

Towards a Typology of Austronesian Pronominal Clisis

Loren Billings & Daniel Kaufman
National Chi Nan University Cornell University

This paper reports ongoing work on the following aspects of Austronesian clisis: the relative ordering of multiple clitics, the position of the clitic cluster.

Clitic phenomena have captured a great deal of attention because they constitute a primary testing ground for the various conceptions of modularity in the architecture of Universal Grammar. Specifically, the linearization of clitics in languages such as South Slavic (Rudin et al. 1999) and Chamorro (Chung 2003) requires access to prosodic phrasing. Clitics in languages such as these therefore represent the most clear-cut cases of movement in PF. However, unlike other such cases, morphosemantic features are still required at this level to determine the relative order of multiple clitics within the cluster. Capturing precisely this interplay between prosody and morphosyntax has been one of the major challenges in work on the subject. The variation among Austronesian languages as to cluster-internal ordering is shown to represent different compromises between morphosemantic and prosodic requirements.

An area that has not received much attention at all in the theoretical Austronesianist literature is the morphosyntax of verb-adjacent clisis. We discuss in this regard several languages of Central Sulawesi, focusing on pre- and post-verbal alternation. A paradigm-based Optimality-theoretic analysis is offered to account for the distribution of pre-verbal positioning in these languages. In addition, we offer observations concerning portmanteaux, co-occurrence restrictions, clitic doubling, and non-pronominal clitics.

An area which has been explored less in the literature is the boundary between verb-adjacent clisis and person agreement. This is, in part, a problem of definitions, as there are still residual problems with the difference between affix- and clitic-hood (despite the good progress in this direction). We therefore seek a robust distinction between verb-adjacent clitics and verbal-agreement affixes based on Austronesian data. In this regard, the variation among the languages of Sulawesi, Indonesia, appears to be highly promising.

1. Internal ordering

Like normal syntactic constituents, clitic pronouns can be ordered based on morphosemantic factors. These include morphological case, semantic role, or ϕ -features. Unlike normal syntax, their order can also be based on prosody.

1.1 Orderings based on morphological case or semantic roles

It is often difficult to determine whether it is morphological case or semantic roles which determines order. This is because in many Austronesian languages, multiple pronominal clitics appear in only a limited number of voice configurations because pronominal (more generally, definite) patients cannot be realized in the GEN case. Generally, in non-Actor voice constructions, the Actor is realized in the GEN. As such, it is impossible to tease case apart from roles. In at least one language that we have come across, however, it is possible to show that the semantic role and not case is the determining factor.

In a number of languages, a GEN (Actor) precedes a NOM (non-Actor) pronoun. The Central Luzon subgroup (Kapampangan, Sambal, Bolinao, and Aita) is categorical in this respect. In Central Philippines, Mamanwa is likewise categorical, but other languages have a mere preference as long as more highly ranked prosodic constraints are satisfied (e.g., Tagalog and Bikol).

Other languages position the NOM (non-Actor) in front of the GEN (Actor). Although we know of no languages that require this ordering, some of them appear to favor this order if other factors (such as prosody or person) are controlled for: Agusan Manobo (Weaver and Weaver 1964), Sarangani Manobo (Dubois 1976: 48–50), and Mayrinax Atayal (Huang 1995: 28–36).

We know of no language where case and crucially not semantic roles determines the order of two *clitic* pronouns. If such a language exists, it would not be a problem, because the pronouns would be ordered using normal syntax.

The only language of this kind we have found in which case and roles can be teased apart is Isbukun Bunun, a language that seems to order the Actor before the other pronoun regardless of case (NOM vs. OBL) or the verb's voice:

- (1) a. Masaiv *-ik* *-su* tasa ahil. (Isbukun Bunun)
 give.AV 1SG.NOM 2SG.OBL one book
 'I gave you a book.' [Huang et al. 1999: 186]
- b. Tahun *-ku* *-as* bunun tu halinga. (Isbukun Bunun)
 teach.NAV 1SG.OBL 2SG.NOM Bunun LNK word
 'I taught you the Bunun language.' [Huang et al. 1999: 187]

Though (1a–b) suggest that grammatical person may be the factor, Huang et al. (1999: 188) add that person (as well as case or number) do not affect ordering.

It is not clear how these Isbukun Bunun data can be handled by prevailing theories, where the syntax tends to have access to a nominal expression's case but not its semantic role as such. Assuming that the syntax precedes the morphological component derivationally, it would be even harder for a morphological theory of clitic ordering to have access to the pronouns' semantic roles. Clearly, this phenomenon deserves further investigation.

1.2 Orderings based on ϕ -features

Unlike case or semantic role, ϕ -features do not change to reflect the predicate's voice. The features of this kind relevant to Austronesian are person and number. (The languages we are aware of do not exhibit pronominal genders.)

Quite a number of languages utilize person in ordering clitic pronouns. It is interesting to note in this regard that all three descriptions of Manobo and Atayal languages mentioned above, in which a NOM (non-Actor) precedes a GEN (Actor), point out explicitly an overriding ordering constraint requiring a third-person clitic pronoun to follow a first- or second-person clitic pronoun.

At least two accounts of the Squiliq dialect of Atayal argue additionally for person to be used to order first- and second-person pronouns relative to each other. Using different combinations of such pronouns, Rau (1992: 146–147) argues for a first-second order, whereas Huang (1995: 34–35) proposes the opposite order. Liao (2004) shows convincingly that prosody rather than person decides the ordering of such clusters. Indeed, if person is used as a factor, the only conclusive examples in our view are of a third-person clitic pronoun having to follow either a first- or a second-person clitic pronoun.

Nor does number appear to be a clear factor in ordering clitic pronouns. Although Huang (1995: 28–36) proposes number as a factor in ordering clitic pronouns for Mayrinax Atayal, prosody is just as valid an explanation.

To summarize, of all the ϕ -feature factors, the only one that seems to matter is person (and only third person as opposed to first or second person). This factor is also noteworthy in that it frequently interacts with both NOM-before-GEN and prosody-based systems. Number does not seem to be crucial.

1.3 Orderings based on prosody

In several Austronesian languages, linguists have argued that number of syllables is the factor determining the order of two clitic pronouns. In each of these languages, if prosody is a factor, the order is light before heavy.

As has been widely publicized for Tagalog (e.g., Schachter 1973), monosyllabic clitic pronouns precede disyllabic ones. Other languages—all from Central Philippine—such as Bikol and Cebuano (Billings and Konopasky 2002) and Tagakaulo, Kaagan, Mansaka, and Davawenyo (Lee 2004), use the number of syllables to order clitic pronouns within the cluster. As mentioned above, Liao (2004) makes the same argument about Squliq Atayal, and these ideas could be extrapolated to Huang's (1995) Mayrinax Atayal data as well.

In some of these languages the prosodic constraint is categorical; only if the pronouns have the same number of syllables do morphosemantic factors emerge. Tagalog, Bikol, Mansaka, and Davawenyo are languages of this kind. All the aforementioned accounts of Atayal varieties—Rau (1992), Huang (1995), and Liao (2004)—argue for a morphosemantic constraint requiring a third-person clitic pronoun to follow first- or second-person one. Liao further argues that the prosodic factor emerges in clusters of first- and second-person clitic pronouns. Tagakaulo, heavily influenced by Sarangani Manobo, also shows a strong preference to position third-person forms last.

1.4 Generalizations about cluster-internal ordering

To summarize the factors that influence the order of two clitic pronouns in a cluster, several types emerge. In each of these, there is just one possible order.

First are the languages that order a GEN-case Actor before a NOM-case non-Actor: the Central Luzon group and Mamanwa. Isbukun Bunun is similar in ordering an Actor before the other role, regardless of voice/case. We assume as a hypothesis that the crucial ordering factor in the other languages as well is the semantic role and not case. Our main reason for developing this notion is that each of the factors—role, case, person, and even prosody—have a single order; with regard to semantic roles, it is the Actor that precedes any other role.

Next are the languages with the opposite order: NOM-case non-Actor preceding GEN-case Actor. On strictly conceptual grounds, we assume that case is the relevant factor here, with the subject preceding non-subject cases. Our hypothesis predicts that if case is the factor, then NOM is leftmost.

Similarly, it is quite common cross-linguistically for third-person pronouns to follow any others in the cluster; see, for example Rudin et al. (1999) regarding Slavic. As with the preceding two factors, if person is the relevant factor, then a third-person pronoun is predicted to follow any other.

Finally, if prosody is the relevant factor, then our hypothesis predicts that a lighter pronoun precedes a heavier one, not vice versa. Billings and Konopasky (2003) offer a rationale for this ordering; only a monosyllabic clitic

can adjoin prosodically to a preceding word in Tagalog. It remains to be seen how this rationale could be extended to languages (e.g., Tagakaulo) in which trisyllabic clitic pronouns behave differently from disyllabic ones.

So far, we have compiled four necessarily pre-theoretic constraints on cluster-internal ordering. These represent a hypothesis in and of themselves. We leave it to future work to determine if the languages under discussion can continue to be accounted for by alternative rankings of these constraints.

2. External ordering

As the preceding discussion of particles has shown, there are two basic positions for the clitics relative to the rest of the clause. We give only a brief overview of Wackernagel systems in order to concentrate on the verb-adjacent systems, which have attracted less attention in the Austronesianist literature.

2.1 Wackernagel systems

The vast majority of Philippine languages instantiate Wackernagel clisis: following some initial element. Several restrictions on clitic movement are common. Initial topics do not constitute the initial element, in (2a), and initial OBL-case foci generally behave as impermeable constituents for clitics, in (2b):

- (2) a. [Buhat (**siya*) sa= Maynila =ay] (**siya*) mag-la-lakad *siya*. (Tagalog)
 from OBL Manila TOP AV-IRR-walk 3SG.NOM
 ‘He will walk [_{Topic} from Manila].’
- b. [Buhat (**siya*) sa= Maynila] *siya* mag-la-lakad. (Tagalog)
 from OBL Manila 3SG.NOM AV-IRR-walk
 ‘He will walk [_{Focus} from Manila].’
- c. Buhat *siya* sa= Maynila. (Tagalog)
 from 3SG.NOM OBL Manila
 ‘She/He is from Manila.’ [(b–c) from Schachter and Otnes 1972: 190]

Another common feature is that the Case Phrase (KP) represents an island for clitic movement. Exemplifying again with Tagalog, clitics can never precede the *ang* ‘NOM’, *ng* ‘GEN’ or *sa* ‘OBL’ phrases they are associated with.

- (3) a. Hindi (**ko*) iyon [_{KP} ang= problema *ko*]. (Tagalog)
 NEG 1SG.GEN that NOM problem 1SG.GEN
 ‘That’s not my problem.’ (= ‘That’s not [the problem I have].’)
- b. Hindi *ko* iyon problema. (Tagalog)
 NEG 1SG.GEN that problem
 ‘That’s not my problem.’ (= ‘That’s not a problem to me.’)

Clitic climbing from a lower clause raise to a higher one, is also rarely attested in Philippine languages; (4a–b) are typical examples. However, modals and serial verbs generally do not block clitic movement, as (5a–b) both show.

- (4) a. Hindi *ko* (**siya*) s<in>abi [na taksil *siya*]. (Tagalog)
 NEG 1SG.GEN 3SG.NOM PV.PRF-say COMP traitor 3SG.NOM
 ‘I didn’t say he’s a traitor.’

- b. Achi *-na* aforot-on [ay fiakon *-ak* om-ali]. (E. Bontoc)
 NEG -3SG.GEN believe-PV COMP NEG -1SG.NOM AV-come
 'He does not believe that I am not coming.' [Fukuda 1997: 44]
- (5) a. Hindi *ka* dapat (**ka*) um-alis (**ka*). (Tagalog)
 NEG 2SG.NOM should 2SG.NOM AV-leave 2SG.NOM
 'You shouldn't leave.'
- b. Kaman-*ak* om-oy ay ma-rqos. (E. Bontoc)
 seems 1SG.NOM AV-go LNK STA-pass
 'Seems like I go passing by.' [Fukuda 1997: 48]

What qualifies as the initial element is subject to minor variation. In a number of languages, the constituent relevant for hosting Wackernagel clitics is the PrWd, as exemplified by the following sentences, from Tagalog.

- (6) a. L<um>utas *na ako* ng= isa -ng suliranin. (Tagalog)
 AV.PRF-solve PRF 1SG.NOM GEN one LNK problem
 'I already solved one problem.'
- b. Hindi *ka* *ba* l<um>utas ng= bigla -ng suliranin? (Tagalog)
 NEG 2SG.NOM Q AV.PRF-solve GEN sudden LNK problem
 'Haven't you ever solved a sudden problem?'

The NOM-case pronominal set may also appear in clause-initial position in cleft-like constructions and topicalizations. However, for the 2SG pronoun (whose clitic version *ka* is the only monosyllabic form in the NOM paradigm), there exists a non-clitic disyllabic form *ikaw*. Compare examples (7a–b):

- (7) a. *Ako* ang= guro. (Tagalog)
 1SG.NOM NOM teacher
 'I am the teacher.'
- b. *Ikaw* ang= estudyante. (Tagalog)
 2SG.NOM NOM student
 'You are the student.' (cf. **Ka* ang estudyante.)

The 2SG form is thus the only formal diagnostic to distinguish between the nearly identical clitic and nonclitic paradigms of the NOM personal pronoun.

Languages differ with regard to what constituent can serve as the initial element. All the languages of the Philippines that we have encountered treat negation, fronted adverbs, and fronted foci as a potential first element:

- (8) a. Hindi *ka* mag-lu-luto ng= itlog. (Tagalog)
 NEG 2SG.NOM AV-IRR-cook GEN egg
 'You won't cook (the) eggs.'
- b. Ya' *ku* pilay pa luma' bi batna'a. (Mapun)
 NEG 1SG.NOM go DIR house 2SG.GEN now
 'I'm not going to your house now.' [Collins et al. 2001: 591]

Languages differ as to whether complementizers may serve this purpose. As seen from (9a–b), Tagalog does not allow this but Mapun (Sama Bajaw) does:

- (9) a. Kung hindi *ka* mag-lu-luto ng= itlog, ... (Tagalog)
 if NEG 2SG.NOM AV-IRR-cook GEN egg
 'If you didn't cook (the) eggs, ...'
- b. Bong *ko* ya' pa-tagong nangis, ... (Mapun)
 if 2SG.NOM NEG AV-stop AV.cry
 'If you don't stop crying, ...' [Collins et al. 2001: 115]

2.2 Verb-adjacent systems

Moving southwards from the Philippines, we find a major change in clitic placement upon reaching Central Sulawesi. This difference accompanies a distinct kind of verbal paradigm for the Actor voice/Patient voice distinction. Because this represents an important divergence in the linearization of clitics, we will analyze clitic behavior in Kulawi, a representative Kaili-Pamona language, and offer an Optimality-theoretic account of the attested variation. (The case labels in the rest of section 2 are synchronically better analyzed as ERG and ABS. The more neutral labels of GEN and NOM, resp., are used in order to maintain the historical linkage between the paradigms across Austronesian.)

Like most Kaili-Pamona languages (Mead 2002), Kulawi possesses only two diatheses: Actor voice and Patient voice. There are only two moods: realis and irrealis, indicated in the actor voice by *nang-* and *mang-*, respectively. The patient voice, by contrast, is marked overtly only in the realis, by the *i-* prefix (cognate to the *-in-* affix commonly found in the Philippines and Taiwan). The irrealis aspect of the patient voice is marked by the pre-verbal attachment of a pronominal Actor or by *ra-* (3PL.GEN) if the Actor is suppressed. The relevant paradigms can be seen in (10).

(10) Kulawi verbal paradigms (partial) [Adriani and Esser (1939)]

| | ϕ -features | Realis | Irrealis |
|-------------------------------|------------------|-----------------------|-----------------------|
| Actor voice (intransitive) | 1SG | <i>nang-STEM-a</i> | <i>mang-STEM-a</i> |
| | 2SG | <i>nang-STEM-ko</i> | <i>mang-STEM-ko</i> |
| | 3SG | <i>nang-STEM-i</i> | <i>mang-STEM-i</i> |
| | 1PL.IN | <i>nang-STEM-ta</i> | <i>mang-STEM-ta</i> |
| | 1PL.EX | <i>nang-STEM-kami</i> | <i>mang-STEM-kami</i> |
| | 2PL | <i>nang-STEM-komi</i> | <i>mang-STEM-komi</i> |
| | 3PL | <i>nang-STEM-ra</i> | <i>mang-STEM-ra</i> |
| Patient voice (transitive) | 1SG | <i>i-STEM-ku</i> | <i>ku-STEM</i> |
| | 2SG | <i>i-STEM-mu</i> | <i>mu/nu-STEM</i> |
| | 3SG | <i>i-STEM-na</i> | <i>na-STEM</i> |
| | 1PL.IN | <i>i-STEM-ta</i> | <i>ta-STEM</i> |
| | 1PL.EX | <i>i-STEM-kami</i> | <i>ki-STEM</i> |
| | 2PL | <i>i-STEM-mi</i> | <i>mi-STEM</i> |
| | 3PL | <i>i-STEM-ra</i> | <i>ra-STEM</i> |

In the Patient-voice realis, we see that the *i-* prefix marks the mood while a pronominal Actor is expressed as a verb-adjacent GEN enclitic. The corresponding irrealis, on the other hand, is signalled solely through the proclitic attachment of the pronominal. What is generally found in the Kaili-Pamona subgroup is that the NOM pronominals are expressed as either free

pronouns or bound forms. (Free pronouns may be ordered more or less freely within the sentence, whereas the bound sets appear to be verb-adjacent.)

Unlike the Wackernagel positioning observed above in the Philippine languages, the GEN pronominals with verbs prefixed by (n)i- do not encliticize to negation or any preverbal material in Kulawi and related languages. (In the Da'a examples, it's not clear whether *ni* should be analyzed as a PV prefix.)

- (11) Position of the GEN pronominals: invariably after the verb
- a. Moma i-_____ epe -ku ka= rata -mu. (Kulawi)
 NEG PV.REAL hear 1SG.GEN NOM come 2SG.GEN
 'I haven't heard about your arrival.' [Adriani and Esser 1939: 30]
- b. Da'a ninjani -ku. (Da'a)
 NEG know 1SG.GEN
 'I don't know.' [Abas and Friberg 1989: 21]
- c. Wei-wei ninjani -ku -mo. (Da'a)
 now know 1SG.GEN already
 'Now I already know (it).' [Abas and Friberg 1989: 21]

By contrast, NOM clitics may encliticize to preverbal elements such as negation, in (12a), and adverbials, (12b), as well as certain post-verbal elements, (12c):

- (12) Position of the NOM pronominals: Wackernagel clitics
- a. Moma -'a t<um>ai. (Kulawi)
 NEG 1SG.NOM AV-come.here
 'I didn't come here.'
- b. Iwengi -'a no-_____ towa kau i= bone. (Kulawi)
 yesterday 1SG.NOM AV.REAL chop tree OBL field
 'Yesterday, I chopped down a tree in the field.'
- c. Ba i-_____ tudu tina -mu -da -ko? (Kulawi)
 Q PV.REAL send mother-2s.GEN -EMPH -2SG.NOM
 'Are you sent by your mother?' [all from Adriani and Esser 1939: 30]

Although a full treatment of NOM clitics cannot be given here, it appears that postverbal positions such as the one in (12c) may be a result of the syntactic integrity of the GEN argument and the verb. Such sensitivity to syntactic constituency is expected if the NOM forms are phrasal clitics (Billings 2002).

2.3 Formalizing the transition from Wackernagel to verb-adjacent clisis

As is evident from this paper so far, languages of Taiwan and the Philippines tend toward the Wackernagel type of clisis, whereas in Sulawesi we begin to see verb-adjacent systems. Because the verb is usually initial or right after an initial element (such as NEG), the Wackernagel position often coincides with one side or the other of the verb, making it difficult for linguists to determine with certainty whether the language has a Wackernagel or a verb-adjacent clisis. To complicate matters, in some languages, once the crucial environment is found, speakers will often accept more than one clitic position, as in (13):

- (13) Kailan (ka) sa= Maynila (ka) p<um>unta? (Tagalog)
 when 2SG.NOM OBL Manila 2SG.NOM AV.PRF-go
 'When did you go to Manila?' [NB: only one *ka* clitic or the other]

Billings and Konopasky (2003) take the position that Tagalog clitic pronouns are verb-adjacent. Anderson (to appear) assumes a Wackernagel analysis.

Returning to the verb-adjacent GEN set in Kulawi and related languages, there has been some debate among Austronesianists as to the genesis of so-called conjugated verbs (van den Berg 1996, Mead 2002, Wolff 1996, Zobel 2002). Wolff offers the following reanalysis scenario (for Malay):

- Stages: I. NEG=pron.GEN PV.verb *Wackernagel enclitic*
 II. NEG pron.GEN=PV.verb *Wackernagel proclitic ~ verb-adjacent clitic*
 III. pron.GEN=PV.verb *verbal prefix*

Though agreeing overall that pronouns shift through stages such as these, we see several problems with such an analysis. We discuss each in turn.

Our first concern is with part of Wolff's characterization of stage II. While we agree that there exists a diachronic stage in which pronouns are verb-adjacent, we dispute the existence of pronouns that are positioned as Wackernagel clitics, following the first element, while affiliating prosodically to the following word. Billings (2002: 59–65), correcting Klavans (1985) and others, argues that Wackernagel clitics can be prosodically affiliated with only the preceding PrWd. Thus, we correct stage II to just "verb-adjacent clitic".

A related problem arises also with stage II. In a clause with multiple clitics preceded by a preverbal non-clitic element, a verb-adjacent pronoun can, under certain conditions, be separated from the verb. In a number of Philippine languages, in structures such as (14), although clitic pronouns are required to follow the verb, adverbial clitics can—under specified conditions that need not concern us here—follow the first PrWd of an initial phrasal constituent.

- (14) [Bukas *ba* ng= gabi] 'y *sa-sayaw sila* ...? (Tagalog)
 tomorrow Q GEN night TOP IRR-dance 3PL.NOM
 'Will they dance ... tomorrow night.' [Schachter and Otnes 1972: 429]

Billings and Konopasky (2003) argue that this is due to an intrinsic difference in the positioning of the two clitic types: the pronouns being verb-adjacent and the adverbials being Wackernagel clitics. If the Wackernagel position happens to be immediately pre-verbal, then it is possible for a (verb-adjacent) pronoun to precede one or more (Wackernagel) adverbial clitics; (6b) is repeated here:

- (15) Hindi *ka* *ba* l<um>utas ng= bigla -ng suliranin? (Tagalog)
 NEG 2SG.NOM Q AV.PRF-solve GEN sudden LNK problem
 'Haven't you ever solved a sudden problem?'

Despite being a verb-adjacent clitic, *ka* (2SG.NOM), as a monosyllabic pronoun, is still prosodically suffixal: [PrWd [PrWd Hindi] *ka*] [PrWd l<um>utas] ... As such, the Wackernagel adverbial clitic *ba* (Q) is placed after the first (matrix) PrWd, resulting in the attested form: [PrWd [PrWd Hindi] *ka*]-*ba* [PrWd l<um>utas] ... Similarly, Rudin et al. (1999: 562–66) discuss a non-pronominal Wackernagel clitic of sorts in Bulgarian that can appear right after a clitic pronoun, which itself ordinarily must be adjacent to the (following) verb. The generalization in Bulgarian is that only other clitics—even Wackernagel ones—can separate a verb-adjacent clitic from its verb. Indeed, Wackernagel clitics can appear

between pronominal clitics in both languages (Billings and Konopasky 2003: 9, Rudin et al. 1999: 564). This concern turns out to be just a red herring. It shows, however, the way these types of clitics can interact in the same clause.

Our final concern with Wolff's stages is that there is no clear connection between the so-called conjugation in (10) and other morphosyntactic features of the languages in question. We address this problem by combining Wolff's general proposal with a suggestion made by Himmelmann (1996): that the morphological asymmetry between the irrealis and realis triggers rebracketing. Specifically, the fact that the patient-voice realis is marked by a prefix while the irrealis is morphologically unmarked creates a mood slot of sorts which may be filled in the irrealis by a pronominal clitic. This can then be linked to the fact that all languages in question have externalized the Proto-Austronesian perfective infix **-in-* to a prefix (*ni-* or *i-*) and employ the bare form of the stem for the irrealis/non-past. This situation, in combination with the presence of Wackernagel clitics, can be seen as a primary factor in the move towards verb-adjacent clisis. Furthermore, following recent work in Optimality Theory on paradigmatic and systemic markedness (McCarthy 2002/2003, Itô and Mester to appear, and Kaufman 2003) we can take these motivations to be present in the synchronic grammar:

- (16) McCarthy's Optimal Paradigms (OP) approach to markedness
- a. A candidate consists of an entire inflectional paradigm.
 - b. Markedness and input-output constraints evaluate all members of the candidate paradigm cumulatively. The violation-marks incurred by each paradigm member are added to those incurred by all the others.
 - c. The stem (shared lexeme) in each paradigm member is in a correspondence relation \mathfrak{R}_{OP} with the stem in every other paradigm member. (That is, for every candidate paradigm P there is a relation \mathfrak{R}_{OP} on $P \times P$.) There is no distinctive base. Rather, every member of a paradigm is a base of sorts with respect to every other member.
 - d. On \mathfrak{R}_{OP} there is a set of output-output faithfulness constraints.

The paradigm which is most harmonic according to both the OP constraints and the regular faithfulness and markedness constraints wins out. Thus, it is not necessary to stipulate a base form/attractor to which all forms in a paradigm are compared. Attractors are epiphenomenal in that the member that can influence other members to satisfy a given OP constraint in the most harmonic way possible will naturally do so. The following constraints are employed:

ANCHOR (STEM, L, PRWD, L) OP: Violated when a stem's left edge coincides with a PrWd's left edge in one paradigm member but not in another.

REALIZE-MORPH: A morpheme in the input must have some phonological exponent in the output (Kurusu 2002).

DEP: Every segment in the output has a corresponding segment in the input.

Given a grammar that contains a REAL prefix but no IRR affix, if ranked sufficiently high, then ANCHOR-OP is satisfied through another means. In Kulawi, this is done by placing the pronominals on the left edge, producing a symmetrically prefixed inflectional paradigm: <REAL-STEM, pron-STEM>.

(17) Tableau

| Realis input: STEM <i>i</i> - PV, REAL | Irrealis input: STEM | ANCHOR -OP | DEP | REALIZE MORPH |
|--|----------------------------|---------------|-----|------------------|
| a. PrWd[<i>i</i> -STEM-pron] | PrWd[STEM-pron] | * | | |
| b. PrWd[<i>i</i> -STEM-pron] | PrWd[<i>i</i> -STEM-pron] | | * | |
| c. PrWd[<i>i</i> -STEM-pron] | PrWd[pron-STEM] | | | * |
| d. PrWd[STEM-pron] | PrWd[STEM-pron] | | | * |
| e. PrWd[STEM-pron] | PrWd[pron-STEM] | * | | * |

The ANCHOR-OP constraint is satisfied either by uniform prefixing, as in candidates (b) and (c), or its uniform absence across an inflectional paradigm, as in (d). As such, candidate (a) is ruled out because the left edge of the stem is aligned with the left edge of the PrWd in one member of the paradigm but not in the other. Candidate (b) is ruled out because the realis prefix is employed without corresponding realis semantics in the input. By contrast, candidate (d) is ruled out because the realis semantics in the input does not correspond to a realis marker in the output. Thus, the winning candidate is one which employs an element whose presence is independently licensed by the input but whose position may be determined by the prosodic constraint ANCHOR-OP.

In other Kaili-Pamona languages, only local (i.e. 1SG or 2SG) pronominals become pre-verbal while the rest remain post-verbal. This is implemented here by positing person-specific constraints to be in a fixed relationship to each other within a sub-hierarchy (cf. Aissen 1999).

Fixed sub-hierarchy: ALIGN CLITIC 3sg > ALIGN CLITIC 2sg > ALIGN CLITIC 1sg

Other constraints may be interpolated between the person constraints, yielding differential treatment of salient persons along a variety of dimensions. For the Central Sulawesi languages under discussion, the relevant dimension is proclitic versus enclitic attachment. The regular enclitic alignment of the GEN clitics can only be violated for those persons whose ALIGN constraint is dominated by the interpolated ANCHOR-OP. In Da'a, ANCHOR-OP is inviolable; when the irrealis verb occurs with a non-local pronominal, the 3pl. clitic is recruited to become proclitic while the argument pronominal remains enclitic. This is analyzed here as morphological epenthesis but an analysis where *ra* is treated as an irrealis morpheme in competition with the local pronominals is also possible. The ranking in (18) derives the relevant Da'a facts.

(18) Da'a constraint ranking:

ALIGN CL 3SG > ANCHOR-OP > DEP *ra* > ALIGN CL 2SG > ALIGN CL 1SG

(19) Bara ku=po-balu ri potomu. (Da'a)
perhaps 1SG.GEN=CAU-buy OBL market
'Perhaps I'll sell it in the market.' [Abas and Friberg 1989: 47]

(20) Nuapa ra=powia ira ri saa? (Da'a)
what *ra*-do 3PL.GEN OBL there
'What are they going to do there?' [Abas and Friberg 1989: 104]

To summarize this section, we have shown some common constraints in the second position systems of the Philippines and some of the major issues regarding verb-adjacent clisis in Sulawesi. Proclisis was shown to be derivable from an ANCHORING constraint which compares irrealis and realis forms and enforces a kind of uniformity between them. This constraint is furthermore tempered by animacy hierarchy effects which determine what clitics are available for procliticization.

3.0 Ancillary issues

This last section sketches several remaining issues. These are portmanteau forms, co-occurrence restrictions, clitic doubling, and non-pronominal clitics.

To begin, quite a number of Austronesian languages have developed fused forms standing in place of certain combinations of pronominal clitics. Most commonly such a portmanteau pronoun represents the combination of a 1SG.GEN and a 2SG.NOM pair, as in (21a). Languages of this kind range from every Atayalic language (all the ones mentioned above plus both major Seediq dialects), every Central Luzon language we have looked at (Kapampangan and Sambalic), nearly every member of Central Philippines, and even languages as far flung as Kimarangang Dusun (Sabahan). The Central Luzon subgroup appears to attest the most such forms. The following are from Kapampangan:

| (21) | Features involved | Clitics in isolation | Portmanteau form |
|------|----------------------|----------------------|------------------|
| a. | 1SG.GEN + 2SG.NOM | ko + ka | daka |
| b. | 1SG.GEN + 2PL.NOM | ko + kamo | dakayú |
| c. | 1SG.GEN + 3SG.NOM | ko + ya | ke |
| d. | 1PL.GEN + 2SG.NOM | mi + ka | daka |
| e. | 1PL.GEN.EX + 2PL.NOM | mi + kayu | dakayu |
| f. | 1DL.GEN.IN + 3SG.NOM | ta + ya | te |
| g. | 2SG.GEN + 3SG.NOM | mo + ya | me |
| h. | 2PL.GEN + 3SG.NOM | moyo + ya | ye |
| i. | 3SG.GEN + 1PL.NOM.IN | na + tá:mu | nakatamu |
| j. | 3SG.GEN + 3SG.NOM | na + ya | ne |
| k. | 3PL.GEN + 1PL.NOM.IN | da + tá:mu | dakatamu |
| l. | 3PL.GEN + 3SG.NOM | da + ya | de |

[Gonzalez 1981: 177, among others]

This is the most extensive list from a single language. In other language groups, most notably Northern Philippines, pronominal forms are fused to the verb (Reid 2001). The primary significance of such forms is what they can tell us about the syntactic history of the languages. For example, at least two Subanon languages (spoken in Western Mindanao), instead of exhibiting a 1SG.GEN + 2SG.NOM portmanteau, as in (21a), attests a 2SG.GEN + 1SG.NOM pair (William Hall and Ryan Galorport, p.c.): Northern and Central Subanon *mau*; cf. Western Subanon *mu akon* '2SG.GEN 1SG.NOM'. In addition, Northern Subanon attests 3SG.GEN + 1SG.NOM *naun* and 2PL.GEN + 1SG.NOM *niyau*. Why this group of languages would have portmanteaux involving the 1SG.NOM rather than the 1SG.GEN may help explain the history of its morphosyntax.

Another type of phenomenon occasionally found in Austronesian clitic systems is co-occurrence restrictions. Here we demonstrate using the E.

Formosan language Kavalan, as reported in Huang et al. (1999: 193–96). Like quite a number of languages (cf. §1.2 above), in Kavalan, the only possible clitic-pronoun cluster is that containing a GEN Actor and a NOM non-Actor (in that order, as in the Central Luzon group and Mamanwa). The verb is typically first with the cluster following directly. Details like these are found in numerous other Austronesian languages. The unique co-occurrence restriction in Kavalan becomes evident when a non-clitic element (NEG or an auxiliary) aside from the verb is initial. In such environments, although it is apparently possible for the two to remain after the verb, one of the two clitic pronouns can appear in front of the verb. The choice of which pronoun moves is based on case and person. Huang et al. (1999: 195) write that “when the two bound pronouns are the first and second person, the Genitive pronoun manifesting the Agent participant may be added to the negator/auxiliary. However, when one of the two bound pronouns is the third person, then the first/second person pronoun gets to attach to the negator/auxiliary.” Co-occurrence restrictions of a different kind in N. Tsou (Tsouic) and Agutaynen (Calamianic, Philippine) have reached our attention (Elizabeth Zeitoun and J. Stephen Quakenbush, p.c., resp.). In these languages, the presence of one pronominal form can cause deletion of a second one or replacement by an OBL pronoun. In N. Tsou there can be no overt cluster. Even if a NOM-case pronoun is possible with a verb, the presence of a GEN-case pronoun displaces the other pronoun. Similarly, in Agutaynen, a first- or second-person NOM-case Patient causes its otherwise-GEN counterpart to appear in the OBL case instead.

One important issue which we are not able to examine here in any depth is that concerning the parameters of clitic doubling. We only mention here two general observations. There appears to be a close connection between verb-adjacent clisis and clitic doubling. We have uncovered only two examples of clitic doubling in second position systems: Kapampangan and Mandar. Sasak (doubled) clitics appear to show some second position effects but the available data is insufficient. In Kapampangan, clitics double the subject and object of non-Actor voice verbs but only the subject of Actor voice verbs. Abstract nouns are also typically not doubled.

- (22) a. E ya masanting ing igu. (both from Kapampangan)
 NEG 3SG.NOM pretty NOM rattan.basket
 ‘The rattan basket is not pretty.’ [Gonzalez 1981:161]
- b. Seli ne nitang tau ing bale.
 buy.PV 3SG.GEN+3sg.NOM that.GEN-LNK man NOM house
 ‘The house was bought by that man.’ [Gonzalez 1981:168]

In Mandar, the Actor of intransitive and transitive verbs is doubled by a second position clitic as in (23). Patient voice verbs in Mandar seem to not require clitic doubling although more information is needed here.

- (23) a. Manarang=i ma’-elong i Murni. (Mandar)
 skilled-3SG.NOM VRB-sing PM Murni
 ‘Murni is good at singing.’ [Abd. Muis Ba’dulu 1990:69]
- b. Andiang=i pole i Ali. (Mandar)
 NEG-3SG.NOM return PM Ali
 ‘Ali isn’t arriving.’ [Abd. Muis Ba’dulu 1990:70]

c. Ma'ande=i loka i Kaco' (Mandar)
 eat-3sg.NOM banana PM Kaco [Abd. Muis Ba'dulu 1990:71 ex.2b]
 'Kaco eats a banana.'

Kulawi offers interesting evidence for the second generalization concerning clitic doubling. In Kulawi only proclitics (appearing in the irrealis patient voice) may double arguments. Enclitics are not indexed with full NPs in the same clause. The extent of this pattern requires more investigation.

One final issue that bears on the typology of Austronesian clitic systems is clitics that are not personal pronouns. These include demonstrative pronouns, which only optionally appear in the clitic cluster in several Central Philippine languages (Lee 2004), and adverbial clitics (Kaufman to appear). Optional clitics tend to appear at the end of the cluster and therefore are not of as much interest as adverbial clitics, which are often positioned according to rather complicated and less than categorical rules. For example, in Tagalog the adverbial clitics—aside from four optional ones (McFarland 2001: 8)—all appear after a monosyllabic clitic pronoun, as in (6b)/(15) above, and before any disyllabic clitic pronoun, as in (6a). If there are both types of pronouns, the adverbial appear in between. (See also the reference above to Kavalan in this regard.) However, as (14) shows, it is occasionally possible for adverbial and pronominal clitics in the same clause not to cluster. How to position adverbial and pronominal clitics together in the same cluster has remained a challenge to generative linguistic theories for decades (Schachter 1973, Billings and Konopasky 2003, and Anderson to appear). Schachter points out that any purely syntactic theory of clitic placement would have trouble handling the portmanteau 1SG.GEN + 2SG.NOM form in Tagalog. Instead of 1SG.GEN *ko* + 2SG.NOM *ka*, the disyllabic form *kita* is used; cf. (21a) above. The problem, as Schachter points out, is that both *ko* and *ka*, being monosyllabic appear before adverbials, but disyllabic *kita* appears after adverbials. Purely syntactic models would likewise have difficulty handling such data. Our database of adverbial clitics is relatively small at this point, so we cannot make claims about distribution. Still, any theory of Austronesian clitics must consider adverbials.

ACKNOWLEDGEMENT: We are grateful for the help of the following colleagues: Robert Blust, Mark Donohue, Ryan Galorport, William Hall, Paul Kroeger, J. Virginia Larson, Celeste Lee, Helen Miller, Gary Persons, Stephen Quakenbush, Elizabeth Zeitoun, and the audience at AFLA-ilf. Despite their valued assistance, only we are responsible for errors in this paper.

References

- Abas, Husen, and Barbara Friberg. 1989. *Tesa-tesa eo-eo ri ara basa Daa*. Palu: Summer Institute of Linguistics.
- Adriani, Nicolaus, and Samuel Esser. 1939. *Koelawische taalstudien*. Vol. 1. *Overzicht der spraakkunst, gesprekken en verhalen met vertaling*. Bandoeng: Nix.
- Aissen, Judith. 1999. Markedness and subject choice in Optimality Theory. *Natural Language and Linguistic Theory* 17. 673-711.

- Anderson, Stephen R. To appear. An a-morphous account of Tagalog second position clitics. *The Nature of the Word: Essays in Honor of Paul Kiparsky*, ed. by Sharon Inkelas and Kristin Hanson. Cambridge, MA: MIT Press.
- Ba'dulu, Abd. Muis. 1990. Frase verbal bahasa Mandar Majene. *Bahasa-bahasa daerah Sulawesi dalam konteks bahasa nasional*, ed. by Husan Abas and T. David Andersen, 69–84. Ujung Pandang: Universitas Hasanuddin and Summer Institute of Linguistics.
- Berg, René van den. 1996. The demise of focus and the spread of conjugated verbs in Sulawesi. *Papers in Austronesian linguistics* 3, ed. by Hein Steinhauer, 89–114. Canberra: Pacific Linguistics.
- Billings, Loren A. 2002. Phrasal clitics. *Journal of Slavic Linguistics* 10.53–104.
- Billings, Loren A., and Abigail Konopasky. 2002. The role of morphology in ordering verb-adjacent clitics: from syntax to prosody in Bulgarian and Tagalog. *Linguistics in Potsdam* 19.1–26.
- Billings, Loren A., and Abigail Konopasky. 2003. Reassessing the role of syntax inside the morphological word: verb-adjacent clitics in Tagalog and Bulgarian. *Cornell Working Papers in Linguistics* 19.20–34.
- Chung, Sandra. 2003. The syntax and prosody of weak pronouns in Chamorro. *Linguistic Inquiry* 34.547–99.
- Collins, Millard A., Virginia R. Collins, and Sulfilix A. Hashim. 2001. *Mapun-English dictionary*. Manila: Summer Institute of Linguistics.
- Dubois, Carl D. 1976. *Sarangani Manobo: an introductory guide*. Manila: Linguistic Society of the Philippines.
- Fukuda, Takashi. 1997. *A discourse-oriented grammar of Eastern Bontoc*. Manila: Linguistic Society of the Philippines.
- Gonzalez, Andrew B. 1981. *Pampangan: towards a meaning-based description*. Canberra: Pacific Linguistics.
- Himmelman, Nikolaus P. 1996. Person marking and grammatical relations in Sulawesi. *Papers in Austronesian Linguistics* 3, ed. by Hein Steinhauer, 115–36. Canberra: Pacific Linguistics.
- Huang, Lillian M. 1995. *A study of Mayrinax syntax*. Taipei: Crane.
- Huang, Lillian M., Elizabeth Zeitoun, Marie M. Yeh, Anna H. Chang, and Joy J. Wu. 1999. A typological study of pronominal systems in some Formosan languages. Selected papers from the Fifth International Conference on Chinese Linguistics, ed. by Wang Hsu, Feng-fu Tsao, and Chin-fa Lien, 165–98. Taipei: Crane.
- Itô, Junko, and Mester, Armin. To appear. Systemic markedness and faithfulness. *Chicago Linguistic Society* 39.
- Kaufman, Daniel. 2003. Paradigm effects and the affix-shape/position generalization. *Proceedings of the 22nd West Coast Conference on Formal Linguistics*, ed. by Gina Garding and Mimmi Tsujimura, 273–86. Somerville, MA: Cascadia Press.
- Kaufman, Daniel. To appear. Rigidity versus relativity in adverbial syntax. *Clause Structure and Adjuncts in Austronesian Languages*, ed. by Hans-Martin Gärtner, Paul Law, and Joachim Sabel. Berlin: Mouton de Gruyter.
- Klavans, Judith L. 1985. The independence of syntax and phonology in cliticization. *Language* 61.95–120.

- Kurisu, Kazutaka. 2001. The phonology of morpheme realization. Santa Cruz: University of California dissertation.
- Lee, Celeste Chia-Yen. 2004. The ordering of clitic pronouns in the languages of Southeast Mindanao. Shalu: Providence University M.A. thesis.
- Liao, Hsiu-chuan. 2004. Transitivity and ergativity in Formosan and Philippine languages. Honolulu: University of Hawai'i dissertation.
- McCarthy, John J. 2002/2003. Optimal paradigms. Amherst: University of Massachusetts, ms. [Rutgers Optimality Archive 485.]
- McFarland, Curtis D. 2001. Filipino Enclitics. *Philippine Journal of Linguistics* 32. 1–15.
- Mead, David. 2002. Proto-Celebic focus revisited. The history and typology of Western Austronesian voice systems, ed. by Fay Wouk and Malcolm Ross, 143–77. Canberra: Pacific Linguistics.
- Rau, Der-Hwa Victoria. 1992. A grammar of Atayal. Taipei: Crane.
- Reid, Lawrence A. 2001. On the development of agreement markers in some Northern Philippine languages. *Issues in Austronesian morphology: a focusschrift for Byron W. Bender*, ed. by Joel Bradshaw and Kenneth L. Rehg, 235–57. Canberra: Pacific Linguistics.
- Rudin, Catherine, Christina Kramer, Loren Billings, and Matthew Baerman. 1999. Macedonian and Bulgarian *li* questions: beyond syntax. *Natural Language and Linguistic Theory* 17. 541–86.
- Schachter, Paul. 1973. Constraints on clitic order in Tagalog. *Parangal kay Cecilio Lopez*, ed. by Andrew B. Gonzalez, 214–31. Manila: Linguistic Society of the Philippines.
- Schachter, Paul, and Fe T. Otanes. 1972. Tagalog reference grammar. Berkeley: University of California Press.
- Weaver, Dan, and Marilou Weaver. 1964. Ranking of personal pronouns in Agusan Manobo. *Oceanic Linguistics* 3.161–70.
- Wolff, John U. 1996. The development of the passive verb with pronominal prefix in Western Austronesian languages. *Reconstruction, classification, description: Festschrift in honor of Isidore Dyen*, ed. by Bernd Nothofer, 16–40. Hamburg: Abera.
- Zobel, Erik. 2002. The position of Chamorro and Palauan in the Austronesian family tree: evidence from verb morphosyntax. *The history and typology of Western Austronesian voice systems*, ed. by Fay Wouk and Malcolm Ross, 405–34. Canberra: Pacific Linguistics.

Loren Billings

Department of Foreign Languages and Literature
National Chi Nan University
Puli, Nantou County 545
Taiwan
billings@ncnu.edu.tw

Daniel Kaufman
Department of Linguistics
Cornell University
Ithaca, NY 14853–4701
U.S.A.
dak37@cornell.edu