

# Clitics in Sasak, eastern Indonesia

Prof Peter K. Austin  
Endangered Languages Academic Program  
Department of Linguistics  
SOAS, University of London  
pa2@soas.ac.uk  
29 April 2004

## 1 Introduction<sup>1</sup>

This paper is a discussion of the distribution of clitics in Sasak, an Austronesian (Western Malayo-Polynesian) language spoken by approximately two million people on the island of Lombok, eastern Indonesia. I will outline the types of clitics found in Sasak and then show that there are interesting interactions between clitic placement and focus constructions that result in the violation of a number of canonical word orders in Sasak. I will argue that these violations can be seen as arising from competition for linear positions within the sentence; this can be accounted for within an optimality-theoretic syntax framework (Bresnan, 2000, Grimshaw, 1999) which describes sentence structures in terms of violable ranked constraints, the interaction of which accounts for observed structures as being the most optimal result of constraint competition.

## 2 Background to Sasak

Sasak is a Western Malayo-Polynesian language spoken on the island of Lombok, Nusa Tenggara Barat province in eastern Indonesia (Austin, 2000, 2001). It is closely related to Samawa (spoken on the western half of Sumbawa Island to the east of Lombok) and Balinese, and sub-groups with them as a member of the Western-Malayo-Polynesian branch of Austronesian (see (Adelaar, 2002) for detailed discussion). It shows a wide range of local dialect variation in lexicon and syntax, which is the topic of ongoing investigation. There is also sociolectal variation with high, middle and low speech styles (Nothofer, 2000). Data for this paper is drawn from the Central and Southern varieties identified by their speakers as *Menó-Mené*, *Meriq-Meriku* and *Menu-Meni Sasak*<sup>2</sup>. I have done a limited amount of work on this topic for the eastern *Ngenó-Ngené* variety, but further depth research is required. Other dialects are yet to be investigated at all.

---

<sup>1</sup> I am grateful to Herman Suheri, Lalu Dasmara, Lalu Hasbollah, Aozar Zawad, and Yon Mahyuni for generous and patient assistance with the Sasak data, and to Joan Bresnan, Aaron Broadwell, Simon Musgrave, and Rachel Nordlinger for helpful comments on an earlier draft of the materials on optimality theory; any errors of fact or interpretation are my own. Research was supported by the Australian Research Council ARC grant A59803558.

<sup>2</sup> Sasak ethnolinguistic labels for regional variation encode the shibboleth terms for ‘like that – like this’, thus *menó* ‘like that’, *mené* ‘like this’ etc. These labels correspond in a complex way to linguistic features of dialect variation.

### 3 Clitics

There is a very large literature on clitics and they have been intensively investigated cross-linguistically (see (Napoli, 1996, Scalise, 1984, Zwicky, 1977, Zwicky, 1983) and the bibliography by Zwicky and Nevis). It is generally agreed in the literature that we can characterise clitics as follows:

- words whose distribution is syntactically determined. Unlike affixes clitics may attach to words of many different classes, including words like prepositions or adverbs that don't usually take affixes
- clitics are not independent prosodic units and may phonologically attach to a **host** which may be a word or a phrase. There is usually no lexically conditioned idiosyncratic forms or distributions, and word-level phonology, eg. **sandhi** rules, also applies to clitics. A clitic that precedes its host is **proclitic** while one that follows is **enclitic**. The rules for clitic location can be quite complex (eg. in Italian object pronoun clitics are proclitic to finite verbs but enclitic to non-finite forms)
- clitics may appear in **clusters** with multiple clitics in different functions
- clitics typically code **nominal features** (person/number/case), auxiliaries or verbal features (**tense/aspect/mood/polarity**), or pragmatic functions (**question, information status**)
- clitics may collapse distinctions made in non-clitic forms, eg. Italian 1<sup>st</sup> and 2<sup>nd</sup> person clitics collapse accusative/dative/reflexive together; they are distinguished in 3<sup>rd</sup> person (but 3<sup>rd</sup> person dative collapses gender and 3<sup>rd</sup> person reflexive collapses gender and number contrasts). Thus, *mi* codes first person singular accusative/dative/reflexive, cf. *lo* for 3<sup>rd</sup> singular masculine accusative, *gli* for all 3<sup>rd</sup> singular datives, while *si* codes all 3<sup>rd</sup> person reflexives (Grimshaw, 1997).
- clitics can be classified into different types. Thus, ((Napoli, 1996: ) gives the following three classes:

Simple clitic	phonologically reduced allomorph of full lexical item, eg. 'll in English
Special clitic	<b>weak</b> or <b>atonic</b> form attached to a host, not phonologically reduced and often formally different from strong form, eg. Italian object clitics
Bound word clitic	independent phonological word (not atonic) with no non-clitic counterpart, eg. Serbo-Croatian <i>li</i>

The syntactic distribution of clitics often needs to refer to a special clausal position called **Wackernagel position** (P2) (Wackernagel, 1892) – second position in the clause, defined as **after first word** or **first syntactic constituent**. P2 can apply to any kind of clitic. An example is the Serbo-Croatian question particle *li*.

The following are examples of the different types of clitics:

#### Simple clitic (English)

The English future auxiliary 'will' can appear as a full or reduced enclitic form:

- (1) I will go to Naples tomorrow  
 (2) I'll go to Naples tomorrow

The reduced form encliticises to the last word of the preceding subject noun phrase, regardless its category (noun in (3), verb in (4) or adverb in (5)):

- (3) That guy'll go to Naples tomorrow  
 (4) That guy you met'll go to Naples tomorrow  
 (5) That guy you met yesterday'll go to Naples tomorrow

### Special clitic (Italian)

In Italian pronominal reference for objects can be achieved by independent pronouns, as in:

- (6) Elena torna, e trova proprio **me** a casa  
 Elena return:3sg and find:3sg certainly 1sgacc loc house  
 'Elena returns and finds me at home'

If there is no adverb, a proclitic pronoun is used with a finite verb<sup>3</sup>:

- (7) Elena torna, e **mi** trova a casa  
 Elena return:3sg and 1sgacc find:3sg loc house  
 'Elena returns and finds me at home'

and an enclitic pronoun with a non-finite verb, as in the following purpose clause:

- (8) Elena torna a trovar**mi** a casa  
 Elena return:3sg to find:1sgacc loc house  
 'Elena returns to find me at home'

Indirect object (dative) and direct object (accusative) proclitics appear in a cluster before the finite verb, as in:

- (9) **Glielo** dico  
 3dat:3sgmascacc speak:1sg  
 'I speak to him about it'

---

<sup>3</sup> In Italian orthography proclitics are written as a separate word, enclitics are written together with their host.

### Bound word clitic (Serbo-Croatian) — Wackernagel's position

The bound word question marker clitic *li* follows the first word of a clause, regardless of its category (examples from (Napoli, 1996)):

- (10) Hocêš      li      docî  
 aux:2sg      Q      come?  
 'Will you come?'
- (11) Jeste      li      joj      se      predstavili      u      sali  
 aux:2pl      Q      3sgf.dat      reflex.acc      introduce      in      hall  
 'Did you introduce yourselves to her in the hall?'
- (12) Da      li      mi      ih      je      dao      Jovan  
 comp      Q      1sg.dat      3pl.acc      aux:3sg      give      Jovan  
 'Did Jovan give them to me?'

Notice that a clitic complex consisting of a pronoun and auxiliary may follow the first **constituent** of a clause (the NP in (13)) or else the first **word**, interrupting the syntactic constituent, as in (14):

- (13) [Taj      pesnik]      mi      je      napisao      knjigu  
 that      poet      1sg.dat      aux:3sg      wrote      book  
 'That poet wrote me a book?'
- (14) [Taj      mi      je      pesnik]      napisao      knjigu

## 4 Sasak basic clause structure

Sasak phonological words have a single stress which falls on final syllable. Roots that end in *a* change this phonetically to mid-central tense vowel in word-final position in most dialects. Compound words take a nasal linker –N (realised as an velar nasal *ŋ* in Menu-Meni and an alveolar nasal *n* in other dialects) attached to the first element of their constituent parts if it ends in a vowel; examples are:

- (15) *ime* 'hand'      *inaq* 'mother'      *imen inaq* 'thumb'  
*mate* 'eye'      *bulu* 'hair'      *maten bulu* 'eyelash'  
*tai* 'excrement'      *mate* 'eye'      *tain mate* 'sleepdust'

There are a limited number of suffixes all of which code derivational morphology (Sasak has no inflectional morphology – no tense, aspect, mood, case), eg. *-ang* 'applicative verb formative', *-an* 'nominaliser'. These are all part of word stress domain (they will carry stress) and never show *a* ɿ *e* or nasal linker:



**PP**

P NP	<i>kance Ali</i>	‘with Ali’
	<i>léq balé beléq nó</i>	‘in that big house’

**AP**

Adj PP	<i>senang léq hal nó</i>	‘pleased about that matter’
Adj PP	<i>puas isiq kakenan ni</i>	‘satisfied with this food’
Adj Adv	<i>beléq lalòq</i>	‘very big’

**VP**

V PP	<i>bedait kance Ali</i>	‘meet with Ali’
	<i>tòkòl léq góbuk</i>	‘stand in the village’
V NP	<i>mpuk acòng nó</i>	‘hit that dog’
V NP PP	<i>bèng hadiah ni jòk guru nó</i>	‘give this present to the teacher’
	<i>jauq kedelé jòk peken</i>	‘carry soybeans to market’

The order of heads and modifiers and heads and specifiers is also strict, so that X’ elements must precede modifiers and specifiers. Note that this clearly violates the proposed ‘universal’ of (Kayne, 1994) “specifier positions must invariably appear to the left of their associated head, never to the right” (see also (Broadwell, 2001)).

**Modifiers**

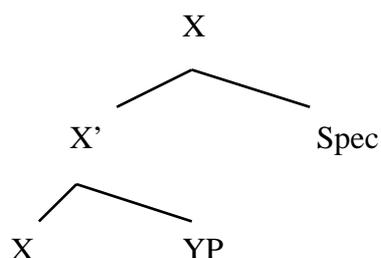
N AP	<i>kanak beléq lalòq</i>	‘very big child’
N PP	<i>dengan éléq Penujaq</i>	‘person from Penujaq’
	<i>balé léq góbuk nó</i>	‘house in that village’
N NPposs	<i>balén Ali</i>	‘Ali’s house’
	<i>semetòn dengan beléq nó</i>	‘brother of that big man’
N relC	<i>buku saq beléq lalòq</i>	‘book which is very big’
	<i>buku saq kanak bace</i>	‘book which the child is reading’
	<i>kanak saq mpuk acòng nó</i>	‘the child who hit the dog’

**Specifiers**

N Dem	<i>dengan ni</i>	‘this person’
	<i>balé nó</i>	‘that house’
	<i>dengan sai</i>	‘which person?’
	<i>balé saq mbé</i>	‘which house?’

Note that in some dialects of Sasak such as Menó-Mené, demonstrative specifiers are **simple clitics** that encliticise to the previous constituent in the NP, regardless of its category.

We can represent these facts by a tree structure:



Alternatively, if we separate hierarchical structure from linear order this can be stated in terms of the following linearisation principles:

$$X < YP$$

$$X' < YP_{[ADJ]} < SPEC$$

#### 4.1 Ngenó-Ngené Sasak

This variety of Sasak is spoken widely throughout central and eastern Lombok and serves as a kind of ‘standard language’ for educational materials. I have worked on this variety with a few speakers, but much more checking of sources remains to be done. There may be village level variation within this variety. The following examples come from Selong and Ketara village.

The basic clause structure for this dialect can be summarised as follows. Simple sentence consists of verb<sup>5</sup> plus one or more arguments, or just a verb to express imperative:

- (17) *Laló!*                      *Makan!*  
       go                              eat

Verbs fall into three classes: intransitive (taking a single ACTOR<sup>6</sup> argument), transitive (taking two arguments, an ACTOR and an UNDERGOER) or ditransitive (with three arguments: ACTOR, UNDERGOER and THEME). Transitive verbs have two forms: an **unprefixed** verb or a verb with a **nasal prefix**. For unprefixed verbs the UNDERGOER precedes the verb and the ACTOR usually precedes it, except if it is third person in which case it is marked by the preposition *isiq* and follows the verb like an oblique PP. When the verb takes a nasal prefix then the ACTOR precedes the verb and the UNDERGOER follows it:

---

<sup>5</sup> We exclude non-verbal predicates from consideration here.

<sup>6</sup> For convenience I use the terms ACTOR and UNDERGOER from the Role and Reference Grammar framework (see Foley and Van Valin 1984, Van Valin and La Polla 1999) to refer to the more agent-like semantic role of a transitive verb and the more patient-like role respectively. For ditransitive verbs with a goal and a theme NP the UNDERGOER will be the goal. These terms are thus equivalent to LOGICAL SUBJECT and LOGICAL OBJECT in other frameworks. There is evidence that Menó-Mené Sasak does not distinguish between grammatical relations (other than core versus oblique) — see Austin 2002 for further discussion. The issue of grammatical relations is not of concern here.

- (18) *Ime-n*            *Aózar*            *kaput=ne*            *isiq*            *Ali*  
 hand-link            Aozar            bandage=3            by            Ali  
 ‘Ali bandaged Aozar’s hand’
- (19) *Ali*                    *ngaput=ne*            *ime-n*                    *Aózar*  
 Ali                    N.bandage=3            hand-link                    Aozar  
 ‘Ali bandaged Aozar’s hand’

## 4.2 Menó-Mené Sasak

The Sasak dialect termed Menó-Mené is spoken throughout central Lombok; the data reported here is from Praya and Puyung villages. The basic clause structure of this dialect shows a number of differences from Ngenó-Ngené. Although intransitive, transitive and ditransitive verbs are also distinguished, the normal word order is ACTOR V (UNDERGOER) (THEME). Transitive verbs have both an unprefix and a nasal prefixed form, but unlike Ngenó-Ngené their use is rather different. Nasal prefixed verbs are used with an unspecified (non-referential) undergoer that is usually not expressed, eg.

- (20) *Kanak=nó*            *bace*                    *buku=ni*  
 child=that            read                    book=this  
 ‘That child reads this book’
- (21) *Kanak=nó*            *m-bace*  
 child=that            N-read  
 ‘That child is reading’

Unprefixed verbs are used with referential undergoers. Menó-Mené shows the effect of a global animacy hierarchy (Sliverstein 1976) in encoding semantic roles: if ACTOR is 3<sup>rd</sup> person and UNDERGOER is 3<sup>rd</sup> person ACTOR appears in a PP with the preposition *isiq* at the end of the sentence:

- (22) *Iaq=n*            *gitaq*                    *kanak-kanak=nó*                    *isiq*            *Herman*  
 fut=3            see                    reduplicated-child=that                    by            Herman  
 ‘Herman will see the children.’

If ACTOR is 3<sup>rd</sup> person and UNDERGOER is first or second person **passive** must be used: the verb takes *te-* prefix, is intransitive and the agent if expressed is in an optional oblique prepositional phrase with the preposition *isiq*:

- (23) *Aku te-gitaq isiq Ali*  
 1sg pass-see by Ali  
 ‘I was seen by Ali’

If ACTOR is 1<sup>st</sup> or 2<sup>nd</sup> person then ACTOR V UNDERGOER order is found.

Optional PP adjuncts follow the verb, adverbs follow the verb or come at the beginning of sentence. Co-ordinate or subordinate clauses are introduced by conjunctions. Relative clauses take an invariant marker *saq* and occur in modifier position, following the head noun and preceding any demonstratives. The coreferential element inside the relative clause is gapped.

Topicalised elements may appear at the beginning of the clause, after adverbs but before the verb: (cf. King 1995):

- (24) *Kance guru=m iaq=k bedait lèmaq*  
 with teacher=2 fut=1sg meet tomorrow  
 ‘It’s with your teacher that I will meet tomorrow’

Note that adverbs may precede the topicalised item, as in:

- (25) *Lèmaq kance guru=m iaq=k bedait*  
 tomorrow with teacher=2 fut=1sg meet  
 ‘It’s with your teacher that I will meet tomorrow’

There are a number of auxiliary particles that encode tense/aspect/mood/polarity categories, eg. *iaq* ‘future’, *jangke* ‘continuous’, *taó* ‘can’, *ndéq* ‘not’, *déndéq* ‘don’t’ – these are bound word clitics and occur in Wackernagel position following the first constituent of the clause. If only AUX and V occur the AUX must be first:

- |      |                               |   |                |                           |
|------|-------------------------------|---|----------------|---------------------------|
| (26) | <i>Kanak=nó</i><br>child=that | <i>jangke=n</i><br>cont=3<br><i>iaq=n</i><br>fut=3<br><i>taó=n</i><br>can=3<br><i>ndéq=n</i><br>not=3 | <i>mancing</i> | ‘That child is fishing’   |
|      |                               |   | N.fish         | ‘That child will fish’    |
|      |                               |   |                | ‘That child can fish’     |
|      |                               |   |                | ‘That child doesn’t fish’ |
|      |                               |   |                |                           |

There are special clitics that mark nominal categories and functions: a Wackernagel enclitic that encodes ACTOR is attached to the first non-NP constituent of the clause (Conjunction, Adverb, Preposition, AUX) as in:

**Subordinate Conjunction**

- (27) *Guru iaq=n tulak malik sèngaqm mpuk=k*  
 teacher fut=3 return again because=2 hit=1  
 ‘The teacher will come back again because you hit me’

**Adverb**

- (28) *Terus=k iaq bedait kance guru=nó*  
 then=1 fut meet with teacher=that  
 ‘Then I will meet that teacher’

- (29) *Telu jam=k antih=m wah*  
 three hour=1 wait=2 already  
 ‘I already waited for you for three hours’

**AUX particle**

- (30) *Inaq iaq=n laló jòk peken lèmaq*  
 mother fut=3 go to market tomorrow  
 ‘Mother will go to market tomorrow?’

**Preposition** (focus on a PP fronts and reverses the order of P and NP constituents (this is discussed in detail below, recall that N is a host for possessors **not** the ACTOR enclitic)):

- (31) *Mbé éléq=n tulak*  
 where from=3 return  
 ‘Where is he coming back from?’

If the verb is intransitive and there is no other potential host in the clause then it may take the enclitic ACTOR marker:

**Intransitive verb**

- (32) *Tulak=n éléq peken*  
 return=3 from market  
 ‘He comes back from the market’

If there are several possible hosts the enclitic follows **first host in a sequence**

- (33) *Kanak=nó ndéq=n iaq mancing*  
 child=that not=3 fut N.fish  
 ‘That child will not fish’

See also (28) above. We discuss this further below.

Note that special clitics are only used when there is specific reference. They cannot occur if the ACTOR is non-individuated. A clitic is also not possible within a relative clause when the ACTOR has been relativised (see below).

In Menó-Mené a special clitic also attaches as an enclitic to a transitive or ditransitive verb to encode a referentially the specific UNDERGOER (goal participant for ditransitives):

- (34) *Iaq=m*                      *gitaq=n*              *léq*      *peken*  
 fut=2                              see=3              loc      market  
 ‘Will you see him at the market?’

- (35) *Iaq=m*      *bèng=k*                      *kèpèng*  
 fut=2              give-link=1sg              money  
 ‘Will you give me some money?’

Note that nasal prefixed verbs cannot take an undergoer enclitic since their undergoer always has non-specific reference. Similarly, when an undergoer has been relativised the verb inside the relative clause cannot bear an undergoer enclitic.

If the verb is transitive or ditransitive and there is no potential host for the ACTOR clitic then a special AUX particle *muq* (in Meriaq-Meriku *bi* or *si*) occurs to support the ACTOR special clitic:

- (36) *Muq=k*                      *gitaq=n*              *léq*      *peken*  
 AUX=1sg                              see=3              loc      market  
 ‘I saw him at the market’

- (37) *Muq=n*      *bace*                      *buku=ni*              *isiq*              *kanak=nó*  
 AUX=3              read                      book=this              by              child=that  
 ‘That child reads this book’

In Menó-Mené Sasak any argument of a verb can be relativised by a gap strategy; inside the relative clauses we cannot get a clitic agreeing with the relativised NP – we can determine the function of that NP however from the remaining clitic: if it is attached to the Verb then ACTOR must have been relativised, if it is attached to another host or *muq* then the UNDERGOER must have been relativised:

- (38) *Kanak*    *saq*              (\**muq=n*)              *gitaq=n*    *léq*              *peken=nó*  
 child    rel    see=3              loc              market=that  
 ‘That child who saw him at the market’ [relativised ACTOR]

- (39) *Kanak*    *saq*              *muq=k*              *gitaq*    (\**gitaq=n*)    *léq*              *peken=nó*  
 child    rel                              AUX=1              see    loc              market=that  
 ‘That child who I saw at the market’ [relativised transitive UNDERGOER]

- (40) *Kanak*    *saq*              *muq=k*              *bèng*    (\**bèng=n*)    *kèpèng=nó*  
 child    rel                              AUX=1              give    money=that  
 ‘That child who I give money to’ [relativised ditransitive UNDERGOER]

The following is a phrase structure summary for this dialect:

CP	‡	(Comp) (Topic) (Focus) S
S	‡	(Aux)*, V, NP <sub>[SUBJ]</sub> , NP <sub>[OBJ]</sub> , NP <sub>[INDOBJ]</sub> , PP* <sub>[ADJ]</sub>
PP	‡	P NP
NP	‡	N' (NP <sub>[POSS]</sub> ) (XP <sub>[MOD]</sub> ) (Dem)
N'	‡	N (N)

### Linearisation principles

SUBJ < OBJ < ADJ [violated by 3<sup>rd</sup> person SUBJ]

Aux < V

Aux is bound word clitic in P2 — follows first **word** or **constituent**

UNDERGOER clitic attaches to V, ACTOR clitic attaches to first available host otherwise to *muq*

### 4.3 Menu-Meni Sasak

The variety of Sasak spoken in south-central Lombok is referred to by its speakers as Menu-Meni; this data comes from Ganti village. Basic sentence structure is similar to Menó-Mené Sasak except that animacy does not play a role and there is no forced passive or use of *isiq* with 3<sup>rd</sup> person ACTOR.

The major difference between the dialects is in special clitic use: clitics are only used for specific (referential) NPs and there are different forms for ACTOR and UNDERGOER functions in second and third person:

	ACTOR	UNDERGOER
1sg	= <i>kò</i>	= <i>kò</i>
1pl	= <i>te</i>	= <i>te</i>
2	= <i>ò</i>	= <i>kem</i>
3	= <i>ng</i>	= <i>è</i>

The ACTOR clitic is typically attached to a preverbal host and the UNDERGOER clitic to the Verb. If there is no non-verb host for the ACTOR special clitic then the Verb will host it (unlike Menó-Mené and Meriaq-Meriku where a special AUX particle appears as the host). If UNDERGOER is non-specific then the Verb just takes actor forms. If UNDERGOER is specific then a clitic complex of ACTOR+UNDERGOER occurs on the verb — if 1<sup>st</sup> **and** 2<sup>nd</sup> person are both involved this complex loses the specification of **number** for 1<sup>st</sup> person (both 1<sup>st</sup> ACTOR and 1<sup>st</sup> UNDERGOER neutralise the number distinction in the presence of 2<sup>nd</sup> person):

ACTOR	UNDERGOER			
	1sg	1pl	2	3
1sg			= <i>kem</i>	= <i>kè</i>
1pl			= <i>kem</i>	= <i>tè</i>
2	= <i>òkò</i>	= <i>òkò</i>		= <i>òkè</i>
3	= <i>ngkò</i>	= <i>ngte</i>	= <i>ngkem</i>	

Recall that inside a relative clause no clitic appears coding the relativised NP (it is treated as if it were non-specific) — this results in ambiguity for first person because =*kò* could be relativised ACTOR + 1sgUNDERGOER **or** 1sgACTOR + relativised UNDERGOER:

- (41) *nu ie kanak saq gitaq=kò nuqng*  
 that 3 child rel see=1sg that  
 (a) ‘That is the child who I saw’  
 (b) ‘That is the child who saw me’

#### 4.4 Origins of the dialect difference

How can we explain the difference in clitic behaviours between the dialects? There is some evidence for historical chance in Menó-Mené and Meriaq-Meriku Sasak that has been moving the UNDERGOER enclitic more towards an affix and making it less like a clitic. This evidence is as follows:

- the special clitics have lost any syllabic character and are only a single consonant in these dialects (so are irrelevant for stress assignment)
- the special clitics show a Nasal linker however it is optional (thus, 1pl is either =*nt* or =*t*)
- the undergoer special clitic has become fused with the applicative suffix, so we get –*angk* ‘applicative 1sg’, –*ant* ‘applicative 1pl’, –*am* ‘applicative 2’ and –*an* ‘applicative 3’. In Menu-Meni the suffix and clitic are clearly distinct, so we find –*ang=kò* ‘applicative 1sg’, –*ang=ò* ‘applicative 2’ etc.

These developments suggest that the V=clitic sequence is being treated more like a combination of verb and affix, which is then less open to being broken up with another clitic. The special AUX host has been innovated to solve this problem and keep the two clitics in different places; interestingly the auxiliary as variant forms depending on dialect, thus *muq*, *bi*, and *si* all occur in different villages.

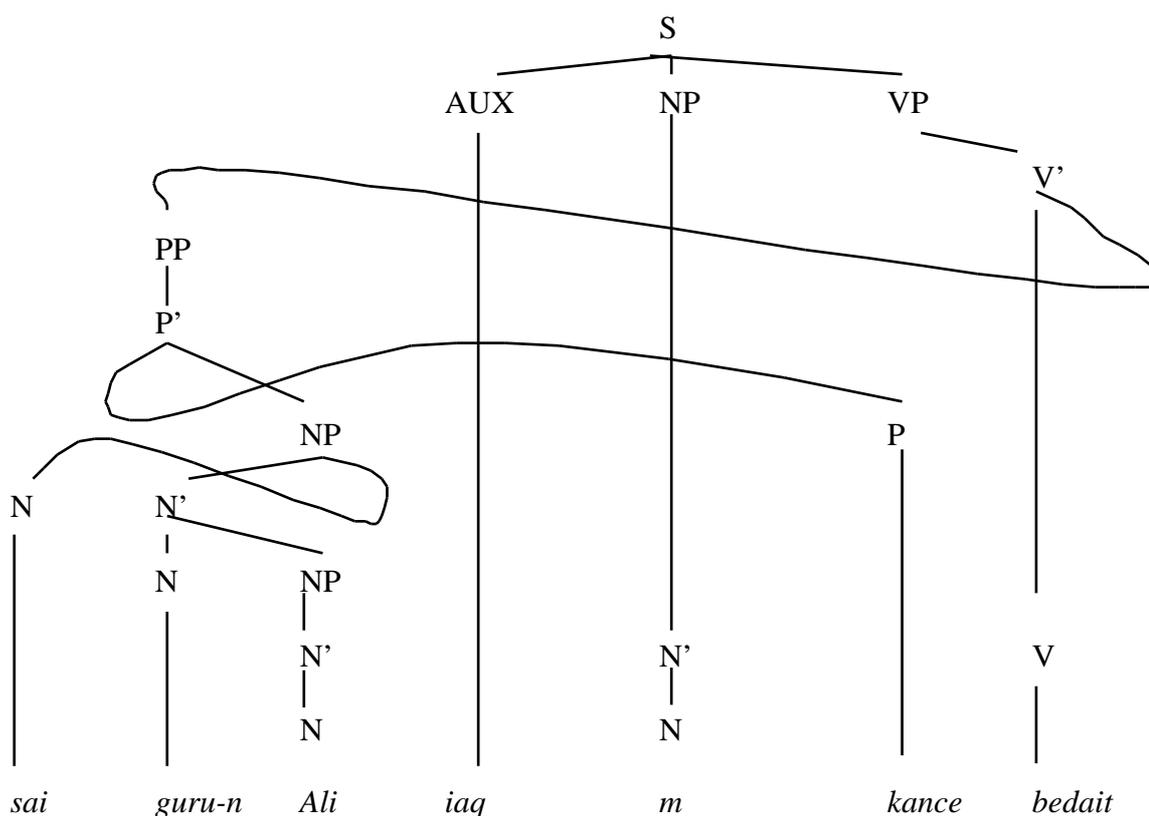
## 5 Focus constructions — content questions and interactions with clitics

Sasak sentences such as the following raise interesting issues for an understanding of constituency and word order. This sentence involves a content question: questioning the identity of a particular person. The question word is a specifier of an NP that is within a PP complement of an intransitive verb. Notice that the Specifier precedes its head noun

within the NP and this object complement NP precedes its head preposition, totally the opposite to usual phrasal word order<sup>7</sup>:

- (42) *Sai guru=n Ali iaq=m kance bedait?*  
 which teacher-link Ali fut=2 with meet  
 PP[ NP[Spec N-link NPI<sub>POSS</sub>]NP AUX=encl<sub>ACTOR</sub> P ]PP V  
 ‘Which teacher of Ali’s will you meet with?’

The following diagram shows dependency relations in this sentence and the resulting crossing of dependencies:



Content questions in Sasak are formed according to the following principles:

1. the focus (wh-) element must appear in clause initial position
2. if the focus is within a PP then the PP is fronted and order flipped to NP-P, the P can then be a host for the ACTOR enclitic:

- (43) *Sai kance=m bedait léq peken?*  
 who with=2 meet loc market  
 ‘Who do you meet with in the market?’

<sup>7</sup> The following examples are in Menó-Mené; examples illustrating identical behaviour have been checked for Ngenó-Ngené and Menu-Meni dialects.

- (44) *Ape isiq=m mpuk=n?*  
 what with=2 hit=3  
 ‘What did you hit him with?’

Notice that PP fronting with flip applied to both complements, as in (43) and adjuncts, as in (44). Similar phenomena are found in Zapotecan and Mayan languages (Broadwell, 2001).

3. if there is an AUX clitic particle it occupies Wackernagel position and follows the first constituent, either the PP as in (45) or the NP as in (46), in the latter case it now hosts the ACTOR enclitic:

- (45) *Sai kance=m iaq bedait léq peken?*  
 who with=2 fut meet loc market  
 ‘Who will you meet at the market?’

- (46) *Sai iaq=m kance bedait léq peken?*  
 who fut=2 with meet loc market  
 ‘Who will you meet at the market?’

4. NPs containing post-head modifier focus items like possessors are clause initial but the actual focus word itself is not and remains in post-head position:

- (47) *Guru-n sai iaq=m dengah léq masjid?*  
 teacher-link who fut=2 hear loc mosque  
 ‘Whose teacher will you hear at the mosque?’

- (48) *Guru éléq mbé harus=m dengah?*  
 teacher from where must=2 hear  
 ‘The teacher from where must you listen to?’

5. if the focus word is the wh-specifier *sai* ‘which’ (used for human reference only) or *saq mbé* ‘which’ (alternatively also *siq mbé*) it must be clause initial and precede the remainder of its NP:

- (49) *Sai guru-n Ali kance=m iaq bedait?*  
 which teacher-link Ali with=2 fut meet  
 ‘Which teacher of Ali’s will you meet with?’

- (50) *Sai gurun Ali iaq=m kance bedait?*

- (51) *Saq mbé bije kance=m iaq bedait?*  
 which child with=2 fut meet  
 ‘Which child will you meet with?’

- (52) *Saq mbé bije iaq=m kance bedait?*

So, in example (50) the focussed specifier is first in the clause, followed by rest of its NP. The AUX particle is in Wackernagel position following this NP, and the ACTOR clitic follows the AUX (since it is the first available host) and the preposition follows all these constituents. We have non-canonical constituent order of specifier and N' and P and NP because word order reflects functional roles.

## 6 An OT syntax account

Optimality theory (presented by (Prince, 1992) and elaborated in much following work) takes the view that language structures are the result of interactions between sets of linguistic constraints. Crucially:

- all constraints are universal
- constraints can be violated in well-formed sentences
- a grammar is a ranking of the universal constraints which determines which constraint is satisfied in case of a conflict
- the optimal form (the one best satisfying the language specific ranking) is grammatical, all non-optimal candidates are ungrammatical

As Grimshaw (Grimshaw, 1997: 169) points out: “an optimal output form for a given input is selected among the class of competitors in the following way: a form which for every pairwise competition involving it, best satisfies the highest-ranking constraint on which the competitors conflict, is *optimal*”.

The constraints which we need for Sasak to account for the data discussed above are as follows. There are two constraints that apply to clitics. For the bound word AUX clitics we set up a constraint P2-AUX that places the AUX particle after first **constituent**. This can be interpreted as the minimal or maximal XP. Secondly, for the actor special enclitic we need a constraint P2-ACTOR to place the enclitic for ACTOR after the first available host **word**. Next, we have a constraint that prevents maximal projection constituents (NP and PP) being split up and made discontinuous: \*DisconXP. This is reflected in the fact that in Sasak:

- prepositions cannot be stranded in focus and content question constructions but must be ‘pied-piped’ with their NP complement
- content question specifiers cannot be removed from their containing NP but must be ‘pied-piped’ with their NP head

The constraint \*DisconXP is violable and ranks below the clitic constraints because we find AUX particles and actor special clitics appearing **inside** PPs in content questions. The next constraint is an linearisation restriction that the minimal XP containing a content question word must be in focus position, aligned with the left edge of S: AlignFocusLeft. Finally, there is a violable constraint that lexical heads must precede their complements: X < YP. This constraint is violated in PP flip but because AlignFocusLeft outranks it in Sasak. In English and Italian this constraint outranks AlignFocusLeft: (cf. In whom do you trust? Con qui hai parlato?)

This gives us the following Tableau:

	candidates	P2-AUX P2-ACTOR	*DisconXP	AlignFocusLeft	X < YP
	<i>*sai kance iaqm bedai</i>	*!			
	<i>*sai iaq kancem bedait.</i>	*!			
	<i>*sai iaqm bedait kance</i>		*!		
	<i>*iaqm bedait kance sai</i>			*!	
	<i>*kancem sai iaq bedait</i>			*!	
<b>F</b>	<i>sai kancem iaq bedait</i>				*
<b>F</b>	<i>sai iaqm kance bedait</i>				*

Note there are two alternative successful candidates because P2-AUX can treat PP as P1 or NP as P1.

Consider now the constraint  $X' < YP$  that heads of phrases precede their modifiers. This constraint outranks AlignFocusLeft because a modifier focus phrase is not in leftmost position

	candidates	P2-AUX P2-ACTOR	$X' < YP$	AlignFocusLeft	X < YP
	<i>*gurun sai kance iaqm bedait</i>	*!			
	<i>*gurun sai iaq kancem bedait</i>	*!			
	<i>*saigurun kancem iaq bedait</i>		*!		*
	<i>*sai gurun iaqm kance bedait</i>		*!		*
	<i>*sai iaqm bedait kance gurun</i>		*!		
	<i>*kancem guru sai iaq bedait</i>			**!	
<b>F</b>	<i>gurun sai kancem iaq bedait</i>			*	*
<b>F</b>	<i>gurun sai iaqm kance bedait</i>			*	*

Again, there are two alternative successful candidates because P2-AUX can treat PP as P1 or NP as P1. Note that violation of P2-ACTOR is bad.

Now consider the constraint of  $X' < SPEC$  that heads of phrases precede their specifiers. This constraint is ranked below AlignFocusLeft because the specifier focus words *sai* and *saq mbé* (or *siq mbé*) ‘which’ must be in leftmost position even though this means they now precedes the head and associated modifiers:

	candidates	X' < YP	AlignFocusLeft	X' < SPEC	X < YP
	<i>*Ali sai gurun kancem iaq bedait</i>	*!	*		
	<i>*sai Ali gurun kancem iaq bedait</i>	*!			
	<i>*gurun Ali sai kancem iaq bedait</i>		*!		
<b>F</b>	<i>sai gurun Ali kancem iaq bedait</i>			*	*
<b>F</b>	<i>sai gurun Ali iaqm kance bedait</i>			*	*

There are two alternative successful candidates because P2-AUX can treat PP as P1 or NP as P1. Note that violation of P2-ACTOR is bad: *\*sai gurun kance iaqm bedait* or *\*sai gurun iaq kancem bedait*.

This account explains nicely the violations of standard constituent order in Sasak content questions and the linear order of elements in terms of their discourse and grammatical functions. Violations of canonical constituent ordering can thus be seen as arising from competition between elements for positions early in the clause.

## 7 Conclusions

Sasak has a variety of clitics that have different distributions in its various dialects. We find the following types:

1. **simple clitics** — demonstratives in some dialects (especially in Menó-Mené) have short forms that can encliticise to the final element of the X' that precedes them.
2. **special clitics** that encode nominal categories of person and number. With nouns these encode inalienable possessors, and with other hosts they encode ACTOR and UNDERGOER semantic roles. In Menó-Mené the UNDERGOER clitic attaches to the verb while the ACTOR clitic is attached to the first available host, otherwise to a special auxiliary that exists just to support the clitic if the verb is transitive and already takes an UNDERGOER clitic. The same pattern is found in Meriaq-Meriku dialect. In Menu-Meni dialect there are different ACTOR and UNDERGOER clitic forms and these attach to transitive verbs to form clitic clusters. These clusters reduce number specification for certain combinations and also result in ambiguity in relative clauses.
3. **bound word clitics** — auxiliary particles in Sasak follow the first constituent of the clause and can occur in sequences of up to three. The first in a sequence can host special clitics for ACTOR.

There are interesting interactions between the clitics, and also between focus constructions and the location of clitics. Under certain circumstances bound word clitics and attached special clitics intervene within a prepositional phrase. This results in violations of unmarked word orders, but can be seen as resulting from competition between constituents for positions early in the clause. Related but not identical phenomena have been found in Central American languages.

The patterns reported here for clitics in Sasak need to be supplemented by further research investigating other dialects and other types of constructions. It will also

be interesting to compare the analysis of Sasak with its neighbouring and related languages, Balinese and Samawa.

## 8 References

- Adelaar, K. Alexander. 2002. Classification of Sasak. *Papers in Languages of Bali, Lombok and Sumbawa* 1:000-000.
- Austin, Peter K. 2000. *Working Papers in Sasak, Volume 2*. Melbourne: University of Melbourne.
- Austin, Peter K. 2001. Verbs, valence and voice in Balinese, Sasak and Sumbawan. In *Explorations in valency in Austronesian languages*, ed. Barry Blake and Margaret Florey Peter Austin, 47-71. Bundoora: La Trobe University.
- Bresnan, Joan. 2000. Optimal Syntax. In *Optimality Theory: Phonology, Syntax and Acquisition*, ed. Frank van der Leeuw and Jeroen van Weijer Joost Dekkers, 334-385. Oxford: Oxford University Press.
- Broadwell, Aaron. 2001. Optimal order and pied piping in San Dionicio Zapotec. In *Formal and Empirical Issues in Optimality Theoretic Syntax*, ed. Peter Sells, 199-225. Stanford: CSLI.
- Grimshaw, Jane. 1997. The best clitic: constraint conflict in morphosyntax. In *Elements of Grammar: Handbook of Generative Syntax*, ed. Liliane Haegeman, 169-196. Dordrecht: Kluwer Academic Publishers.
- Grimshaw, Jane. 1999. Projection, heads, and optimality. *Linguistic Inquiry* 28:73-422.
- Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge: MIT Press.
- Napoli, Donna Jo. 1996. *Linguistics*. Oxford: Oxford University Press.
- Nothofer, Bernd. 2000. A preliminary analysis of the history of Sasak language levels. In *Working Papers in Sasak, Vol. 2*, ed. Peter K. Austin, 57-84. Melbourne: University of Melbourne.
- Prince, Alan & Paul Smolensky. 1992. *Optimality Theory: Constraint Interaction in Generative Grammar*. Piscataway, NJ: Rutgers University.
- Scalise, Sergio. 1984. *Generative Morphology*. Dordrecht: Foris.
- Wackernagel, Jacob. 1892. Über ein Gesetz der indogermanischen Wortstellung. *Indogermanische Forschungen* 1:333-436.
- Zwicky, Arnold. 1977. *On Clitics*. Bloomington: Indiana University Linguistics Club.
- Zwicky, Arnold & Geoffrey Pullum. 1983. Cliticization vs. inflection: English n't. *Language* 59:502-513.