (7b).

7. a. *el laváva* : *jeu laváva* :: *el lava* : X X = *jeu láva*
   b. *rog'el* : *el róga* :: *rog jeu* : X X = *jeu róga*
   c. *jeu selável, ti selávas, ella seláva, nus selavéin, vus selavéis, ellas selávan*
      'I wash myself, you wash yourself, etc.'

At any rate, it is interesting to note that, as with /+el/, a morphological variant again affects the 1S.PRES.IND, where the etymological loss of a person marker may create a favorable environment for some kind of reconstructive adjustment. Identity of form across morphosyntactic function is not a problem, since this situation obtains rather frequently in Surselvan. Referential identity is usually guaranteed by the obligatory subject pronoun. One also observes the leveling of secondary personal reference in the reflexive conjugation, where a single reflexive marker /sə+/ stands for all three persons, singular and plural (7c).

The 1S form is subject to non-etymological alteration through processes that appear to be analogical in nature. A functional need for person marking cannot be the motivating force behind the developments of *–a* and *–el*, since the first one leads to homophony (already multiply present in the conjugation patterns) and the second one creates a non-harmonic form with unnecessary expressivity in the morphological domain. 1S may well be clearly marked, but it is not a central point of morphological uniqueness, in contrast to 2S and/or 2PL in Bündnerromanisch or the Romance languages in general. Through the non-null subject status of Surselvan, morphological identification of person carries reduced weight.

6. The expansion of Surselvan /+əl/

To sum up so far, the eventual Surselvan 1s affix derives from a phonetically delimited punctual form that, due to its salience, undergoes reinterpretation ('is perceived') as a freely available hypercharacterization of the 1s function. It is eventually upgraded from a phonological to a morphological condition, i.e. from *áffel* to *áfflel*, *cóntel* and *afflável*. In spite of the limited textual documentation, the chronological phases for the expansion of Surselvan /+əl/ to reach its modern near-regularity permit interesting insights into the mechanics of this expansion. The spread of /əl/ as a 1s ending exhibits the following distinct historical phases.

(8) a. Purely phonological condition as resyllabification of a stem and word final /Cl#/ cluster in PRES.IND (17th c., earliest documentation): *jeu affel* vs. *jeu lav.*

b. Extension within 1S.PRES.IND to some (regular?) verbs beyond phonological conditioning (18th c., with somewhat broader documentation): *jeu lavel, jeu dromel.*

c. Appearance of /-əl/ in 1S.PRES.IND forms of verbs with stem final /Cl+/ cluster: *jeu afflel* (18th c.). The effect of thereby creating an isosyllabic present indicative inflection (1a) may have been a factor in favor of such an extension (Stimm 1980: 653).

d. Broad use of /-əl/ in regular present indicative morphology and extension to 1S.IMPF.IND (for all verbs and classes?) completed by the early 20th c. as documented for various local dialect descriptions (e.g. Jaberg & Jud 1928–40): *jeu lavável.*

e. Regularization of 1S.PRES.IND and 1S.IMPF.IND belongs to the written language of the Surselva in the 20th c., with variable use in auxiliary, modal, irregular and otherwise marked verbal conjugations. No penetration of /-əl/ into other tenses and moods takes place in the standardized (and written) Surselvan practice. Local dialects may differ with regard to the extension of /əl/, especially for verbs with variable exposure in the standard (see note 2 again).

f. Locally in the Sutselva, and historically in sporadic documents, the alternative filiation *jeu rog-a, jeu affl-a* persists in the spoken language. This variant has a historical presence since the earliest documentations (Stimm 1980: 649; see Section 5 for discussion of this variant).

g. The artificially created common formal/written standard Rumantsch grischun does not take over the morpheme /+əl/ as too strictly localized in the Surselva and not anchored in Engadinian or other varieties of Bündnerromanisch.

Like the previous accounts, this discussion has passed over in silence the parallel configuration of verbs with stem-final /Cr/. The loss of word-final *–o* in 1s produced the same problem of resyllabification for verbs like *entrar* 'to enter', *cumprar* 'to buy' and many more: *jeu énter, jeu cúmper* (already mentioned in Meyer-Lübke 1890(1): §315),

compared to modern *jeu éntrel, jeu cúmprel*. The string [-ər] is as disharmonic as a 1s morph as is [-əl]. Yet there is no mention in the literature of an analogical extension of this salient solution in the inflection, neither historically nor currently: **jeu éntrer*, **jeu cónter* parallel to *jeu ánflel, jeu cóntel*.[5]

7. Accounting for the analogical changes

How can the rather surprising development best be characterized and understood? The foundation is obviously the operation of some kind of analogy understood as the antagonistic and complementary force to rule-like changes, an otherwise unexplained phenomenon that 'goes against standard'. But in which way? And for what reasons?

7.1. *Frequency*

Even though the documentation of the spread over four hundred years is sporadic, it is quite clear that the result of the development does not support an interpretation based on rule-like generalization and ultimate simplification of the obviously incomplete and partially variable distribution. In the beginning, the motivation for the emancipation of a morph [-əl] leading to an eventual morpheme /+əl/ does not respond to clear lexical or textual prominence of this string. However, the non-etymological addition of an unstressed, eccentric morpheme /+əl/ for 1s eventually results in a phonetically and lexically very well anchored pattern in Surselvan, regardless of any referential implications. The inverse dictionary of modern Surselvan (Lutz & Strehle 1988: 215–234) gives long lists of /-'VC*əl/ as nouns or adjectives, in addition to the one verbal ending. These items mainly represent the productive suffix –'V:BILE > *-ábel,* , *-éivel,* e.g., *miserábel* 'miserable', *nuschéivel* 'harmul' (with some 600 N, A listed for the modern language). But this is the end point of the development, not its original condition, so that a presumably limited weight of the verb stems in /Cl/ should not have operated as a trigger for modifying the conjugational pattern in a major way. If the existence in the language of a considerable lexical trove of words in /-'VC*əl/ supports the inflectional forms in /+əl/, the same situation should hold for a hypothetical expansion of */ər/ given the equally present pattern /-'VC*ər/ in nouns and adjectives; e.g. *catscháder* 'hunter', *svízzer* 'Swiss' (some 500 items). Here there are also many verbs of the conjugation class with short stem /e/; e.g. *precèder* 'to precede' (some 300 entries; Lutz & Strehle 1988: 386–421). The strong asymmetry in analogical thrust between /+əl/ and */+ər/ might find a possible explanation in the clash of functions created by a syncretism of 1S.PRES.IND and the infinitive. In the older language, at the time when the analogical extension must have taken place, the presence of frequent recessively stressed infinitives (*precèder* vs. *cumprá(r)*) was

[5] Note that modern 'to find' is *anflár*. – Other stem-final, and thus secondarily word-final clusters in 1S.PRES.IND requiring resyllabification could also be considered as potential sources for a demarcating 1s allomorph. But the language is tolerant of final consonant clusters that do not violate major sonority constraints. Stem-final /C+liquid/ clusters stand out as requiring adjustments.

already true, supporting the hypothesis that this condition may have influenced the choice between the available models for analogical extension.

7.2. *Proportionality*

The standard version of proportional analogy, a : b :: c : d, will not by itself yield a convincing model for the Surselvan development. The choice of terms to produce the original extension is almost entirely arbitrary. The most immediate candidate for such a proportion as shown in (9a) is imprecise due to the stem-final shape /fl/ vs. /f/ that is not reflected in the target forms. The proportion is furthermore unconvincing, since almost any morphological function in the paradigm could be put into positions a and c (e.g. *anflas, anflan*, in parallel to *contas, contan*). This underdetermination is a general problem with proportional analogy. On the other hand, a second stage (9b) extending the already established morph /+əl/ (8c) to the original forms (6a) looks more plausible, but does not solve the issue of its origin.

9. a. *affla*$_{3S.PRES.IND}$: *affel*$_{1S.PRES.IND}$ :: *conta*$_{3S\ PRES\ IND}$: x x = *contel*$_{1S.PRES.IND}$
   b. *chonta*$_{3S.PRES.IND}$: *chontel*$_{1S.PRES.IND}$ :: *affla*$_{3S\ PRES\ IND}$: x x = *affl-el*$_{1S.PRES.IND}$

7.3. *Pattern congruity*

Similarly, the analogical mechanism permitting the spread of /-əl/ cannot find its motivation as a characteristically good fit of the morphological marker for the function. Note the prevalent distracting association of /-əl/ with some 3S reference based on ILL- (elevated to an improbable solution in Lausberg 1963–1965(2): §798). The discrepancy in functional fit for the parallel /-ər/ is apparently so strong that there is no evidence of an expansion of such a morpheme .

8. Analogical Modeling & Dynamical Grammar

To gain some additional perspective on the expansion of /el/, I propose to think of a different kind of analogical force. The concept of analogy assumed here[6] has its antecedents in the framework of Analogical Modeling described in the contributions contained in Skousen et al. 2002 (see in particular Skousen 2002) and in the general model of contact assimilation postulated in Dynamical Grammar (Culicover & Nowak 2003).

The essence of analogy is the judgment of similarity between two (or more) forms arrived at in the speaker's mind, be it during acquisition or in regular linguistic practice. Dynamical Grammar proposes a sociological model applied to language acquisition and operation where any contact between speakers qua operators of grammar creates the context where one interlocutor may influence the other by sheer (linguistic) contact. Various dimensions of strength (closeness, prestige, incidence) will modulate the

[6] For more extensive discussion and motivation, see Wanner 2006.

probability of actual influence. At the same time, receptivity for influence by a speaker may vary according to parameters of attention and openness. Any linguistic event could result in an analogical assimilation at the base level of forms coming into contact, i.e. a minimal, local nonce change in a speaker's linguistic behavior and knowledge. At the same time, this contact could also not take place, resulting in default invariance. Over the long haul, the accumulation of singular successful and repeated analogical events may create some degree of regularity construed from the bottom up. Analogical changes will internally spread in the grammar of a speaker at differential speed depending on the level of analysis: individual lexical impact, broader classifications such as 1S.PRES.IND, etc. Externally, such changes progress at an independent rate in the social context. The generalizing effect of such analogy series is always weak and contingent, reaching completion only for closed form groups and leaving potentially divergent fringes around a more coherent core. Completion of one line of analogy (e.g. 1S.PRES) may intersect with a different line of analogical expansion (e.g. reformation of present subjunctive endings, where 1S = 3S) and thus, the first thrust of 1s could be stopped short of a possible classificatory goal (all of 1S). Activation of any event of analogy is random; a chain of analogical events follows an unpredictable path without forced closure.

Analogical Modeling (Skousen 2002) specifies ways in which the linguistic material co-determines the possible paths of analogy. Forms and functions are internally organized by associative binding of forms based on the same kinds of similarities driving analogy. The analogical set consists of forms construed as similar by the speaker and thereby cohering harmoniously at least in the speaker's analysis. The existence, composition, and extension of a set is again unpredictable for the individual, and even more so for the speech community (the language), hence leaving open whether a specific analogical assimilation will take place or not. While the dimensions of similarity in the analogical set exert pressure on the shape of potential analogical events and their extensions, there always exists the option of a random choice among all the options for change guaranteeing the unpredictability of analogical instances. Linguistic practice and development exist between motivated regularities and random developments without clear motivation.

This conceptualization contrasts with the formally derivative (top-down) perspective of globalizing regularity. The rule-governed formal approach and the analogical perspective advocated here do, however, converge in the observed preponderance of regularities in language: sweeping and abstract for the rule-based view, much smaller and often working at counter-purpose for the analogical perspective. Unlike the top-down conception, the analogical stance is at ease with incomplete generalizations, complex variation, and downright exceptions understood as not (yet) consummated analogical options.

9. Conclusion

The unique shape in its associative context, i.e. the presumable referential mismatch and thus saliency of the secondarily segmented marker [-əl] in a very limited lexical range of 1S.PRES.IND may have served as a motivation for the initial random choice

expanding the occurrence of [-əl] (8a,b). Its saliency over the non-expressive etymologically derived zero ending for 1s (-O:# > Ø) in 1s.PRES.IND apparently motivated repeated extensions to further verb stems without the original phonological motivation of /Cl_{STEM}+/. By the 18th century, [-əl] will have mutated to a variant morph /+əl/, now strengthened by this reclassification and thus accelerating its expansion. As a consequence, the forms at the origin will undergo a secondary analogy and regularize the outcome to *jeu anflel* (8d). But /+əl/ will not be able to reach all of 1s.PRES.IND due to the hold-out by irregular and some very common verbs[7], an incomplete generalization comprised by the analogical approach. On the other hand, /+əl/ is able to expand its range by analogy of shape and, more importantly, through a functional assimilation from 1s.PRES.IND to 1s.IMPF.IND, yielding a more or less complete applicability for 1s.IND (there is no synthetic future in the language).

The failure of [-ər] to enter into competition with [-əl] may in part be due to the random factor of choosing one over another option. But the functional clash with the infinitive through syncretic [-ər] can be understood as the effect of two different harmonic sets for [-əl] and [-ər], where only the latter set contained an important inhibiting factor in the form of the interfering infinitive, i.e. of a clash within the same morphosyntactic category.

The unexpected moments of the deployment of /-əl/ are the two random choices, first, of this string over other options and, second, the elimination of [-ər] in favor of [-əl]. Its accidental nature fits well into the analogical approach, but this acceptance of indetermination also imposes the obligation to continue the search, to go beyond what can at present be understood and to look at the surprises of language with open eyes and mind, treating our imperfect analytical instruments with critical spirit.

References

Bybee, Joan L., & Hopper, Paul. 2001. Introduction to frequency and the emergence of linguistic structure. In Bybee and Hopper (eds.) *Frequency and the emergence of linguistic structure*, 1–24. Amsterdam: Benjamins.

Culicover, Peter W., & Nowak, Andrzej. 2003. *Dynamical Grammar: Minimalism, acquisition, and change*. Oxford: Oxford University Press.

Decurtins, Alexi. 1958. *Zur Morphologie der unregelmässigen Verben im Bündnerromanischen: Historisch-deskriptive Studie mit besonderer Berücksichtigung der Sur- und Sutselvischen Mundarten* (Romanica Helvetica 62). Bern: Francke.

Grünert, Matthias. 2003. *Modussyntax im Surselvischen: Ein Beitrag zur Erforschung der Morphosyntax des Verbs im Bünderromanischen* (Romanica Helvetica 122). Tübingen: Francke.

Hacks, Franziska Maria & Kaiser, Georg A. 2013. Zur Syntax der Fragesätze im Rätoromanischen. In Darms, Georg & Riatsch, Clà & Solèr, Clau (eds.) *Akten des V. rätoromanistischen Kolloquiums*, 137–162. Tübingen: Narr-Francke-Attempto.

[7] The reticence of special and frequent forms is a common theme of conditions on analogy (see e.g. Mańczak (1958) and Bybee & Hopper (2001) for discussion of how such forms may push the change as models or retard a change due to their inertial weight).

Jaberg, Karl & Jud Jakob. 1928–1940. *Sprach- und Sachatlas Italiens und der Südschweiz* (8 vols.). Zofingen: Ringier.

Lausberg, Heinrich. 1963–1965. *Lingüística románica* (2 vols.). Madrid: Gredos.

Liver, Ricarda. 2010. *Rätoromanisch*. 2nd edn. Tübingen: Narr.

Lutz, Florentin & Strehle, Dieter. 1988. *Rückläufiges Wörterbuch des Surselvischen. Dicziunari invers dil romontsch sursilvan* (Romanica Monacensia 29). Tübingen: Narr.

Mańczak, Witold. 1958. Tendances générales des changements analogiques. *Lingua* 7. 298–325.

Meyer-Lübke, Wilhelm. 1890–1902. *Grammatik der romanischen Sprachen* (4 vols.). Leipzig: Reisland.

Prader-Schucany, Olga. 1970. *Romanisch Bünden als selbständige Sprachlandschaft* (Romanica Helvetica 60). Bern: Francke

Schmid, Heinrich. 1982. *Rumantsch grischun. Eine Schriftsprache für ganz Romanischbünden: Voraussetzungen, Probleme, Erfahrungen*. Chur: Lia Rumantscha.

Skousen, Royal. 2002. An overview of analogical modeling. In Skousen, Royal & Lonsdale, Deryle & Parkinson, Dilworth B. (eds.) *Analogical modeling: An exemplar-based approach to language,* 11–26. Amsterdam: Benjamins.

Skousen, Royal & Lonsdale, Deryle & Parkinson, Dilworth B. (eds.) 2002. *Analogical modeling: An exemplar-based approach to language*. Amsterdam: Benjamins.

Spescha, Arnold. 1989. *Grammatica sursilvana*. Chur: Casa editura per mieds d'instrucziun.

Stimm, Helmut. 1980. Zur Entstehung der Verbalendung *-el* (1. Pers. Sing.) im Surselvischen. In Schmidt, Gerhard & Tietz, Manfred (eds.) *Stimmen der Romania: Festschrift für W. Theodor Elwert zum 70. Geburtstag*, 633–658. Wiesbaden: Heymann.

Stimm, Helmut & Linder, Peter. 1989. Bündnerromanisch. In Holtus, Günter & Metzeltin, Michael & Schmitt, Christian (eds.) *Lexikon der romanistischen Linguistik*, vol. 3, 764–785.

Wanner, Dieter. 2006. *The power of analogy*: *An essay on historical linguistics* (Trends in Linguistics. Studies and Monographs 170). Berlin: Mouton de Gruyter.

Dieter Wanner
The Ohio State University
wanner.2@osu.edu

Wellington Working Papers in Linguistics (WWPL)

Policy guidelines

1. The WWPL are a production of the School of Linguistics and Applied Language Studies at Victoria University of Wellington.

2. The primary objective of the WWPL is to provide an initial avenue of publication for high quality material in the area of linguistics produced by students and staff of the department.

3. Contributions may take the form of papers, reviews, squibs, critical commentary etc.

4. Issues of the WWPL will be produced as opportunity to publish appropriate material arises, with the aim of producing at least one issue per year.

5. Any member of staff (including Research Fellows) may act as editor of an issue of WWPL. Guest editors may be appointed on occasion.

6. The content of the WWPL will usually include material produced by post-graduate students and staff, but, in special circumstances, may include contributions from people outside the School.

7. Copyright remains with the author(s) of contributions. Authors are welcome to publish contributions to WWPL in other journals. It would be appreciated if a re-published article contained an acknowledgment that the article had been first published in WWPL.

8. The School should be informed of issues planned by individual staff members.

Contents of volumes 1–22

Volume 23 of the WWPL is dedicated to Liz Pearce on her retirement from many years of service in the School of Linguistics and Applied Language Studies at Victoria University of Wellington. The Festschrift brings together papers from the wide range of subject areas spanned by Liz's research interests: historical linguistics, Romance linguistics, syntactic theory, Austronesian linguistics, and language documentation. The diversity of the contributions is a fitting tribute to the many ways in which Liz has enriched our lives and the field of linguistics, both as a scholar and a friend.

ISSN 1170-1978 (Print)
ISSN 2230-4681 (Online)

TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI
VICTORIA
UNIVERSITY OF WELLINGTON