Non-peripheral cliticization and second position in Udi and Sorani Kurdish
Peter W. Smith, University of Connecticut

1. Introduction

This paper deals with a number of cases of clitics which do not fit the usual patterns of cliticization seen across natural languages. Overwhelmingly, we observe that clitics tend to be either *proclitics,* where they linearly precede their host word, or *enclitics*, where they follow their host word. Instances of each are shown below in (1) and (2) respectively:

(1) uraŋ   **k-mpu**-ŋa-tkam-t   (Yimas, Foley 1991:208)
     coconut 3SG-3PL-1SG-show-PERF
     'They showed me the coconut.'

(2) Karinganta-**rna** kuyu-jarra   yampi-ja-rni.   (Warlpiri, Legate 2008:14)
     fact-1SG.SUBJ   meat-DL   leave-PAST-hither
     'The fact is I left two animals (I speared) and came here.'

The overwhelming majority of the world's clitics can be categorized as enclitic or proclitic, and it is often assumed that these are the only two types of clitics that exist. Indeed, Zwicky & Pullum (1983), in trying to provide a uniform set of criteria for distinguishing clitics from affixes, propose that it is a property of clitics that they attach to the peripheries of words, and never inside affixes, since their criterion F states "clitics can attach to material already containing clitics, but affixes cannot (Zwicky & Pullum 1983:504)." In fact, this observation that clitics appear outside of affixes has guided a lot of analyses on clitics and it is commonly assumed that word-internal clitics do not actually exist, and that cases that are found ought to be reanalyzed as other phenomenon. For instance, Nevis (1984, 1988, and Macauley (1989) both show that purported instances of a clitic which does appear word-internally in Estonian, Northern Saame and Karok respectively, are all in fact not cases that don't involve clitics at all. Furthermore, word-internal cliticization seems to cause serious theoretical problems for a variety of frameworks and approaches. As noted by Harris (2002), approaches which assume Lexical Integrity (that no syntactic operation can affect the internal structure of a word, DiSciullo & Williams 1987) would struggle to accommodate clitics being placed inside a word. Similarly, the idea that clitics are phrasal affixes (Anderson 1992, Everett 1996) also appears incompatible with this observation (though see Anderson 2005).

However, the notion that clitics can appear only at the peripheries of words seems to undersell the possibilities of clitics. We might expect that clitics show the same variation of affixes, given that both categories are in some way bound morphemes, and they represent the same information cross-linguistically. And whilst there are of course prefixes and affixes found in natural language, we also find cases of infixation, whereby an affix appears internal to another morpheme, such as in the following data from Oaxaca Chontal (from Yu 2007). Here, the plural marker *-ł* infixes before the final syllable of the singular form:

(3) SINGULAR   PLURAL   MEANING
     akanˀoʔ   →   akałnˀoʔ   'woman'   Oaxaca Chontal
     teʔa   →   tełʔa   'elder'

Since infixation exists, then *prima facie* we might expect to find cases of clitics which appear internal to their host. In fact, it seems that we do find some cases of these, however, they are extremely rare. We find two examples of clitics which appear internal to their host. Firstly, there are cases of so called *mesoclisis*, which for the purposes of this paper I define as follows:

(4) A mesoclitic is a clitic which is non-peripheral, but appears intermorphemically within its host.

Thus, by (4), we see mesoclisis when a clitic appears between the head of its host, and affiliated affixes. We find mesoclisis in European Portuguese (from Anderson 2005), (some) dialects of Italian (Manzini & Savioa 2011) and Sorani Kurdish (Thackston 2006:47):

(5) dár-**te**-íamos   (European Portuguese, Anderson 2005:158)
     give-2SG-1PL.COND
     'We would give it to you.'

(6) tʃirka-**'m**-itə   le   (Senise, Manzini & Savoia 2011:1104)
     ask-1SG-2PL   it/them
     'Ask me for them/it.'

(7) dît-**yân**-im   (Sorani Kurdish, Thackston 2006: 47)
     see.PAST-3PL-1SG
     'They saw me.'

All of these morphemes appear internal to where we might otherwise expect them to be. None of these morphemes are in a word peripheral position, and all appear in the position between the verb stem and its inflectional affixes.[1] This runs counter to the observation of Zwicky & Pullum noted above, since the affixes will have attached to the [head+clitic] item. In contrast to the case of infixation given above, the clitic does not in fact disrupt the integrity of any single morpheme.
     A different type of non-peripheral cliticization that will be discussed in this paper is where a clitic does interrupt the integrity of another morpheme. This type of cliticization I will term *endoclisis* in order to distinguish it from mesoclisis.

(8) An endoclitic is a non-peripheral clitic, but appears intramorphemically within its host.

Thus, in order to be an endoclitic, a clitic must appear in a position away from the peripheries of the word, but also internal to another morpheme, causing that morpheme to be realized discontinuously. This type of clitic is much rarer than mesoclisis, which is rare to begin with, and is only found in a handful of cases around the world. The most established proponent of this type of clitic comes from Udi (Harris 2000, 2002), where the clitics that mark the features of the subject can at times appear internal to the verb root. In both (9a,b), we see that the subject clitic forces the verb root to be realized discontinuously:[2]

(9) 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 thief stole my money.'

_____

[1] All of the final morphemes in question are affixes. Sorani Kurdish will be discussed in some detail below, where it will be shown that, in the above example, *yân* is a clitic whilst *im* is an affix.

[2] In the course of this paper, I adopt Harris' glossing conventions where the parts of discontinuous morphemes are subscripted to indicate that they form separate parts of the overall morpheme. Thus, in *baš-**q'un**-q'-e* 'steal$_1$-3PL-steal$_2$-AORII', the verb 'to steal' is *basq'*.

b. kaɣuz-ax   a-**z**-q'-e
   letter-DAT   receive₁-1SG-receive₂-AORII
   'I received the letter.'

The subject clitics of Udi do not always do this, and at times they behave as an enclitic, (10):

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

However, in the cases of (9), the verb root is simplex, and not composed of distinct morphemes. Therefore, the clitics in Udi very clearly are endoclitics in these examples. Another purported case of endoclisis comes from Degema (Niger-Congo), as detailed in Kari (2003). In this language, the factative clitic is sometimes able to tuck inside another morpheme, as seen in (11):

(11) a. mɪ-tá-ꜜ**á**-m
        1SG-chew₁-FACT-chew₂
        'I chewed (sth.).'

   b. mi-ɓó-ꜜ**ó**-l
      1SG-hold₁-FACT-hold₂
      'I held (sth.).'

In (10), like in the Udi cases, the clitic appears inside the verbal roots, which are discontinuous. Without the incursion of the clitic, the roots in question are *tám* and *ɓól* respectively. The Degema facts are complex, and will be discussed in greater detail below, but what is interesting about this clitic is, like the subject clitics in Udi, we find that it does not always behave as an endoclitic, and in other cases is an enclitic:[3]

(12) mi-bí-ꜜ**ín**
     1SG-be.black-FE
     'I am black.'

(13) mɪ-sɪsɛ́-ꜜ**ɛ́n**
     1SG-remove-FE
     'I removed something.'

The interest from endoclisis comes from the fact that Udi and Degema seem to be the only cases that exist of a clitic appearing internal to another morpheme. Pashto is often held up as a third case of endoclisis, but it seems that this clitic too is a mesoclitic (see Tegey 1977, Kaisse 1981, Yu 2007). Endoclisis is therefore strikingly rare in natural language. This then raises the question of

---

[3] In all Degema examples, I follow the glossing convention of Kari (2003), where 'ꜜ' represents a downstep tone on the following vowel.

how it could possibly arise to begin with. Should we analyze endoclisis directly, allowing clitics to be directly placed inside other morphemes, or instead say that morpheme internal clitic placement is not directly generable, but endoclisis can arise indirectly by means of surface readjustment? In other words, can endoclisis arise due to the requirements of the clitic, or does it arise due to independent factors in the environment? The first option has been argued for by Harris (2002) on the basis of the Udi data (see also Yu 2007). One of the claims of this paper is that the second option is the correct analysis of endoclisis. Whilst mesoclisis will be treated as a possible clitic attachment site (though in a limited way), attachment inside another morpheme is not *directly* generable by UG. What this means is this is that *clitics cannot break the integrity of another morpheme when they are initially positioned*; they cannot subcategorize to appear internal to another morpheme. Thus, the central claim of this paper is as follows:

(14) The internal structure of a morpheme is opaque to a clitic when it satisfies its own positioning requirements.

This claim, taken here to reflect a universal part of UG, strongly restricts the potential places where a clitic can occur. Taken in its strongest form this entails that *all* clitics will appear at the peripheries of morphemes; and there cannot be a clitic equivalent to infixation (notably *contra* Anderson 2005). As it turns out, this is slightly too strong, since there do exist a couple of reported cases where a clitic does break the integrity of another morpheme as discussed above. Note though the qualification in (14) that the internal structure of a morpheme is opaque to clitics when satisfying *their* positioning requirements. I will show that clitics can be displaced from the position in which they are placed in order to satisfy demands that follow the initial placement of the clitic, for example the morphotactic requirements of the host word (in accordance with Arregi & Nevins' 2012 proposal for Basque). Thus, any time that we find a clitic that *does* break the integrity of a morpheme, it will necessarily be the case that something has forced the clitic into that position. I support this claim primarily with discussion of the only clearly reported cases of endoclisis that I am aware to have been reported, Udi (Harris 2002) and Degema (2003). Whilst the discussion centers around Udi, I also further support he claims made in the paper with discussion of two other cases of non-peripheral cliticization, Sorani Kurdish (Samvelian 2007) and Pashto (Tegey 1977).

The paper is organized as follows. In section 2 I discuss to the first main case study of this paper, Udi, and provide an analysis of the endoclisis in this language as being mesoclisis followed by an instance of morphological metathesis, which in certain contexts places the clitic inside the verbal root in order to appease the morphotactic requirements of Udi. Mesoclisis is seen in Udi since the clitics target the second morphemic position within the verb (in the default case). That is, Udi endoclitics will be shown to be a special case of second position clitics. In section 3, I turn to the second major case study of the paper, Sorani Kurdish, to motivate the view that there are clitics which are able to target the second morphemic position within the word, and create mesoclisis in this way. Section 4 discusses the issues in the paper in a general context. The crucial claim of the paper that endoclisis can *only* be derived through post-syntactic surface readjustments, whilst *mesoclisis* is directly generable (in the sense that some clitics can inherently target the internal structure of words) will be discussed with respect to two further languages that show apparent word-internal clitics, Pashto and Degema. Finally, I conclude the paper in section 5.

2. Case Study 1: Udi[4]

In this section I present the first case study of a non-peripheral clitic. The language in question is Udi, a North East Caucasian language spoken primarily in Azerbaijain. For any work that looks at the nature of cliticization, Udi is important because it contains an extremely complicated system of subject clitic placement, where the clitics in question alternate between being enclitics in some contexts, mesoclitics in other contexts and endoclitics in other cases still. As defined in the previous section, I make a difference between mesoclisis and endoclisis for exposition. Both of these clitics are clitics that appear non-peripherally inside the word, but the difference between the two is that mesoclitics are those that appear *inter*morphemically, that is, between identifiable morphemes within a complex word, whilst endoclitics are those that appear *intra*morphemically, breaking up the integrity of a single morpheme.

Udi is furthermore important because it contains the best established case of endoclitics that has been documented in the literature. As will be shown, Harris (2002) carefully documents the positions of the clitics and notes that they do split up individual morphemes, causing them to be realized discontinuously. Furthermore, she carefully distinguishes the behavior of these elements from affixes, and concludes that they are indeed clitics. Udi thus is a language involving non-peripheral cliticization *par excellence*, and clearly has great relevance for the discussion on possible and impossible positions for clitics.

2.1. The seven rules of clitic placement

The clitics that we will be discussing in Udi agree with the person, number and case features of the subject. As shown in (15), these clitics show a curious distribution. Two things are of note. Firstly, the clitics are not that picky in what they attach to, since they can cliticize to verbs, nouns and negation, showing the characteristic lack of selectional requirements that is familiar from other clitics. More curiously, the clitics also alternate between different positions within their host. As shown in (15a,b) the clitics are at times enclitics, but they also appear as a mesoclitic (15c) and an endoclitic (15d) (following Harris' convention, the parts of a discontinuous root will be glossed with a subscript). Not only do we see that the subject clitics can in principle be enclitic, mesoclitic or endoclitic, we see that all three of these types manifested within a verb, as can be seen by comparing (15a,c,d).:

(15) 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**   bųγa-b-e   p'ạ   ačik'alšey
          mother-ERG   NEG-3SG   find-DO-AORII   two   toy.ABSL
          'Mother did not find two toys.

      c. 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.'

---

    d.   q'ačaγ-γ-on   bez   tänginax   baš-**q'un**-q'-e
         thief-PL-ERG   my   money.DAT   steal₁-3PL-steal₂-AORII
         'The thieves stole my money.'

The clitics have different allomorphs, as shown in (16):

(16)

	General	**Inversion**	**Possession**	**Question**
1SG	-zu, -z	-za	-bez, -bes	
2SG	-nu, -n, -ru, -lu	-va	-vi	
3SG	-ne, -le, -re	-t'u	-t'a	-a
1PL	-yan	-ya	-beš	
2PL	-nan, -ran, -lan	-va, -vạn	-ẹf	
3PL	-q'un	-q'o	-q'o	

Within the cells of the table, the choice of allomorph is phonological, and will not be considered further here. The class labeled 'inversion' refers to clitics which mark subject experiencers, since they appear with psych-verbs. Thus, it may be possible to relabel the category as 'dative', since the subjects of these verbs in Udi are generally in the dative case. We could then further analyze the general category as 'absolutive/ergative'. However, it is important to stress that this does not refer to the case of the argument that the clitic agrees with, but the argument position of the argument that the clitic is associated with, since clitics can in principle show disagreement in case with their associated argument. Harris (1984) shows that with experiencer verbs, the clitic must always be drawn from the 'inversion' class, even if the experiencer subject is ergative, (17), leading her to claim that the allomorphy of the clitics is based on thematic role, and not case (Harris 2002:27,fn5). For the time being I leave the issue of whether the categories are best analyzed in terms of case or thematic role for future research:

(17)   Zu   a-**za**-ḳ-sa   šel   läzättu   pak.
        I.ERG   see₁-1SG-see₂-PRES good   pretty   garden
        'I see a good, pretty garden.'

Though the data look messy at first sight, Harris (2000, 2002) goes through in detail that there are predictable rules governing where the clitic will appear in a given sentence. She gives a list of seven hierarchically ordered rules that allow us to predict where the clitic will show up in the Udi sentence, which I repeat here (with minor modifications for the sake of exposition, see Harris 2002:130 for the non-abbreviated list):

(18) Rule 1: Clitics are final in the V*x*[5] 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 mesoclitic[6] in a complex verbstem, occurring between the Incorporated element (IncE) and the light verb or verb root.

Rule 5: For verbstems of class M[7], in the intransitive, clitics are endoclitic occurring between the verbstem and the present tense marker.

Rule 6: With verbs 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.

Rule 1 is exemplified by the following data below:

(19) 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. 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.'

Rule 2 is shown by the following data. It should be noted that Harris uses 'focused' constituent pragmatically, to refer to a variety of things that are essentially new information. Things which are taken to be in focus by Harris (and as she claims, are universally in focus) are focused arguments, question words and negation.

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

 b. met'in-al burim sa yärävgä q'əzəl-**le** aqsa
  this.ERG-AND apparently one bag.ABSL gold.ABSL-3SG take
  'She also apparently takes a bag of gold.'

 c. zu aba-za, šin-**a** be met'o
  I.ERG know-1SG who.ERG-Q do this.ABSL
  'I know who did this.'

There is a dedicated position for focused elements in Udi, and they appear immediately before the verb, as shown in (21), where the information that answers the question (thereby being focus, since

---

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

[6] Harris uses the term endoclitic here, but since it's used in a manner inconsistent with my terminology here, I have changed it to mesoclitic.

[7] 'Class M' will be discussed below. It is used as a grouping term, without further significance. Likewise for category A and B referred to in Rule 6.

it's new information in the sentence) must obligatorily appear *before* the verb. Simple adjacency is not sufficient, as shown by the ungrammatical A' in (21):

(21) Q: xinär-mux   ma-q'un   taysa?
         girl-PL.ABSL   where-3PL   go
         'Where are these girls going?'

     A: ok'tomber-a-q'un   taysa
         Ok'tomber-DAT-3PL   go
         'They are going to Ok'tomber.'

     A': *taysa Ok'tomber-a-q'un

The clitic is placed on the argument focus only if the verb is not in any of the TAM categories referred to in Rule 1, the imperative, subjunctive I and II or future II. When the verb is in any of these categories, it must obligatorily appear at the end of the verb, as described in Rule 1. Thus, Rule 1 takes precedence over Rule 2, as shown in (19) above.

    Rule 3 is shown in the following, but will not be discussed much throughout the paper. I give the data for completeness:

(22)   nana   k'wa-**ne**
       mother.ABSL   house.DAT-3SG
       'Mother is at the house.'

Rule 4 is the first rule in the list where the clitic begins migrating inside its host. Udi verbs come in two types, they are either complex, in which case they are formed by a light verb which incorporates some element, or they are simplex, where they simply consist of a verb. In complex cases, as Rule 4 states, the clitic appears between the light verb and the incorporated element (IncE). The structure of a complex verb is given in (23), and (24) shows the clitic in the intermorphemic position. Note that it is not important what category the incorporated element (IncE) is; a variety of elements can be incorporated by light verbs to form complex verbs, but the clitic appears uniformly in between the IncE and the light verb (in (24a) the IncE is an adjective, a verb in (24b) and a noun in (24c)):

(23)   IncE-(clitic)-light verb-TAM suffix

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

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

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

Rule 5 is shown by the intransitive/transitive alternations below. The term 'Class M' in Harris' rules refers only to the relevant group of verbs, and does not hold theoretical import:

(25)   <u>Transitive</u>

    a. a-**t'u**-k'-sa
        see$_1$-3SG-see$_2$-PRES
        'he sees'

                <u>Intransitive</u>

    b. ak'-**ne**-sa
        see-3SG-PRES
        'it shows, is visible'

(26) a. bo-**ne**-<u>x</u>-sa
        boils$_1$-3SG-boils$_2$-PRES
        'he boils, cooks'

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

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

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

As Harris shows, Rule 5 simply refers to what you see on the surface. The difference in transitivity on these verbs is formed syntactically by the addition of a light verb, which incorporates the simplex transitive verb. Thus, whilst it appears as though the difference in transitivity is marked in these verbs by moving the clitic, the real difference is one of simplex versus complex verbs. The simplex verbs in the transitive forms regularly follow the placement of clitics described in Rule 7, whilst the syntactically complex, intransitives regularly follow Rule 4. The underlying structure of 'boil (intr.)' is thus not (26b), but in fact (28). This structure is hidden because the light verb that forms the intransitive is suppletively null in the present tense. Changing the tense specification of this form reveals the hidden light verb, as shown in (29) (note that the clitic shifts in (29a) according to rule 1):

(28) box-ne-∅-sa
      boils-3SG-LV-PRES
      'it boils'

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

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

Rule 6 describes the situation where a simplex verb has the clitic appearing at the end of the verb form instead of inside the verb root, as would otherwise be expected by Rule 7. Illustrative examples are given in (40). Note that these verbs are different from the (b) cases in (25-27) above in that the verb is simplex, and there is no light verb in the structure. Note the AORII form in (30b), which would uncover the presence of the light verb if these also involved the same suppletive light verb, as the cases discussed above (c.f. 29b):

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

      b. k-e-ne
        eat-aorII-3sg
        'she ate'

   c. bi-esa-**zu**
      die-PRES-1SG
      'I am dying'

It is worthwhile contrasting these with the cases of Rule 7, which are given below, where the clitics appear internal to the verb root, causing the root to be realized discontinuously:

(31) a. q'ačaγ-γ-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. kaγuz-ax a-**z**-q'-e
        letter-DAT receive$_1$-1SG-receive$_2$-AORII
        'I received the letter.'

It is important to stress that these cases are different from the complex verb cases described above. Harris outlines that whilst there may be some simplex verbs which are analyzable as being historically complex, this is not the case today, synchronically they cannot be analyzed as such (contrary to a proposal by Luis & Spencer 2006).

2.2. The clitics are really clitics

Before getting into the analysis of these clitics, it is important to establish that they are indeed clitics and not affixes. Harris (2002) takes great care to establish that in all the applicable tests, they pattern as we would expect clitics to pattern, and not as we would affixes. There are of course no established fixed criteria for knowing what constitutes a clitic as opposed to being an affix, and so Harris goes through the tests that have been proposed by Zwicky & Pullum (1983), Klavans (1995) and Scalise (1984). For reasons of space I do not discuss the full set of data, and refer the reader to Harris (2002, chapter 5), but with respect to, for instance Zwicky & Pullum's criteria, Harris notes that the subject markers in Udi are unselective with respect to the category of their host (Zwicky & Pullum's condition A), do not show arbitrary gaps with respect to what they attach to (Condition B), do not have real morphophonological idiosyncrasies (Condition C), attachment of the clitic to the host does not give idiosyncratic meanings (Condition D), fail to behave like affixes in compounding (showing Condition E, effect of syntactic rules) and they follow other clitics in the word (Condition F). All of these tests show that the person markers show the prototypical properties of being clitics, and not affixes. In addition to these, the subject markers behave as clitics in all the other tests of Klavans (1985) and Scalise (1984). Thus, it is clear that they are clitics, and not affixes.

2.3. Rules 1-2: Leaning on something higher[8]

Udi presents two major challenges in providing an analysis of its clitic system. Firstly, it needs to be understood why the clitics will at times appear enclitic and at other times meso/endoclitic. Secondly, it must be understood why there is such a hierarchy of rules. Why for instance, does the

---

[8] Since rule 3 contains no competition with the verb, I omit discussion of it here. The position of the clitic suggests that in predicate constructions, the predicate nominal is merged above the clitic, which goes on the end of it. Either that, or the final placement is a true default given that there is no verb in the structure for the clitic to target if it hasn't found anywhere else.

verb being in the subjunctive form cause the clitic to not move onto negation, but verbs in the aorist do not? Here I propose that the answer lies in movement of the verb, which picks up the clitic and bleeds any leaning further leaning on of the clitic to focus. With respect to the difference between the Rule 1 TAM categories versus the rest of the verbs, I propose that the differences stem from the position of the verb; where the clitic encliticizes to the TAM categories (i.e. Rule 1), this is because these verbs have been forced to move higher than the clitic in order to enter into an Agree relation with a mood operator that is high in the tree. They are forced to move higher in order to comply with the PIC (Chomsky 2000, 2001). This feature is not present on verbs which are not of the relevant TAM categories. In these cases, the verb stays low, the clitic moves onto it post-syntactically and is positioned within the verbal form. This will be discussed in the next section.

With respect to the position of the clitic, I assume, for reasons that will become clear shortly, that clitics syntactically lie immediately above VP, below vP, in the head position of a CliticP (ClP) (see Sportiche 1996). I further assume that the clitics are base generated in this position away from their associated argument, and enter into an Agree relation with the subject prior to the subject's movement into the higher domain.[9,10] Focus in Udi has a close relationship with the verb, with focalized elements always appearing left adjacent to the verb, like as is the case in Turkish (Sener 2010). Thus, I assume that there is a low focus position in Udi (without making any commitment to whether or not there is also a high focus position in the language - for instance in CP). This focus position lies above ClP, but beneath vP, and elements in focus move syntactically to the specifier of FocP.[11] A low focus position has been independently proposed in other languages, see e.g. Belletti & Shlonsky 1995, Bošković 1997, 2014 Belletti 2004 and Bastos 2011 *a.o.* The position of objects in the structure is difficult to determine in Udi, due to the lack of data available to me and the fact that the basic order of the Udi is SOV. For clarity, I assume that during the syntax the object moves from the complement of V to some object position in the structure, either Spec,AgrO or Spec,vP. The nature of this movement is not clear at this point, and I leave further investigation to future research. For clarity, in what follows I assume that the object moves to Spec,vP.[12] The structure is thus:

---

[9] Note that this proposal is then inconsistent with the Big-DP analysis of cliticization.

[10] I assume that agreement of the clitic is post-syntactic (Bobaljik 2008, Author 2013) and therefore can go either up or down in the structure.

[11] As will be discussed later on, this must be coupled with an extra assumption that constrains PF representations such that focus must absolutely be left adjacent to the verb in Udi. Usually this is handled by the low position of focus, which ensures that any element will always be close to the verb in the syntax. However, any movement of the verb to a higher position will interfere with this. Thus, I assume that there is also a condition at PF that ensures that focused elements are left adjacent to the verb, which moves . Admittedly, this introduces a certain amount of redundancy, in that there are thus two overlapping processes which serve the same goal of getting focalized elements to the left of the verb. This conclusion however appears to be forced by empirical considerations; thus I take the criticisms of redundancy on the chin and proceed without further discussion.

[12] Note that SOV is not the only order that Udi permits. Harris (2002) notes that this is the basic order, but the position of the arguments relative to the verb is quite flexible, such as in the following, where the object follows the verb. SVO word order could arise from the objct not moving to the higher position in the tree, but there isn't enough data to comment on this. Harris herself notes that it is not clear what the conditions are that allow non-SOV word orders. Other word orders in the language, such as PP etc suggest that Udi is a head-final language, and so it is tempting to conclude that Udi is underlyingly SOV, but I leave the issue open. What is important for our purposes is the assumption that the object vacates its base position, wherever that is, and moves higher.

(32)
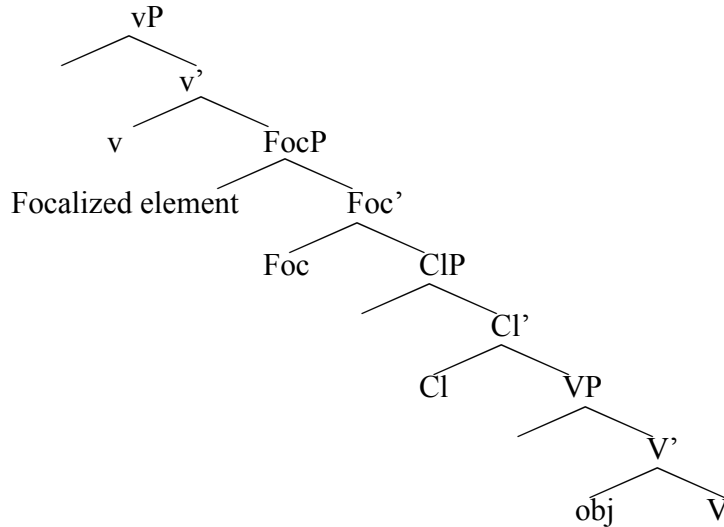
The clitic, generated as the head of ClP therefore is not generated on its host, nor does it move directly to the host from a big-DP. With the complement of *v* constituting a spell-out domain this means that the clitic will be sent to spell-out with any elements in focus, as well as the verb. For now, I focus attention on Rule 2, the relevant configuration being (33b). I return to discussion of (33a) in section section 3.4 below.

(33) a. [clitic [(XP) verb]]
     b. [XP$_{foc}$ [Foc [clitic [verb]]]]

In (43b) the clitic is left free standing at spell-out, and must therefore find a host to lean on. Since the clitics in Udi are suffixal elements (recall that they never appear at the beginning of a word, and that proclisis is not seen in Udi), I assume that if the clitics are freestanding at the point of linearization and there is an element to their left, they lean on that element, finding a prosodic host. Thus, after linearization, the clitic simply attaches onto the element that lies on its left thereby satisfying its need to attach to another element, as clitics do. To see how this works, consider the following derivation for the sentence in (44). At the point of spell-out of the complement of *v*, the object *p'ạ ėš-ne* 'two apples', the verb, which has remained *in situ* and the clitic are all spelled out together. I assume that the object does move to Spec,vP as mentioned above, but a lower copy is selected for pronunciation (Bobaljik 1995, 2002). Selection of the lower copy is forced by the close relationship between focus and the verb. Recall that the verb and focus are adjacent in Udi. We can view this as a requirement that they be spelled-out in the same domain. If the verb stays in its lower position and the object pronounced in Spec,vP, then they will be spelled out in different domains. Lower copy selection however allows the verb and focus to be spelled out together. *Left* adjacency of focus is not a principle of focus positioning, but merely a byproduct of the fact that Spec,FocP is linearized to the left of VP. Thus, the structure which is sent along the PF branch is as given in (35) (the clitic *-ne* is boldfaced):

(34) äyel-en   p'ạ ėš-ne   aq'-e
     child-ERG two apple-3SG take-AORII
     'The child took two apples.'

(35) [$_{FocP}$ [$_{DP}$ p'a ėš] [$_{Foc'}$ FOC [$_{ClP}$ [$_{Cl'}$ **ne** [$_{VP}$ [$_{V'}$ aq'-e]]]]]]

Since the DP that is in focus lies to the left of the clitic, the clitic is able to lean on *pa ėš*. Thus, I assume that the clitic essentially rebrackets under adjacency (by morphological merger, Marantz 1988, see also Bošković 2011 Embick & Noyer 2007), shown below with a simplified bracketing structure (labels are omitted from this point on):

 (36) [[[p'a ėš] **ne**] [aq'-e]]

Before moving onto the cases where the clitic ends up being internal to the word (and Rule 6, which will be shown to come from the same mechanism as the meso/endoclisis cases), what remains to be explained is the competition between Rule 1 and Rule 2, and why a subjunctive/ imperative verb attracts the clitic and not focus. The competition between the two I propose is resolved by movement of the verb to a higher position. The subjunctive/imperative verb moves higher via head movement, This movement I assume to be forced by the requirement to license a [uF:mood] feature against a Mood head higher in the tree. I assume that this licensing can be accomplished through an Agree relation, which allows [uF:mood] to be deleted (see Chomsky 2000, 2001). However, since Mood is in the higher domain of the tree, subjunctive/imperative verbs are compelled to move to the edge position of the aspectual domain through considerations of the Phase Impenetrability Condition (PIC, Chomsky 2001):

(37) The domain of H is not accessible to operations outside HP; only H and its *edge* are accessible to such operations (Chomsky 2001:13).

Given the PIC, if a verb which has the [uF:mood] feature on it remains in V then it will not be able to have the feature licensed by a head which is outside of the aspectual domain; it must move up to a phase head in order to see the higher Mood. It then follows from this, that *all* subjunctive and imperative verbs move up at least as high as the edge of the phase, in order to be able to see into the higher phase and check their mood feature Therefore, verbs that carry the mood feature must move up at least as far as *v* in the tree, and potentially even higher. Movement of the verb to the edge of the phase will pick up the clitic as the verb moves up, and the clitic will then appear on the verb. Recall that the rebracketing onto a focalized element discussed above was when the clitic was left free-standing after the syntax. Clitics require a host word and leaning on a host word allows them to do this. However, in the cases where the subjunctive and imperative verbs move higher, they will pick up the clitic, thus after the syntax the clitics have a host and do not need to undergo any further rebracketing.

Two things require further clarification at this point. Firstly, with the subjunctive and imperative verbs requiring movement up to (at least *v*), then if the object remains in Spec,FocP, then this would mean that the focused element would appear to the right of the verb in these cases contrary to fact. I assume therefore that movement of the higher verb is also accompanied by an additional movement of the element that is in focus. This could either be within the syntax, with movement up to Spec,vP or an inversion at PF to switch positions with the verb (without empirical evidence to choose between the two, I assume that this happens at PF). Recall from the discussion earlier that there is a requirement in Udi that focalized elements and the verb must be spelled-out together, since they share a close PF-relationship. This then forces the movement of the focalized element when the verb moves to the edge of the phase - focus must also move up.

The second issue which requires some attention is the position of the TAM marker on the verb. Since I am assuming that verbs in Udi generally do not move as high as T or Mood, then it must be the case that the affix lowers (Embick & Noyer 2007) onto the verb post-syntactically, and that the verb does not (usually) form a complex head with TAM in the syntax. Thus, we might

expect that post-syntactic lowering would place the TAM affix outside the clitic when the verb is subjunctive or imperative, giving rise to the following:

(38)   verb$_{subj}$ - clitic - TAM

As will be discussed in greater detail below, (48) is not a legitimate representation in Udi. Verbs and TAM suffixes have an extremely close relationship in Udi, and nothing ever intervenes between them (as noted by Harris). Below, this observation will be used to motivate the morphotactic rule given in (49) that disallows anything appearing between the verb and TAM. In the spirit of this rule, we can assume that the TAM suffix wants to get as close to the verb as possible when it lowers. Thus, I assume that the lowering rule that applies to TAM doesn't just simply move it generally to the complex V^0 head, but actually next to the verb itself. This pushes the clitic to the end of the entire verb form in subjunctive, hence in the cases where the verb is subjunctive, the subject clitics are enclitic to the entire form.

(39)   *verbal head - X - TAM

In sum, in this section I have shown that Harris' rules 1 and 2 result from the clitic being merged into the structure in a relatively low position. When the clitic appears on the end of the verb in subjunctives, this was shown to arise from the verb moving to a position higher than the clitic, picking it up as it moves. Lowering of the TAM affix pushes the clitic to the final position. When there is no verb movement, which was forced only in verbs that need to license a [uF: Mood] feature, then the clitic will be spelled out without a host. However, when an element is spelled out in Spec,FocP, then this allows the clitic to lean on that element, undergoing local dislocation into the host word. Thus, as predicted, when there is an element in focus, we see that clitic will appear enclitic to the focalized element. Importantly, since this is rebracketing under adjacency, we predict that when there are multiple focalized elements, as is the case with multiple wh-questions, the clitic should attach to the rightmost one. This prediction is borne out, (40).

(40)   a.   šin   šux-**a**   k'al-exa
            who.ERG   who.DAT-Q   call-SAY.PRES
            'Who is inviting whom?'

       b.   merab-en   šu   ek'a-**a**   ta-d-e
            Merab.ERG   whom.DAT   what.ABSL-Q   thither-LV-AORII
            'What did Merab give to whom.'

Finally, we see that the competition between the two rules is not really a competition at all. Subjunctives and imperative verbs bleed clitic attachment to focus because they pick up the clitic as the verb moves in the syntax. M-merger happens only when there is a need to satisfy some requirement, and since the clitic already has a host word when it appears in the complex head of the verb, then it is inapplicable when there is both focus and a subjunctive/imperative verb. Therefore, there is no real competition, but rather syntactic movement of the verb takes away the need for the clitic to undergo m-merger onto focus. With Rules 1 and 2 sorted, I now proceed to discussion of when the clitic is spelled out on its own, but has no element to its left within the spell-out domain, i.e. there is no focus.

2.4. Rules 4 - 7: Endoclisis, mesoclisis and enclisis

Above I proposed that the clitics are generated above VP, and are attracted by an element that merges or moves higher, covering the cases where there is a mood operator or focus. These accounted for Rules 1-3, on the assumption that the clitic find a host if at all possible. However, the explanation of these cases relied on the clitic being able to move to a c-commanding element during the syntax. Clearly, this explanation is not possible with Rules 4-7 since we then would not expect to find any difference between the rules as to the place of the clitic. In this section I propose that all the remaining rules, 4,6 and 7[13], where the clitic alternates between being a mesoclitic, an enclitic and an endoclitic all stem from the same original placement of the clitic, but that the differing positions are the result of a morphological metathesis operation that, under the circumstances of each rule's environment, repositions the clitic from its original position. The position which I claim the clitic to target is the second position within the verb. That is, Udi subject clitics, when not attracted by any higher element are a type of second position clitic.

2.4.1. Verb-internal second position placement

In the cases under discussion, there is no focus, and there is no reason for the verb to move higher in the structure. Recall that the movement of the verb was forced in order to license an uninterpretable mood feature against an mood operator in a higher phase of the tree. Since the PIC prevents Agree from happening between the elements in the lower aspectual domain, and the higher phases, the verb must move to the edge of the phase, taken here to be v in order to be able to enter into an Agree relation with the higher mood operator. This movement picked up the clitic on the way past, and thus provided the clitic with a host word. Similarly, when there was a focalized element in the tree but not a subjunctive/imperative verb, the clitic would lean on the element in focus and undergo m-merger with that element, satisfying its need to be suffixal.

However, there remains a spell-out configuration in which we still need to consider, namely when the spell out of the complement of v yields a structure like the following, where the object has moved to Spec,vP and there is no focalized element (I omit FocP in the structure). Thus, there is only the verb and the clitic:

(41) $[_{ClP} [_{Cl'} Cl [_{VP} [_{v'} V]]]]$

Since there is no element to the left of the clitic, it cannot simply m-merge onto the end of a left adjacent element. In fact, within the spell-out domain, there is only one element that gets spelled-out with the clitic, the verb.[14] M-merger with the adjacent verb takes place, but this does not allow the clitic to satisfy its need to be a suffixal element. There are no proclitics at all in Udi, reflecting a

---

[13] The Rule 5 cases recall from above are completely regular instances of Rules 4 and 7, with their regularity obscured by the presence of a suppletively null light verb in the present tense

[14] The situation with ditransitive verbs is not clear. At some times an indirect object appears to the left of the verb, and at other times to the right. Furthermore, it does not seem to occupy a fixed position relative to the direct object, as they both can precede the verb but apparently in either order (see examples below). This all suggests that there isn't a fixed position for the indirect object, but it has some flexibility to move around and potentially vacates VP. In the absence of good evidence, I omit discussion of ditranstive constructions, and focus solely on intransitives and transitives. If it turns out to be the case that the indirect object of a ditransitive is able to remain *in situ,* or indeed other material intervenes between the clitic and the verb once the structure is linearized, then a modified account to the one presented here would need to be pursued. In such an instance, we could assume that there is a rule that lowers the clitic to the initial position in the complex head that contains the verb. Its suffixal requirement would be met by a repair that moves the clitic inside the word, after the first morpheme.

general prohibition of the language against proclisis. Thus, in order to find some material to its left edge, the clitic occupies the second position within the verb. Thus, second position should be seen as the clitic undergoing m-merger to the verb that lies on its right, and subsequently tucks inside the first element within the complex head in order to avoid being a proclitic (*cf.* Halpern 1995, Bošković 2001, Embick & Noyer 2001).

Having the clitic target the second position within the verb allows for two advantages; (i) we do not need to define some element that the clitic is positioned relative to; and (ii) we can legitimately analyze movements away from a given position. The first advantage is fairly obvious in itself, since although the clitic displays some relationship to the verb in Rules 4-7 environments, there isn't any element that it is consistently next to. If there were, we could assume that the clitic moves down to that element and nothing more needs to be said. The closest element that we get would be the head of the verb (so either verbal root or the incorporating light verb), since with complex verbs the clitic immediately precedes the light verb and with simplex verbs the clitic lies internal to the verb root. However, with the cases of Rule 6, we see the clitic is non-adjacent to the verb root, with the TAM suffix intervening between the two. Therefore, the clitic cannot be claimed to attach to the root, at least without further qualification.

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

      b. q'ačaɣ-ɣ-on bez   tänginax   baš-**q'un**-q'-e
          thief-PL-ERG   my   money.DAT   steal₁-3PL-steal₂-AORII
          'The thieves stole my money.'

(43) b-esa-**ne**
      make-PRES-3SG
      'she makes'

A second position analysis however avoids these shortcomings, since we do not need to define the position of the clitic relevant to another element as that does not matter; what is important is the linear position of the clitic. We can see that second position within the verb is transparent in the case of complex verbs, where we see that the clitic always follows the incorporated element. Schematically then, the clitic occupies the second position within the verb:

(44) sa   adamar-en   fikir-**re**-b-sa   te...
      one   person-ERG   think-3SG-do-PRES   that
      'A person thinks that...'

(45) IncE-clitic-light verb-TAM

Note that when the incorporated element is itself internally complex, the clitic treats it as a whole unit, following the entire constituent. Such an example is found in the causative construction (Harris lists this as Rule 4a, but notes that these are another case of complex verb formation, when what is incorporated by the causative light verb is the infinitive non-causative verb):

(46) nana-n   äyel-ax   ak'-es-**ne**-d-e   k'uč'an
      child-ERG child-DAT see-INF-3SG-CAUS-AORII   puppy.ABSL
      'The mother showed a puppy to the child.'

The clitic in this sentence lies in the third morphemic position in the verb, since it follows the incorporated verb and its infinitive marker. This may be taken as evidence against second position being the correct generalization, however, second position is correct if whatever is incorporated is treated as one unit.[15]

2.4.2. Second position placement, morphotactic repairs and articulated VI

Motivating second position for simplex verbs is harder, but as will be shown it is indeed the position where the clitics initially occupy in the verb, though not the correct surface position. Second position is clearly not the surface position with simplex verbs, since then we'd expect them uniformly to appear between the verb root and the TAM suffix. If second position is therefore the starting point of these clitics, as is claimed here, then something must move them away. The movement operation is clearly of a post-syntactic nature, since it can target positions within the phonological information of the host, as in the Rule 7 cases where the clitic is positioned inside the final consonant of the root. Assuming that phonological exponents are inserted late in the morphology (Halle & Marantz 1993 and other work in DM, see Halle 2001 for discussion on this issue with respect to infixation), then it follows that the syntax cannot provide the position of the clitic (contrary to Harris' analysis of Udi, discussed below). I therefore propose that for simplex verbs, the clitic leans on the verb but moves after the verb in order to satisfy its need to be suffixal, and then undergoes a metathesis operation that moves it away from this place (see Arregi & Nevins 2012 for relevant discussion of morphological metathesis with respect to clitics in Basque). This metathesis operation is able to move the clitic either to the right or left within the verb, with the choice leading to the cases described by Rules 6 and 7, shown schematically in (47a) and (47b) respectively. For concreteness I adopt the version of morphological metathesis proposed by Calabrese & Pescarini (2013), opposed to the more complicated generalized reduplication approach given by Harris & Halle (2005) and adopted by Arregi & Nevins (2012).

(47) a. verb-clitic-TAM → verb-~~clitic~~-TAM-clitic   (Rule 6)
      b. verb-clitic-TAM → verb$_1$-clitic-verb$_2$-~~clitic~~-TAM   (Rule 7)

(48) Morphological metathesis (Calabrese & Pescarini 2013)
          X α Y W
      a. Generate a copy ($α_1$) of α: X αα$_1$ Y W
      b. Move α$_1$ after Y: X α Y α$_1$ W
      c. Delete α: X Y α$_1$ W

---

[15] Second position is in fact further supported in these instances, since there is apparently some ambiguity of where the clitic goes; Harris (2002:124) states that in these instances the clitic can appear before the infinitival marker a, in second morphemic position. The following sentences are acceptable, though less preferred than when the clitic follows the entire IncE. This shows that when the clitic shifts into the word in order to not be a prefix, then it can go in two places, both consistent with the notion of second position:

(i)   baba-n   mzia-x   arux-ne-b-es-t'e
      father-ERG   Mzia-DAT   fire-3SG-DO-INF-CAUS-AOR
      'Father had Mzia build a fire.'

(ii) baba-n   äyel-ax   ači-ne-p-es-t'-e   nard
      father-ERG   child-DAT   play-3SG-SAY-INF-CAUS-AORII   backgammon.ABS
      'Father had the child play backgammon.'

17

What are the conditions that force this movement of the clitic? The answer is that there is a morphotactic restriction in Udi that requires the TAM suffix and head of the verb to be adjacent. If true, we then expect to find no element whatsoever, including clitics, appearing between the verbal head and TAM suffix in Udi, a prediction which is borne out without counterexamples.[16] Far from representing an apparent coincidental gap, I propose that this is a true morphotactic restriction of Udi, with the following condition on surface forms active in the language:

(49) *verbal head - X - TAM

Thus, when the verb is simplex a conflict arises in Udi, the clitic is positioned between the verb and the TAM suffix, but if it stays there the resulting surface form would violate the morphotactics of the language. This is the motivation of the metathesis of the clitic; it allows the verbal head and the TAM suffix to be adjacent.

 Why then the difference between Rules 6 and 7? Both rules make reference to simplex verbs, but the choice of whether the clitic becomes an enclitic or an endoclitic by repair is not arbitrary, but rather conditioned by the phonological shape of the verb. The verbs that are relevant to Rule 6 are those that are formed of either a single consonant or are CV in shape. Verbs which give rise to Rule 7 cases are those that end in a coda. Clearly then, as identified by Harris, the clitic placement is sensitive to the phonology - open syllables are unable to be hosts for clitics. Now we begin to understand the differences between Rule 6 verbs and Rule 7 verbs. Consider the derivation for the verb form in (50), where the clitic appears internal to the verb root. The clitic is first placed in second position within the verb.[17] If it remains in this position it will violate (49) and so must be metathesized away. Metathesis moves the clitic inside the final consonant of the verb allowing the verb and TAM to be adjacent on the surface. The derivation is shown in (51).[18]

(50) kaγuz-ax   a-**z**-q'-e
  letter-DAT receive$_1$-1SG-receive$_2$-AORII
  'I received the letter.'

(51) i. after cliticization: √RECEIVE-**[1SG]**-[AORII]
  ii. VI of root: /aq'/-**[1SG]**-[+AORII]
  iii. VI of clitic: /aq'/-/**z**/-[+PRES]
  iv. metathesis repair: /a-**z**-q'/-[+PRES]
  v. VI of TAM: /a-**z**-q'-e/

For Rule 6 cases I assume that leftward metathesis fails, and must therefore go rightwards in order to produce a surface form consistent with (49). Leftward metathesis fails for two reasons. Firstly, following Harris (2002) I assume that CV roots are not large enough to host a clitic; in other words,

---

[16] With respect to the rule 5 cases, it must be the case that the incorporating light verb is what counts as the verbal head in (49), since there is clearly no problem with the clitic appearing between the incorporated verb and the TAM marker. This makes a certain amount of sense if the rule makes reference to the *head* of the verb, as the light verb intuitively is the head of the complex verb, even though it does not supply (much of) the lexical meaning. This is true because the category of the resulting form when a light verb incorporates some element is always a verb, no matter what is incorporated (see discussion of complex verbs above).

[17] The necessity of this step will be discussed below.

[18] The metathesis step iv has been simplified in (51). The steps of copy, move and delete in the metathesis algorithm of Calabrese & Pescarini (2013) have all been conflated. Also, in the labels of the steps, cliticization refers to movement of the clitic into second position.

the size of the phonological rhyme is crucial to whether something can host a clitic or not, and suitable syllables must have a coda. CV roots therefore are unable to host clitics. The second reason that leftward metathesis fails is that the only other place to put the clitic would be at the front of the verb. I assume that the clitics are inherently suffixal (see Wojdak 2005 on Nuu-Chah-Nulth predicate affixes, which are also inherently suffixal), that is, it is an inherent property of the clitics that they cannot be at the beginning of the word (this is similar to Klavans 1985, but in a more limited form). Udi also does not in general allow proclitics, since there are none in the language. Therefore, the only derivation that will satisfy (49) *and* the suffixal requirement of Udi subject clitics is one where the clitic has shifter rightwards, as is shown in the following derivation for *k-e-ne* 'eat-PRES-3SG':

(52)　ii. after cliticization:　　　√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/

There is further evidence that the rightwards metathesis approach is on the right track, instead of an approach where the clitic goes to the end if endoclisis fails (as Harris is more or less forced to assume). We should expect that when there is another element outside of the TAM suffix, the clitic should be able to tuck in between that element and TAM. This is correct, as shown below where the verb also has the past clitic *-y*, creating the imperfect form from the present stem. Here we see the subject clitic nestle between TAM and *-y*, and not at the end of the verb, as shown in the derivation in (54):[19]

(53)　bi-esa-ne-y
　　　DO-PRES-3SG-PAST
　　　'She was doing.'

(54)　i. after cliticization:　　　√DO-[3SG]-[+PRES]-[+PAST]
　　　ii. VI of root:　　　　　　/bi/-[3SG]-[+PRES]-[+[PAST]
　　　iii. VI of clitic:　　　　　/bi/-/ne/-[+PRES]-[+PAST]
　　　iv. metathesis repair:　　　/bi/-[+PRES]-/ne/-[+PAST]
　　　v. VI of TAM:　　　　　　/bi-esa-ne/-[+PAST]
　　　vi. VI of past clitic　　　/bi-esa-ne-y/

In the derivations in (51,52,54) I assume that evaluation of the morphotactic violation is evaluated at the point of vocabulary insertion (VI) of the clitic, though in principle it could happen before or after any VI has taken place. Arregi & Nevins (2012) assume that although there is an elaborately structured set of operations that happen within the morphology, clitic metathesis in Basque occurs prior to VI. This however gives the wrong result for Udi, since then the phonological information of the root will not be present in the derivation for the clitic to target, and there is no way to capture the asymmetry between C/CV roots and -VC roots.

The phonological asymmetry *can* be captured if clitic metathesis happens after VI, since the phonological information of the roots would be in the derivation and able to be targeted. However, this does not allow for us to capture another asymmetry, namely that the phonological structure of

---

[19] The past clitic -y, although superficially it may look like one of the TAM affixes, is in fact a clitic and not a regular TAM affix. (54) is thus not applicable.

roots can be targeted, but not the phonological structure of affixes. We see this in the form *bi-esa-zu* 'die-PRES-1SG'. Here, the exponent of present tense *-esa* could arguably host the clitic, since there is a consonant that the clitic could be placed before. However, we don't find the form *bi-e-zu-sa*. If clitic metathesis happens after VI then there is nothing to rule this out. However, if VI is interleaved with repair rules that are evaluated at the point of VI of the offending element - here the clitic - then we find a natural explanation for the above two asymmetries. Taking the standard assumption within DM that VI proceeds cyclically from the root outward (see Bobaljik 2001, Embick 2010, Moskal 2013) then the phonological information is always going to be present in the derivation before VI (and morphotactic evaluation) of the clitic. But, since VI of the clitic will follow VI of the root, but crucially *precede* VI of TAM, then at the point that the clitic comes to move, then all that TAM is is a feature bundle. Thus, the clitic cannot target its phonological structure as it's simply not yet in the derivation. This is shown with the derivation of *bi-esa-zu* in (55). Interleaving VI with morphotactic repairs allows us to very simply capture these two asymmetries without saying anything further. If metathesis repairs precede or follow VI, then something extra needs to be stated to capture these asymmetries, and it is not clear to me what this would be.

(55) i.   after cliticization: √DIE-[1SG]-[+PRES]
      ii. VI of root: /bi/-[1SG]-[+PRES]
      iii. VI of clitic: /bi/-/zu/-[+PRES]
      iv. metathesis repair: /bi/-[+PRES]-/zu/
      v.   VI of TAM: /bi-esa-zu/

2.4.3. Is second position necessary?

In the preceding discussion, it has been assumed that the clitic leans on the verb to its right, but must come to be positioned one morpheme inside the verb in second position to best satisfy its demand for not being a proclitic. Moving one morpheme in allows it to do satisfy both of these conditions. However, I have also proposed that morphotactic violations are evaluated at the point of spell-out of the offending element, which allows us to understand why the clitic goes inside the root, but not TAM suffixes. Yet, I have also assumed that the positioning in second position happens before VI, in that at the point of m-merger, the clitic comes to occupy second position within the verb, and then other morphological processes happen.

We could potentially subsume the placement of the clitic under a repair to avoid being a proclitic. Thus, in all of the cases of Rule 4-7, the clitic could remain a proclitic up to the point that it undergoes VI, and then move inside the verb to avoid being at the beginning of the word. If such an analysis turned out to be viable, then we could avoid the step of postulating that there are word-internal second position clitics altogether, and simply have a case of a clitic moving around to avoid being a proclitic. Thus, the question is, do we need the step of moving into second position within the verb at all?

Without taking the extra step to move the clitic into second position in the verb, we run into problems. Firstly, with respect to simplex verbs, if the clitic went rightwards at the point it undergoes VI, then it follows that the phonological information of the root will already have been inserted, with VI standardly taken in DM to proceed outwards from the root as discussed above. We know that insertion into the root is dependent on the root ending in a consonant, with CV roots not allowing for endoclisis, and the clitic going to the end of the verb form. Suppose that clitics will move inside the root if there is a closed syllable, as I have claimed above. The difference between the two accounts is the direction from where the clitic is coming from. On the account above, the clitic is coming from a position to the right of the root and so makes the minimal movement inside the root to allow the root and the TAM suffix to be adjacent. However, on the other analysis under

discussion, the clitic is in fact coming from the *left* of the root, and avoiding the edge. With CVC roots, there is nothing to distinguish the two cases, since they both target the position before the consonant. However, with CVCC roots, things are different. We know from forms like (52b) that the clitic moves inside the *final consonant*, not the entire coda. On the account above, this was handled straightforwardly; the clitic makes the smallest possible metathesis and tucks inside the final consonant. On the other hand, the account that we are considering here, cannot claim that any minimal movement takes place. In fact, it must be the case that the clitic moves as far inside the root as it can, deliberately skipping past a possible position before the first consonant of the coda. Considerations of parsimony seem to favor placement of second position placement then.

A second argument favoring second position placement comes from complex verbs. Here, considerations of parsimony are not that important; it simply becomes extremely difficult to analyze it otherwise. Suppose that the clitic were to remain in the proclitic position until it is spelled-out. With VI proceeding cyclically, it must be the case that the incorporated element undergoes VI before the clitic. If then the clitic moves inside the verb at the point of its own VI, the the phonological information of the incorporated element should be available for clitic to move inside of. But, this doesn't happen; the clitic does not target the final consonant of the incorporated element but rather appears between the incorporated element and the verb. A way to save the proposal would be to assume that the repair rule that moves the clitic inside the verb form is as follows:

(56) Make the shortest possible move inside the verb form, respecting the integrity of morphemes if possible.

However, the above statement is little more than a fudge designed to capture the facts by force. On the other hand, the approach given above doesn't need to resort to such statements. Nothing needs to be said about complex verbs; the clitic simply goes in the second position, and as shown above, we get a very simple and natural explanation for the divergent behavior of simplex verbs that or C/ CV (Rule 6) or CVC(C) (Rule 7).

2.5. Comparison with Harris (2002) and Yu (2007)

Having provided an analysis of Udi clitics, I now turn to comparing the proposed analysis against two others. First is the proposal of Harris herself, who gives an account within the framework of Optimality Theory (OT). Harris essentially rewords the descriptive rules as OT alignment constraints, proposing the following constraints in (57 - 60) and the ranking in (61).

(57) Align-PM-*al/a*
        Align (PM,L,-*al/-a*,R)
        Read as: "align the left edge of the person marker to the right edge of -*al/-a*"

(58) Align-PM-FocC
        Align (PM,L,FocC,R)

(59) Align-PM-IncE
        Align (PM,L,IncE,R)

(60) Align-PM-Verbstem
        Align(PM,R,Verbstem,R)

(61)   Align-PM-*al/a* » Align-PM-FocC » Align-PM-IncE » Align-PM-Verbstem
= constraint for (TAM)   (focus)   (complex verbs)   (simplex verbs)

These constraints, and the proposed ranking ensure that the clitic appears in the positions of the descriptive rules. The exception to this is Rule 6, which Harris must handle through a default positioning of the clitic to the end when all the other constraints fail to produce a proper output. It is hard to criticize Harris' proposal on empirical grounds, since the OT constraints do little more than place the clitic exactly in the places where Harris describes them to be. However, there is reason to be suspicious that this account turns out to be the correct one for Udi. The reason is constraint (60), which puts the clitic directly inside a morpheme. The way that the constraint is formulated, is that the OT candidate that best satisfies (60) is one where the clitic goes inside the final consonant of the verb. The candidate will incur one violation mark, however all candidates would do this since the constraint must line up the *right* edge of the clitic with the *right* edge of the verbstem.

Now, the problem with this constraint is one of overgeneration. It is an extremely powerful addition to UG, since it encodes within UG the ability of a clitic to be positioned by the grammar *directly* into another morpheme, and allows clitics to effectively subcategorize to appear always internal to another morpheme (if a constraint like (60) is the only one positioning the clitic). If this is true, and constraints of the type in (60) are a possible clitic placement rule of UG, then we can reasonably expect to find other cases in languages around the world. As it happens, we do not find this. To my knowledge, there is no other language which has a clitic that goes inside a morpheme (Degema will be discussed below). All other cases of non-peripheral clitics are such that the clitics appear intermorphemically within the word. There is also to my knowledge no example of a clitic which appears always internally to another morpheme.

Another account which proposes the same thing - that clitics can be positioned directly within morphemes is Yu (2007) who claims that his theory of infixation can be extend to include the Udi case of endoclisis. Yu picks up on a suggestion of Harris that (60) can be reformulated as referring to the final consonant of the verbstem (Align(PM, R, C]$_{st}$, L)). Yu then claims that this fits in well with his theory of phonological pivots that infixes subcategorize for. The thrust of Yu's approach is that there are a fixed number of cognitively salient pivots which infixes can subcategorize for. They are crucially all phonological in nature, and the infix may look for consonants, vowels and syllables at the edges of words (first/last consonant etc), or stressed feet, syllables and vowels. The new formulation of (60) then would fit the pivot of being the last consonant pivot, and thus providing a unification of infixation and the only case of endoclisis.

However, Yu's approach faces the same problem as Harris', in that if clitics are really subject to the same positioning/subcategorization rules as affixes, then we expect more endoclitics to arise in languages around the world. Whilst it might not be the case that we would expect them to be at the same frequency of infixes (higher level constraints forcing clitics to the peripheries of words may exist, though I'm not aware of any proposals), we should expect more than one case to have been reported in the world's 6000 or so languages. Furthermore, we would surely expect to see at least one case where a clitic invariably goes inside another morpheme, since we do find such infixes (see Blevins 1999 on Leti). But again, such a case is unattested though predicted by both Yu and Harris.

The approach advocated for here, it may be claimed, apparently faces the same problem. Clitics *can* appear internal to other morphemes since readjustments of the type in Udi serve to place the clitics there. However, note that on the proposal that I am making, we do expect the phenomenon to be strikingly rare. The endoclisis seen in Udi comes about only through the confluence of various factors; (i) the clitic must be placed internal to the word itself; (ii) that placement of the clitic must interfere with a language specific, arbitrary morphotactic rule; (iii) metathesis of the clitic must be the chosen repair as opposed to deletion; (iv) the metathesis

operation must move the clitic inside a morpheme as opposed to around it. Factor (iv) in Udi is actually composed of two separate preferences - a preference for leftward metathesis and the inability of the clitic to be a proclitic. Once you multiply the odds of each thing of these factors happening together, the odds of getting an endoclitic become very small very quickly. Therefore, though we still predict other cases to exist, we don't predict the frequency that Yu and Harris seem to. Furthermore, we don't also expect a clitic to intrisically target an internal position of a morpheme, since I do not make any claim regarding the selectional requirements of the clitic. Clitics are placed according to linearization rules and their own suffixal/prefixal nature. Any further movements are forced by the environment that the clitic finds itself in, not the clitic itself.

2.6. Conclusion

In this section I have presented an analysis of the position of the subject clitics in Udi. I argued that the somewhat erratic positioning of the subject clitics is the result not of different placement rules, but rather the the search for a clitic position that best satisfies the rest of the surface requirements of the language. Clitics were shown to find a host in the syntax whenever possible, but can also find a host post-syntactically. However, the variable post-syntactic positionings follow from second position placement plus any necessary repairs to satisfy morphotactic demands. In the next section I provide further evidence for two of the major claims in this section, that clitics can target the second position within the word and that they can be moved around from this position in order to satisfy morphotactic requirements of other elements. I also provide further evidence for the aspectual domain structure in (32) above, where the clitic is generated between VP and vP, rather than with its associated argument.

3. Sorani Kurdish

Moving on from Udi, I turn to the second case study of non-peripheral clitics that I will look at in this paper, Sorani Kurdish. Sorani Kurdish is interesting not only from the point of view that it is another language that contains non-peripheral clitics, but also because it provides further motivation that there are clitics that become systematically positioned in the second position within the word. Sorani is also interesting because it has a pattern of split ergativity within the clitic system which bears on the question of where clitics are generated within the syntax, as will be discussed below.

3.1. Clitics vs affixes in Sorani Kurdish

The clitics that interest us in Sorani Kurdish are again clitics that mark the features of an argument, however the matter of which argument is slightly more complicated than in Udi. The reason is that Sorani Kurdish, like its Northern Kurdish counterpart Kurmanji Kurdish (see Atlamaz & Baker 2013) has a pattern of split ergativity based on tense. In the present tense, the language takes a nominative-accusative pattern of agreement, with the verb agreeing with the subject of intransitive verbs (referred to as the S argument) and the subject of transitive verbs (A). Clitics, if present mark the object of transitive verbs (O), however they are not obligatory. A nominative-accusative alignment is indicated by the S and A arguments patterning together to the exclusion of O. However, in the past tenses in Sorani, the verb agrees with the object of transitives (O, note that this is not obligatorily shown on the verb) and the subject of intransitives (S). Clitics are obligatory, and mark the subject of transitives (A). This is an absolutive-ergative alignment, with S and O arguments patterning together to the exclusion of A arguments. In short, the pattern of which argument is marked by what element is as follows:

(62)   Present tense   Past tense
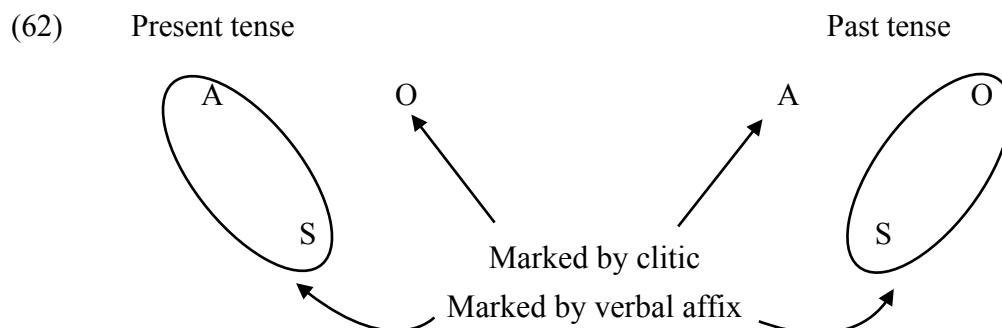
The above patterns of agreement are illustrated by the following examples:

(63) a. (min) kitêb-êk   bo   Narmînda-kir-**im**
          (I)   book-INDEF   for   NarminPROG-buy.PRES-1SG
          'I am buying a book for Narmin.'

      b. bo   ci   pê-m   pê   da-kan-**î**?
          for   what   to-1SG to   PROG-dig.PRES-2SG
          'Why are you laughing at me?'

      c. Narmin   ba   Sirwan-**î**   dâ-n
          Narmin   to   Sirwan-3SG   give.PAST-3PL
          'Narmin gave them to Sirwan.'

      d. mâw-**în**
          remain.PST-1PL
          'We remained.'

In (63) I have boldfaced the markers of subject agreement. We see that in the present tense transitive (63a), the verb agrees with the subject, as does the intransitive verb in (63b). In (63c) however we see that the marker on the verb is this time marking the features of the direct object, whilst in the intransitive (63d), there is no clitic on the verb, but there is a affix that appears on the verb marking the subject. The pattern is clearly complex, and it could be immediately argued that we are simply dealing with clitics that move around in the clause. However, the two classes can be distinguished on the basis of morphological shape. The markers which lie on the verb are referred to by Thackston (2006) as *personal endings* and I assume that these correspond to affixal agreement of the verb (the equivalent of T-agreement, though we will see later on that the probe is *voice* and not T). The elements I have taken to be clitics are referred to as agent affixes by Thackston. He notes however that they are identical to enclitic possessive pronouns, and they clearly form the same class of elements that mark the direct object in present tense, a class that he says are enclitics.

(64)   <u>clitics</u>   <u>affixes</u>

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

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

As mentioned in section 2, we need to be careful that what we are dealing with really are clitics as opposed to affixes. Whilst I do not have the wealth of arguments that Harris (2002) gives for the Udi subject markers to show that they are clitics, it is possible to show that the elements I am taking to be clitics have a much higher degree of freedom than the elements I take to be affixes. Since clitics are distinguished from affixes through their lack of selectivity with respect to the host, and relative mobility within the syntax, these will function as diagnostics for this study.

The placement of the clitics is complex, just like in Udi, and also just like in Udi, the clitics will at sometimes target some position on the verb and at other times they will appear elsewhere in the sentence. However, despite appearing complicated given the number of different elements that the clitics will attach to, their distribution is actually fairly regular. The clitics will only appear as a mesoclitic when they attach to the verb. Thackston (2006:42) notes that "if anything other than the verb is expressed, then the agent is affixed to the first available pre-matter." This quote needs some untangling, but what it boils down to is the following. Clitics will attach to the leftmost element within the VP, and if there are no lexical elements to the left of the the verb within VP, then the clitic will attach to the leftmost morpheme within the verb itself.

XP attachment is shown by the following, and as can be seen, the clitic predictably is enclitic to the leftmost XP. It should be noted that the fact that the clitic stays within the VP is suggested by the fact that it never appears on the subject. I assume that the clitic, like the ones in Udi cannot raise out of VP. Further assuming that the subject raises to Spec,TP, then the clitic will not be able to appear on the subject:

(65) a. Narmin   ba   Sirwan-**î**   dâ-n
       Narmin   to   Sirwan-3SG   give.PAST-3PL
       'Narmin gave them to Sirwan.'

   b. bâzirgân-akân   asp-akân-**yân**   da-kirî
      merchant-DEF.PL   horse-DEF.PL-3PL   PROG-buy.PAST
      'The merchants were buying the horses.'

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

However, sentences in Sorani can consist solely of the verb (McCarus 1958). In cases such as these, then there is obviously nothing to the left of the verb within the VP. When this happens, the distribution of the clitics becomes more erratic. We can see in the following data, which will be considered further in section 3.3 below:

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

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

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

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

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

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

Though this behavior is erratic, I will show below that it can be accounted for in the same way as with Udi; the clitic gets positioned in the second position within the verb and at that point is subject to further readjustments based on the surface requirements of Sorani verb forms. At this point however, I move on to providing an account of the split-ergativity.

3.2. Split ergativity

The pattern of split-ergativity that we find in Sorani Kurdish presents an extremely interesting, and complex, challenge for the theory of clitics. The problem that split-ergativity presents is fundamentally one of look-ahead. Clitics in Sorani, like Udi, are limited to appearing within the lower domain of the sentence. That clitics are confined to this domain is suggested by the fact that they cannot attach to the subject or temporal adverbs. However, if we assume that temporal adverbs lie outside of the lower domain, that is, they are elements that lie in the TP domain, and assume that clitics never get that far, then the absence of clitics targeting them falls out naturally. If this is true, then it must be the case that the subject lies in the TP domain, since subjects always appear to the left of temporal modifiers (McCarus 1958). Taking it to be true that clitics are confined to the lower domain, then it follows that they must merge into the derivation before the tense information is introduced, which under standard assumptions is introduced in TP.

At this point, a tricky problem presents itself for theories of cliticization where the clitic merges into the derivation with its associated argument. The argument which is the subject in Sorani must obligatorily be marked, whether it is by affixal agreement on the verb or by a clitic. However, at the point at which the subject is merged into the derivation, it is not known whether a clitic is to be needed or not. The same goes for the direct object, which can be marked by a clitic in the present tense. The direct object is clearly introduced into the derivation prior to any tense marking, therefore it is not known whether a big-DP is needed to introduce a clitic.[20]

Here I give an analysis that is consistent with the one given above for Udi, where it is assumed that the clitic is generated away from the argument. To begin, I follow Atlamaz & Baker's (2013) (A&B henceforth) conclusion about morphological case in Kurdish. Their analysis is based

---

[20] Note that I'm not claiming that the Big-DP hypothesis is wrong in every instance of cliticization, merely that it is inappropriate for Sorani. It may be the case that the clitic systems of Romance languages, which primarily have motivated Big-DP analyses, are a different type of clitic constructions and hence stem from a separate source. I leave this open for future research, since the data needed to make a thorough comparison of Sorani to Romance languages aren't available to me.

on the Kurmanji dialect of Kurdish, where case marking is transparent unlike in Sorani, we do not find the same case markers as we do in the Kurmanji dialect. However, as will be discussed, the patterns of affix vs clitic agreement allow us to draw the same conclusions. A&B show that Kurmanji Kurdish has the same type of split-ergativity that we see in Sorani. In the present tense transitives, agreement of the verb is with the subject of the sentence, which appears without a case marker and the object is marked by what is traditionally called the *direct* case (72), marked here as oblique as per their notation. However in past tense transitives, it is the object which agrees with the verb and now appears without the case marker, whilst the subject is marked by the direct case (73). Intransitives of both tenses show a non-case marked subject which agrees with the verb (contrast the form of the 1SG pronouns in (74) versus when it is not-nominative, (75)):

(72) Ez   Eşxan-ê   dı-vun-ım-e   [Kurmanji]
     I.NOM   Eşxan-OBL   IMPF-see.PRES-1SG-PRES.COP
     'I am seeing Eşxan.'

(73) Eşxan-ê   ez   di-m
     Eşxan-OBL   I.NOM   saw.PAST-1SG
     'Eşxan saw me.'

(74) a. Ez   rivi-m
         I.nom   run.PAST-1SG
         'I ran.'

     b. Ez   dı-rv-ım-e
         I.nom   IMPF-run.PRES-1SG-PRES.COP
         'I am running.'

(75) Tı   mı   dı-vun-ê
     you.nom   I.OBL   IMPF-see.PRES-2SG.PRES.COP
   'You are seeing me.'

Note that we see the same pattern of split-ergativity in Sorani Kurdish. If we compare with the examples above, we see that all the arguments that A&B mark as nominative (A and S in the present tense, S and O in the past tense) are those which are marked by affixal agreement on the verb in Sorani. Given the widespread cross-linguistic tendency for verbal agreement to track the nominative argument (see Bobaljik 2008, Preminger 2011), it is reasonable to assume that the arguments that are marked by affixal agreement in Sorani are in the nominative case. By way of contrast, those arguments which correspond to oblique arguments in Kurmanji are the ones that are (at least potentially, since object marking is not obligatory) marked by clitics in Sorani, namely O in the present tense, and A in the past tense. I therefore assume that the pattern of Case marking in Sorani is equivalent to that of Kurmanji Kurdish, with case differences not morphologically realized.

Returning to the issues under discussion, it needs to be explained not just why the clitics mark the arguments that they do (which as we have seen is different according to tense), but why they appear where they do. We noted earlier that the clitics show a varied distribution in Sorani, as summarized below:

(76) a. Enclitic to the leftmost XP that is within the lower domain (i.e. never on the subject or manner adverbs).
  b. If no material is available, the clitics are mesoclitic within the verb, appearing after the first morpheme, unless, either of the following hold:
    i. The clitic is 3SG, in which case it goes to the end.
    ii. The participle marker *bû* is present, in which case the clitics do not intervene between this and the verb.

It should be already clear from the wording in (76b) that the clitics in Sorani Kurdish seem to be the same type of clitics that we see in Udi, namely that they come to occupy the second morphemic position within the verb. This provides further support for the analysis for Udi given above, in that the novel proposal that there are clitics which are word-internal second position finds support from an unrelated (both genetically and geographically) language.

The similarities between Udi and Sorani do not stop at the existence of word-internal second position clitics, but rather the entire system of clitic placement in Sorani has a great deal in common with Udi. Consider (76a), which states that the clitics appear enclitic to the leftmost element within the lower domain of the verb. It's clear that the clitic remains in the lower domain, since it never attaches to the subject or adverbs which I have assumed to lie outside of vP above. I have also argued that the split ergativity in Sorani renders any Big-DP analysis of cliticization unsuitable for Sorani, since it suffers from a significant look-ahead problem. Therefore, we can assume the same that we did for Udi above, that the clitic is base generated as the head of ClP in a position underneath vP. I assume that this is the same as it was in Udi, with the clitic appearing higher than VP, but beneath vP However, given that whenever there is material available in the lower domain, the clitic attaches to the end of it, I assume that ClP has some type of EPP feature that forces material to appear in its specifier if material is available.

(77)

```
vP
├─
└─ v'
   ├─ v
   └─ ClP
      └─ Cl'
         ├─ Cl
         └─ VP
            └─ V'
               ├─ object
               └─ V
```
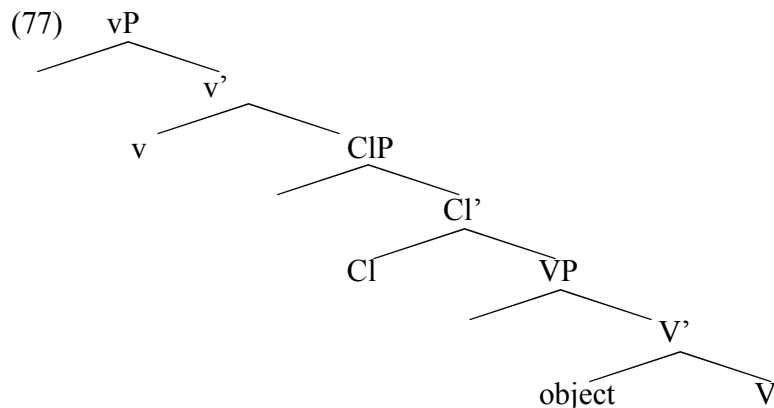
Using the same set of assumptions as in Udi, we can see that the clitic undergoes m-merger to the left-adjacent material, giving the clitic a host word to lean on, and allowing the clitic to satisfy its need to be a suffixal element (again, like in Udi, these clitics never appear in a proclitic position). Consider the following sentence in (78). The EPP feature of ClP forces movement of the indirect object to XP, giving the structure in (79).

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

(79) [TP Narmin [T' T [vP [v' v [XP [PP ba Sirwan]i X [ClP [Cl' **-î** [VP [v' ti dâ-n]]]]]]]]]

When the lower phase is spelled out, it spells out the following bit of structure, as the complement of *v*:

(80) [$_{XP}$ [$_{PP}$ ba Sirwan]$_i$ X [$_{ClP}$ [$_{Cl'}$ **-î** [$_{VP}$ [$_{V'}$ t$_i$ dâ-n]]]]]]]]]]

Since there is material to the left of the clitic, the clitic is able to undergo m-merger and rebracket inside the PP that lies in Spec,XP. We end up with the resulting configuration (lables an irrelevant structural details omitted):

(81) [$_{XP}$ [ba Sirwan-**î]** [dâ-n]]

Thus we see why the clitics in Sorani are at times enclitic to some phrasal element. They are attracted to this position since their host has moved to their left within the syntax. The clitics are spelled out without a host word, and so lean on any left adjacent material, which is there by virtue of the movement to Spec,XP to satisfy the EPP requirement that that position be filled.

3.3. Second position placement and jigging around

I now turn to the cases where the clitics of Sorani get positioned internal to the verb. The data that exemplify this are examples such as the following:

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

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

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

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

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

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

If we exclude (83) and (84), we can see the generalization that the clitic is not bound to a particular element within the verb form, say for instance the verbal root. This is shown in particular by (87), where the clitic is not adjacent to either the left edge of the root or the right edge. Rather, it is

positioned between the progressive and negative prefixes. The position of the clitic *is* however able to be analyzed in linear terms - in the second position within the word.

These cases only arise when the VP internal material is either absent or covertly realized. Sorani sentences can happily consist of just the verb alone, with the arguments expressed by means of clitics and affixes on the verb. Again, just as was the case in Udi we can analyze these cases as the result of the clitic being left freestanding at spell-out and searching for a host to lean on. When there is no VP internal material then the clitic cannot lean on anything to the left, therefore, the clitic will need to lean on the verb. However, since the verb appears to the clitic's right, then simply leaning on the verb is not enough to fully satisfy the positioning demands of the clitic, which, like the clitics in Udi, is inherently suffixal. Rather, the clitic must also shift inside the verb itself in order to ensure that it has material to its left. This results in the clitic being positioned in the second morphemic position in the verb

What about the exceptions to second position, notably the third position that is seen in (83) and the fourth position that is seen in (84). Third position is potentially easily explainable on the grounds of morphotactics; nothing can intervene between the verb and the participle marker *bû* aside from a passive marker *-re* (Thackston 2006, Walther 2013). I assume that *bû* must form a close relationship with the verb, but so must the passive morpheme. Given that morphotactically both must be adjacent to the verb, and only one slot is available to the right of the verb, the passive 'wins'. However, since the clitics do not show any particular affinity toward the verb, they are happy to move away so that *bû* can be next to it.

The case where the clitic goes in fourth position in the verb in (84) is more curious. However, it can be seen that this is the result of the clitic being 3SG. Thackston (2006) notes that the 3SG clitic always appears after the verbal affix, no matter what other elements are in the verb form.[21] However, it seems that it is also appropriately stated that it goes at the end of the verb. At this point it is not clear to me why 3rd singular clitics behave this way. The fact that it is the 3rd singular clitic, the most unmarked morphological combination seems to play some role, however anything beyond this would be speculative at this stage. What is important however is that it is a systematic departure from second position, therefore it can be viewed as an additional condition that moves the clitic away from its original place when it was the second position within the verb. For now though, I leave the speculations there.

3.4. Conclusion

In this section I have provided an analysis of the clitics in Sorani Kurdish, as well as the pattern of split ergativity. In doing so, I have provided independent evidence supporting the high position of clitics in the aspectual domain of Udi, as well as evidence that there are some clitics that come to lie in the second position within a word. Furthermore, as was the case in Udi, arbitrary morphotactic rules about Sorani Kurdish at times force the clitic to be metathesized away from a position where it originally is placed, a finding that is in accord with Arregi & Nevins (2012) for Basque.

4. Further discussion

One of the central claims that I have made in this paper is that clitics can never be placed directly inside another morpheme. When we therefore see a clitic that appears inside another morpheme, the positioning there must be indirect - the clitic placed in some position other than its surface position,

---

[21] Ž. Bošković (p.c.) points out that this is similar to a pattern seen in Serbo-Croatian (SC) 3SG auxiliary clitics, which follow all other clitics in SC, a different position to the other auxiliary clitics. The extent to which these two phenomena share a related analysis I leave open for future research.

and then moved inside the morpheme in order to satisfy some placement-external requirement. Along the way, I have criticized the approaches of Harris (2002) and Yu (2007) for being too permissive, in predicting patterns of clitic placement that we do not see attested in the world's languages. In this section I wish to briefly discuss two further cases of purported non-peripheral cliticization that support the proposal here. The first case is Degema, using data reported by Kari (2003), who argues that Degema is another example of Udi-style endoclisis, with a clitic moving internal to another morpheme. The second case is Pashto, a long-studied example of a non-peripheral clitic (Tegey 1977, Roberts 1997, Anderson 2005 *a.o.*), where it appears as though a clitic subcategorizing for phonological information can appear internal to another morpheme. If this were the case, it would be problematic for the proposal here, since I have assumed that clitics are bound by morpheme structure.

4.1. Degema

Another example of purported 'endoclisis' comes from Degema, a Niger-Congo language spoken in Nigeria. A description of the clitic system of this language can be found in Kari (2003)[22], where all the following data are taken from (unless cited otherwise). As Kari shows, Degema has a variety of clitics expressing various categories. Agreement with the subject for instance, is marked by means of proclisis, as shown below in (88):

(88) **mó**-kpéɲ   ísama
       3SG-wash   shirts
       '(S)he washes shirts.'

Similarly, there are a range of enclitics in the language, that mark verbal information, such as the following 'discontinuation' enclitic, which "suggests that somebody/something will stop or has stopped doing something (Kari 2003:96)":

(89) ə́-ji-**munu**.
       2PL.NEG-come-DE
       'You are not coming again.'

The clitic that is of interest to us, and analyzed by Kari as an instance of an endoclitic, is the factative enclitic (FE in the glosses). This clitic is generally an enclitic, which attaches to the final element within the clause. This is shown below, in (90) and (91) below:

(90) mi-bí-ꜜ**ín**
       1SG-BE.black-FE
       'I am black.'

(91) mɪ-sɪsɛ́-ꜜ**ɛ́n**
       1SG-remove-FE
       'I removed something.'

In the above, we can see that the form of the clitic is -Vn. The vowel is analyzed as being unspecified for segmental content, and so harmonizes with the adjacent vowel. The vowel also contains a downstep tone, which is retained even after the harmonization process. Kari notes that

---

[22] For further discussion on the clitics of Degema, see also Kari (2001a,b) and Kari (2009).

the factative enclitic will remain an enclitic, as long as the host ends in a vowel. However, if the host ends in a consonant, then the clitic apparently undergoes displacement, and surfaces inside the final consonant of the host:

(92) mɪ-tá-ꜜá-m
       1SG-chew$_1$-FE-chew$_2$
       'I chewed (sth.).'

(93) mi-ɓó-ꜜó-l
       1SG-hold$_1$-FE-hold$_2$
       'I held (sth.).'

In (103) and (104), we see that the factative clitic is located internal to the host verb. The verbal roots are made discontinuous by the addition of this clitic, with the verbs in (103) and (104) being *tám* and *ɓól* respectively. In both of these cases, we see that the vocalic part of the clitic tucks inside the final consonant of the host. This is seen by a lengthening of the vowel in the root, and crucially, the retention of the downstep tone of the clitic. Since the tone is kept, Kari concludes, that the factative morpheme does not simply lengthen the vowel in the host but the downstep tone is a real part of the clitic, which shifts inside the host. This is not simply a floating tone migrating leftwards either, since the vocalic part of the clitic is retained. We can assume, along with Kari, that the factative clitic is -ꜜVn (see Bye & Svenonious 2011 for discussion on unspecified segmental content being a possible exponent):

(94) [factative] ⇔ -ꜜVn

Kari assumes that endoclisis arises in the following steps. First, the final consonant of the host and the vocalic part of the clitic reorder. Secondly, there is harmonization, which makes the adjacent vowels agree in content. Finally, the [n] deletes at the end of the word, in order to avoid an illicit CC coda. Kari concludes that a metathesis operation is the best way of accounting for the data, however he assumes a form of generalized metathesis operation.

Whilst I agree with Kari's intuition that some form of metathesis is at play here, it is not concrete exactly what Kari assumes. If we follow his steps, then it seems as though a purely phonological metathesis is needed, not the switch between a morpheme and a phoneme that Kari claims. If [n] moves with the clitic and then deletes, then it seems as though we do have an instance of a morpheme switching places with a phoneme. Kari is not clear on the details of what triggers the metathesis rule. It is clearly related to the presence of a consonant in the final position of the host, since that is the only real split between the enclitic and the endoclitic cases. What is not clear is why the clitic needs to move when there is a consonant in the way. Degema has a process of vowel harmony that happens legitimately across consonants and so it is not simply the case that the clitic needs to be adjacent to a verb for this to happen.

What is clear about the Degema case is that the clitic initally attaches to the right edge of the word and migrates inwards if there is a consonant for it to tuck inside of. This may be held up as an instance of a clitic which is at its most basic a true endoclitic since it will move internal to another morpheme if it can. However, it is not clear what the reason is for the clitic moving inside of the word. It could plausibly be related to the requirement that the downstep tone of the clitic be properly associated with a vowel. If the exponent of the clitic really is a vowel-slot without segmental content to begin with then the tone will not be pronounceable unless the tone either

moves to another vowel, or the vowel gets a segmental value from elsewhere. In Degema, what could be happening is that a tone associated simply with V could trigger a phonological readjustment rule like the following, which copies the content of an *adjacent* vowel:

(95) ꜜV → ꜜα / α __

Importantly, this explanation requires two things. Firstly, the rule must only work under adjacency, and so phonological metathesis applies in order to get the unspecified vowel close enough to another vowel in order to allow the copying process to work. This movement of the clitic strands the /n/ part of the clitic, which is further deleted to avoid a CVCC syllable. Secondly, this rule must happen before the point of vowel harmony happens in the language, since the harmonization process would grant segmental content to the vowel, and there would be no need for the rule to apply.

4.2. Pashto

Now I turn my attention to Pashto, one of the earliest, and most widely discussed cases of a non-peripheral clitic. The central observation in Pashto is the following data set from Tegey (1977), who observes that the clitic which marks the subject of the verb can apparently appear in two different places on certain verbs in the imperfective aspect. The pattern is illustrated as in (107) and (108) below:

(96)   a.   axistә́lә-**me**   b.   á-**me**-xistәlә
            buy-1SG   buy$_1$-1SG-buy$_2$
            'I was buying them.'   'I was buying them.'

(97)   a.   aɣustә́-**me**   b.   á-**me**-ɣustә
            wear-1SG   wear$_1$-1SG-wear$_2$
            'I was wearing it.'   'I was wearing it.'

In (96a) and (97a), we see that the clitic is attached to the end of the verb, whereas in (96b) and (97b) the clitic has moved internal to the verb, apparently to follow the position of stress. The shift in stress is optional in these forms and apparently comes without any difference in meaning. Relevant here is the observation that if these data are as they seem, then the proposal that I am making faces a serious counterexample. I have been claiming that clitics placement is bound by morphemic structure, and consequently there cannot exist a clitic that gets placed directly inside another morpheme, but all morpheme internal positionings arise independently, and crucially *after* the clitic has been placed elsewhere. Pashto seems to fly in the face of this proposal, since the placement of the clitic seems to be conditioned by stress, and the clitic seems to be breaking up the integrity of a morpheme.

However, as has been discussed, things aren't what they seem to be in Pashto. Despite Tegey's claim that "it is important to bear in mind that in such instances the clitics are placed after a phonological segment which constitutes part of the root (i.e. '*a'*- PWS), and which is not a separate morpheme. (Tegey 1977:89)" there is sufficient reason to doubt that *a-* does not constitute a morpheme in and of itself. Kaisse (1981) shows that the grounds for analyzing *-a* as part of the root are flawed. Firstly, there are only nine verbs in Pashto that begin with a vowel, and these all have the *a-* morpheme at their beginning. Analyzing this as a prefix, Kaisse claims allows us to formulate a general condition on verbstems in Pashto that they do not start with a vowel. Secondly, Tegey considers *-a* as part of the root since there is no synchronic meaning to it. However, this does

not entail that it is not a morpheme in itself, but rather it could simply be a bound prefix. Furthermore, there are other verbs in Pashto which Tegey does claim are bimorphemic on the grounds that the clitic can intervene between the two morphemes. If *-a* is a prefix, then it fits into a wider pattern of mesoclisis that is already seen in the language. Kaisse eventually concludes that the verbs in Pashto that begin with *-a* are bimorphemic, consisting of a bound prefix and the verb root. Stress can go on either morpheme and the clitic will follow stress, accounting for the variable clitic placement. The examples in (96b) and (97b) are then more accurately glossed as follows in (98a) and (98b) respectively, with there being a compound verb in place of the discontinuous root seen above:

(98)   a.   á-**me**-xistələ   b.   á-**me**-ɣustə
             PREFIX-1SG-buy   PREFIX-1SG-wear

What is relevant for our purposes here is that the clitic placement is sensitive to stress. Therefore, like Yu predicts, we have a clitic which can seemingly subcategorize for phonological in nature. The question is what happens when there is a multi-syllabic root that has stress on the first syllable? This question is interesting, because if we see the clitic going inside the morpheme, then we have a clear counterexample to the (strongest form of the) proposal I make here, since there would be a clitic placement directly inside another morpheme. However, we see that as predicted on the account here, the clitics in Pashto remain bound by morpheme structure, in that they can only attach to the *morpheme* that is in stress, not the syllable. This is shown by the following verbs, which permit the same stress shift in the imperfective, where the regular stress shifting to the initial syllable. Clitics, always appear to the right regardless of where the stress is, indicating that these truly are monomorphemic verbs:

(99) a.   pərebdə́ **me**   b.   pə́rebdə   **me**
           beat   1SG   beat   1SG
           'I was beating him.'   'I was beating him.'

4.3. Summary

In this section I have remarked on two further purported cases of endoclisis, Degema and Pashto. Both of these cases are in accord with the analysis presented here. Pashto showed us that even though clitics *can* be placed according to phonological information, they are still bound by the morphemic structure of the word which is their host. We saw that even though the clitics in Pashto are placed according to word stress, they still can't move any closer to the stressed syllable than morpheme boundaries of the word; that is, they appear after the first stressed morpheme, and not internal to any morphemes in themselves. This shows striking support for the proposal given here that clitics can never break the integrity of a morpheme when they get placed. The only circumstances in which a clitic will break the integrity of a morpheme is to satisfy requirements external to the placement of the clitic.

Degema, like Udi, gave us another example of a clitic that is positioned in one place, but moves to another position on the surface in order to satisfy requirements that are not directly related to the placement of the morpheme. Whilst the extra movement in Udi was shown to be forced by the need for the verb root and the TAM suffix to be adjacent, in Degema the movement of the clitic is more phonological in nature. The exponent of the clitic needs to be associated with a vowel. This must be done under adjacency with a particular vowel, and so if the clitic adjoins to a stem that ends in a vowel, the clitic copies the features of the vowel onto its V, and an epenthetic consonant is added to satisfy the requirement that words end in a closed syllable. On the other hand, where we

see endoclisis, the V exponent of the clitic metathesizes in the phonology in order to be adjacent to a vowel to allow the vowel copying process to happen.

5. Conclusion

In the course of this paper, I have investigated cases of clitics that appear non-peripherally within their host, a position that is extremely rare across languages. I have made, and defended, the proposal that cliticization cannot in general target the internal structure of a word, but clitics can be moved there as a result of being positioned in a position that violates their own requirements. This has led to the proposal that clitics that appear internal to their host word should appear no further in than one morpho-syntactic element in their word (since I take the strong stance that cliticization precedes insertion of phonological exponents). This proposal has been shown to provide a restrictive analysis of the extremely complex clitic system in Udi, as well as giving us an insight as to why the endocliticization seen in Udi is so rare typologically. It is not a primitive option that Universal Grammar allows, contrary to the proposals of Harris (2002) and Yu (2007), but rather the particular placement in Udi requires a number of different factors to come together. It is true that even on the approach taken here for Udi we expect endoclisis patterns to arise in the language but in contrast to Harris and Yu's approaches, we predict its extreme rarity relative to infixes. Each of the assumptions needed for Udi - word internal second position plus morphotactic rearrangement - were independently needed for Sorani Kurdish, providing evidence that the confluence of factors seen in Udi were no fluke. The difference between the languages is that Sorani chooses to enact its metathesis repair outwards to the right as opposed to inwards to the left.

There are various questions which remain open at this stage, which have not been fully addressed in this paper through considerations of space and scope. Firstly, Udi and Sorani Kurdish were both analyzed as clitics that face a conflict with respect to positioning. They want to be leftmost within the verb, but must have something separating them from the word edge, explaining why they come to occupy the second position within the word. This makes the prediction that there should be clitics which want to be positioned rightmost within their host, but need something separating them from the edge. Thus, we expect to find clitics which come to be penultimate within their host. However, no clear cases of this type exist. Degema was discussed in section 5, however at this stage the inward movement of the clitic is more likely phonological rather than involving the morphosyntax like Udi and Sorani. It remains to be seen whether this typological gap will be filled by some language, however it should be noted that this is also a well-known typological gap in sentence level clitics (Halpern 1995, but The outcome of this outstanding issue will have ramifications for analyses of 'second position.'

Another question which has been left open, and which will attract debate for many years to come, is the difference between clitics and affixes. Positing a true distinction between the categories is simultaneously conservative and controversial, with there being good arguments either way for treating clitics and affixes differently and the same. In this paper, I proposed that clitics and affixes differ descriptively in that clitics never position themselves within another morpheme consistently. That is, there is no true correlate of infixation to be found with clitics. Where this difference stems from is not entirely clear. What I have shown in this paper is that there is no known case that forces us to concede this descriptive generalization; endoclisis in Udi can be fruitfully analyzed as a clitic that begins life intermorphemically but moves to a position intramorphemically under pressure from other elements. There is no sense that the intramorphemic placement is an inherent property of the clitic itself, despite Harris and Yu's suggestions. This finding is not expected if clitics comprise the same things as affixes, since there do seem to be affixes which *are* true infixes, and always move inside another morpheme. To the extent that the proposals put forward in this paper hold their weight, it seems that we can identify a deep

difference between clitics and affixes; affixes can apparently be positioned internal to another morpheme, whereas clitics can never be.

References

Anderson, S. (1992). Amorphous morphology. Cambridge, Cambridge University Press.

Anderson, S. (2005). Aspects of the theory of clitics. Oxford, Oxford Studies in Theoretical Linguistics.

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

Atlamaz, A. & Baker, M. (2013). On the relationship of case to agreement in split-ergative Kurmanji. Ms, Rutgers University.

Bastos, A. (2011). Information structure within the traditional nominal phrase: The case of Brazilian Portuguese.

Belletti, A. (2004). Aspects of the low IP area. In Rizzi, L. (ed.) *The Structure of CP and IP.* Oxford Studies in Comparative Syntax, Oxford University Press.

Belletti, A., & Shlonsky, U. (1995). The order of verbal complements: A comparative study.

Blevins, J. (1999). Untangling Leti infixation. *Oceanic Linguistics, 38.2,* 383-403.

Bobaljik, J.D. (1995). *Morphosyntax: the syntax of verbal inflection.* Doctoral dissertation, MIT.

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.

Bobaljik, J.D. (2008). Where's phi? Agreement as a post-syntactic operation. In Harbour, D., D. Adger & S. Bejar, (eds.), *Phi-theory, phi-features across modules and interfaces.* Oxford Studies in Theoretical Linguistics, Oxford University Press, pp. 295-328.

Bobaljik, J.D. (2012). Universals in Comparative Morphology.

Bobaljik, J.D. and Wurmbrand, S. (2005). The domain of Agreement. *Natural Language and Linguistic Theory* 23:4, pp. 809-865.

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

Bošković, Ž. (1997). Fronting wh-phrases in Serbo-Croatian.

Bošković, Ž. (2002). *On the nature of the syntax-phonology interface: Cliticization and related phenomena.* Oxford, Elsevier.

Bošković, Ž (2014). Now I'm a phase, now I'm not a phase. *Linguistic Inquiry 45(1),* 27-89.

Bye, P. & Svenonious, P. (2012). Non-concatenative morphology as epiphenomenon. In Trommer, J. (ed.), *The morphology and phonology of exponence.* Oxford Studies in Theoretical Linguistics, Oxford University Press.

Caha, P. (2009). *The nanosyntax of case*. PhD dissertation, University of Tromsø.

Calabrese, A. & Pescarini, D. (2013). Clitic metathesis in the Friulian dialect of Forni. Ms.

Chomsky, N. (2000). Minimalist Inquiries: The framework. Martin, R., Michaels, D. & J. Urigareka, *Step by step: Essays on minimalist syntax in honor of Howard Lasnik.* MIT Press, Cambridge, MA, pp. 89-115.

Chomsky, N. (2001). Derivation by phase. In In *Ken Hale: A life in language.* MIT Press, Cambridge, MA.

Chomsky, N. (2008). On phases. In *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, ed. by Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta, 133–166. Cambridge, MA: MIT Press.

Condoravdi, C. & Kiparsky, P. (2004). Clitics and clause structure. *Journal of Greek Linguistics 5,* 159-183.

Di Sciullo, A. & Williams, E. (1987). On the definition of Word. Cambridge, MA: MIT Press.

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

Embick, D. & Noyer, R. (2001). Movement operations after syntax. *Linguistic Inquiry 32,* 555-595.

Everett, D. (1996). *Why there are no clitics: An alternative perspective on pronominal allomorphy.* The Summer Institute of Linguistics.

Foley, (1991). *The Yimas Language of New Guinea.* Stanford, CA: Stanford University Press.

Halpern, A. (1995). *On the placement and morphology of clitics.* CSLI Publications, Stanford, CA.

Halle, M. (2001). Infixation versus onset metathesis in Tagalog, Chamorro, and Toba Batak. In *Ken Hale: A life in language.*

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. (1982). Case marking, verb agreement, and inversion in Udi. In D. M. Perlmutter and C. G. Rosen, eds., *Studies in Relational Grammar* 2, pp. 243–258, University of Chicago Press, Chicago.

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.

Kari, E. (2003). Clitics in Degema: a meeting point of phonology, morphology and syntax. Tokyo : Research Institute for Languages and Cultures of Asia and Africa (ILCAA), Tokyo University of Foreign Studies, 2003

Klavans, J. (1979). On clitics as words. *The Elements: A parasession on linguistic units and levels.*

Klavans, J. (1985). The independence of syntax and phonology in cliticization. *Language 61,* 95-120.

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

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

Luís, A. & Spencer, A. (2012). *Clitics: An introduction.* Cambridge Textbooks in Linguistics, Cambridge University Press.

Macaulay, M. (1989). A suffixal analysis of the Karok 'endoclitic'. *Lingua 78,* 158-180.

Manzini, M. & Savioa, L. (2012). Mesoclisis in the imperative: Phonology, morphology or syntax? *Lingua 121,* 1101 - 1120.

Marantz, A. (1988). Clitics, morphological merger and the mapping to phonological structure. In Hammond, M. and Noonan, M. (eds.), *Theoretical Morphology: Approaches in modern linguistics, 253-270.* San Diego, Academic Press.

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

Moskal, B. (2013). A case study in nominal suppletion. Ms. University of Connecticut, Storrs.

Nevis, J. (1984). A non-endoclitic in Estonian. *Lingua* 64, 209-224.

Nevis, J. (1988). A morphotactic paradox in Northern Saame: Comitative *-guim.* Ural-Altasische Jahrbücher 8, 38-50.

Preminger, O. (2009). Breaking agreements: Distinguishing agreement and clitic doubling by their failures. *Linguistic Inquiry 40.4,* 619-666.

Preminger, O. (2012). *Agreement as a fallible operation.* Doctoral dissertation, MIT.

Roberts, T. (1997). *The optimal second position in Pashto.* Rutgers Optimality Archive #174. http://roa.rutgers.edu/files/174-0297/174-0297-ROBERTS-0-0.PDF

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

Scalise, S. (1984). *Generative Morphology.* Dordrecht: Foris.

Sportiche, D. (1996). Clitic constructions. In Rooryck, J. & Zaring, L. (eds.), *Phrase Structure and the Lexicon.* Dordrecht, Kluwer.

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

Wojdak, R. (2005). *The linearization of affixes: evidence from Nuu-chah-nulth.* Doctoral dissertation, University of British Columbia.

Woolford, E. (2003). Clitics and agreement in competition: Ergative cross-referencing patterns. In Carpenter et al (eds), *Papers in Optimality Theory II.* Amherst, MA: GLSA.

Yu, A. (2007). *A natural history of infixation.* Oxford Studies in Theoretical Linguistics.

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