The (relatively) simple workings of complex clitic systems*

Peter W. Smith, University of Connecticut
peter.w.smith@uconn.edu

1. Introduction

Clitics generally appear at word peripheral sites, either as proclitics at the beginning of a certain word/phrase, or enclitic at the end of a word/phrase.

(1) Karinganta-**rna** kuyu-jarra yampi-ja-rni. (Warlpiri, Legate 2008)
fact-1SG.SUBJ meat-DL leave-PAST-hither
'The fact is I left two animals (I speared) and came here.'

(2) **Ku**-hili mài (Kambera, Klamer 1997)
1SG-again-come
'I come again/I'll come again.'

Word peripheral position has been one of the traditional diagnostics for being a clitic as opposed to an affix, for instance in Zwicky & Pullum (1983:504), **"clitics can attach to material already containing clitics, but affixes cannot."**

In this talk, I wish to discuss a couple of cases of clitics which appear to contravene this matter and are cases of endoclitics.

Endoclisis is used in a loose sense (as used by Harris 2002) to describe clitics which are neither proclitic or enclitic. That is, they appear at a non-peripheral position within the word.

An example comes from Pashto (Tegey 1977, Kaisse 1981):

(3) á-**me**-xistələ
PREFIX-1SG-wear
'I was wearing it.'

Such clitics are rare cross-linguistically, and there are only a few (clearly) documented cases.[1]

The two cases that I will concentrate on will be Udi (Nakh-Dagestanian) and Sorani Kurdish (Indo-Iranian). These languages both involve clitics which go inside a word, therefore having non-peripheral clitics.

* Thanks to Jonathan Bobaljik, Alice Harris, Beata Moskal and Susi Wurmbrand for discussions and comments on the material here. All mistakes are mine.

[1] Cases to my knowledge include Pashto, Sorani Kurdish and Udi as mentioned here, as well as Old Irish (Kern 2011), European Portuguese (Luís 2004), (some) Italian dialects (A. Calabrese, p.c.).

Both systems are further complicated, since they have a complex system of clitic placement, where the clitics have a great deal of mobility, exhibiting various different attachment sites in the sentence.

Udi

(4) baba-n ęš nut eč-al-**le** k'wa
father-ERG apple.ABSL NEG bring-FUTII-3SG house.DAT
'Father will not bring apples to the house.'

(5) nana-n ten-**ne** bụγa-b-e p'ạ ačik'alšey
mother-ERG NEG-3SG find-DO-AORII two toy.ABSL
'Mother did not find two toys.'

(6) q'ačaγ-γ-on bez tänginax baš-**q'un**-q'-e
thief-PL-ERG my money.DAT steal$_1$-3PL-steal$_2$-AORII
'The thief stole my money.'

Sorani Kurdish

(7) ba dûrbin-**mân** dît-in
with binoculars-1PL see-3PL
'We saw them with binoculars.'

(8) na-**m**-xwârd
NEG-1SG-eat.PAST
'I did not eat'

(9) nard-in-**î**
send.PAST-3PL-3SG
'He sent them'

(10) xward-bû-in-**î**
eat.PAST-PART-3PL-3SG
'He had eaten them'

2. The predictable origins of chaos part 1: Sorani Kurdish[3]

2.1. What clitics (and affixes) mark

Sorani Kurdish marks agreement of both subject and object. An agreement affix on the verb marks one and a clitic marks the other.

In the present tense, agreement of the subject is marked by a suffix on the verb. Agreement features of the direct object are marked by a clitic:

(11) (min) bo Narmîn-**î** da-kir-im
(I) for Narmin-3sg prog-buy.pres-1sg
'I am buying it for Narmin.'

In the past tense, subject agreement is marked by the clitics and object agreement by the suffix on the verb:

(12) Narmin ba Sirwan-**î** dâ-n
Narmin to Sirwan-3sg give.past-3pl
'Narmin gave them to Sirwan.'

This is all the result of a system of split ergatvity present in the language. What is relevant is the agreement of the subject is obligatory in all tenses. Therefore, in the past tense we will see clitics.

I will abstract away from the present/past tense dichotomy, and will focus on past tense forms.

To see that clitics are different from verbal suffixes, below are the exponents of the various morphemes (taken from Samvelian 2007):

clitics

	SINGULAR	PLURAL
1	-(i)m	-mân
2	-(i)t	-tân
3	-î/-y	-yân

verbal suffixes

	SINGULAR	PLURAL
1	-(i)m	-în
2	-î	-(i)n
3	-ê(t)/∅	-(i)n

[3] Data in this section are taken from Thackston (2006) Samvelian (2007), Walther (2012). Other data can be found in Bonami & Samvelian (2008) and McCarus (1958).

2.2. Different positions of clitics

As pointed out earlier, there are a number of different positions that the same clitic can appear in.

As a basic, default position, subject clitics will be enclitic to the first element within the VP:

(13) Narmîn u Sirwan bâng-**mân** da-ka-n
Narmîn and Sirwan voice-1PL PROG-do.PRES-3PL
'Narmin and Sirwan are calling us.'

But they never attach to the subject of the clause:

(14) *Narmin-**yan** da-kuj-ê
Narmin-3PL PROG-kill.PRES-3SG
Intended: 'Narmin is killing them.'

When the VP consists solely of the verb and its agreement markers (clitics and affixes), things start to get more complicated, and the clitic position becomes more flexible, occupying various positions within the verb form:

(15) dît-**yân**-im
saw-3PL-1SG
'They saw me'

(16) xward-bû-**man**-in
eat.PAST-PART-**1PL**-3PL
'We had eaten them'

(17) xward-bû-in-**î**
eat.PAST-PART-3PL-**3SG**
'He had eaten them'

(18) ná-**yân**-dît-im
NEG-3PL-saw-1SG
'They didn't see me'

(19) da-m-xwârd
PROG-1SG-eat.PAST
'I was not eating.'

(20) na-m-da-xwârd
NEG-1SG-PROG-eat.PAST
'I was not eating.'

➡ In (15), the clitic immediately follows the verb root.
➡ In (16), the clitic follows the participle marker but precedes the object agreement suffix.
➡ In (17), the clitic follows both the participle marker *and* the object agreement suffix.
➡ In (18), the clitic precedes the verb root, following the negative prefix *na-*.
➡ In (19), the clitic follows the progressive prefix *da-*.
➡ In (20), the clitic follows the negative prefix *na-,* but precedes the progressive prefix *da-*.

⇨ Importantly, we cannot capture the position of the clitic simply by positing that it attaches to some position in the V^0 complex, since the position is highly variable.

There are predictable patterns in the chaos however, which make it possible to account for the position of the clitic in a relatively simple manner. The following generalizations hold:

- The clitic generally appears in second position within the verb form (Samvelian 2007).
- Clitics never appear between the verb and participle marker *-bu* (Samvelian 2007)
- Clitics that are 3rd singular appear after all object markers (Samvelian 2007, Walther 2012)

⇨ Samvelian (2007) notes all the above patterns, but rejects any kind of an analysis based on 'second position' since it would fail to account for cases where the clitic lies away from second position, such as in (16) and (17).

I propose that we take 'second position' seriously, since being in the second position in some domain is a pervasive pattern of clitic placement found in languages worldwide.

The challenge then becomes how to account for this *and* account for the cases where the clitic appears away from second position.

We can give the following descriptive rules for the placement of Sorani Kurdish clitics:

(21) a. When there is material available (eg, a DP or PP), clitics attach to the first phrasal element within VP.
b. If no material is available, then *all else being equal,* the clitic will appear in the second position within the verb form.

Departures from second position can be seen in two ways. Firstly, there is a morphotactic requirement in Sorani Kurdish where adjacency of the verb stem and the participle marker cannot be interrupted.

(22) *STEM-X-PARTICIPLE

Secondly, there is a requirement that 3rd singular subject markers follow object agreement.[4]

(23) 3SG subject markers follow object markers

[4] This rule can't be written as saying that 3rd singular clitics go to the end, since in the present tense, 3rd singular clitics appear happily in second position (when, because of the ergative split pattern, they mark the object agreement).

2.3. Preliminary account.

I assume that the (relevant) clitics in Sorani Kurdish are placed quite late on in the derivation. They are placed only after all other elements within V^0 have undergone linearization.

Concretely, I assume that the rule that linearizes clitics is of the following type:

(24) Linearization of subject clitics: elsewhere case
clitic + $[_{V^0}$ X ≫ (Y)...] → $[_{V^0}$ X ≫ clitic ≫ (Y)...]

This rule basically states that when a subject clitic comes to be linearized within V^0, then it linearizes to the right of whatever element is leftmost within the linearized complex head.

This rule can be overridden under certain circumstances. For instance, if there is a more specific ordering rule, as is the case when the clitic is 3rd singular and there is an object marker. Recall, that in this case, the clitic will follow the object marker:

(17) xward-bû-in-**î**
eat.PAST-PART-3PL-**3SG**
'He had eaten them'

(24') Linearization of subject clitics: 3sg/obj
$clitic_{3sg}$ + $[_{V^0}$ X ... Y_{AgrO}] → $[_{V^0}$ (X) ≫ Y_{AgrO} ≫ clitic]

Thus, when the complex V^0 element comes to be linearized as in (17), the more specific rule of (17) will apply, placing the clitic after the object agreement, instead of second place in the verb.

Thus, we have the following derivations:

(25) dît-**yân**-im
saw-3PL-1SG
'They saw me'
yân + $[_{V^0}$ dit ≫ im] → $[_{V^0}$ dît-yân-im]
- by application of (24).

(26) da-m-xwârd
PROG-1SG-east.PAST
'I was not eating.'
m + $[_{V^0}$ da ≫ xwârd] → $[_{V^0}$ da-m-xwârd]
- by application of (24).

(27) xward-bû-in-**î**
eat.PAST-PART-3PL-**3SG**
'He had eaten them'
î + $[_{V^0}$ xward ≫ bû ≫ în] → $[_{V^0}$ xward≫bu≫in≫î]
- by application of (24')

In order to capture the cases where the clitic appears in third position within the word, since appearing in second position would interrupt the adjacency of the verb and its participle marker, I take a slightly different tack.

I assume that what underlies this is the following morphotactic condition that is operative on surface representations in Sorani Kurdish:

(28) *STEM - X - PARTICIPLE

This forces the stem and participle to be adjacent in the word. Anything placed in this position will need to be moved before the form is pronounced, otherwise it will violate (28).

We see that this is operative in (16) above (repeated below):

(16) xward-bû-**man**-in
eat.PAST-PART-**1PL**-3PL
'We had eaten them'

The rule that linearizes the clitic places it between the stem *xward* and its participle *bu.*

This would give us the configuration, which violates (28):

(29) *xward-man-bû-in

In these instances, I assume that a rule of morphological metathesis applies and moves the clitic one morpheme over, allowing *xward* and *bu* to be adjacent.

This gives us the following derivation for *xward-bû-man-in*:

(30)	i. linearization of elements:	xward-bû-in
	ii. cliticization:	xward-man-bû-in
		****VIOLATION OF (28)****
	iii. violation repair:	xward-~~man~~-bû-man-in
	iv. output:	xward-bû-man-in

3. Interim summary

In the above account of Sorani Kurdish I proposed that there is competition among different linearization rules in the language, and that more specific rules win out over less specific ones, in the spirit of Distributed Morphology (Halle & Marantz 1993).

I also proposed that there is a derivational component to the spell out of words, and that certain morphemes can be placed in one position before moving to another position based on competing requirements on the output.

Nothing in Sorani Kurdish forced us to the conclusion that there was a derivational component to the spell-out of words; everything said above could be just as easily handled in a representational framework such as Optimality Theory (Prince & Smolensky 1993)

Questions to consider:
1. Is there anything to prefer a derivational theory over a representational one?
2. Is the morphotactic rule of (28) a desirable rule to have in the grammar.

In the following section, I show that Udi bears on the discussion of both questions, and that the answer to both is yes.

4. The predictable origins of chaos part 2: Udi[5]

Udi (Nakh-Dagestanian), as described in Harris (2000, 2002) has an even more complicated system of placing subject clitics than Sorani Kurdish.

To give some indication of what we're up against when trying to predict the position, consider the following four sentences:

(31) a. baba-n ęš nut eč-al-**le** k'wa
father-ERG apple.ABSL NEG bring-FUTII-3SG house.DAT
'Father will not bring apples to the house.'

b. nana-n ten-**ne** bụγa-b-e p'ạ ačik'alšey
mother-ERG NEG-3SG find-DO-AORII two toy.ABSL
'Mother did not find two toys.'

c. äyel kala-**ne**-bak-e
child.ABSL big-3SG-BECOME-AORII
'The child grew up.'

d. q'ačaγ-γ-on bez tänginax baš-**q'un**-q'-e
thief-PL-ERG my money.DAT steal$_1$-3PL-steal$_2$-AORII
'The thief stole my money.'

[5] For alternate analyses of the Udi data, in a variety of frameworks, see Cryssman (2000), Spencer & Luís (2006) and Wescoat (2009). All Udi data are taken from Harris (2000, 2002).

- In (31a) the clitic appears enclitic to the verb.
- In (31b) the clitic appears enclitic to the negative marker *ten.*
- In (31c) the clitic appears as an endoclitic within the verb, between the incorporated adjective and the verbal root.
- In (31d) the clitic appears as an endoclitic inside the verbal root, which is realized discontinuously.

The offenders:

	Absolutive/Ergative	Dative
1SG	-zu, -z	-za
2SG	-nu, -n, -ru, -lu	-va
3SG	-ne, -le, -re	-t'u
1PL	-yan	-ya
2PL	-nan, -ran, -lan	-va, -vạn
3PL	-q'un	-q'o

Amongst the chaos though, there is some order. Harris gives the following list of ranked rules that serve to predict the position of the clitic:

Rule 1: Clitics are final in the V*x*[6] if the verb is in the future II, the subjunctive I, the subjunctive II, or the imperative.
Rule 2: Clitics occur enclitic to a focused constituent.
Rule 3: In clauses with zero copulas, clitics are enclitic to predicate nominals.
Rule 4: Clitics are endoclitic in a complex verbstem, occurring between the Incorporated element (IncE) and the light verb or verb root.
Rule 5: For verbstems of class M, in the intransitive, clitics are endoclitic occurring between the verbstem and the present tense marker.
Rule 6: With verbs forms of category A and category B, clitics are enclitic to the entire verb form.
Rule 7: Clitics are endocliticized immediately before the final consonant in monomorphemic verbstems.

Application of rule 1 prevents rule 2 from applying.

So, if the verb is in the future II tense, subjunctive I,II or imperative form, then the subject marker appears enclitic to the verbal complex (hence application of rule 1):

[6] Harris uses the notation V*x* to mean the complex consisting of the verb and negative.

(32) a. baba-n ẹš nut eč-al-**le** k'wa
father-ERG apple.ABSL NEG bring-FUTII-3SG house.DAT
'Father will not bring apples to the house.'

b. nu aq'-a-**n** box-ala k'ok'oc'-ax
NEG take-SUBJI-2SG boil-PTCPL chicken-DAT
'You should not take the chicken that it to be cooked.'

If none of these TAM categories are present, then the clitic attaches to the constituent in focus (rule 2):

(33) nana-n ten-**ne** bụγa-b-e p'ạ ačik'ališey
mother-ERG NEG-3SG find-DO-AORII two toy.ABSL
'Mother did not find two toys.'

And so on...

4.1. Udi endoclitics....

What I will focus on in this talk are the contexts described by rules 4 through 7, since they (will be shown to) all derive from the same default rule of clitic placement:

Rule 4:	Clitics are endoclitic in a complex verbstem, occurring between the Incorporated element (IncE) and the light verb or verb root.
Rule 5:	For verbstems of class M, in the intransitive, clitics are endoclitic occurring between the verbstem and the present tense marker.
Rule 6:	With verbs forms of category A and category B, clitics are enclitic to the entire verb form.
Rule 7:	Clitics are endocliticized immediately before the final consonant in monomorphemic verbstems.

Data showing the constraints at work are given in the following. Firstly, rule 4:

Rule 4:	Clitics are endoclitic in a complex verbstem, occurring between the Incorporated element (IncE) and the light verb or verb root.

(34) äyel kala-**ne**-bak-e (incorporated adjective)
child.ABSL big-3SG-BECOME-AORII
'The child grew up.'

(35) nana-n tur-ex oc'-**ne**-k'-e (incorporated verb)
mother-ERG foot-DAT wash-3SG-LV-AORII
'Mother washed her foot.'

(36) pasčaɣ-on ɣar-muɣ-on lašk'o-**q'un**-b-esa (incorporated noun)
king-GEN boy-PL-ERG wedding-3PL-DO-PRES
'The king's son's married.'

(34-36) show that in the absence of focus and relevant TAM suffixes, the clitic places itself in between the incorporated element and light verb.

When none of the other alignment rules apply, rule 7 kicks in and the clitic is placed inside the verbal root by the Align-PM-verbstem constraint:

> Rule 7: Clitics are endocliticized immediately before the final consonant in monomorphemic verbstems.

(37) a. q'ačaɣ-ɣ-on bez tänginax baš-**q'un**-q'-e
thief-PL-ERG my money.DAT steal$_1$-3PL-steal$_2$-AORII
'The thieves stole my money.'

b. kaɣuz-ax a-**z**-q'-e
letter-DAT receive$_1$-1SG-receive$_2$-AORII
'I received the letter.'

The paradigm governing rule 5 is the alternation between transitive and intransitive forms. Some examples are listed below. The *a* examples are transitive, *b* examples intransitive:

> Rule 5: For verbstems of class M, in the intransitive, PMs are endoclitic occurring between the verbstem and the present tense marker.

(38)

Transitive	Intransitive
a. a-**t'u**-k'-sa see$_1$-3SG-see$_2$-PRES 'he sees'	b. ak'-**ne**-sa see-3SG-PRES 'it shows, is visible'
(39) a. bo-**ne**-x-sa boils$_1$-3SG-boils$_2$-PRES 'he boils, cooks'	b. box-**ne**-sa boils-3SG-PRES 'it boils (intr.)'
(40) a. bǫ-**ne**-q'-sa gather$_1$-3SG-gather$_2$-PRES 'he gathers'	b. bǫq'-**ne**-sa gather-3SG-PRES 'it gathers, is gathered'

Finally, contexts in which rule 6 are applicable are the following:

> Rule 6: With verbs forms of category A and category B, PMs are enclitic to the entire verb form.

Some examples of these forms are given below; note the final position of the clitic (boldfaced) within the verbal complex:

(41) a. b-esa-**ne**
make-PRES-3SG
'she makes'

b. k-e-**ne**
eat-AORII-3SG
'she ate'

(42) bi-esa-**zu**
die-PRES-1SG
'I am dying'

4.2. ....are like Sorani Kurdish endoclitics

In what follows, I will show that the cases covered by Harris' rules 4-7 are all covered by a default rule of placement, shown in (16d) below:

(43) a. Clitics are enclitic to the TAM categories Future II, subjunctive I, subjunctive II and imperative. (= Harris' Rule 1)
b. Clitics are enclitic to focus. (= Harris' Rule 2)
c. Clitics are enclitic to predicate nominals. (= Harris' Rule 3)
d. Elsewhere, clitics appear second position within the verb.

Note that (43d) is the same as the placement of Sorani Kurdish clitics given above.

I concentrate on how (43d) interacts with other aspects of Udi morphotactics in order to produce the illusion of endoclisis.

(43d) immediately accounts for the placement of the clitic in complex verbs, since it will be positioned after the incorporated element:

(44) pasčaγ-on γar-muγ-on lašk'o-**q'un**-b-esa
king-GEN boy-PL-ERG wedding-3PL-DO-PRES
'The king's son's married.'

(45)

Input to cliticization		Output of cliticization
lašk'o-b-esa	→	lašk'o-**q'un**-b-esa
wedding-DO-PRES		wedding-3PL-DO-PRES

Second position placement in the case of the simplex root *bak* (‘be’) would give the (ungrammatical) following:

(46) *bak-**ne**-sa sa paščʼaγ-kʼena adamar.
be-3SG-PRES one king-like person.ABSL

⇨ Something more needs to be said here.

The clitic goes obligatorily inside the root, so we find (47) instead:

(47) ba-**ne**-k-sa sa paščʼaγ-kʼena adamar.
be$_1$-3SG-be$_2$-PRES one king-like person.ABSL
‘[Once upon a time, there] is a person like a king.’

⇨ Here I propose that the clitic gets placed intramorphemically due to a confluence of three factors:

i. second position placement of the clitic.
ii. second position placement in these cases causes a morphotactic violation.
iii. the violation triggers a repair.

The morphotactic violation caused by clitic placement is that it interrupts a requirement of Udi that verb root and TAM suffix are adjacent:

(48) *root-X-TAM

⇨ This requirement is surface true in Udi, since there exists no case where the verb and TAM suffix are non-adjacent.

Placing the clitic in second position violates this adjacency, as we see above:

(49) *bak-ne-sa
root-clitic-TAM

When this adjacency is violated, the morphology enacts a repair to allow convergence. The chosen repair for Udi is metathesis, which moves the clitic to allow the right edge of the root to be adjacent to the TAM suffix.

⇨ Morphological metathesis as a repair has been recently argued for in Arregi & Nevins (2012) for Basque clitics (see also Harris & Halle 2005).

➩ I assume that metathesis moves the clitic the minimal amount required to allow the (right edge of the) root and TAM suffix to be adjacent, thus adhering to the morphotactic requirements of Udi.

➩ Further I assume that this violation is evaluated at the point of spell out of the clitic (I return to this below).

➩ Assuming that vocabulary insertion (VI) proceeds from the root outwards (Embick 2010, Bobaljik 2012 *inter alia*), this means that at the point where the morphotactic violation is seen, the phonological features of the root are in the derivation.

For a sample derivation consider how *ba-ne-k-sa* (be_1-3sg-be_2-pres) in (20) above is derived:

(50)	i. input to cliticization:	√BE-[+PRES]
	ii. second position placement:	√BE-[3SG]-[PRES]
	iii. VI of root:	/bak/-[3SG]-[+PRES]
	iv. VI of clitic:	/bak/-/ne/-[+PRES]
	v. metathesis repair:	/ba-ne-k/-[+PRES]
	vi. VI of TAM:	/ba-ne-k-sa/

Below is the derivation for the verb form *a-z-q'-e* ($receive_1$-1SG-$receive_2$-AORII) in the following:

(51) kaγuz-ax a-**z**-q'-e
letter-DAT $receive_1$-1SG-$receive_2$-AORII
'I received the letter.'

(52)	i. input to cliticization:	√RECEIVE-[+AORII]
	ii. second position placement:	√RECEIVE-[1SG]-[+AORII]
	iii. VI of root:	/aq'/-[1SG]-[+AORII]
	iv. VI of clitic:	/aq'/-/z/-[+AORII]
	v. metathesis repair:	/a-z-q'/-[+AORII]
	vi. VI of TAM:	/a-z-q'-e/

4.3. Transitive/intransitive alternations are the same.

The elsewhere placement rule of the clitic allows us to make sense of the intransitive/transitive alternations in (38-40), repeated below:

(53)

Transitive	Intransitive
a. a-**t'u**-k'-sa	b. ak'-**ne**-sa
see_1-3SG-see_2-PRES	see-3SG-PRES
'he sees'	'it shows, is visible'

(54) a. bo-**ne**-x-sa
boils$_1$-3SG-boils$_2$-PRES
'he boils, cooks'

b. box-**ne**-sa
boils-3SG-PRES
'it boils (intr.)'

(55) a. bọ-**ne**-q'-sa
gather$_1$-3SG-gather$_2$-PRES
'he gathers'

b. bọq'-**ne**-sa
gather-3SG-PRES
'it gathers, is gathered'

Harris shows that the intransitive variants are formed with a suppletive light verb *go*, which is phonologically null in the present tense. Its presence can be seen by changing the tense of the verb:

(56) a. box-**eɣ**-al-le
boil-GO-FUTII-3SG
'it will boil'

b. box-ne-**c**-e
boil-3SG-GO-AORII
'it boiled'

Clitic placement is then in the regular second position and follows the complex verb analysis of section 4.1 (27 below the derivation of (10b)):

(57)

Input to cliticization		Output of cliticization
ak'-∅-sa	→	ak'-ne-∅-sa
see-GO-PRES		see-3SG-GO-PRES

The transitive variants then involve second position placement, and leftward metathesis as in section 4.2.

(58)

Input to cliticization		Adjacency violation and repair		Output of cliticization
ak'-sa	→	*ak'-t'u-sa	→	a-t'u-k'-sa
see-PRES		*see-3SG-PRES		see$_1$-3SG-see$_2$-PRES

⇨ The cases covered by Rule 5 then are completely regular, following the same derivations as the complex verbs for the intransitive forms, and simplex verbs for the transitive.

4.4. The final bit of the Udi puzzle

The final thing which we need to explain is why in some cases the clitic appears at the end of the verb form altogether, and not in third position.

We saw this above in Sorani Kurdish, however there it was the result of the linearization rule (24'), which placed 3sg clitics after object markers.

Such a linearization rule is not much help here; the controlling environments are phonological in nature.

➪ Rule 6 only applies in cases where the verbal root is either C or CV in shape.

Under DM assumptions, linearization happens *before* vocabulary insertion. Therefore, the phonological shape of the exponents shouldn't make a difference. Anything that makes reference to phonology must be done after VI of the relevant items.

Therefore, it is not possible to write a rule as follows, where X^{CV} indicates that it has the phonological shape CV:

(59) **<u>Linearization of subject clitics:</u>
**clitic + $[_V{}^0\ X^{CV}] \rightarrow [_V{}^0\ X^{CV} \gg$ clitic ...]

To return to Udi, here are the relevant contexts

(60) a. b-esa-**ne**
make-PRES-3SG
'she makes'

b. k-e-**ne**
eat-AORII-3SG
'she ate'

(61) bi-esa-**zu**
die-PRES-1SG
'I am dying'

None of these forms involves a light verb, so they are all simplex verb forms. In the approach above, we seem to predict that the clitic would go inside the verb root.

This clearly isn't possible with roots that are formed of a single consonant.
Nor is it apparently possible with roots that are CV in shape.

The question is, why does the clitic go right to the end of the verb?

The answer, I propose lies in how the repairs forced by second position placement operate when confronted with C or CV roots.

Assume:
1. There is a general prohibition on proclisis in Udi.
   - This is at least surface true in Udi, since there are no proclitics in the language.
2. Also, assume that metathesis cannot apply to open syllables.[7]

[7] This amounts to saying that open syllables cannot host clitics, which is an equally required constraint in Harris' approach.

Firstly consider the derivation for *k-e-ne* ('he eats') (from (29b) above):

(62)	i. input to cliticization:	√EAT-[+PRES]
	ii. second position placement:	√EAT-[3SG]-[+PRES]
	iii. VI of root:	/k/-[3SG]-[+PRES]
	iv. VI of clitic:	/k/-/ne/-[+PRES]
	v. metathesis repair:	/k/-[+PRES]-/ne/
	vi. VI of TAM:	/k-e-ne/

At the point where the metathesis repair would be triggered, moving the clitic leftward in the regular manner would cause it to be a proclitic. Therefore, metathesis instead applies rightwards moving the clitic outside the TAM suffix.

⇨ Crucially this leaves the right edge of the root to be adjacent to the TAM suffix, satisfying the morphotactic requirements of Udi, with the output being /k-e-ne/ (eat-pres-3sg).

An interesting consequence of the above is that we can see why the clitic goes to the final position in the verbal complex, even when the exponent of TAM is in principle large enough to host the clitic.

⇨ For instance in (61) above, where *-esa,* the exponent of present tense is the correct syllable structure to host a clitic (it has a consonant), yet we don't find *bi-e-zu-sa* (= die-pres$_1$-1sg-pres$_2$)

This is because at the point that the metathesis repair is enacted, only the morphosyntactic features of the TAM suffix are present within the derivation. As the phonological exponent of TAM is missing, the operation literally cannot place the clitic inside the phonological information of TAM (in the spirit of Bobaljik 2000).

This is shown in the derivation of *bi-esa-zu* (die-pres-1sg) below:

(63)	i. input to cliticization:	√DIE-[+PRES]
	ii. second position placement:	√DIE-[1SG]-[+PRES]
	iii. VI of root:	/bi/-[1SG]-[+PRES]
	iv. VI of clitic:	/bi/-/zu/-[+PRES]
	v. metathesis repair:	/bi/-[+PRES]-/zu/
	vi. VI of TAM:	/bi-esa-zu/

4.5. Discussion

The analysis presented here for Udi crucially relies upon the phonological information of the root being available before the phonological information of some affixes.

As shown, this is easily captured within a derivational framework where the clitic is placed in a certain position before being moved elsewhere. Due to the nature of VI, such that it proceeds outwards from the root, we can explain not only why Udi clitics appear at the edge of the verb form, but also why they never appear internal to the TAM morpheme.

The key features of Udi, which allow us to conflate rules 4-7 of Harris into one default rule of clitic placement are as follows:

- All else being equal, Udi subject clitics are placed second position within the verb.
- Udi TAM suffixes must be adjacent to their verbal root.
- In cases where the root-TAM adjacency is interrupted, clitics are metathesized away.
- Proclisis is not allowed in Udi.

All the above conspire to produce a highly mobile clitic system.

Furthermore, they conspire to produce a pattern that is elsewhere unattested - clitics appearing internal to another morpheme.[8]

5. Conclusions

In this talk I have outlined two complex systems of clitic placement, which each involve the rare phenomenon of endoclisis.

I have shown that in both Sorani Kurdish and Udi the same phenomenon is in operation - second position placement within a word.

Deviations from second position in the word in both languages are forced by (i) competition among linearization rules (Sorani Kurdish) and (ii) morphotactic considerations (both languages).

Questions for future research:
How to relate the Udi and Sorani Kurdish data into general theories of second position placement.
What is the impact of the analysis given here on theories of cliticization in general, for instance Anderson's (1992) proposal that clitics are phrasal affixes.

[8] This is turn has desirable consequences, as is discussed in Smith (to appear), where it is argued that Harris' system of OT constraints posited for Udi adds into the theory of UG a massively powerful device - the ability of the grammar to place clitics within another morpheme. It is argued that this is too powerful since there exists no other attested case of intramorphemic clitic placement. The conclusion drawn is that without further cases of intramorphemic clitic placement, the theory proposed here of intramorphemic placement being essentially an epiphenomenon is to be preferred.

Peter W. Smith
University of Connecticut
peter.w.smith@uconn.edu
http://homepages.uconn.edu/~pws10003

References

Anderson, S. (1992). Amorphous morphology.

Arregi, K. and Nevins, A. (2012). Morphotactics: Basque auxiliaries and the structure of spellout. Springer.

Bobaljik, J.D. (2000). The ins and outs of contextual allomorphy. In K. Grohmann and C. Struijke, eds., *University of Maryland Working Papers in Linguistics,* volume 10, 35-71.

Bonami, O. & Samvelian, P. (2008). *Sorani Kurdish person markers and the typology of agreement.* Paper presented at the 13th International Morphology Meeting, Vienna.

Crysmann, B. (2000). On the placement and morphology of Udi Subject Agreement. http://www.coli.uni-sb.de/~crysmann/papers/UdiAgr.html

Embick, D. (2010). Localism versus globalism in morphology and phonology. MIT Press.

Bonami, O. & Samvelian, P. (2008). *Sorani Kurdish person markers and the typology of agreement.* Paper presented at the 13th International Morphology Meeting, Vienna.

Halle, M. and Marantz, A. (1993). *Distributed morphology and the pieces of inflection.* In Hale, K. and Keyser, S. (eds.), The view from building 20. Cambridge, MA: MIT Press.

Harris, A. (2000). Where in the word is the Udi clitic? *Language 76, 593-616.*

Harris, A. (2002). *Endoclitics and the origins of Udi morphosyntax.* Oxford, Oxford University Press.

Harris, J. & Halle, M. (2005). Unexpected plural inflections in Spanish: Reduplication and metathesis. *Linguistic Inquiry 36,* pp.195-222.

Kaisse, E. (1981). Separating phonology from syntax: a reanalysis of Pashto cliticization. Journal of Linguistics 17, 197-208.

Kern, G. (2011). Incliticization? Apparent word-internal clitic placement in Pashto and Old Irish. Ms. MIT.

Klamer, M. (1997). Spelling out clitics in Kambera. *Linguistics 35,* 895-927.

Legate, J.A. (2008). Warlpiri and the theory of second position clitics. *Natural language and linguistic theory 26,* 3-60.

Luís, A. (2004). Clitics as morphology. Doctoral dissertation, University of Essex.

Luís, A. & Spencer, A. (2005). Udi clitics: A generalized paradigm function morphology approach. Available at http://privatewww.essex.ac.uk/~spena/papers/UDI.PDF

McCarus, E. (1958). A Kurdish Grammar: descriptive analysis of the Kurdish of Sulaimaniya, Iraq. Doctoral dissertation, American Council of Learned Societies.

Prince, A. & Smolensky, P. *Optimality theory: constraint interaction in generative grammar.*

Samvelian, P. (2007). What Sorani Kurdish absolute prepositions tell us about cliticization. *Proceedings of the Texas Linguistic Society IX.*

Tegey, H. (1977). *The grammar of clitics: evidence from Pashto and other languages.* Doctoral dissertation, University of Illinois at Urbana-Champaign.

Thackston, W., M. (2006). *Sorani Kurdish: A reference grammar with selected readings.* Available online.

Walther, G. (2012). Fitting into morphological structure: accounting for Sorani Kurdish endoclitics. http://www.linguist.univ-paris-diderot.fr/~gwalther/homepage/Publications_(fr)_files/MMM8_GWalther.pdf

Wescoat, M. (2009). *Udi person markers and lexical integrity. Proceedings of the LFG09 conference.*

Zwicky, A. & Pullum, G. (1983). Cliticization vs inflection: English n't. *Language 59, 502-513.*