Differential inalienable marking in linguistic Wallacea

Laura Arnold

The linguistic area of Wallacea encompasses Nusa Tenggara, Maluku, the Bird’s Head and Neck of New Guinea, and Cenderawasih Bay (Schapper 2015). In this paper, I will present results from the most comprehensive typological survey of possessive constructions in Wallacea and its surrounds to date. Drawing on data from 106 Austronesian and 67 Papuan languages spoken from Sulawesi and Sumbawa in the west to central New Guinea in the east, I will discuss the distribution and development of differential inalienable marking in the region – a feature which has thus far not been discussed in any typological, areal, or theoretical literature.

Differential inalienable marking (DIM) is found in some languages with a morphosyntactic alienability distinction in adnominal possessive constructions. Languages with DIM make a further morphosyntactic distinction in ‘inalienable’ constructions (i.e., constructions more closely associated with expressing inalienable relationships between the possessor and possessee, such as body parts and kin terms) – typically, with two distinct paradigms marking the possessor.

DIM may be semantically conditioned: for example, all kin terms may be marked with one paradigm, all body parts another. For example, in Ambai, a 3sg possessor is predictably marked on kin terms with the suffix -na, and on body parts with -n, as shown in (1). Alternatively, the distinction may be lexically specified. For example, in Kula, the possessor is marked on most body parts and kin terms with one paradigm; however, there is a subset of body parts which are unpredictably marked with a different paradigm. This is exemplified in (2): a 1excl possessor is marked on the body part nikwa ‘eye’ with the prefix ng-, but on kárik ‘finger’ with nge-.

(1) Ambai (Austronesian):  (2) Kula (Timor-Alor-Pantar):
   a. ina-na
      mother-3sg
      ‘his/her mother’
   b. awe-n
      foot-3sg
      ‘his/her foot’
   a. ng-nikwa
      1excl-eye
      ‘my/our eye’
   b. nge-kárik
      1excl-finger
      ‘my/our finger’

DIM appears to be extremely rare cross-linguistically – in a recent request for data on the lingtyp mailing list, only 14 potential cases outside of the surveyed area were reported. DIM, however, is relatively common in the surveyed area: in the 173 languages looked at in this study, 20 are attested with DIM (10 A, 10 P). All 20 of these languages are spoken within Wallacea. This suggests that DIM is an additional, previously unidentified feature of this linguistic area.

Within Wallacea, DIM further clusters in ways that produce some interesting historical insights. For example, in northern Wallacea, semantically-conditioned DIM is found in several languages around the Bird’s Head peninsula: six languages belonging to the South Halmahera-West New Guinea (SHWNG) branch of Austronesian; and the Papuan languages Moskona and Meyah (East Bird’s Head; EBH) and Hatam (Hatam-Mansim). Owing to a non-cognacy of forms, DIM cannot be reconstructed to proto-SHWNG, suggesting that DIM independently developed several times in SHWNG, as the result of contact with the Papuan languages of the area. Data from the EBH languages additionally indicate that historically, only kin terms were inalienably possessed; and that DIM in EBH may have developed through contact with the South Bird’s Head languages, which only possess body parts in inalienable constructions. Finally, lexically-specified DIM clusters genetically in the Timor-Alor-Pantar (TAP) family, in southern Wallacea. Data from these languages suggest that DIM can be reconstructed to proto-TAP – and that at an earlier stage, DIM was semantically conditioned, as it is in northern Wallacea. The historical implications of these distributions will be explored, drawing on additional evidence from linguistic and archaeological correlates and oral histories to further our understanding of the history of this little-known region.

References

Reciprocal constructions in Northern Amis.

Isabelle Bril (CNRS-LACITO)

In Austronesian languages, reciprocal relations are generally marked by polysemous prefixes that are cognate with Proto-Austronesian *maR-/*paR-. The prefixes marking reciprocal relations also denote collective and plural relations (Pawley 1973, Lichtenberg 2000, Zeitoun 2002); they also have various other meanings (Middle, comparative, intensive, iterative), which further diversify in many Oceanic languages, taking on for instance dispersive, distributive, reflexive meanings (Bril 2005, Moyse-Faurie 2008).

In Northern Amis, mal(a)- (cognate with *maR-) marks reciprocal relations, dyadic reciprocal kinship and comparison, while the prefix ma- together with obligatory Ca- reduplication (ma-Ca-) tends to encode more weakly symmetrical reciprocal relations, in which plural participants are engaged in collective actions, done in turn or repeatedly (ex.2), as well as chaining and mode of grouping (the planks are piled on top of each other). Both markers appear in (1).

1) **Ma-ca-curuk k-uhni a ma-paliw.**
   - MA-Ca-take.turn NOM-3PL COMP REC-collaborate
   - ‘They took turns to help one another.’

   Only mal(a)- can be affixed to stems denoting kinship or social relations, these can be symmetrical (‘they’re REC-siblings’) or asymmetrical dyadic kinship ‘they’re REC-mother/daughter’ (Evans 2005). Languages vary as to which term (the higher or the lower) of the dyad is chosen, in Amis, it is always the higher term.

   Extended, plural reciprocal relations are marked by stem reduplication, as mal-CVCV or ma-(Ca)-CVCV-. While mal-CVCV forms denote plural relations conceptualized as one whole, ma-Ca-CVCV- reduplication denote pluractional reciprocal actions done in turn (ex.2); they then give rise to intensive meanings.

2) **Ma-ka-kiti.kiting k-uhni a ma-keru.**
   - MA-Ca-(CVCV)link NOM-3PL COMP MA-dance
   - ‘They dance holding each other’s hands.’ (as a group)

   The semantics of the constructions will be shown to result from compositional effects between the prefixes and (i) the type of part of speech involved (i.e. nouns, verbs), (ii) the verbs’ semantic feature (active, motion, stative, property-denoting), (iii) as well as lexical semantics. Relations may be (i) symmetrical in strictly reciprocal relations (they (2) hold each other’s hands); (ii) weakly symmetrical in extended plural relations, collective actions and chaining (they walk behind each other); (iii) asymmetrical in mode of grouping involving plural entities with some inherent orientation (the logs are on top of each other).

   It will also be shown that reciprocal constructions generally display low transitivity, thus favouring the evolution of these prefixes into middle markers, though rarely into reflexive markers. In Amis, reciprocal and reflexive constructions are different.


Variation in Ende Word Order
Megan Brown
Boston University

This paper outlines a set of preliminary findings related to word order variation in Ende, a Pahoturi River language spoken by at least 600 people in southern New Guinea (Eberhard et al., 2019). In doing so, this paper serves as the first analysis of word order in a language of southern New Guinea, a language family which previous research has suggested would have distinct syntactic features from other Papuan languages given that they do not have verb-chaining or switch reference, two features that are very common to the Papuan languages of other regions (Evans et al., 2018). This raises the question of exactly what the underlying structure is of a southern Papuan language such as Ende, which this paper addresses.

Across the language, Ende demonstrates use of both OV and VO word orders. It has been argued that in cases where a language makes use of two or more word order options, one order functions as the unmarked or default ordering of the language (Dryer, 1995). In many cases, this default is the most common word ordering found in the language.

This analysis considers four texts from the Ende corpus (Lindsey, 2015) in order to determine the frequency of these word order patterns. In addition to considering OV vs. VO ordering, frequencies of (a) pre- vs. post-nominal adjectives, (b) prepositions vs. postpositions, (c) main verb-aux verb vs. aux verb-main verb orderings, and (d) predicate-copula vs. copula-predicate orderings were also considered. The analysis found that 74% of all transitve sentences with overt objects used OV word order, as exemplified in (1) (n=60), 81% of adjectives were post-nominal as shown in (2) (n=16), and 90% of all adpositions were postpositional (n=10) (2). Additionally it was found that auxiliary verbs always followed the main verb (3) (n=75), and and copula constructions are always ordered predicate-copula (4) (n=25).

While the frequency of OV vs. VO orderings in this data is not in and of itself a reliable indicator of the default order (Dryer, 1995), Dryer (1991, 1992) has outlined a set of correlations between OV vs. VO word order and other word order phenomenon across languages. More specifically, OV languages most often have, among other features: (a) postnominal adjectives (b) postpositions (c) main verb-aux verb orderings (d) predicate-copula orderings, all of which were found in this analysis. Taken together, the frequency of OV orderings and the predominance of other word-order phenomena correlating with OV word order support this author’s claim that SOV is not simply the most common word order in Ende, as suggested by Lindsey (2019) but that SOV also functions as the default Ende word order.

That being said, this leaves the remaining 26% of the transitive sentences from the data set mentioned above requiring explanation. These less common occurrences include orders such as OVS (5) and OVS (6). One possible explanation is that the use of object-initial word orders in Ende can be brought about by topicalizing the object as in the OSV and OVS orders shown below.

Additionally, since Ende’s default SOV word order would not allow for this same type of sentence-initial movement for a clear topicalization of the subject, a similar emphasis can be placed on the subject using SVO.

I argue that despite the surface appearance of free variation with respect to Ende word order, Ende has an overall set syntax and variations found in the data follow systematic rules. These rules involve a default SOV word order where factors such as topicalization can allow for alternative orderings. Finally, I will present other word order-related findings in the language, such as the presence of both pre-nominal and post-nominal adjectives, in order to emphasize the wealth of other potential findings related to word order in this language.

The possessive classifier systems of Oceanic languages can provide a unique insight into the origin and nature of gender. Typically, a noun can occur with different classifiers, depending on how the possessed item is used by the possessor (Lichtenberk, 1983). For example, *wi ‘water’ in Lewo (Vanuatu) occurs with either the drinkable or the general classifier.

1a. *ma-na wi b. sa-na wi
   CL.DR-3SG water CL.GEN-3SG water
   ‘her (drinking) water’ ‘her (washing) water’

(Early, 1994:216)

In marked contrast, North Ambrym’s (Vanuatu) cognate for water – *we – occurs only with the drinkable classifier (2a), not the general classifier (2b):

2a. *ma-n we b. *mwena-n we
   CL.DR-3SG water CL.GEN-3SG water
   ‘her water (for any purpose)’ intended: ‘her water’

(Franjieh, 2016:95)

We argue that North Ambrym’s innovative system resembles a gender system: a noun occurs with a particular classifier regardless of contextual interactions. We ask whether gender systems can indeed emerge from possessive classifiers in this way. If so, we must then uncover how and why languages would relinquish a useful, meaningful classificatory system, and adopt a rigid, apparently unmotivated gender system.

We have designed a suite of novel experiments to compare possessive classifier systems. We chose these six Oceanic languages: Merci, Lewo, Vatlongos, North Ambrym (Vanuatu), Nélémwa and Iaai (New Caledonia), for two main reasons: (i) each has a different inventory size of classifiers, from a two-way distinction to a more complex inventory of twenty-three; and (ii) these classifier systems represent varying degrees of informativeness: some have transparent semantic motivation, whereas others have opaque assignment. Now effective categorisation needs to be simple, to minimise cognitive load, and informative, to maximise communicative efficiency (Hawkins, 2004). Our sample languages allow us to investigate the trade-off between these two principles of simplicity and informativeness.

We have conducted three experiments so far, involving 122 speakers, on our sample languages: (i) free listing, (ii) card sorting, and (iii) video vignettes. Free listing establishes central members of a classifier’s semantic domains, which vary from language to language. Card sorting reveals how speakers categorise relevant nouns and whether conceptual groupings map onto classifiers. Finally the video vignettes depict typical and atypical interactions with different items; they determine whether speakers are free to use different classifiers (as is generally believed) or whether there is a rigid assignment of nouns to classifiers.

We present data, suggesting that the classifier systems tested represent different stages of grammaticalisation from classifier to gender marker. They reveal intriguing inter-speaker variation in classifier choice, which helps establish the relative optimality of each system.

This combination of typology with psycholinguistics promises to shed new light on the development and functioning of systems of nominal classification. We are keen to have feedback before conducting the second round of psycholinguistic experiments in the field.

References:

New subgrouping evidence for the Bima-Lembata languages
(Austronesian, east Indonesia)

Hanna Fricke
Leiden University

This paper contributes to the genealogical history of the Austronesian Central-Eastern-Malayo-Polynesian languages (CEMP) by providing new evidence for the Bima-Lembata subgroup within CEMP. The analysis is carried out using the comparative method with lexical data from the LexiRumah database online (Kaiping, Edwards & Klamer 2019).

The Bima-Lembata subgroup includes all Austronesian languages on the eastern Indonesian islands from Bima in the west to Lembata in the east. The lenition of initial Proto-Malayo-Polynesian (PMP) *b- > PBL *w- in a specific set of lexical items (see table) provides new evidence for this subgroup. The table shows reflexes in a representative set of the Bima-Lembata languages, i.e. Bima on Sumbawa, Kambera on Sumba, Proto-Central Flores (PCF) (Elias 2018) and Proto-Flores-Lembata (PFL) (Fricke 2019). These lexical items do not show this lenition in other Austronesian languages of the region, exemplified by the Proto-Rote-Meto (PRM) reconstructions (Edwards in prep) which have reflexes in a large group of languages in western Timor and on the island of Rote.

<table>
<thead>
<tr>
<th>PMP</th>
<th>PBL</th>
<th>Bima</th>
<th>Kambera</th>
<th>PCF</th>
<th>PFL</th>
<th>PRM</th>
</tr>
</thead>
<tbody>
<tr>
<td>*babuy ‘pig’</td>
<td>*wawi</td>
<td>*wai</td>
<td>*vai</td>
<td>*wai</td>
<td>*wai</td>
<td>*bafi</td>
</tr>
<tr>
<td>*batu ‘stone’</td>
<td>*watu</td>
<td>*vatu</td>
<td>-</td>
<td>*vatu</td>
<td>*vatu</td>
<td>*batu</td>
</tr>
<tr>
<td>*buaq ‘fruit’</td>
<td>*wua</td>
<td>*vua</td>
<td>*vua</td>
<td>-</td>
<td>*vua</td>
<td>*bua-k</td>
</tr>
<tr>
<td>*bulan ‘moon’</td>
<td>*wulan</td>
<td>*vulaŋ</td>
<td>*vula</td>
<td>*vulan</td>
<td>-</td>
<td>*bulan</td>
</tr>
<tr>
<td>*bahi ‘woman’</td>
<td>*wai</td>
<td>-</td>
<td>-</td>
<td>*vai</td>
<td>*vai</td>
<td>*fee</td>
</tr>
<tr>
<td>*bujaq ‘foam’</td>
<td>*wud[a?]a</td>
<td>-</td>
<td>wura</td>
<td>*voda</td>
<td>*vuda</td>
<td>*fu[dʒə]</td>
</tr>
</tbody>
</table>

A subgroup encompassing the Bima-Lembata languages has been suggested by Blust (2008: 48) who proposed that the Sumba-Hawu group may include languages of Western and Central Flores, but only can include Bima if also languages further east are included. The present study suggests that these languages further east must be the Flores-Lembata languages. Gasser (2014) already proposed phylogenetic evidence for exactly this hypothesis. The present paper provides additional evidence of a different nature, and thus confirms the proposal of Bima-Lembata with the shared sound change of PMP *b > PBL *w in a specific set of lexical items as clear evidence for this subgroup.

References
Child-directed speech in Qaet: A multimethod-approach

Henrike Frye, University of Cologne

Qaet is a non-Austronesian language, spoken by about 15,000 people in East New Britain, Papua New Guinea. In the remote inland, children acquire Qaet as their first language.

Much of what we know about child-directed speech (CDS) stems from W.E.I.R.D. contexts (Western, Educated, Industrialized, Rich, Democratic; Henrich, Heine and Norenzayan 2010). Anthropological research (e.g. Ochs 1988; Schieffelin 1990) indicates that CDS is a register rather found in Western middle class societies. This talk combines evidence from different methods showing that the features typical for speech to children in such contexts are also found in Qaet CDS.

Preliminary insights from naturalistic audio-recordings suggest that Qaet children are infrequently addressed directly. In interviews, Qaet caregivers express the view that children “pick up” the language on their own. Still, they have clear ideas about how to talk to children in a way that makes it easier for them to understand what is said.

In order to compare Adult- and Child-Directed Speech in Qaet, 20 narrations of the pear film (Chafe 1980) have been analyzed, half of them told to adults and half to children. The data show that talk directed to children differs from talk directed to adults for several features, among them utterance type, mean length of utterance, amount of hesitations and intonation. Despite this clear tendency, there is a huge amount of variation found between the individual speakers.

Language documentation, acquisition and socialization: The sketch acquisition manual

Birgit Hellwig, University of Cologne

Over the past few years, there has been a growing interest in documenting child language and child-directed language – an interest shared by both language documentation and by language acquisition and socialization research. From the perspective of language documentation, the language used by and with children constitutes “observable linguistic behavior” (Himmelmann 1998: 166), i.e., its documentation falls within the scope of the language documentation paradigm. From the perspective of language acquisition and socialization, there is an urgent need to expand the empirical database: our theories of linguistic and cognitive development are grounded in data that is strongly biased towards the major European languages, with acquisition data being available for only 1-2% of the world’s languages (Lieven & Stoll 2010: 144). Yet, despite this common interest, studies on language acquisition and socialization in underdocumented languages are still few and far between, and the Austronesian and Papuan language are no exception here. It is likely that the small number of studies is due to numerous methodological and ethical challenges (see, e.g., Kelly et al. 2015). Following up on an original idea by Slobin et al. (1967), this talk reports on an on-going joint project (Defina et al., in prep.) that combines the construction of manageable corpora of natural interaction with children (of 5 hours of analyzed data) with a sketch description of the corpus data, exemplifying the possibilities and challenges by means of a case study of Qaqt (a Baining language of Papua New Guinea).


Title: Voice Choice in Äiwoo

Sindre Holmen

University of Oslo, Department of Linguistics and Scandinavian Studies

Äiwoo is an Oceanic language with a symmetrical voice system (Næss 2015). This is a highly unusual case as most Oceanic languages have accusative alignment, while symmetrical voice is typically found in Western-Austronesian languages. Most transitive verbs in Äiwoo show an actor voice and an undergoer voice form. In texts, however, the undergoer voice is vastly more frequent than the actor voice, by a factor of roughly 1 to 10. It thus seems as though the undergoer voice is the default transitive construction in Äiwoo. In this talk, I will investigate which circumstances condition the use of the actor voice form.

Previous studies of voice and transitivity in Austronesian languages (e.g. Walters 1994, Wouk 1999, Quick 2005, Huang & Tanangkingsing 2011) have looked at topicality as defined by Givón (1983), in which referents are tracked over larger units of discourse by means of two measurements of continuity: referential distance and topic persistence. While the former measures how far back the previous mention of a given referent is found, the latter investigates how many mentions the referent has in subsequent discourse. Næss (2015) puts forth a hypothesis that the referential distance of the voice selected argument (VSA) is a governing factor for the choice of actor voice over undergoer voice in Äiwoo. However, preliminary results from my statistical corpus study seem to refute this hypothesis.

In this talk, I will investigate several factors that have been looked at in previous studies (e.g. McDonnell 2016) and attempt to determine what governs voice selection in Äiwoo. These factors all concern information structure and the role of prominence in grammar. Among them are transitivity as defined by Hopper and Thompson (1980), and the different activation states (given, accessible, new) of Chafe (1987). Preliminary results suggest that activation status of the undergoer argument in particular may play a role, but as one of multiple factors influencing the choice of voice.


Revisiting Givon’s generalization – On the prosodic integration of SVCs in Daakaka

Jens Hopperdietzel (Leibniz-ZAS Berlin) & Nicola Klingler (ÖAW Wien)

Introduction: In a serial verb construction (SVC), two verbs function together as a single (but complex) predicate in a monoclausal environment without any form of coordination or subordination (1/2). Since Givon (1991), most descriptions make reference to the generalization that are pronounced under a single intonation contour (cf. Aikhenvald 2018). In this paper, we revisit Givon’s generalization by conducting a pilot study on the prosody of two types of SVCs – single- and multiple-marking – in Daakaka (Central Vanuatu, Oceanic) by using phonetic standards. Investigating the distribution of phrase-final boundary tones, our results suggest that while single-marking SVCs are uttered within the same intonational phrase (IP), multiple-marking SVCs are prosodically less integrated. Thus, our study not only supports Givon’s generalization for single-marking SVCs, but also suggests that multiple-marking SVCs reflect a different type of syntactic composition (cf. Cleary-Kemp 2015).

Background: Daakaka exhibits two syntactic types of SVCs: (a) single marking with a single realization of Mood (REAL) before the initial verb (1); (b) multiple marking with each verb independently marked for Mood (2). Such a split is described in other Oceanic languages as well (Crowley 2003).

(1) *Ma ta wu-wuo* (nge) **DAAKAKA**
   REAL cut RED-open 3SG.PRON
   ‘He cut it open and ate it.’

(2) *Ya=ma te lee ma mwelli*
   3PL=REAL cut tree REAL be small
   ‘They cut the tree small.’

To add a prosodic perspective, we focus on intonation patterns in both types of SVCs. Daakaka exhibits two types of boundary tones marking the right boundary of an IP that roughly corresponds to the syntactic unit of a clause (von Prince 2015). Following Givon’s generalization, one expect both verbs to be uttered within a single IP – i.e. the presence of single phrase-final boundary tone (RBT). However, as syntactic and semantic studies have argued for a looser integration of multiple marking SVCs (e.g. Foley & Olsson 1985), we predict an additional boundary tone (PBT) before the second Mood marker; i.e. every Mood marker denominates its own IP.

Methods: For this pilot study, we examined data from 12 native speakers of Daakaka (L2: Bislama), balanced for both gender and age (7m/5f; age: >20). The semi-spontaneous data (storyboard elicitation; Burton & Matthewson 2015) were analyzed using PRAAT (Boersma & Weenik 2014). In the pre-analysis, we adjusted the pitch-range (in Hz) as well as the voicing thresholds and segmented the data. As our aim was to test the integration of verbal structures in intonation units, we examined the pitch contours of 12 utterances and mapped the presumed syntactic clause boundaries with pitch movements.

Results: In the control condition (coordination), we replicated von Prince’s observation that boundary tones occur at IP boundaries that roughly correspond to the syntactic size of clauses, e.g. before <te> ‘and, then’ in the examples below. For single marking SVCs, the sole mood marker bears the only pitch rise (RBT) with no boundary tone between the verbs (Fig. 1; above). For multiple marking SVCs, mood markers tend to align with pitch rises: In three of six cases, we found additional PBTs before the second tense marker (Fig. 1; below). In two cases, we did not find PBTs (one case remains uncertain).

Discussion: Based on the results, we argue that our data on single marking SVCs support Givon’s generalization: Both verbs are part of a single IP with a sole boundary tone (RBT) indicating the monoclausality of the whole construction. In contrast, every verb marked for Mood seems to form its own IP in multiple marking SVCs. This pattern suggests a multiclausal structure for multiple marking SVCs.
SVCs (≠ multiple exponentence; *pace* Rolle 2018 on Degema). Thus, our phonetic findings mirror syntactic considerations that verbs in single and multiple marking SVCs may not only differ in their syntactic size, but also in their compositional type (Cleary-Kemp 2015, also Foley & Olsson 1985).
**On the composition of manner and result: Causative secondary predicates in Oceanic**

Jens Hopperdietzel (Leibniz-ZAS Berlin)

**Introduction:** In resultatives (e.g. *Peter hammered the metal flat*), an initial predicate (*hammer*) denotes the manner of an action that causes a change-of-state of an object with the resultant state expressed by a secondary predicate (SP; *flat*). Thus, two (independent) predicates together form a complex predicate (*‘hammer-flat’*) of a single clause (cf. Levin 2019, Williams 2015, Beavers 2012, Halliday 1967). In Oceanic languages, this type of complex predication is usually expressed within a serial verb construction (RSVC) (Verkerk & Frostad 2013; see also Gast, König & Moyse-Faurie 2014, Næss 2012, Bradshaw 1982 for more detailed discussions). Crucially, many Oceanic languages show a cross-linguistically rare type of resultatives in which a transitive/causative verb functions as the SP (1-3). This pattern contrasts intransitive/stative SPs in more prominent languages such as English or Mandarin (e.g. Hu 2018, Kratzer 2005). As transitive/causative SPs do not only diverge in their transitivity but also in their basic event type from intransitive/stative SPs (accomplishments vs. states; cf. Dowty 1979), this talk investigates the syntactic and semantic composition of the two predicates in RSVCs with causative SPs in Oceanic. As a result, the data suggests that the different types of complex resultative predication differ essentially in their headedness and argumenthood.

**Two types of causative SPs:** Conducting a cross-linguistic survey of Oceanic RSVCs, I classify two basic types of causatives SPs: (i) bare (lexical) causatives like Daakaka (1) and (ii) morphological causatives derived by a reflex of the Proto-Oceanic causative prefix *pa/-paka- like in Samoan (2) (Mosel 2004). While in some languages such as Daakaka and Samoan, only one of the two types is available, both types may co-occur within a single language (3a/c) (e.g. Saliba; Margetts 2005). However, this distribution is barely understood yet (but see Verkerk & Frostad 2013, Bradshaw 1982).

(1) *Bong mwe tas tiwiye etastas.* DAAKAKA
   'Bong break the bench by sitting.'

(2) *Sā lamu fa’a-malū e Malia le mea ai.* SAMOAN
   ‘Mary softened the food by chewing it.’

(3) a. *Ye-koi-kesi-o.*
   3SG-hit-break-3SG.OBJ
   ‘He broke it.’

b. *Ye-koi-mwalo-i-o.*
   3SG-hit-dead-3SG.OBJ
   ‘He hit it dead.’

c. *Ye-koi-he-mwalo-i-o.*
   3SG-hit-CAUS-dead-3SG.OBJ
   ‘He hit it dead.’ (Margetts 2005)

**The distribution of causative SPs:** Based on novel data from Daakaka and Samoan, I demonstrate that the derivational type of causative SPs is determined by the lexical semantics of the respective verbs: While stative property concept verbs (e.g. *malū ‘soft’ or *mwalo ‘be.dead’) must be causativized in RSVCs, inherently causative verbs (e.g. *tiwiye ‘break.TR’, *kesi ‘break.TR’) appear in their bare form. Note that bare stative SPs are grammatical in all three languages (3b). Moreover, the complementary distribution of bare and causativized SPs in Daakaka and Samoan arise from language specific morphosyntactic and semantic properties: (i) Daakaka lost its Proto-Oceanic causative prefix and lacks morphological causatives altogether; (ii) Samoan seems to lack lexical causatives based on several diagnostics on manner/result entailment of verbs (e.g. combinatorial restriction with instruments/resultatives; cf. Beavers & Koontz-Garboden 2012 et seq., Rappaport Hovav & Levin 1998 et seq.). Instead, causatives must be derived by the causative prefix *fa’a-*. In contrast, Saliba has both lexical and morphological causatives and as predicted, exhibits both types of causative SPs.

**The compositional type of RSVCs with causative SPs:** Assuming a bi-eventive structure of causatives (event + state), I propose that causative SPs are actually the head of RSVCs with the initial predicate as a manner modifying adjunct, i.e. the initial predicate modifies the yet unspecified action within the complex causative SP (cf. Dowty 1979). This contrasts, resultative constructions with stative SPs in which the initial predicate has been shown to be the head of the construction that takes the stative SP as a syntactic argument (Christie 2013, Larson 1991, Simpson 1983, etc.). Additional evidence for this claim comes from the availability of a narrow reading of ‘again’ that solely scopes over the initial predicate; a reading that is notably absent in resultative construction with stative SPs (cf. Lechner et al. 2015, Beck & Snyder 2001). Thus, Oceanic RSVCs with causative SP instantiate a
distinct type of resultatives diverging both in their semantic and syntactic composition from resultatives with stative SPs. Instead, they rather resemble structures like causative by-phrases in English in which a manner-denoting PP is adjoined to a causative verb (e.g. *I killed him by hitting him*).
Transitivity and Valency-changing devices in Motu

Huade Huang
Australian National University

Stephanie Yam
Australian National University

Motu (ISO 639-3: MEU) is a Western Oceanic language of the Papuan Tip linkage. Similar to other Oceanic languages, Motu has several morphological devices that change the valency of verbs but the full breadth of their functions and distribution has not been closely examined in previous research (Taylor 1970). This paper, based on contemporary data collection, will investigate the expression of transitivity in Motu, exploring not just the morphological marking of transitivity on verbs but also the varying correlations between root valence and clause-level transitivity (Margetts 2008).

The morphological marking of transitivity in Motu presents a system similar to that found in other Oceanic languages, including the set of valency-changing devices summarised in Table 1, as well as the indexing of the object argument on verbs.

Table 1: Valency-changing devices and derivations in Motu

<table>
<thead>
<tr>
<th>Valency-changing device</th>
<th>Valency-changing derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduplication</td>
<td>A → S</td>
</tr>
<tr>
<td>prefix he-</td>
<td>A/O → S</td>
</tr>
<tr>
<td>suffixes –lai and -heni</td>
<td>S → A</td>
</tr>
<tr>
<td>prefix ha-</td>
<td>S → O</td>
</tr>
</tbody>
</table>

The prefix he- can be used to demote either the A or O argument, leaving the other as the S argument of the derived intransitive verb, while reduplication can only demote the O argument. The difference between the two valency-increasing derivations lies in whether the introduced argument is an A argument (causative, prefix ha-) or an additional non-A argument (applicative, -lai and -heni). The differences between the valency-changing derivations allow verbs in Motu to be classified as Actor-subject or Undergoer-subject, aligning with the semantic role of the S argument (cf. Dixon 1988).

In some contexts, we have found a discord between verb-level and clausal level transitivity, as seen in (1), where the causative prefix is attached to the monovalent verb makohi.

(1) niu  e  ha-makohi=mu
coco   nunt    CAUS-break=PROG
‘(the man) is breaking the coconut’ (SY1-VIDNR001-20190922)

In this example, the derived verb does not carry an object marker, as expected of transitive verbs in Motu, and indeed many Oceanic languages, where transitive verbs are typically defined as verbs that carry transitive suffixes and/or indexing of the object argument (Lynch, Crowley & Ross 2002: 35, 44-45). Rather ha-makohi is intransitive, and (1) is an example of discord between verb-level and clause-level transitivity as the derived verb lacks a morphological object marker but co-occurs with niu, which appears to be an object argument.

Discord constructions, such as (1), provide support for a tripartite representation of transitivity features that allow for a more nuanced understanding of the role of transitivity in Motu grammar. This paper will present a more detailed description of transitivity in Motu and provide a further empirical study on which to base typological and historical accounts of transitivity in Oceanic languages.

‘Iamitives’ are discourse markers: Evidence from Bunun
Siwei Chen and Haowen Jiang
Academia Sinica and Zhejiang University

Iamitives refer to markers commonly found in languages of Southeast Asia whose function overlaps with both the English present perfect and ‘already’ (Olsson 2013). This paper investigates a prototypical example of iamitives, the enclitic =in in Isbukun Bunun, an Austronesian language spoken in Taiwan, and argues that, in the spirit of unifying inchoativity and speaker expectation as proposed by Soh (2009), =in is best analyzed as a discourse marker that signals the truth values of p have been updated (i.e., changed) at the speech time. Our findings suggest that “iamitives” are not a mixed or ambivalent category but a distinct type of discourse-structuring elements.

A salient feature of =in is that it consistently gives rise to a transition in out-of-the-blue contexts, with the reading varying in predicate types (completion with achievements and inception with activities and statives) (Huang 1997; Huang & Shi 2006). Stative predicates marked by =in appear to be only interpreted as inchoative, (1). Likewise, a universal-perfect reading cannot be expressed with =in. However, a homogeneous state with =in is attested in contexts where the hearer believes the negation of the asserted state is true, (2–3). It is evident that the change of state associated with =in cannot be attributed to a propositional component.

(1) Ma-diai(#=in) a bunbun=a. [Context: You are teaching kids the color: AF-yellow=NOM banana=DIST.NOM “Look. Those bananas are yellow.”]
‘Those bananas are yellow.’ (With =in: ‘Those bananas become yellow.’)

(2) Nau tu maisikauma-ikit saia hai, masmuh=in. [Context: You think that EPIST AF.be.from-little 3SG.NOM TOP AF.fat=COS he recently becomes fat, ‘(Contrary to what you thought) Since he was young, he has been fat.’ but I say:]

(3) Na-masabah=in saia mais minaunin. [Context: You’ve been trying hard FUT-AF.sleep=COS 3SG.NOM when.NPST afternoon to put your kid to sleep Your mom comforts you, “It’s okay…”]
‘He will (then) sleep this afternoon.’

Furthermore, we present evidence against analyzing =in as a perfect aspect or a counterpart of ‘already’. Firstly, in out-of-the-blue contexts, =in conveys an imminent future reading with a future reference time. The temporal relation is different from that of either a future perfect or the present perfect. Secondly, =in involves a certain “current relevance” effect that is different from the English present perfect: it is compatible with a past-time adverbial and the past tense (Jeng 1999). Thirdly, =in does not carry an earliness implicature as ‘already’ typically would do (Vander Klok & Matthewson 2015; cf. Lübner 1989; a.o.). Notably, =in is preferred with ‘finally’, which requires ‘a state to be evaluated as later than expected’ (Olsson 2003:11).

We propose that much like the Mandarin sentence-final le, the presence of =in is to indicate that the state of affairs is contrary to what the hearer assumes (Soh 2009), yielding a change of belief at the current stage of discourse (cf. Li et al. 1982; Ebert 2001). That is, =in is a marker that updates the information in the current discourse. We suggest that the line of reasoning explains why =in lacks the adverbial restrictions as with the English perfect as well as why it can have a non-inchoative or non-imminent-future reading. Data such as these are crucial for =in and potential iamitives to be not analyzed as a temporal marker that entails a result state.

A recently-discovered translation from Dutch into the Formosan Language, Siraya

In Amsterdam in 1661, the Dutch missionary Daniël Gravius published a volume comprising his translations of the Gospels of St. Matthew and St. John in the Formosan language, Siraya. Until recently, it was thought that only the translation of the Gospel of St. Matthew had survived. However, this author recently identified a copy of the 1661 publication which contains both Gospel translations. As Siraya became extinct in the nineteenth century, these two Gospel translations are important sources for this language. Extensive work has already been carried out on Siraya by Li (2010) and Adelaar (2011). However, the Gospel of St. John differs in both content and lexis from the Gospel of St. Matthew, and an analysis of its language offers the opportunity to increase our understanding of Siraya. This contribution analyses the lexical items which occur in the Gospel of St. John, but not in other surviving Siraya texts. The lexical items that we can add to those in the glossary of Adelaar (2011) include ‘water jar’, ‘mud’, ‘cave’ and ‘to spit’. The paper will identify cases where these words appear to be reflexes of reconstructed proto-Austronesian forms and cognates of words in other Formosan languages. For example, it will argue that pangarei, the Siraya verb ‘to spit’, is related to the proto-Austronesian word for ‘saliva’ *ŋalay, with pa- functioning as a causative prefix (Wolff 2010: II, 917, 1039). The paper will then analyse borrowings in the Siraya Gospel of St. John using the loanword typology of Winford (2003) based on that of Haugen (1953). These include ‘pure loanwords’, ‘loanblends’ and ‘loanshifts’. In this regard, the paper will pay close attention to the influence of the source text, the Dutch Statenbijbel (1st ed. 1637), on the borrowings in the Siraya target text.

Bibliography


Author:

Professor Christopher Joby
Visiting Professor,
Dept. of Dutch and S.A. Studies
Poznan, Poland
http://wa.amu.edu.pl/wa/joby_christopher
Decoupling the relation-marking and transitivity-marking functions of Austronesian voice

Daniel A. Kaufman, City University of New York
Danield.Kaufman@qc.cuny.edu

It is generally agreed upon that the Austronesian voice system fulfills both (i) a transitivity-marking function in the basic Actor Voice/Non-Actor Voice (AV/NAV) dichotomy as well as (ii) an important role in marking the grammatical relation of an extracted argument. Focusing on (i) has led to typologizing Austronesian languages by alignment type along an ergativity-accusativity axis while focusing on (ii) has led to a variety of proposals that view the voice system as a type of case agreement. Here, I put aside questions of morphosyntactic alignment to ask whether these two functions can be fruitfully decoupled or whether they must be linked, as is often presupposed. Treating them as (semi-)independent variables, we arrive at a rough typology where voice plays a role both in relation-marking in extraction contexts and transitivity marking (Tagalog-type languages), only relation marking (Malagasy type languages), and only transitivity marking (Selayarese-type languages), as shown:

Tagalog-type: Relations, Transitivity
Malagasy-type: Relations, *Transitivity
Selayarese-type: *Relations, Transitivity

In languages where the historical voice morphology has been completely lost, other syntactic reflexes of the old AV/NAV distinction arguably continue a role in transitivity marking, as in Tongan-type languages. In Tetun type languages, on the other hand, the historical Austronesian voice system plays no role in either transitivity marking or relation marking:

Tongan-type: *Relations, Transitivity
Tetun-type: *Relations, *Transitivity

While the two functions above are relatively independent, they come into clear conflict in cases of agent extraction with a definite patient. Here, the relation marking function requires AV while the transitivity function requires NAV. By viewing these principles as constraints in conflict, we can explain the rare Selayarese-type among the voice-preserving languages, in which the transitivity-marking function takes precedence over the relation-marking function. I argue that this type provides a very valuable insight into the development of a pattern seen more widely in the Oceanic languages. Philippine languages on the other hand, which tend strongly to prioritize the relation-marking function over the transitivity-marking function of voice when these two come into conflict, employ case-based strategies to resolve the tension, an option which is not available to those languages south of the Philippines which have largely lost the historical case markers.
What is “natural” speech? Comparing free narratives and Frog stories in Indonesia

Marian Klamer & Francesca R. Moro
Leiden University & Leiden University/Università degli Studi dell’Insubria, Como

Using visual prompts to elicit narratives has benefits, but also comes with a cost: it elicits less natural language from speakers. While there is overall consensus that narratives obtained by means of visual stimuli contain less natural language than free narratives (Himmelmann 1998, Foley 2003, de Leon 2009), it has been less clear what exactly is “natural” speech, and how the naturalness of narratives can be measured in a crosslinguistically meaningful way.

In this paper we study the naturalness of narratives along the following three dimensions: (1) lexical density (measured as noun-pronoun ratio, noun-clause ratio, and noun-verb ratio); (2) narrative style (measured as the amount of direct speech and tail-head linkage used), and (3) speech rate (measured as words uttered per minute). We use these dimensions for a quantitative comparison of the language of free narratives (traditional narratives and narratives about personal experiences or daily activities) and narratives that were elicited using the Frog story picture book (Mayer 1969). We compare the free and prompted narratives of eight speakers of Alorese and Teiwa, see map 1. Both these languages are spoken in the same region and neither is commonly written. However, as they belong to different families (Alorese is Austronesian, Teiwa is Papuan), they are lexically as well as typologically very different.

Our study finds significant differences between free and prompted narratives along the three dimensions mentioned above. For example, one difference we find is the higher degree of lexical density in Frog stories, with significantly more overt nouns and fewer pronouns. Generally, the presence of many nouns is related to explicit style which is typical of context-independent written language, while the use of pronouns is more typical of spoken language that is contextually embedded (Norby & Håkansson 2007). Frog stories are delivered orally, but they lack a context because they are new to the speaker, and are not part of his/her cultural practice. Therefore, instead of tracking referents by using pronouns, speakers prefer to be explicit by using full NPs. We also found that style and speech rate varies systematically per narrative type: Frog stories, as picture descriptions, make less use of specific stylistic devices as direct speech and tail-head constructions than free narratives, and have a slower speech rate because the words and structures of the Frog stories are not entrenched in the speakers’ repertoire so that they need more time to plan utterances.

Overall, we show that using the Frog story picture book as a stimulus to elicit narratives has clearly measurable linguistic effects. The three dimensions used in the present study proved to be useful to measure these effects crosslinguistically.

Introduction. In 2003, Warama Kurupel (Suwede) founded the Ende Language Committee (ELC) to produce a written representation of Ende (Papuan; Pahoturi River). Since then, the aspirations of the ELC have grown to include documentary and descriptive goals that aim to capture both the structural complexity of the language and the sociolinguistic variation within the community. To meet these goals with limited time to collect data, we wanted to maximize the utility of our efforts. Therefore, we trialed three semi-controlled experimental methodologies—a picture task, an interview task, and a video stimuli task—and found that they generated sufficiently natural discourse to reveal a diverse array of linguistic structures in a controlled-enough framework to allow comparison across speakers, dialects, and even languages.

Methods. We found success with three experimental strategies: (1) the Family Problems Picture Task (FPPT; Barth & Evans, 2017), in which participants were asked to describe and organize a set of 16 picture cards into a narrative; (2) the Sociolinguistic Questionnaire (Lindsey, 2019, p. 301), in which participants engaged in a monolingual interview with 60 standard questions concerning their social and linguistic background; and (3) the Verbal Number Stimuli Task, in which participants described a series of videos that contrasted a range of aspectual and pruralional event types known to be important in Pahoturi River verbal morphology (Lindsey & Schokkin, 2018). All three methods included visuals, discussion topics, and tasks that were familiar to participants and were video-recorded and transcribed on-site by native Ende speakers. These data collection strategies were successful at meeting the goals of the ELC because they were semi-controlled, adaptable, and natural. These characteristics allowed us to collect spontaneous speech from a large sample of speakers of Ende and other languages, ultimately supporting the breadth and depth of the documentation of Ende.

Findings. The use of these experiments resulted in 200% more community representation than the use of narrative collection and grammatical elicitation strategies alone (N=103 speakers vs. 51 speakers), making this data valuable for sociolinguistic analysis. For example, the Sociolinguistic Questionnaire corpus was instrumental in investigating word-final /n/-elision and retroflex stop affrication, two salient sociophonetic patterns observed across the Pahoturi River family. Lindsey (submitted) and Strong et al. (2019) determined that the use of the standard forms correlated significantly with the speaker’s status as an orator, a position of prestige within the community.

These methods also enhanced the descriptive depth of the corpus. For example, the FPPT data illuminated an elaborate possessive marking system in Ende. The FPPT is part of the SCOPIC project, meaning that this task has run on multiple languages, including related Idi (Gast, p.c.). While Gast reports a startling absence of possessive morphology in the Idi FPPT corpus (Evans & Barth, 2017), I observed at least five possessive marking strategies in the Ende FPPT corpus, including (1) possessive pronouns, (2) possessive pronominal clitics marked on the possessor (e.g., lla=bo ngonongg ‘man=3.SG.POSS thought; the man’s thought’), (3) a possessive kinship suffix marked on the possessum (e.g., mulla-da=bom ‘wife=3.kin.POSS=ACC; [they saw] his wife’), (4) a combination of strategies 2 and 3 resulting in an inalienable type possessive suffix (e.g., lla-da=bo eka ‘man=3.kin.POSS=3.SG.POSS story; the man’s story’), and (5) word order, possessor followed by possessum. These strategies were observed across multiple iterations of the task; no single run of the experiment would have been sufficient for capturing all the marking strategies.
Finally, the Verbal Number Stimuli Task not only allowed for an in-depth investigation of aspectual and pluractional number indexing in Ende verbs, but the data corpus also showcases the variation within the community and across the language family, as the task was run in both Ende and Idi (Schokkin and Lindsey, submitted). For example, Idi and Ende ditransitive verb roots differ in that Idi only indexes plurality of the theme, while Ende indexes plurality whether it is the recipient or the theme that is plural.

**Summary.** This talk details three methodologies that greatly enhanced the documentation of Ende, both in uncovering structural depth and in providing comparable data for studying inter- and intra-speaker variation.

In this talk, I will describe the use of the particle *asang* ‘new situation’ and the TAM marker *daa* ‘irrealis; future’ as grounding strategies in narrative texts in Lakurumau, a Western Oceanic language of New Ireland (Papua New Guinea). All data come from my own fieldwork (2017-2019).

Following Hopper (1979), I refer to the events that bring forward the actual story line as ‘foreground’ and to the events that support the foreground and do not themselves narrate the story as ‘background’. In Lakurumau, both can be in the unmarked realis, which can express both past and present reference and is the default choice for narratives. Often, however, a clear distinction is made: foreground events are marked by the particle *asang* and background events by the irrealis/future *daa*. *Asang*-marked predicates bring forward the storyline, while *daa*-marked predicates expand on it, encoding events seen as habitual or persistent and referring deictically to the present of the narration (the ‘speaker-now’; Fleischman 1972). In (1) the background provides information about the magician-warrior Malaxon and his magic plant (*a Cordyline australis*); in (2) the narrator interrupts the flow of narration to comment on his use of Tok Pisin *baak* instead of Lakurumau *xupxup* (backgrounded events are in square brackets).

(1) *Orait, nam Malaxon asang ka ramazik, di ramin asang a yaan aa xa xus a mu piu zana. Ka gonin asang naadi wana mu mos pena waan lo vezop.* [Aa naan, nam Malaxon akam, a vuna roi moxom a waxaana si. Taim o kain taatai urukaam… taim gu *daa rapin* maan a ron sim, ka *daa tef tepin* kam pana si.]

‘All right, Malaxon *got up ASANG*, they *decided ASANG* a date (for the battle, LM) and he told it to his warriors. He *prepared ASANG* them to go to the battle. [And him, that Malaxon, he used to hold a leaf of *Cordyline australis*. This kind of men… when you will throw your spear, he *will fan it away* with the *Cordyline australis*.]'

(2) *Okay maadi valaa asang. Maadi walongin asang anga wiu xa baak.* [Ga *daa pet* oy anga wiu xa xupxup, ah.] *Nanga walongin asang anga wiu xa xupxup.*

‘Okay, the two of us *went ASANG*. We *heard ASANG* a dog barking. [I *will say* that a dog ‘xupxup’, huh.] I *heard ASANG* a dog barking’

I argue that the discourse functions of *daa* and *asang* can be directly derived from their semantics. The basic function of *daa* is to encode typical irrealis functions, including habituality, and to express future tense reference: the latter is re-analysed as ‘speaker-now’ reference in discourse. *Asang* expresses meanings typically related with *already* and perfect grams, such as an earliness implication, change-of-state with stative predicates and a ‘new situation’ with dynamic predicates (Krajinović 2019). Even though its functions resemble those of
perfects, *asang* does not appear in the slot reserved for TAM markers, can co-occur with aspectual markers such as *nga* ‘progressive’ and it can also modify nouns and adverbs, marking contrastive focus and unexpectedness. I analyse it therefore as a ‘new situation’ and ‘contrast’ marker. Its function as a foregrounding device is consistent with this semantics: *asang*-marked events are seen as contrasting with the previous ones and thereby inherently marking a development in the story line (cf. the analysis about Tokelauan, where *kua* ‘perfect; new situation’ is also analysed as a foregrounding device (Hooper 1998).

**References**


Bringing psycholinguistics to the field: experiences from the Solomon Islands

Åshild Næss, University of Oslo
ashild.nass@iln.uio.no

Sebastian Sauppe, University of Zürich
sauppe.s@gmail.com

Existing research on the cognitive and neurobiological basis of language covers only a very small sample of the world’s languages, with a heavy bias towards major European languages (Anand et al. 2010, Norcliffe et al. 2015). As tools commonly used in the cognitive and neurosciences (such as eye tracking and EEG) become more portable, this may change. However, bringing experimental psycholinguistics to the field comes with its own set of challenges, «from simple issues such as working in a difficult climate or with limited access to power supplies, to more complex issues such as negotiating cultural expectations or establishing a collaborative research environment» (Hellwig 2019). As Hellwig points out, for such challenges to be met, it is necessary to attempt to reconcile the methods and values of experimental psycholinguistics and language documentation, two approaches which from the outset have very different perspectives and goals.

In this talk, we will report on our experiences carrying out an EEG experiment with speakers of the Äiwoo language in Solomon Islands in the southwest Pacific. We will mention specific technical challenges that arose in designing suitable stimuli and implementing the experiment. Our main focus, however, will be on culturally-related challenges involved in recruiting participants and explaining not only the purpose of the experiment, which differed substantially from what the participants expected from ‘language research’, but also its value as understood from a local perspective. Ultimately, our experiences show that experimental psycholinguistic research can be carried out relatively successfully in a fieldwork setting, but it depends crucially — and unsurprisingly — on how the local community is approached and informed. We had planned to carry out the experiment in the Reef Islands where Äiwoo is the main language, and where one of us has extensive experience and a large local network, but for various practical reasons ended up collecting data instead in Solomon Islands’ capital Honiara. Our strategy for engaging with the community therefore had to be rethought and largely improvised, and succeeded in the end mainly due to an individual speaker who took an interest in the project and acted as liaison to the local community of Reef Islanders.

References

Hellwig, Birgit. 2019. Linguistic diversity, language documentation and psycholinguistics: The role of stimuli. in Aimée Lahaussois and Marine Vuillermet (eds), Methodological tools for linguistic description and typology, 5-10. Language Documentation and Description Special Publication 16.

Syntactically relevant morphology: The Coastal Marind fourth gender
Bruno Olsson & Matthew Carroll (CoEDL, Australian National University)

A common assumption in most approaches to morphosyntax is that purely morphological facts, such as inflectional class membership or patterns of syncretism, do not have effects in the syntax (Zwicky 1987). In this paper, we argue that a potential counterexample to morphology-free syntax is provided by syncretism in the exponence of gender agreement in Coastal Marind, a Papuan language of the Anim family.

Coastal Marind has a system of four genders (labelled with the Roman numerals I–IV; Olsson 2017). Nouns are assigned to a gender on the basis of animacy and sex: male humans are in Gender I, female humans and all animals in Gender II; inanimates are assigned to Genders III and IV. The morphological pattern that we claim is relevant to syntax is found in targets that agree according to gender. We show that a pattern of syncretism runs through all targets, regardless of part of speech and lexical idiosyncrasies, whereby the exponent of the plural of Genders I & II (which make a number distinction) is identical to the exponent of Gender IV (which, like all inanimate nouns in the language, lack number). This systematic syncretism is morphomic (Aronoff 1994), most clearly because there are no semantic features grouping the categories (Gender IV and animate plural) that map onto the shared set of exponents. We discuss, and reject, descriptions in term of semantic commonalities, which would have deprived the pattern of its morphomic status, as morphomes per definition have no semantic basis.

We then demonstrate some unexpected repercussions of the morphomic pattern in two areas of Coastal Marind grammar: verb suppletion according to participant number and the availability of applicative constructions. Several verbs employ unrelated stems depending on whether the absolutive participant is singular or plural, as in the pair *man* ‘one come’ vs. *na°am* ‘many come’. We show that the morphomic syncretism reoccurs with such verbs, and we systematically find the suppletive plural stems triggered by absolutive arguments belonging to Gender IV. This alternation is remarkable as the semantic concept of participant number is a well-known trigger of suppletion cross-linguistically, but here the verb suppletion turns out to obey a completely morphomic pattern found elsewhere in the grammar. We then turn to two applicative constructions used to add a comitative participant to a motion verb. One construction is used to add an animate object, while the other adds an inanimate object. We show that in addition to involving unrelated morphology, these constructions differ structurally and semantically and must be treated as separate syntactic configurations (rather than different variants of the same construction). Surprisingly, we find that the construction associated with inanimates is unavailable to nouns in Gender IV, and that despite their inanimate status, these nouns select for the construction associated with an added animate object—again, as predicted by the morpheme grouping Gender IV nouns with (plural) animates. The data is remarkable, and we argue that it constitutes a serious challenge to the claim that syntax is morphology-free.

References

Verbal focus marking in Coastal Marind and elsewhere
Bruno Olsson (CoEDL, Australian National University)

Despite an explosion of interest in the cross-linguistic expression of focus in recent decades, the phenomenon of VERBAL FOCUS MARKING remains woefully underexplored. Verbal focus marking means that the presence and/or role of a focussed constituent is marked not primarily by resources such as prosodic prominence or focus particles on the focussed constituent itself, but rather by a special form of the verb. Such systems have been described most prominently for Bantu languages, but also in two Papuan languages, Watam (Lower Sepik-Ramu; Foley 1999) and Coastal Marind (Anim; Olsson 2017). In all of these languages, the focussed constituent is obligatorily placed in a verb-adjacent position (immediately post-verbal in the Bantu cases, and immediately pre-verbal in the Papuan ones), while the verb carries morphology distinguishing it from verb forms used in focus-less clauses.

In this presentation, I zoom in on the verbal focus marking system of Coastal Marind, and its relevance for the cross-linguistic picture. Examples of focus marking in a ditransitive clause are given in (1). In (a), the constituent expressing the theme (T) participant is focussed, and therefore obligatorily pre-verbal (contrasting with the otherwise flexible constituent order of the language). The prefixal complex marking the verb contains the ‘Object Orientation’ prefix ma-marking the role of the focussed constituent. Compare this with (b), in which the recipient (R) is focussed, now triggering the ‘Directional Orientation’ ka- on the verb. Appropriate contexts for these utterances would be e.g. as answers to content questions (‘What did they feed the dogs?’) or as replacive/contrastive focus (‘They fed meat to the pigs.—No, they fed meat to THE DOGS.’).

(1) a. patul muy ma-n- y-akoh nggat
    boy [ meat ]T OBJ-3pl.A- 2|3pl.u-feed dog
    ‘The boys fed MEAT to the dogs.’

b. patul nggat ka-n- y-akoh muy
    boy [ dog ]R DIR-3pl.A- 2|3pl.u-feed meat
    ‘The boys fed meat to THE DOGS.’

The Coastal Marind system of verbal focus is more complex than those previously described, involving five prefixes (two of which are illustrated above) marking different roles of the focussed constituent, and exhibiting various interactions with other subsystems of the grammar. In the presentation, I will use corpus data to demonstrate how speakers put the focus system to use, and how syntax, morphology and prosody interact in the expression of focus. I will discuss some striking similarities between verbal focus marking systems in different languages, but also address some fundamental differences. The Coastal Marind data show that rather than forming a tight-knit ‘focus system’, it is probably best to describe the various patterns discussed here (preverbal placement, special verb prefixes, prosodic prominence) as independent semiotic resources that serve several functions in the language, although speakers need to harness all three of them to achieve focus interpretations.

References

The phylogenetic status of the Kaure languages of northeastern West Papua

Bill Palmer

University of Newcastle

The eastern Lakes Plain of West Papua is a phylogenetically complex and poorly understood linguistic region: home to 7 unrelated families, 7 isolates, and members of several larger families (Foley 2018). Little historical research exists on these languages, and most phylogenetic groupings are tentative and disputed. I present findings of a comparative study of one grouping. Five languages have been associated with a Kaure stock within Trans New Guinea (TNG) - a Kaure family of Kaure itself, Narau and Kosare, and 2 family-level Kaure Stock isolates: Kapauri and Sause (Voorhoeve 1975; Wurm 1982). However, Silzer & Heikkinen (1984) place Sause in a Tor-Lakes Plain Stock unrelated to Kaure. Foley (2018) regards Kapauri as a likely true isolate unrelated to Kaure, and Sause as a possible member of the Tor family, but more likely also an isolate. Pawley & Hammarström (2018), regard the evidence on even Kosare as insufficient to posit a link with Kaure, which they do not include in TNG. They do not mention Sause.

I compare 9 doculects identified as Kaure, from Lereh, Kasu, Wes and Soar villages; Guai and Yafi districts (2 each); and 1 of unknown origin; along with the Narau doculect. With lexical evidence and sound correspondences and I show that all but Yafi are varieties of a single language. Narau is not a separate language, but a Kaure dialect most closely akin to Kasu and Soar (e.g. *nai ‘eat’ vs *di in Lereh and Wes). However, with low cognate counts (66%) and regular sound correspondences (e.g. *h>k/̃+back), I show that Yafi is a separate language.

I then compare Kaure and Yafi with 3 doculects of Kosare, 3 of Kapauri, and 7 of Sause, showing that all are related, with regular sound correspondences. Sause has undergone several changes not seen in the other varieties, including *k>ʔ/V_V; *w>0/C_V; *n>ŋ/#; *s>h/#_ (neutralizing with *h); and loss of some final vowels. Other correspondences include *k>h/#_ in Lereh and Wes; *k>q/V_C in Wes; *k>x medially in Lereh; *r>l in Lereh; *r>l/C_V and *r>0/V_C in the other Kaure dialects; and loss of final consonants in most Kaure varieties. Higher cognate numbers and shared lexical innovations (e.g. the IEXCL pronoun wey) indicate that Kaure, Kosare and Yafi form a closely related group, and I present a set of reconstructed forms for Proto Kaure-Kosare. Shared lexical innovations also link Kapauri and Sause, but less closely than Kaure-Kosare. Overall, the data suggests a tentative branching into a Kaure-Kosare-Yafi subgroup and Kapauri and Sause, possibly forming a subgroup, with evidence of contact between many of the varieties, particularly Kapauri and Kosare (e.g. Kapauri’s innovated form for ‘ear’ has been borrowed into Kosare alongside that language’s regular cognate form). I reconstruct a set of forms for the proto-language. Finally, I explore possible links with TNG, and find no evidence supporting TNG membership.

<table>
<thead>
<tr>
<th></th>
<th>Yafi</th>
<th>Lereh Kaure</th>
<th>Wes Kaure</th>
<th>Kasu Kaure</th>
<th>Kosare</th>
<th>Kapauri</th>
<th>Sause</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>one</code></td>
<td>su tu</td>
<td>soltok</td>
<td>sutu</td>
<td>sutu</td>
<td>sroto</td>
<td>srutu</td>
<td>hor</td>
</tr>
<tr>
<td><code>eye</code></td>
<td>we  we</td>
<td>hwen</td>
<td>hwai</td>
<td>wèi</td>
<td>-</td>
<td>[hu]kwani</td>
<td>kən</td>
</tr>
<tr>
<td><code>ear</code></td>
<td>yo-klu</td>
<td>ho-xelyk</td>
<td>hu-[a]gllyt</td>
<td>-</td>
<td>kəro, wɔru?</td>
<td>[tu]waru</td>
<td>[to]?or</td>
</tr>
</tbody>
</table>

Smits, L. & C.L. Voorhoeve. 1994. The J. C. Anceaux Collection of wordlists of Irian Jaya languages B: Non-Austronesian (Papuan) languages (Part I). Leiden-Jakarta: Department of Cultures and Languages of Southeast Asia and Oceania.


Prior motion serial verb constructions in Vatlongos (Southeast Ambrym, Vanuatu):
same morphosyntactic behaviour, different diachronic outcomes

Eleanor Ridge, Massey University

In prior motion serial verb constructions (SVCs), a deictic motion verb precedes another verb, indicating the direction of motion prior to performing the action of the other verb (see ‘sequential’ SVCs, Lynch et al. 2002; ‘associated motion’ SVCs, Cleary-Kemp 2015).

(1) Inou na-mmei na-pus sista navan (2) muli-ha muli-lau e rin ak
1SG 1SG.NFUT-come 1SG.NFUT-see sister my 2DU.DFUT-go 2DU.DFUT-hunt LOC side PROX
‘I’ve come to see my sister.’ ‘You’ll go hunt on this side.’

In closely-related Paamese prior motion is instead expressed using subordination (Crowley 2002): an affirmative realis motion verb is followed by an immediate future action verb, whereas serialised affirmative verbs must have matching TAM marking (as in the Vatlongos examples above). These were therefore tagged separately in the initial stages of corpus development for Vatlongos, making the quantitative analysis presented in this paper possible.

Vatlongos SVCs are most robustly identified by the behaviour of the negative clitic =ti in future tenses. Negative polarity is marked by both a prefix and a clitic, but in an SVC the negative clitic only occurs after the first verb. The second verb in the prior motion SVC in (3) takes the prefix naa-, but is not followed by the clitic =ti, so would be ungrammatical if a separate clause.

(3) mut-naa-va =ti mut-naa-pat?
2PC.DFUT-NEG-go =NEG 2PC.DFUT-NEG-sleep
‘Aren’t you going to go sleep?’ [20141028a_c01m002_14]

‘Go’ and ‘come’ are the only verbs which participate in SVCs and also have auxiliary verb counterparts: monosyllabic forms lacking verbal prefixes. The auxiliary verbs express prior motion, and these SVCs are the likely bridging context for this diachronic development.

There is also evidence for prior motion SVCs following a distinct diachronic pathway in contemporary variation in the use of SVCs. Prior motion SVCs are very frequent, a factor favouring their reanalysis as auxiliaries. In a corpus of ~48,000 words, they occur once every 98 words, while all other types of SVCs combined occur once every 25 words. This high frequency meant that prior motion SVCs could be analysed separately in a negative-binomial regression analysis that modelled the frequency of tokens of constructions per text, based on speaker community, age, gender, years of education and text genre. There was a significant effect of years of education on the use of other types of SVCs (more educated speakers less likely to use SVCs), and near significant effects of community (speakers in the urban community less likely to use SVCs) and genre (SVCs more likely in animal narratives; less likely in employment texts). However, there were no significant effects on prior-motion SVCs, nor consistent trends in the apparent effects.

This paper describes the morphosyntactic criteria used to identify SVCs and auxiliary verbs in Vatlongos, and the methods of corpus preparation and quantitative modelling used to analyse the distribution of the constructions, using FLEx and R, via the interlineaR package (Loiseau 2018). It also considers possible reasons for the distinctive diachronic outcomes for prior-motion SVCs, such as the relative grammatical simplicity of tokens of prior-motion SVCs, the role of Bislama and English in the language contact situation, and linguistic vitality. Overall the paper aims to demonstrate how tagging constructions at this broad level can be a useful first step to identifying more specific contexts for syntactic variation, while also facilitating careful qualitative analysis throughout all stages of documentation and description.
Voice choice in Totoli discourse
Sonja Riesberg¹,², Maria Bardaji i Farré³, Kurt Malcher¹, Nikolaus P. Himmelmann¹
¹Universität zu Köln,
²ARC Centre of Excellence for the Dynamics of Language, Australian National University

Totoli is a western Austronesian symmetrical voice language, i.e. it has three basic transitive constructions – one actor voice and two formally distinct but functionally equivalent undergoer voices. In some contexts, the choice between these voices can be straightforwardly predicted, as there are number of syntactic operations (e.g. questioning and relativization) that are exclusively accessible to the subject argument. In most cases, however, voice selection does not depend on any syntactic constraints. The question addressed in this paper is thus: What determines voice choice in Totoli discourse?

In the first half of this paper we present the results of an extensive discourse study. We annotated 27 oral texts of different genres (including traditional narratives, stimuli based monologues and dialogues, and natural conversations), using the GRAID (Grammatical Relations and Animacy in Discourse) schema developed by Haig and Schnell (2014). Our corpus comprises 02:50:04 hours natural speech, consisting of 16.160 words, 6.745 intonation units, and 3.052 clauses (1.056 of which are transitive). In this, it is considerably larger and more diverse than most corpora that built the basis for similar research questions in other Austronesian languages¹ (Hopper 1983, e.g., used 150 written clauses, taken from a novel, in his study on voice choice in Malay). The major finding from this study is that the factors claimed to be relevant for voice choice in other languages – primarily discourse transitivity and topicality (cf. Cumming 1991 and Pastika 1999), but also animacy, activation status, collostruction strength, or clause type (McDonnell 2016) – appear to only play marginal role in predicting voice choice in Totoli.

In the second half of the talk we will then report on a considerably smaller pilot-study. In a small sub-corpus consisting of 5 Totoli narratives (1480 words, 396 intonation units, 270 voice marked clauses) we found a strikingly high number of actor voice constructions at episode boundaries, i.e. in the first two clauses of a new episode. That is, while the overall corpus had a ratio of 20% actor voice vs. 80% undergoer voices (roughly the same ratio reported for other Austronesian languages), we found the reverse distribution at episode boundaries: In 85% of all cases, the first transitive voice marked sentence of a new episode was an actor voice constructions. Conversely, only 15% of these clauses were marked as undergoer voice.

We argue that these observations can be taken as evidence that the discourse function of the different voices in Totoli (but probably also in other Austronesian languages with symmetrical voice systems) differ from those of Standard Average European active-passive alternations in that the former operate on a far more global discourse level than the latter.

References

¹ The one exception is McDonnell’s excellent dissertation on Basemah (McDonnell 2016), which investigates voice choice in a corpus of Basemah conversations with similar size to ours.


Numeral and possessive classifiers in Patani: one coherent nominal classification system?

Linn Iren Sjânes Rødvand, University of Oslo

Two types of nominal classification often found in Austronesian languages are numeral classifiers (Blust 2013: 282ff) and systems of possessive classification (e.g. Lichtenberk 1983). However, while numeral classifiers are more often found in Western Austronesian languages, possessive classification systems occur mostly in Oceanic languages, meaning that the two types of system rarely co-occur within the same language. One exception to this is found within the languages of the South Halmahera-West New Guinea (SHWNG) subgroup, where both types of system are described e.g. for Sawai (Whisler 1996). By contrast, Bowden (2001) describes an elaborate system of numeral classifiers for Taba which like Sawai belongs to the South Halmahera branch of SHWNG, but no distinctions in possessive marking.

In this paper, I will present data from the hitherto undescribed SHWNG language Patani, which has both numeral and possessive classification, and investigate, firstly, how Patani’s numeral classification system compares to those of closely related Sawai and Taba; and secondly, to what extent the two systems in Patani use the same principles of classification. On the first point, Patani has the classifying prefixes pi- and si-, which are clearly cognate with Sawai pe- and se- and Taba p- and sis=-; but Patani has a third prefix met- not found in Sawai, and with no obvious parallel in Taba. I will describe the principles underlying the selection between these numeral classifiers in Patani, and compare them to those governing the choice between the three possessive constructions found in Patani: 1) affixation on the possessed noun, 2) preposed possessive article anik/ni (1SG/3SG), or 3) preposed possessive article anak/nɔ (1SG/3SG). While fieldwork on Patani is ongoing, preliminary data suggest a close overlap between the two systems, as shown in Table 1 below. Given that relatively few Austronesian languages display both numeral and possessive classification, an investigation into the relationship between the two in the languages that do have them constitutes a valuable contribution to the typology of Austronesian.

<table>
<thead>
<tr>
<th></th>
<th>si-</th>
<th>pi-</th>
<th>met-</th>
</tr>
</thead>
<tbody>
<tr>
<td>anik</td>
<td>anik</td>
<td>meja</td>
<td>si-lu</td>
</tr>
<tr>
<td>1SG.Poss table</td>
<td>1SG.Poss table</td>
<td>CL-two</td>
<td>'my two tables'</td>
</tr>
<tr>
<td>anak</td>
<td>anak</td>
<td>bet</td>
<td>pi-lu</td>
</tr>
<tr>
<td>1SG.Poss garden</td>
<td>1SG.Poss garden</td>
<td>CL-two</td>
<td>'my two gardens'</td>
</tr>
<tr>
<td>affixation</td>
<td>auye-g</td>
<td>met-lu</td>
<td></td>
</tr>
<tr>
<td></td>
<td>foot-1SG.Poss</td>
<td>CL-two</td>
<td>'my two feet'</td>
</tr>
</tbody>
</table>

Table 1: preliminary data illustrating the overlap between the numeral classifiers (horizontally) and the possessive classifiers (vertically)

Selected references:
What can we learn from historical documentation of the Pahoturi River family?

Linguists studying the Pahoturi River (PR) family often emphasize the newness of documentation efforts for these languages, e.g. Evans (2012) “there is negligible published material on languages of the PR family.” With the exception of the detailed grammar sketch of Ende in Lindsey (2019), none of the PR languages has a published grammar, and a huge amount of research remains to be done. Nevertheless, the earliest published documentation of PR languages is well over a century old. Quite a few word lists are available from the colonial period, covering six language varieties within the PR family. Although far from complete, this data is not negligible. In this talk, I outline some of what we can learn from the early documentation of PR languages.

I begin by contextualizing the early word lists, sketching the history of European interaction with PR-speaking people. I introduce the available documents and researchers and illustrate the names given for the tribes and languages on a map of the region. I will keep my discussion of the sociopolitical dynamics surrounding the creation of these early records brief to save time for linguistic issues and findings. Nonetheless, it is necessary to acknowledge the close and explicit relationship between linguistics, missionization, and imperialism, and I share some primary sources to that effect.

Next, I discuss my methods, from the search for relevant literature and documentation through organizing the data and making sense of the spelling conventions. I discuss how I matched the historical wordlists to the modern varieties, a task that is not as simple as it sounds due to people moving and names changing. In particular, I present evidence challenging the idea that Bugi and Agob refer to the same language variety (Evans et al. 2018), and I discuss a couple of ways that the old documentation may shed light on recent sound changes, such as variation between aspirated and unaspirated stops.

Finally, I take on an academic question raised by these early linguists which has been set aside by academics in the intervening century: the relationship between the languages of the Torres Strait islands and the PR languages. As Ray and Hadson demonstrated as early as 1898, there are many words in the Meriam language of the Eastern Islands of the Torres Strait that closely match words in the PR languages on the New Guinea mainland. Using a combination of recent and historical sources, I present further apparent cognates between the PR languages and the Meriam language of the eastern Torres Strait, providing further evidence of long-term contact between Torres Islanders and people from the south coast of New Guinea, as suggested in oral history and legends (Lawrence, 2004).

In conclusion, there is a great deal to be learned even from sparse colonial linguistic documentation. My next steps will be to take this information to speakers of PR languages and get their perspectives. I am particularly interested in learning more about oral history, legends, and current attitudes regarding the Torres Strait, and earlier tribe, language, and place names that will help confirm or refute my analysis of the relationships between current varieties and those recorded in the early documents.

References


Building the meaning of CAUSED POSSESSION –
the case of Vera’a (Oceanic) and Yali (Trans-New Guinea)

Stefan Schnell (Otto-Friedrich-Universität Bamberg)
& Sonja Riesberg (Universität zu Köln)

The concept of CAUSED POSSESSION is often assumed to be so basic to human social life that we should expect it to be lexicalised in all languages in a verb meaning ‘give’ (Newman 1996:3, 46ff), as has been shown to be the case for English give (and equivalents in Russian and Hebrew) (Levin 2008). Both the Oceanic language Vera’a (vera1421), and the Trans-New Guinea language Yali (pass1247) appear to lack a verb with a meaning similar to ‘give’ in the sense of CAUSED POSSESSION. The Vera’a verb frequently found in sentences expressing CAUSED POSSESSION is le (cf. (1a)); but in contrast to English give (Rappaport-Hovav & Levin 2008), le is compatible with a CAUSED (ACCOMPANIED) MOTION interpretation, and may in some contexts not express CAUSED POSSESSION at all, as illustrated by examples in (1b):

(1) a. \[di=m \text{ le } \_\] [mē die]\[3SG=PFV \text{ LE (it) DAT 3SG} \]
   ‘[He took the hat] and he gave (it) to him.’ [BSPF.102]

   b. \[\text{e } [\text{no}=k \text{ van } [\text{no}=k \text{ le lu } [\text{di]} \text{ ma}] \text{ DISC 1SG =AOR:1SG go 1SG =AOR:1SG LE across 3SG hither} \]
   ‘I’ll go and get her [so you can take a look].’ [JSV.096]

Likewise, the most frequent verb used to express CAUSED POSSESSION in Yali – og – (cf. (2a)), is also used to denote CAUSED (ACCOMPANIED) MOTION (cf. (2b)):

(2) a. \[\text{e } \text{anggen hinahan og } \text{isa-ehek} \]
   \[\text{tree fruit three og 3PL:OBJ:B-3SG.IM.PST} \]
   ‘he gave them three fruits’ [pear_story_Marthen_b 059]

   b. \[\text{og inap-tuk } \text{la-tug ik kume wat inap-tuk lat-pag} \]
   ‘they take them over there and they wash them’ [1st_initiation 400]

We show that both Vera’a le and Yali og have the sole meaning of CAUSED MOTION and that in both languages the CAUSED POSSESSION is a matter of compositional meaning corroborated by different constructional elements, in particular dative prepositions and speaker-directed particles in Vera’a and different types of auxiliary in Yali, as well as pragmatic inference. We hypothesise that more thorough cross-linguistic investigation will reveal that this strategy of creating the meaning of CAUSED POSSESSION is much more widespread than currently perceived in the literature.

AOR = aorist; DAT = dative; DISC = discourse particle; IM = immediate; OBJ:A/B = object auxiliary of paradigm A/B; PFV = perfective; PL = plural; PROG = progressive; PST = past; REM = remote; SEQ = sequential; SG = singular
References
This paper explores sociophonetic variation in Ende, a Pahoturi River (Papuan) language spoken by the Ende tän in southern New Guinea. While there is a growing body of variationist research taking place in the region, we know little about variation and change in these languages. Here, we examine variable affrication of Ende retroflex obstruents (ʂ̪)–(t̪) and (dz̪)–(d̪) and consider what linguistic and social factors are linked with this variation. Specifically, we highlight the locally relevant social factor of community orator, a group of individuals who perform a prestigious daily oration practice called kawa (Lindsey, 2019b). This ethnographic approach (following Eckert, 2012) not only supports the study of pattern-driven variation in Ende – i.e., how this social factor may be linked with multiple variables – but also provides much-needed contrast and comparison with dominant sociolinguistic theories. This is an important consideration given the underrepresented nature of the southern New Guinea context (Evans, 2012) and the potential for differences in social structure (Stanford & Preston, 2009).

To analyze the variable affrication, we used auditory analysis to examine the speech of 16 Ende speakers. This sample is balanced for gender, age, and orator status. Within a corpus of monolingual interviews (archived in Lindsey, 2017), we treated the retroflex obstruent variable (N=1,542) as a binary category (64% affricate, n=981; 36% stop, n=561). The likelihood of affrication is linked with the two linguistic factors of voice and syllable position, as well as a combination of social factors including speaker age, gender, and orator status. Statistical analysis, using mixed-effects logistic regression models, reveals a significant effect of orator status: retroflex obstruents are more likely to be realized as stops when produced by orators. Among the orators, there are also effects of age and gender, where older speakers and women are more likely to produce tokens as stops compared with younger speakers and men. In contrast, no age- or gender-based differences are observed among the non-orators. Acoustic analysis that takes into account gradience of the acoustic signal, and identifies fine-grained similarities and differences within and between social groups, is a critical next step for this study.

Our results are consistent with an interpretation of the stop variants as prestige forms and suggest that their use in Ende is linked with the speaker’s societal position: orators, who hold positions of high status in the community, use language to assert symbolic power. These findings build on Lindsey’s previous analysis of word-final /n/-elision in Ende (2019a), where the variation was also found to be strongly associated with orator position, and points to extant dynamics of power and prestige operating in the community that map onto linguistic variation. These results may further indicate a nuanced relationship between speech style and social categorization in Ende.

This particular example of sociophonetic variation is not limited to Ende. For example in Idi, retroflex consonants are also variably realized as stops and affricates (Schokkin, p.c). Variable affrication of alveolar fricatives (e.g., [zdźzdʒzdɕ]) is also a well-attested areal feature of the South Fly and can be found in other Pahoturi River languages as well as unrelated families such as Trans-New Guinea. It is likely that these variables, among others in the region, are also linked with a range of social factors, and the present study contributes to the foundational research in this understudied area. Future work that continues to draw on ethnographic documentation and explore broader patterns of variation in the Papuan context will enhance, and likely challenge, our understanding of dominant sociolinguistic trends.

Tonal contrasts in Hainan Cham (Tsat) and the role of enhancement in tonal space

Ela Thurgood
California State University, Chico

The analysis of the internal and external paths of Hainan Cham (HC) development led to a reconstruction of its five tones and showed how HC became a fully tonal Austronesian language (Thurgood 1966; 1999). A study of HC history, contact, and phonology followed (Thurgood et al 2014). Building on these foundations, the present paper aims to make further contributions to the research on tones in HC and beyond.

It has been well established that high vowels have a higher fundamental frequency (F0) than low vowels, a difference that increases the salience of vowel contrasts. There is disagreement about whether F0 difference is ever based on a deliberate enhancement of the speech signal (Diehl 1991; Kingston and Diehl 1994; Kingston 2007) or not (Whalen and Levitt 1995; Connell 2002). On the basis of our data consisting of 602 tokens produced by six HC speakers, we show that in HC the differences in fundamental frequency (F0) between high and low vowels in nonmodal phonations (depending on the tone, either falsetto or creakiness) are considerably beyond mechanically expected values pointing to deliberate enhancement.

Of the five tones, it is in the high tone that the F0 difference between high and low vowels is the largest, a pattern also found in other tonal languages. As observed by Kingston (2007:175), in the case of a high tone, speakers “can raise F0 more without pushing that tone’s F0 target into the range of another tone.” In HC, an association with falsetto results in an exaggerated F0 of a high vowel, and consequently, results in an exaggerated F0 difference between high and low vowels. We found the mean F0 difference to be as big as 130 Hz, while cross-linguistically, a big F0 difference is about 25 Hz (cf. Rose 1994; Shi and Zhang 1987; Zee 1980). The excessively high F0 with its accompanying falsetto in HC is a particularly dramatic example; least dramatic but similar in origin is the creakiness of the low tone.

Cross-linguistically, F0 differences between high and low vowels are reduced in words with low tones. In HC, however, the F0 difference between the high and low vowels can be as large as 30 Hz in the low tone. The F0 difference is enhanced not by raising the F0 of a high vowel but by lowering the F0 of the low vowel: The measurements of the amplitude differences between the first and second harmonic of /a/ show an intensity increase at higher frequencies, a feature of creaky vowels. The enhanced F0 difference based on lowering the F0 of /a/ is found in HC low and mid-level tones.

The two contour tones pattern with the level tones. Measured at two points, the rising tone shows the biggest F0 differences between high and low vowels at its onset and the falling tone at its offset. Once again, it is enhanced beyond what is typically found cross-linguistically; the F0 difference at the offset of the falling tone correlates with creakiness of the low vowel.

In HC, the enhancement of F0 differences between high and low vowels beyond the values reported in the literature occurs regularly. This enhancement leads to the enlargement of tonal space that in turn maximizes perceptual contrasts. In our discussion of the role of the tonal space enlargement, we suggest that in the community of HC speakers where the HC language is no longer spoken on regular basis, it helps reinforce the perception of perceived contrasts by, what Kensington (2007:210) in his study of prosodic context in non-tonal languages calls, “independently controlled means of conveying local information content.”

Phasal Polarity in Austronesian

Ljuba Veselinova  Leif Asplund
Stockholm University  Stockholm University
ljuba@ling.su.se  mawonali@gmail.com

The purpose of this paper is to shed light on the encoding of phasal polarity in Austronesian languages while also setting them in a broader cross-linguistic context.

Phases have to do with the beginning, continuation and end of a situation. Van Baar (1997) notes that the linguistic encoding of phasal polarity (PhP) differs along the lexico-grammatical scale. In many languages, phasal meanings are expressed lexically by adverbials, cf. (1). The PhP domain can be also encoded by morphological means in many languages cf. Plungian (1999).

Finally, there are languages where expressions of phasal meanings are morphosyntactically free but close to the grammatical pole in terms of frequency of use and abstractness of function, cf. Gorbunova (2014). Austronesian languages are a case in point. The main cross-linguistic studies – the PhP domain include Van Baar (1997) and van der Auwera (1998). Van Baar (1997) works with a stratified sample of 40 languages. He suggests that languages favor the continuative part of the system, that is expressions for STILL and NOT YET while ALREADY and NO LONGER would be frequently missing, cf (2a). Van der Auwera offers an area-based study of 45 European languages which results in the postulation of an Euroversal which states that if a language has only one PhP expression, it will be NO LONGER, followed by STILL/NOT YET and finally ALREADY, cf (2b). Löffgren (2019) works on a sample of 54 Bantu languages. She postulates a Bantu-versal where expressions for NOT YET top the accessibility hierarchy, cf. (2c).

In this exploratory paper we use grammars to obtain data from 97 Austronesian languages well spread out across the family. We check for the encoding of PhP concepts, including their lexico-grammatical status. When feasible, we also record their diachronic sources.

Some of our findings include the following: (i) The accessibility hierarchy for Austronesian follows the one suggested for the world by van Baar in that concepts related to continuity come on top, cf. (2d). However, ALREADY shows a very high frequency as well while NO LONGER is missing in half the languages of the sample. (ii) STILL and ALREADY are predominantly expressed by single words and are often part of grammatical oppositions in individual languages. NOT YET is most often expressed periphrastically. Furthermore, it is possible to postulate a lexico-grammatical cline for this concept such as periphrastic expressions > transparent univerbations > semi-transparent univerbations > single morphemes. Mono-morphemic expressions of NOT YET form an almost coherent area stretching from Sumatra, Java, Bali, Flores and a few island groups to the East. In addition to outlining their synchronic distribution, we also use these data for formulating hypotheses about the maturation of NOT YET expressions as a grammatical type.

Examples

(1) Phasal adverbials in English, following van Baar (1997: 1)
   a. Peter is already in London: an (expected) situation has come to exist
   b. Peter is no longer in London: a situation has ceased to exist
   c. Peter is still in London: a situation persists
   d. Peter is not yet in London: a situation has not come into existence despite expectations that it would

(2) Accessibility hierarchies for PhP expressions
   a. World-wide sample (van Baar 1997 : 132-3) STILL/NOT YET > ALREADY > NO LONGER
   b. Euroversal (van der Auwera 1998 : 37) NO LONGER > STILL/NOT YET > ALREADY
   c. Bantu-versal (Löffgren 2019: 16) NOT YET > STILL > ALREADY > NO LONGER
   d. Austronesian (the current work) STILL > NOT YET/ALREADY > NO LONGER

References

Abstract for thematic session 2: The interface between language documentation and experimental methodologies

Eline Visser, Lund University - What is useful data for grammar description?

Many of us have built or are building a corpus with annotated video recordings that aims at representing the language and culture we work on, a point often stressed in literature on language documentation and description. However, much data gathering on understudied languages is also done with the publication of descriptive materials in mind. Which video-recorded data is useful for the description of grammar? I made an overview of the origin of the 700 examples in my grammar (in preparation), and calculated correlations between type of recording (conversation, narrative, stimulus-based), year of recording, level of clarity of speech versus the number of cited examples per thousands words for each recording.

A full analysis with correlations will be presented in June. A pilot study of my data shows that while clarity and year of recording play an important role (the older a recording is, the more often it has appeared in my searches for an example, and the more rounds of transcription it has undergone), the most-cited texts are typically conversations or narratives. For conversations, these are always between people that know each other well and can achieve free-flowing conversation with ease. For narratives, the most-cited texts are those where the speaker masters their topic of narration. For example, stories that I urged unexperienced storytellers to record with me are much less-cited than stories that speakers volunteered to tell. Stimulus-based recordings such as director-matcher tasks and the Family Problems Picture Matching task cover the middle ground. While typically not doing what they were designed for (the man and tree task did not elicit many spatial reference vocabulary), they are often well-transcribed, and the discussion between speakers provides for a variety of interesting constructions.

Recordings of people traveling and working recorded with a chest-mounted action camera, though spatially, ethnobiologically and culturally very interesting, proved largely useless for language description due to poor sound quality.

I argue that recording free conversations between friends, as well as narration on topics that the speakers master and feel comfortable narrating, work the best for language description. Coincidentally, these are also genres likely to be of interest for future generations of linguistic researchers.
A multiple methods approach to lexical flexibility in Tongan

(Svenja Völkel)

As characteristic for Oceanic languages, Tongan shows a high degree of lexical flexibility, i.e., lexemes can function as (noun, verb, adjective, and/or adverb) without any morphosyntactic adaptation (e.g., van Lier & Rijkhoff 2013). For illustration, consider the following example from Tongan (Völkel 2017:453): depending on the syntactic context, *faka'ofo'ofa* ‘beautiful’ functions as adjective (modifier within a referential phrase) in (1a), as adverb (modifier within a predicate phrase) in (1b), as verb (head of a predicate phrase) in (1c) and as noun (head of a referential phrase) in (1d).

(1) a. *ha fefine faka’ofo’ofa*
   ART woman beautiful
   ‘a beautiful woman’

   b. ‘*Oku hiva faka’ofo’ofa ‘a e kui fefine.*
   PRS sing beautiful ABS ART grandparent female
   ‘The grandmother sings beautifully.’

   c. ‘*Oku faka’ofo’ofa ‘a e fefine.*
   PRS beautiful ABS ART woman
   ‘The woman is beautiful.’

   d. *hono faka’ofo’ofa*
   O.POSS.3SG beautiful/beauty
   ‘his/her/its beauty’

Based on such examples, typological studies on lexical flexibility (e.g., Hengeveld 1992:66) have classified Tongan as highly flexible with ‘contentives’, i.e., lexemes that can occur in all 4 functions. Broschart’s famous article in which he describes Tongan as a language without word class distinction on a lexical level (Broschart 1997), has also contributed to this evaluation. Indeed, there are numerous highly flexible lexemes, however, by far not all Tongan lexemes show this high level of flexibility. Thus, based on other approaches such as Evans & Osada (2005: 378), who argue that “it is not sufficient to find a few choice examples which suggest word class flexibility [... but] equivalent statements need to hold for all relevant words in the lexicon”, Tongan would not fulfil the criteria for lexical flexibility. These extremely different evaluations of Tongan suggest that flexibility seems to be a matter of gradation. So, how flexible is Tongan?

This paper will illustrate aspects of flexibility in Tongan (Völkel 2013, including the distinction between lexical and grammatical flexibility as proposed by François 2013) and present an empirical approach which has been developed to study the degree of flexibility in more detail. Taking Croft’s prototype framework as a basis (Croft 2000:89), this project investigates the usability of lexemes, their actual usage in terms of frequency, and the cognitive prominence of linked syntactic frames. For measuring these different dimensions of flexibility, a *multiple methods approach* is needed, more precisely, a combination of descriptive fieldwork (elicitation in terms of systematic survey data), documentary fieldwork (building a corpus for natural language analysis), and experimental fieldwork (cognitive tasks, such as a cued sentence production task).

References:


The diversity of extended person name markers in Austronesian and Mek languages

Bernhard Wälchli (Stockholm University)

Many Austronesian languages have special markers for person names as opposed to common names, such as Tagalog pivot/topic *si*, non-pivot *ni*, oblique *kay* vs. common noun pivot/topic */an/, non-pivot */nay/, oblique *sa*. These go under various names including “personal article”, “proper article” and “name determiner”. In a world-wide perspective, person name markers are strongly overrepresented in Austronesian. A specific feature of Austronesian (including contact languages, such as Sulka) is that person name markers can occur in the personal interrogative ‘who?’ and/or in personal pronouns. In many Austronesian languages, person name markers can be extended to at least some kinship terms, sometimes only ‘father’ and ‘mother’, sometimes only to kin of speech act participants. This talk focuses on languages where person name markers are extended to common noun use beyond occasional use with kinship terms and where the extension notably includes grammatical functions such as topic singular (Bolinao), gender (Tinputz, Teop; Owa; Nakanai) and/or definiteness (Ula; Sursurunga; Tinputz, Teop). In addition to Austronesian languages, even the Mek languages (usually considered to be a branch of Trans New Guinea) are considered, where postposed extended person name markers have given rise to a gender system in Nalca (NW-Mek).

A major descriptive result of the investigation is that the extension of person name markers highly differs from language to language. Even very closely related languages, such as Tinputz and Teop, exhibit major differences. In Tinputz, but not in Teop, agentive nouns are included. In Teop, but not in Tinputz inanimate nouns expressing “a piece/bit/slice of” are included. Using data from reference grammars, grammar sketches, dictionaries, original texts and in particular translations of the New Testament, the differences in use of extended person name markers in five groups of closely related languages are described: 1. Bolinao vs. other Sambalic languages of Central Luzon (Tinà Sambal, Botolan Sambal, Abenlen Ayta), 2. Tinputz and Teop vs. Saposa (North Bougainville Oceanic), 3. Owa and Kahua vs. Arosi (Island of Makira), 4. Nakanai vs. Bola (Kimbe, New Britain), 4. Nalca vs. Yale, Nipsan, Eipo, Una and Kimyal (Mek).

In a second step, the diachrony of the extension is reconstructed in each of the five groups to the extent this is possible. It is found that none of the developments can be explained by a single process, such as grammaticalization. Grammaticalization occasionally plays a role as one of many processes involved, as, for instance, when Nalca innovates a feminine gender marker *ge* (< *gel* ‘woman’) after the erstwhile unisex marker *bi* has been narrowed to male reference in Western Mek due to frequent collocation with male nominalizations suffixed by -nye/nya (which are frequent, among other things, in indigenous person names). Furthermore, changes are not restricted to grammar. Sound changes, such as dissimilation, can play an important role, and often extensions are the result of an interaction with specific lexical items or specific derivational patterns. A further finding is that, while the extension of person name markers can entail considerable increase of complexity, individual diachronic changes within chains of diachronic developments typically are simplifications. Their effect is to do away with some anomaly in the stage immediately preceding the language change, which, however, may induce new anomalies.

While the grammaticalization approach emphasizes the importance of cross-linguistically recurrent grammaticalization paths and independent (unconditioned) diachronic change in grammar, what we find in extended person name markers in Austronesian and in Mek (and also elsewhere in the world) is conditioned grammatical change (one change triggers other changes) and highly language-specific developments. It is argued that it is the diversity of diachronic developments and especially their interplay that account for the considerable cross-linguistic diversity between even closely related languages.
Gender, classifiers, alienability, animacy:

Concurrent categorisation systems in the Bird’s Head

Katherine Walker, University of Amsterdam

Contrary to earlier assumptions, gender and classifier systems have been found to co-occur (Fedden & Corbett 2017). Several Papuan languages in the Bird’s Head have the typologically unusual property of having two orthogonal systems that can be overtly marked in numeral constructions. For example, in (1) (Moskona, East Bird’s Head) CL(ASSIFIER)1 is a non-obligatory independent classifier for humans, while CL2 is an obligatory bound morpheme used with semantically varied nouns, but only on numerals one to three (Gravelle 2010: 171).

(1) MOSKONA

\[ \text{Eri-orna ibah y-erg-ak} \]
\[ 3\text{PL-man} \quad \text{CL}1.\text{human} \quad \text{DU-CL}2.\text{generic-two} \]
\[ \text{‘two men’} \quad \text{(Gravelle 2010: 281)} \]

However, the noun in (1) is also categorised in other ways: only human nouns trigger pronominal marking on the classifier-numeral complex (y- ‘DUAL’), and the root-initial non-high vowel of orna ‘man’ indicates that it is inalienable (Gravelle 2010: 8). Viewed as types of nominal classification systems, animacy splits and alienability systems share properties both with canonical gender – obligatoriness, categorisation of all nouns in the lexicon, fixed class membership – and with classifier systems, in which categorisation is not marked through agreement. As such, Moskona can be described as having four concurrent nominal classification systems with distinct formal and semantic properties.

This paper presents an analysis of the morphosyntactic properties of, and interactions between, the four categorisation systems – gender, numeral classifiers, animacy, alienability – attested in numeral constructions in 15 Papuan languages of the Bird’s Head. That is, individual system types are compared, before an analysis is given of the language-specific restrictions on which system types can, cannot, or must co-occur, and in which contexts. Fourteen of 15 languages in the study have at least two concurrent systems, but the functioning of individual systems and the restrictions governing their (co-)occurrence vary even between very closely related languages.

References


Sundaland: What is un-Austronesian about this area and why?
Tobias Weber
University of Kiel

Sundaland is a geographical area consisting of the Malay Peninsula, Sumatra, Borneo, Java, Bali and their surrounding islands. The present paper argues that although this area is dominated by Austronesian (AN) languages in a similar way to the Philippines, it has a past that is different from the Philippines in both linguistic and extra-linguistic terms, and therefore these differences are expected to be reflected by different structural properties.

To establish in what way the AN languages of Sundaland differ from the more conservative AN languages of the Philippines, WALS features and some additional typological features not mentioned in the WALS were examined. The paper discusses the following feature values, which are found in Sundaland with a significantly higher frequency than in the Philippines:

1. vowel inventories consisting of more than five vowels
2. nasal consonant inventories that include a palatal nasal
3. coronal asymmetry, i.e. dental /t̪/ and alveolar /d/
4. sesquisyllabicity
5. numeral classifiers
6. absence of core argument flagging
7. verbal agreement
8. simplified voice systems consisting of maximally two voices
9. absence of obligatory TAM marking
10. verb-medial word order
11. complex honorificity systems

The following explanations are offered for these feature values:

- Most of these values are also found in Mainland Southeast Asia (MSEA), and some are also found in Wallacea, which is to the east of Sundaland. MSEA, Sundaland and Wallacea form what Gil (2015) calls the ‘Mekong-Mamberamo linguistic area’, suggesting that the common values are the result of some substrate influence of pre-AN languages. However, I suggest that this explanation does not hold for the following feature values:
  - (2): the palatal nasal is also reconstructed for Proto-Malayo-Polynesian (cf. Ross 1995) and was thus lost in the languages of the Philippines but retained in the languages of Sundaland.
  - In the case of (6) and also (8) to (10), simplification processes that occurred during L2 acquisition especially in the large languages of Sundaland such as Malay may offer a better explanation.
  - (11) is best accounted for in terms of the complex structure of the societies that speak the large languages of Sundaland.

- Feature values (3) and (7) are also found further east in Wallacea but are rare in MSEA:
  - (3) might also be explained by substrate influence and is notably still found in some Austroasiatic languages of MSEA.
  - (7) is only found in small languages in non-contiguous remote areas of Sundaland and thus likely evolved by substrate influence of pre-AN languages.

The data are drawn from various typological databases, grammars and other language descriptions.

References


Engagement and information structure in interaction: the intersubjective meanings of two discourse markers in Motu
Stephanie Yam – The Australian National University

One of the fundamental functions of language is the coordination of attention and/or knowledge between speech participants to facilitate the constructive exchange of information. The ways in which speakers attain intersubjective alignment in interaction have been examined by studies of engagement and information structure, both typological categories that are concerned with common ground management and how communicative goals influence the lexicogrammatical form of an utterance. In Motu, an Oceanic language of New Guinea, two discourse markers, \textit{na} and \textit{be}, provide speakers with a grammatical strategy for negotiating the intersubjective relations between speech participants. These optional markers are also used to create communicative effects related to information structural categories in varying interactional contexts.

Due to the difficulty in directly observing speaker intentions, diagnostic contexts are an analytical juggernaut in the study of information structure. In particular, question-answer pairs are widely used to test observed phenomena that appear to encode prototypical features related to topic, focus and contrast. An example of the occurrence of \textit{na} and \textit{be} in such a setting is given in (1).

(1) Q: hanua-mu be dahaka? village-POSS.2SG ADDR.(A)SYM what ‘what is your village?’
A: hanua-gu na Hanuabada village-POSS.1SG SPKR.ASYM Hanuabada ‘my village is Hanuabada’

As illustrated in (1), the markers occur at the right-edge of the NP. Typically, in question-answer contexts, the question word is taken to represent the focus constituent. However, as seen in (1), where \textit{be} marks the nominal referent \textit{hanua-mu} rather than the content interrogative \textit{dahaka}, the use of \textit{na} and \textit{be} in question-response sequences does not align with information structural considerations. Instead, the distribution of these markers in interactional speech indicates distinct underlying semantics more closely related to engagement (see Evans et al., 2018). Within this analysis, \textit{na} and \textit{be} appear to formally encode meanings pertaining to (i) the speaker’s assumptions about their addressee’s mental access to a given referent or proposition, and (ii) the subjective anchoring of the content of the marked nominal referent within the epistemic perspective of either the speaker or addressee.

Additionally, these markers are optional and their (non-)use is based on considerations of pragmatic and interpersonal meaning in spontaneous discourse. In some speech genres, particularly narrative texts, \textit{na} and \textit{be} seem to have a wider role than simply encoding the engagement value of a referent. In these contexts, \textit{na} and \textit{be} have functions with interpretive effects similar to that of topicalisation and/or referent tracking. However, these markers do not fit neatly into established notions such as topic and focus (cf. Matić and Wedgwood, 2013). This paper will outline an analysis of the intersubjective meaning and function of \textit{na} and \textit{be} based on an investigation of their distribution in conversational speech settings and discuss the shared role of engagement and information structure in common ground management in Motu.

Definite Obliques in Yami
Chun-Jan Young
University of California, Santa Barbara

It is widely established in Philippine linguistics that core arguments are typically definite. This is accomplished through marking a definite argument using the Nominative (or Absolutive) Determiner: well-known examples are ang in Tagalog and Cebuano, ing in Kapampangan, and ti in Ilokano. However, it has also been observed in spoken data that definite, identifiable arguments are not always coded by speakers as Nominative/Absolutive (Mithun 1994).

I examine the coding of definite and identifiable referents as non-core arguments in Yami, a Batanic language spoken on Orchid Island, Taiwan. Unlike many Philippine languages, Yami exhibits a four-way distinction in its Determiner system, with Nominative, Genitive, Locative, and Oblique markers (Reid & Liao 2004; Rau & Dong 2018). This study focuses on the Oblique Determiner so, which is usually described as marking indefinite nouns as non-core arguments. I present examples from unscripted spoken data in which so is instead found to precede definite and identifiable arguments, challenging the characterization that so should only mark indefinites.

I show that definite arguments in Yami may be marked using Oblique so due to factors arising from not only discourse, but also syntax. Similarly to Mithun’s (1994) findings for Kapampangan, definite arguments in Yami which are peripheral to the current event in discourse may be expressed as Obliques—for example, when serving to describe the actions of another, more attention-worthy argument. However, it is more common for definite Obliques to occur due to the particular syntactic requirements of the nominalizing prefix ka-, which is polysemous and multifunctional. When a transitive verb is marked by ka-, the argument that would normally take Nominative marking is instead thrust into obligatory Oblique marking, as Nominative marking is not permitted within ka-marked verbal clauses. Representative examples are given below.

This study shows that Oblique marking in Yami may be used to mark definite referents under certain conditions. This observation lays the groundwork for a systematic, typological study of how (in)definiteness maps onto the various Determiner systems across Philippine languages.

(1) o mapalolo so baka dang a mehakay am...
NOM AV. bring OBL cow LOC.DEM.MED LK male TOP
‘That man pulling the cow, …’ (the cow appears only as a modifier for the man)

(2) ka-kozong da rana so asi no kayo do yala ori.
NMZ-pack 3PL.GEN already OBL fruit GEN tree LOC basket DEM.MED
‘And they packed the fruit back into that basket.’

References: