Fitting into morphological structure: accounting for Sorani Kurdish endoclitics

Géraldine Walther
Sorbonne Paris Cité, Université Paris Diderot & CNRS
Laboratoire de Linguistique Formelle
geraldine.walther@linguist.jussieu.fr

0. Abstract: In this paper, we investigate the notion of morphological structure in the light of the behaviour of Sorani Kurdish mobile person markers (MPM) (Samvelian, 2007). We propose a formal analysis of Sorani Kurdish transitive verb inflection taking into account the complete Sorani verb paradigm, including the past tense forms that had been neglected so far. We argue that some structure is needed, even within a realisational approach, to account for the otherwise erratic behaviour of Sorani Kurdish MPMs. Our analysis is couched within a variant of *PFM* (Stump, 2001).

Keywords Sorani Kurdish, mesoclisis, endoclisis, morphology-syntax interface, morphological structure, lexeme-based morphology, non-canonical person marking, PFM.

1. Introduction

The question whether or not there should be structure specified within morphology has provoked numerous debates among contemporary morphologists. Lieber (2006) even goes so far as to call it "a veritable war concerning the relationship between syntax and morphology, with positions ranging from the militantly lexicalist to the staunchly syntacticalist." While the lexicalist, non-morphemic position (Anderson, 1992; Beard, 1995; Stump, 2001) argues that there can be no structure in morphology, the opposed syntactic position, building on the notion of morpheme (Lieber, 1992; Baker, 1988), argues for the exact opposite, namely that morphological structure is of the same kind as syntactic structure.

The existence of structure in morphology also questions the *Lexical Integrity Hypothesis* (Bresnan, 1995), usually broadly admitted among lexicalists. In the introduction to her work on Udi endoclitics, Harris (2002) points out the following:

> "Many linguists have assumed the correctness of the Lexical Integrity Principle, the hypothesis that words are composed according to morphological principles that differ in kind from the syntactic principles responsible for the composition of sentences. According to this hypothesis, the morphological composition of a word is not accessible to the rules of syntax, and items manipulated by syntactic rules (i.e., words) cannot be manipulated by the morphology (see, for example (Di Sciullo and Williams, 1987; Bresnan, 1995))."

However, data of the kind described, for example, by Harris (2002) for Udi show that the Lexical Integrity Hypothesis is seriously challenged by some clitic placement phenomena. This is also illustrated by Crysmann (2002) on European Portuguese, where clitics occur inside words, and by Crysmann (2006) who reanalyses the data presented by Kupść and Tseng (2005) for Polish, where elements described as having affixal status "float of" leaving their word domain to attach with other syntactic units.

The clitic placement data in these papers constitutes a particular challenge for a-morphous morphology. In Crysmann's (2002) analysis for example, word-internal, morph-based structure is needed to account for the *endoclitic*[1] behaviour of European Portuguese person markers.

In this paper, we adopt a lexicalist point of view. In particular, we follow an inferential-realisational approach to word-form realisation (Zwicky, 1985; Anderson, 1992; Corbett and Fraser, 1993; Stump, 2001; Brown and Hippisley, 2012). However, drawing on evidence from known (Samvelian, 2007) and original data from Sorani Kurdish endoclitics, we claim that the solution to the question on morphological structure lies on some intermediate golden path between morphemic and a-morphous views of morphology. We show that some morphological structure needs to be at least partly accessible in order to account for Sorani clitics' particular placement properties. In particular, we argue that their seemingly erratic interaction with Sorani perfect tenses, which only superficially resemble periphrasis, can only be explained through internal morphological structure.

2. Sorani Kurdish data

The data used in this paper has been mainly extracted from reference grammars (McCarus, 1958; Thackston, 2006; MacKenzie, 1961). To a lesser extend we have also used information contained in (Blau, 1980, reprinted 2000), and various other descriptions (Garzoni, 1787; Justi, 1880; Soane, 1913; Beidar, 1926; Wahby, 1956; Bedir Khan and Lescot, 1970). Some original data has also been directly obtained from native speakers of the language.

2.1. General presentation

Sorani Kurdish is a Western Iranian language mostly spoken in the Iraq-Iran border regions. It possesses a written standard based on the dialects of the cities of Sulaymaniyah and Arbil (Iraq). We use the more common name *Sorani* to refer to its standardised dialect, corresponding to what Haig (2010) refers to as *Suleimani,* a dialect of the Central Group of the Kurdish languages. Sorani mainly distinguishes itself through its complex verbal morphology and in particular its intricate system of *"endoclitic" or "mobile" person markers,* henceforth MPM (Samvelian, 2007).

Sorani Kurdish resembles most Iranian languages in the fact that it possesses only a very limited amount of synthetic verbal lexemes (around 300). Most traditional verbal meanings are expressed through complex verbal predicates built from a light verbal head and a predicative element which can be either a noun or an adjective, or even an adposition or a preverb or postverb (MacKenzie, 1961; Blau, 2000).

2.2 Sorani Kurdish verbal inflection and person marking

General verbal inflection pattern Sorani verb forms roughly consist of a set of prefixes and suffixes clustered around a given stem. The relative order of the affixes is illustrated by the position classes shown in Table 3.

[1] Crysmann (2006) uses the term *mesoclisis* for insertion of clitics between identifiable morphs, which allows for reserving the term *endoclisis* for insertion inside the stem. Stem-internal cliticisation is described for Udi verbs in (Harris, 2002), alongside inter-morph insertion. However, (Harris, 2002) uses the term endoclisis as a denomination for both types of word-internal placement, as defined by the original coining of the term by Zwicky (1977). In this paper, we thus follow Harris (2002) and Zwicky (1977) in using the term *endoclisis* as a general term for cases where a clitic appears inside the word thereby breaking its phonological integrity. Note that Anderson (2005) mentions both terms as synonyms and then resorts to using *endoclisis* as a general term as well.

Prefix(es) - Stem - Suffix(es)

Stems Most traditional descriptions of Sorani morphology concur in stating the existence of two distinct verbal stems: one for the present tense forms and one for the past and non-finite forms.

Example: KIRDIN "to do" (STEM1 *ke,* STEM2 *kird)*

(1) *da-**kird** =ê*
IND-***do***-P3.PRES
"He does/makes" (PRESENT INDICATIVE)

(2) *da-**ke**-Ø*
IND- ***do*** -P3SG.PAST
"He did/made" (IMPERFECT INDICATIVE)

(3) *bi-**ke**-m-aja□*
SUBJ-***do***-P1SG.PAST-SUBJ.PAST□
"I would have done/made" (PRETERITE SUBJUNCTIVE)

In addition to these two traditional stems, Bonami and Samvelian (2008) suggest the existence of a third stem for the passive forms due to observable irregularities occurring during stem selection for these forms. The relevant data can be found in (McCarus, 1958) of which some examples are given in Table 1.

	INFINITIVE	STEM1	PASSIVE STEM	TRANSLATION
STEM1	*kuštin*	*kuš*	*kuš–râ/rê*	to kill
STEM2	*ûtn*	*l'ĕ*	*ût–râ/rê*	to say
STEM2	*bistin*	*bye*	*bst–râ/rê*	to hear
STEM1 MINU END-VOWEL	*birdin*	*be*	*b–râ/rê*	to carry
STEM1 MINUS END-VOWEL	*kirdin*	*ke*	*k–râ/rê*	to do, to make
STEM1 MINUS END-VOWEL	*dân*	*de*	*d–râ/rê*	to give
MODIFIED STEM1	*xwardn*	*xo*	*xû–râ/rê*	to eat
MODIFIED STEM2	*girtin*	*gir*	*gîr–râ/rê*	to take

Table 1: Irregular stem formation for Sorani passive forms

Moreover, the data in reference grammars (McCarus, 1958; MacKenzie, 1961; Blau, 2000; Thackston, 2006) shows a fourth possible stem for imperative forms. Hence up to two additional stems may be stipulated for Sorani verbs (Walther, 2011). Examples are given in Table 2.

INFINITIVE	STEM1	IMP. SING.	IMP. PL.	TRANSLATION
kirdin	*ke*	*bi–ke*	*bi–ke–n*	to do, to make
roîštn	*ro*	*bi–ro*	*bi–ro–n*	to go
bûn	*bi*	*bi–bi–e*	*bi–bi–(i)n*	to be, to become
čûn	*č*	*bi–č–e*	*bi–č–(i)n*	to go
girtin	*gir*	*bi–gir–e*	*bi–gir–(i)n*	to take, to grasp
nûsîn	*nûs*	*bi–nûs–e*	*bi–nûs–(i)n*	to write
da nîštn	*da nîš*	*da bi–nîš–e*	*da (bi–)nîš–(i)n*	to sit down
birdin	*be*	*bi–be–re*	*bi–be-n*	to carry
dan	*de*	*bi–de-re*	*bi–de–n*	to give
xistin	*xe*	*bi–xe–re*	*bi–xe–n*	to pull
hatin	*yĕ*	*wer–e*	*wer-in*	to come

Table 2: Irregular stem formation for Sorani Kurdish imperative forms

Inflectional affixes Inflectional affixes cluster around the verbal stems: prefixes mostly convey tense, aspect and mood *(*TAM*)* and polarity features, whereas suffixes may encode TAM and person information. Thus the affix distribution can be represented in terms of position classes (as used in PFM (Stump, 2001)). The distribution is illustrated in Table 3.

-2	-1	0	(P)	1	2	3
ma	da	STEM	râ	+û	im	aja
na	bi	(1, 2, 3, 4)	rê	+bû	î(t)	a
nâ				b	ê(t)	
					a	
					în	
					in	
					Ø	
				n		

Table 3. Sorani Kurdish verbal position classes within a PFM analysis

Sorani passives Passives are formed by inserting the sequence *–rê/–râ* between the stem and the other suffixes (shaded column in Table 3). Apart from these markers, the passive inflection exactly matches the paradigms of active intransitives.[2] Active transitive verbs, on the other hand, follow other, specific patterns.

Person marking Sorani is traditionally considered to display three sets of personal endings: PE1 for the present verb forms and the perfect subjunctive derived from STEM1, PE2 for imperfective, preterite and past perfect forms derived from STEM2, and PE3, being identical to the enclitic present forms of the verb *bûn* 'to be', for the remaining perfect forms. In fact those three paradigms only differ in the third person singular, as shown in Table 4.

[2] Descriptive grammars implicitly present *–rê/–râ* as just another set of inflectional suffixes for transitive verbs (McCarus, 1958; MacKenzie, 1961; Blau, 2000; Thackston, 2006). Walther (2011) gives however arguments for a derivational analysis of Sorani passives.

Personal Endings		**PE3** Enclitic personal endings	
PE1 Present personal endings		*–m*	*–în*
–m	*–în*	*–î(t)*	*–n*
–î(t)	*–n*	*–a*	*–n*
–ê(t)	*–n*	**Mobile person markers MPM**	
PE2 Past personal endings			
–m	*–în*	*–m*	*–mân*
–î(t)	*–n*	*–t*	*–tân*
–ø	*–n*	*–î*	*–yân*

Table 4. Person marking

The MPM constitute a fourth set of person markers. They mark the object in the present tense, while the standard person markers are used for subject marking. In the past tense a morphological reversal in the sense of Baerman (2007) occurs: MPMs become subject markers whereas standard person markers are used for object marking.

Moreover, clitic MPMs (Samvelian, 2007) do not appear in a fixed slot among the positions represented in Table 3 when attaching to a verb form; they may insert in various positions following a seemingly erratic pattern.

1. When it is an endoclitic, the MPM inserts after the first prefix, in verb-internal second position.
   a. The first prefix may be the indicative present and imperfect prefix *da–*,
   b. the subjunctive and imperative prefix *bi–*,
   c. or the negative prefix (be it *na–, nâ–* or *ma–*).
2. The third singular person marker (be it standard or MPM) always follows the object marker.
3. Be it a subject or object marker, the first singular person marker always precedes any plural marker.

Tables 5 to 9 show the placement of the endoclitic (marked through the use of angle-brackets <>) for the indicative forms with positive polarity. Other present or past forms containing prefixes *bi–*, *na–*, *nâ–*, *ma–* behave like the present or imperfect respectively.

SUBJECT=PE1 / OBJET=MPM	P1SG	P2SG	P3SG	P1PL	P2PL	P3PL
P1SG		da=◇=STEM1–PE1	da=◇=STEM1–PE1	da=◇=STEM1–PE1	da=◇=STEM1–PE1	da=◇=STEM1–PE1
P2SG	da=◇=STEM1–PE1		da=◇=STEM1–PE1	da=◇=STEM1–PE1	da=◇=STEM1–PE1	da=◇=STEM1–PE1
P3SG	da=◇=STEM1–PE1	da=◇=STEM1–PE1	da=◇=STEM1–PE1	da=◇=STEM1–PE1	da=◇=STEM1–PE1	da=◇=STEM1–PE1
P1PL	da–STEM1–PE1 =◇	da=◇=STEM1–PE1	da=◇=STEM1–PE1	da=◇=STEM1–PE1	da=◇=STEM1–PE1	da=◇=STEM1–PE1
P2PL	da–STEM1–PE1 =◇	da=◇=STEM1–PE1	da=◇=STEM1–PE1	da=◇=STEM1–PE1	da=◇=STEM1–PE1	da=◇=STEM1–PE1
P3PL	da–STEM1–PE1 =◇	da=◇=STEM1–PE1	da=◇=STEM1–PE1	da=◇=STEM1–PE1	da=◇=STEM1–PE1	da=◇=STEM1–PE1

Table 5. MPM placement in the present paradigm

SUBJECT=PE1 / OBJET=MPM	P1SG	P2SG	P3SG	P1PL	P2PL	P3PL
P1SG		da=◇=STEM2–PE2	da–STEM2–PE2 =◇	da=◇=STEM2–PE2	da=◇=STEM2–PE2	da=◇=STEM2–PE2
P2SG	da=◇=STEM2–PE2		da–STEM2–PE2 =◇	da=◇=STEM2–PE2	da=◇=STEM2–PE2	da=◇=STEM2–PE2
P3SG	da=◇=STEM2–PE2	da–◇–STEM2–PE2	da–STEM2–PE2 =◇	da–◇–STEM2–PE2	da–◇–STEM2–PE2	da=◇=STEM2–PE2
P1PL	da–STEM2–PE2 =◇	da=◇=STEM2–PE2	da–STEM2–PE2 =◇	da=◇=STEM2–PE2	da=◇=STEM2–PE2	da=◇=STEM2–PE2
P2PL	da–STEM2–PE2 =◇	da=◇=STEM2–PE2	da–STEM2–PE2 =◇	da=◇=STEM2–PE2	da=◇=STEM2–PE2	da=◇=STEM2–PE2
P3PL	da–STEM2–PE2 =◇	da=◇=STEM2–PE2	da–STEM2–PE2 =◇	da=◇=STEM2–PE2	da=◇=STEM2–PE2	da=◇=STEM2–PE2

Table 6. MPM placement in the imperfect paradigm

SUBJECT=PE1 / OBJET=MPM	P1SG	P2SG	P3SG	P1PL	P2PL	P3PL
P1SG		STEM2 =◇=PE2	STEM2–PE2 =◇	STEM2 =◇=PE2	STEM2 =◇=PE2	STEM2 =◇=PE2
P2SG	STEM2 =◇=PE2		STEM2–PE2 =◇	STEM2 =◇=PE2	STEM2 =◇=PE2	STEM2 =◇=PE2
P3SG	STEM2 =◇=PE2	STEM2 =◇=PE2	STEM2–PE2 =◇	STEM2 =◇=PE2	STEM2 =◇=PE2	STEM2 =◇=PE2
P1PL	STEM2–PE2 =◇	STEM2 =◇=PE2	STEM2–PE2 =◇	STEM2 =◇=PE2	STEM2 =◇=PE2	STEM2 =◇=PE2
P2PL	STEM2–PE2 =◇	STEM2 =◇=PE2	STEM2–PE2 =◇	STEM2 =◇=PE2	STEM2 =◇=PE2	STEM2 =◇=PE2
P3PL	STEM2–PE2 =◇	STEM2 =◇=PE2	STEM2–PE2 =◇	STEM2 =◇=PE2	STEM2 =◇=PE2	STEM2 =◇=PE2

Table 7. MPM placement in the preterite paradigm

SUBJECT=PE1 / OBJET=MPM	P1SG	P2SG	P3SG	P1PL	P2PL	P3PL
P1SG		STEM2–û=◇=PE2	STEM2–û–PE2 =◇	STEM2–û=◇=PE2	STEM2–û=◇=PE2	STEM2–û=◇=PE2
P2SG	STEM2–û=◇=PE2		STEM2–û–PE2 =◇	STEM2–û=◇=PE2	STEM2–û=◇=PE2	STEM2–û=◇=PE2
P3SG	STEM2–û=◇=PE2	STEM2–û=◇=PE2	STEM2–û–PE2 =◇	STEM2–û=◇=PE2	STEM2–û=◇=PE2	STEM2–û=◇=PE2
P1PL	STEM2–û–PE2 =◇	STEM2–û=◇=PE2	STEM2–û–PE2 =◇	STEM2–û=◇=PE2	STEM2–û=◇=PE2	STEM2–û=◇=PE2
P2PL	STEM2–û–PE2 =◇	STEM2–û=◇=PE2	STEM2–û–PE2 =◇	STEM2–û=◇=PE2	STEM2–û=◇=PE2	STEM2–û=◇=PE2
P3PL	STEM2–û–PE2 =◇	STEM2–û=◇=PE2	STEM2–û–PE2 =◇	STEM2–û=◇=PE2	STEM2–û=◇=PE2	STEM2–û= ◇=PE2

Table 8. MPM placement in the present perfect paradigm

SUBJECT=PE1 / OBJET=MPM	P1SG	P2SG	P3SG	P1PL	P2PL	P3PL
P1SG		STEM2–bû=◇=PE2	STEM2–bû–PE2 =◇	STEM2–bû=◇=PE2	STEM2–bû=◇=PE2	STEM2–bû=◇=PE2
P2SG	STEM2–bû=◇=PE2		STEM2–bû–PE2 =◇	STEM2–bû=◇=PE2	STEM2–bû=◇=PE2	STEM2–bû=◇=PE2
P3SG	STEM2–bû=◇=PE2	STEM2–bû=◇=PE2	STEM2–bû–PE2 =◇	STEM2–bû=◇=PE2	STEM2–bû=◇=PE2	STEM2–bû=◇=PE2
P1PL	STEM2–bû–PE2 =◇	STEM2–bû=◇=PE2	STEM2–bû–PE2 =◇	STEM2–bû=◇=PE2	STEM2–bû=◇=PE2	STEM2–bû=◇=PE2
P2PL	STEM2–bû–PE2 =◇	STEM2–bû=◇=PE2	STEM2–bû–PE2 =◇	STEM2–bû=◇=PE2	STEM2–bû=◇=PE2	STEM2–bû=◇=PE2
P3PL	STEM2–bû–PE2 =◇	STEM2–bû=◇=PE2	STEM2–bû–PE2 =◇	STEM2–bû=◇=PE2	STEM2–bû=◇=PE2	STEM2–bû= ◇=PE2

Table 9. MPM placement in the past perfect paradigm

Remnant split ergativity Finally, transitive Sorani verbs display remnant features of split ergativity (Haig, 2010): for transitive verbs in the present tenses, as well as for intransitive verbs, subject marking is handled as described above by using one of the three sets of standard person markers. In present forms of transitive verbs, MPMs only function as optional object marking. However, in the past tenses of transitive verbs, a morphological reversal (Baerman, 2007) occurs: the standard person markers are replaced with the MPMs as subject-verb agreement markers. The standard person markers are then used for object marking. Their presence is optional when another NP already functions as direct object in the sentence.

3. A previous formal analysis of Sorani Kurdish clitic placement

3.1. Placement properties of the MPM (Samvelian, 2007)

A first account of the clitic placement in Sorani Kurdish is given by Samvelian (2007). The author states that MPMs usually adopt (internal or external) second position.

Thus she specifies the following MPM placement rules:

1. If the VP is not verb-initial, the MPM is in enclitic second position in the VP.
2. If the VP is verb-initial, the MPM is an endoclitic within the verb.

These general rules show that Sorani MPMs can be either enclitics or endoclitics.

Concerning verb-internal placement, Samvelian (2007) also notes the following more specific constraints — the first one being the general rule, as illustrated in as presented in Examples (4) and (5), and the second one a more specific override thereof, cf. Examples (6) and (7).

a) If the VP is verb-initial, the MPM is an endoclitic within the V and occurs after its first morph.

(4) xward=**man**=in
eat.STEM=**P1PL**=P3PL
"We ate them."

(5) na=**man**=xward-in
NEG=**P1PL**=eat.STEM-P3PL
"We did not eat them."

b) The third singular person subject marker (be it standard or MPM) always follows the object marker.

(6) xward-in=**î**
eat.STEM-3PL=**3SG**
"He ate them."

(7) na-xward-in=**î**
NEG-eat.STEM-3PL=**3SG**
"He did not them."

The constraints presented by Samvelian (2007) have been formalised by Bonami and Samvelian (2008). The verb-external placement appears to be straightforward, once the VP internal second-position placement has been stipulated. However, the verb-internal placement still needs to be further investigated — especially since the authors do not study the complete paradigm. In this paper, we therefore focus on the sole verb-internal placement.

3.2 Formalisation of the MPM placement (Bonami and Samvelian, 2008)

Bonami and Samvelian (2008) formalise the clitic placement described by Samvelian (2007) for the present and the past non-perfect tenses.[3] For this purpose, they use concepts introduced by Crysmann (2002) for treating the verb internal insertion of European Portuguese clitic person markers. The idea underlying the analysis by Crysmann (2002) is that morphology operates on *morph lists* rather than unstructured phonology. Morph lists are an intermediate structured level between individual morphs and words. They contain elements that are either independent morphs or opaque morph clusters. The assumption underlying the idea of morph-lists is that elements of heterogeneous nature such as clitics do have access to elements of morph-lists (individual morphs or morph-clusters) but not to the structure within morph-list elements, and in particular do not access the inner structure of the clusters within the morph-list.

[3] Bonami and Samvelian (2008) also formalise the imperative forms. We do not give an explicit formalisation of the imperative forms here since, for those forms, the MPM placement is fully covered by the rules for the general case placement presented by Bonami and Samvelian (2008).

Bonami and Samvelian (2008) use the notion of morph list to account for the placement constraints on Sorani clitics in the following way.

As stated by Samvelian (2007) and illustrated in the data contained in McCarus (1958), MPM placement is not phonologically determined. One cannot, for example, deduce the placement of the MPM from the location of word-stress. Thus, contrarily to the behaviour of Pashto endoclitics, for example, as described by Anderson (2005), Sorani MPM display placement properties that appear to be strictly morphologically determined.

Bonami and Samvelian (2008) state that whenever the MPM is verb internal (i.e., the verb is VP initial), the default placement for the MPM is to occur in verb internal second position within the morph list. A simple example thereof is given in (4) and (5).

(4) xward=**man**=in
eat.STEM=**P1PL**=P3PL ▯
"We ate them."

(5) na=**man**=xward-in
NEG=**P1PL**= eat.STEM-P3PL ▯
"We did not eat them."

In this case the morph list can be formalised as simply being a list of monomorphic elements. The endoclitic (typeset in boldscript in the examples), being itself a morph list consisting of a single element, inserts between the first and second element of the verb's morph list.

$$\begin{bmatrix} \text{MORPHS} \left\langle \textit{xward} \right\rangle \oplus \left\langle \boldsymbol{man} \right\rangle \oplus \left\langle \textit{in} \right\rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} - \textit{perfect} \\ \text{TENSE } \textit{preterite} \\ \text{POLARITY } \textit{positive} \end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} \text{MORPHS} \left\langle \textit{na} \right\rangle \oplus \left\langle \boldsymbol{man} \right\rangle \oplus \left\langle \textit{xward} \right\rangle \oplus \left\langle \textit{in} \right\rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} - \textit{perfect} \\ \text{TENSE } \textit{preterite} \\ \text{POLARITY } \textit{negative} \end{bmatrix} \end{bmatrix}$$

Yet, whenever there is a third singular person marker, it always follows the object marker. This is modelled by a morphological compact operation creating opaque morph-sequences whose inner structure is no longer accessible to the endoclitic MPM. When confronted with compact units (or *morphological stems),* MPMs must attach on the outside. They do so in second position, as specified in the general rule *a).* The relevant operation is given below.[4]

[4] The formal representation of the compact operation given in this paper slightly differs from the representation presented by Bonami and Samvelian (2008). In (Bonami and Samvelian, 2008), compact operates on morph-lists. In this paper we propose a simplified representation that allows for representing the operation and the attached constraints in one feature structure (cf. [SUBJECT [[PERSON *3*] ∧ [NUMBER *sing*]]]).

(6) xward-in=**î**
eat.STEM-3PL=**3SG**
"He ate them."

(7) na-xward-in=**î**
NEG-eat.STEM-3PL=**3SG**
"He did not them."

$$\texttt{compact}\left(\begin{bmatrix} \text{MORPHS} \left\langle \textit{xward, in} \right\rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} - \textit{perfect} \\ \text{TENSE } \textit{preterite} \\ \text{POLARITY } \textit{positive} \end{bmatrix} \\ \text{SUBJECT} \begin{bmatrix} \text{PERSON } \mathit{3} \\ \text{NUMBER } \textit{sing} \end{bmatrix} \end{bmatrix}\right) \longrightarrow \begin{bmatrix} \text{MORPHS} \left\langle \textit{xwardin} \right\rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} - \textit{perfect} \\ \text{TENSE } \textit{preterite} \\ \text{POLARITY } \textit{positive} \end{bmatrix} \\ \text{SUBJECT} \begin{bmatrix} \text{PERSON } \mathit{3} \\ \text{NUMBER } \textit{sing} \end{bmatrix} \end{bmatrix}$$

The complete verb form thus corresponds to the following representation.

$$\begin{bmatrix} \text{MORPHS} \left\langle \textit{xwardin} \right\rangle \oplus \left\langle \boldsymbol{\hat{\imath}} \right\rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} - \textit{perfect} \\ \text{TENSE } \textit{preterite} \\ \text{POLARITY } \textit{positive} \end{bmatrix} \\ \text{SUBJECT} \begin{bmatrix} \text{PERSON } \mathit{3} \\ \text{NUMBER } \textit{sing} \end{bmatrix} \end{bmatrix}$$

If the verb form also contains prefixes, as for example the negative verb forms hereafter, the compact operation not only clusters the stem to the person marker but also includes the prefix(es), as illustrated below.

$$\texttt{compact}\left(\begin{bmatrix} \text{MORPHS} \left\langle \textit{na, xward, in} \right\rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} - \textit{perfect} \\ \text{TENSE } \textit{preterite} \\ \text{POLARITY } \textit{negative} \end{bmatrix} \\ \text{SUBJECT} \begin{bmatrix} \text{PERSON } \mathit{3} \\ \text{NUMBER } \textit{sing} \end{bmatrix} \end{bmatrix}\right) \longrightarrow \begin{bmatrix} \text{MORPHS} \left\langle \textit{naxwardin} \right\rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} - \textit{perfect} \\ \text{TENSE } \textit{preterite} \\ \text{POLARITY } \textit{negative} \end{bmatrix} \\ \text{SUBJECT} \begin{bmatrix} \text{PERSON } \mathit{3} \\ \text{NUMBER } \textit{sing} \end{bmatrix} \end{bmatrix}$$

Hence the following representation for the resulting verb form.

$$\begin{bmatrix} \text{MORPHS} \left\langle \textit{naxwardin} \right\rangle \oplus \left\langle \boldsymbol{\hat{\imath}} \right\rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} - \textit{perfect} \\ \text{TENSE } \textit{preterite} \\ \text{POLARITY } \textit{negative} \end{bmatrix} \\ \text{SUBJECT} \begin{bmatrix} \text{PERSON } \textit{3} \\ \text{NUMBER } \textit{sing} \end{bmatrix} \end{bmatrix}$$

3.3. Extending the existing analysis

In addition to the two constraints formalised by Bonami and Samvelian (2008), McCarus (1958) cites the following constraint, which also overrides the general rule *a).*

c) Be it a subject or object marker, the first singular person marker always precedes any plural marker.

(8) xward=m=**in**
eat.STEM=P1PL=**P3PL** ⍰
"I ate them."

(9) xward-im=**tan**
eat.STEM-P1SG=**P2PL** ⍰
"You ate me."

(10) na=**m**=xward-in
NEG=**P1SG**= eat.STEM=P3PL ⍰
"I didn't eat them."

(11) na-xward-im=**tan**
NEG-eat.STEM=P1SG=**P3PL** ⍰
"You didn't eat me."

The second rule is not taken into account in (Bonami and Samvelian, 2008), but doing so only requires two minor modifications of their model and does not invalidate their analysis *per se.*

1) When the present tense verb has a first person singular object and a plural subject, or when a past tense verb has a first person singular subject and a plural object,[5] the placement of the clitic occurs according to the general case placement, i.e., in second position of a morph-list containing solely monomorphic elements. Hereafter we provide examples for a positive, respectively negative, preterite verb form.

[5] Obviously, since the object is plural and thus not a third person singular, rule *b)* does not apply either.

Positive verb form:

$$\begin{bmatrix} \text{MORPHS} \left\langle \textit{xward} \right\rangle \oplus \left\langle \boldsymbol{m} \right\rangle \oplus \left\langle \textit{in} \right\rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} - \textit{perfect} \\ \text{TENSE } \textit{preterite} \\ \text{POLARITY } \textit{positive} \end{bmatrix} \\ \text{SUBJECT} \begin{bmatrix} \text{PERSON } \textit{1} \\ \text{NUMBER } \textit{sing} \end{bmatrix} \\ \text{OBJECT} \begin{bmatrix} \text{NUMBER } \textit{pl} \end{bmatrix} \end{bmatrix}$$

Negative verb form:

$$\begin{bmatrix} \text{MORPHS} \left\langle \textit{na} \right\rangle \oplus \left\langle \boldsymbol{m} \right\rangle \oplus \left\langle \textit{xward} \right\rangle \oplus \left\langle \textit{in} \right\rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} - \textit{perfect} \\ \text{TENSE } \textit{preterite} \\ \text{POLARITY } \textit{negative} \end{bmatrix} \\ \text{SUBJECT} \begin{bmatrix} \text{PERSON } \textit{1} \\ \text{NUMBER } \textit{sing} \end{bmatrix} \\ \text{OBJECT} \begin{bmatrix} \text{NUMBER } \textit{pl} \end{bmatrix} \end{bmatrix}$$

2) Whenever there is a first person singular subject and a plural object in the present or a first person singular object and a plural subject in the past, another compact operation, similar to the one for rule *b),* applies. Stipulating this additional compact operation suffices to account for the cluster consisting of the stem and the first person singular object affix.

$$\texttt{compact}\left(\begin{bmatrix} \text{MORPHS} \left\langle \textit{xward, im} \right\rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} - \textit{perfect} \\ \text{TENSE } \textit{preterite} \\ \text{POLARITY } \textit{positive} \end{bmatrix} \\ \text{SUBJECT} \begin{bmatrix} \text{NUMBER } \textit{pl} \end{bmatrix} \\ \text{OBJECT} \begin{bmatrix} \text{PERSON } \textit{1} \\ \text{NUMBER } \textit{sing} \end{bmatrix} \end{bmatrix}\right) \longrightarrow \begin{bmatrix} \text{MORPHS} \left\langle \textit{xwardim} \right\rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} - \textit{perfect} \\ \text{TENSE } \textit{preterite} \\ \text{POLARITY } \textit{positive} \end{bmatrix} \\ \text{SUBJECT} \begin{bmatrix} \text{NUMBER } \textit{pl} \end{bmatrix} \\ \text{OBJECT} \begin{bmatrix} \text{PERSON } \textit{1} \\ \text{NUMBER } \textit{sing} \end{bmatrix} \end{bmatrix}$$

For the verb forms with prefixes, such as a negative verb form, the compact operation generates clusters containing the prefix(es), the stem, and the first person singular affix, just as for modelling Example (7).

$$\texttt{compact}\left(\begin{bmatrix} \text{MORPHS} \left\langle \textit{na, xward ,im} \right\rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} - \textit{perfect} \\ \text{TENSE } \textit{preterite} \\ \text{POLARITY } \textit{negative} \end{bmatrix} \\ \text{SUBJECT} \begin{bmatrix} \text{NUMBER } \textit{pl} \end{bmatrix} \\ \text{OBJECT} \begin{bmatrix} \text{PERSON } \textit{1} \\ \text{NUMBER } \textit{sing} \end{bmatrix} \end{bmatrix}\right) \longrightarrow \begin{bmatrix} \text{MORPHS} \left\langle \textit{naxwardim} \right\rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} - \textit{perfect} \\ \text{TENSE } \textit{preterite} \\ \text{POLARITY } \textit{negative} \end{bmatrix} \\ \text{SUBJECT} \begin{bmatrix} \text{NUMBER } \textit{pl} \end{bmatrix} \\ \text{OBJECT} \begin{bmatrix} \text{PERSON } \textit{1} \\ \text{NUMBER } \textit{sing} \end{bmatrix} \end{bmatrix}$$

Thus the complete verb forms from Examples (9) and (11) can be represented as follows.

$$\text{Positive verb form: } \begin{bmatrix} \text{MORPHS} \left\langle \textit{xwardim} \right\rangle \oplus \left\langle \textbf{\textit{tan}} \right\rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} - \textit{perfect} \\ \text{TENSE } \textit{preterite} \\ \text{POLARITY } \textit{positive} \end{bmatrix} \\ \text{SUBJECT} \begin{bmatrix} \text{PERSON } \textit{2} \\ \text{NUMBER } \textit{pl} \end{bmatrix} \\ \text{OBJECT} \begin{bmatrix} \text{PERSON } \textit{1} \\ \text{NUMBER } \textit{sing} \end{bmatrix} \end{bmatrix}$$

$$\text{Negative verb form: } \begin{bmatrix} \text{MORPHS} \left\langle \textit{naxwardim} \right\rangle \oplus \left\langle \textbf{\textit{tan}} \right\rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} - \textit{perfect} \\ \text{TENSE } \textit{preterite} \\ \text{POLARITY } \textit{negative} \end{bmatrix} \\ \text{SUBJECT} \begin{bmatrix} \text{PERSON } \textit{2} \\ \text{NUMBER } \textit{pl} \end{bmatrix} \\ \text{OBJECT} \begin{bmatrix} \text{PERSON } \textit{1} \\ \text{NUMBER } \textit{sing} \end{bmatrix} \end{bmatrix}$$

4. Accounting for the perfect tenses

4.1 The yet unaccounted data

Bonami and Samvelian (2008) state that MPMs always adopt a second position, be it internal (i.e., in the morph-list) or external (i.e., in the VP). They explicitly choose to analyse only verb forms in the present and non-perfect past tenses. After extending their analysis with constraint *c)*, the resulting model holds for all transitive verb forms they analyse. Yet, if one takes into account the whole paradigm, including the perfect tenses, the situation gets much trickier. As shown in Examples (12) to (15), in perfect tenses, MPMs seem to appear in third (or fourth) position, after the perfect maker *–û* or the past perfect marker *–bû* respectively.

(12) xward-û=**man**=in
eat.STEM-û=**P1PL**=P3PL
"We have eaten them."

(13) xward-û-in=**î**
eat.STEM-û-P3PL=**P3SG**
"He has eaten them."

(14) xward-bû=**man**=in
eat.STEM-bû=P1PL=**P3PL**
"We had eaten them."

(15) xward-bû-in=**î**
eat.STEM-bû-P3PL=**P3SG**
"He had eaten them."

4.2. A periphrastic analysis of the perfect tenses

This contradiction seems to disappear if the perfect tenses are analysed as periphrastic. Given the apparent, and at least historically confirmed, presence of forms of the auxiliary *bûn* 'to be' in the perfect tenses, this looks plausible. Indeed, the personal endings used for the present perfect (PE3) are equivalent to the enclitic forms of the verb *bûn* 'to be'. Thus the present perfect forms can be re-analysed as the past participle (i.e., STEM2 + *–û)* plus the inflected forms of the auxiliary *bûn*, which match the personal endings used for the present perfect (PE3). In the past perfect, we can make out the verb *bûn* in the past tense: the past perfect appears like a combination of a STEM and the past forms of *bûn* (*bû*– + PE2). Thus the present and past perfect forms, whose features only differ in tense, indeed show auxiliary forms that also only differ in tense, since for building the past perfect the auxiliary *bûn* is simply used in the past tense. 311

In such a periphrastic analysis, we would simply have to say that MPMs attach in auxiliary- second-position for the past perfect and as external second position clitics in the (present) perfect.

(16) [V xward] [V bû=**man**=in]
[V eat.STEM] [V AUX.STEM =**MPM.P1PL**=P3PL]
"We had eaten them." (PAST PERFECT)

(17) [V xward-û]=**man**=[AUX in]
[V eat.STEM-*û*]=**MPM.P1PL**=[AUX AUX.P3PL]
"*We have eaten them."* (PRESENT PERFECT)

4.3. Drawbacks of a periphrastic analysis of Sorani perfect tenses

However, this analysis runs into difficulties when considering the following facts:[6]

[6] While the indicative present perfect tenses seem to constitute a convincing case for canonical periphrasis (due to the presence of synthetic subjunctive forms), nothing in the verbal paradigm raises expectations as to the existence of past perfect forms. Their periphrasticity would therefore be at best non-canonical in the sense of Brown *et al.* (2012). Yet, one would be tempted to state that if periphrastic present perfect forms exist and if there are forms for the past perfect, then the latter

1. In present perfect constructions, MPMs indeed attach to the first element of the VP containing the two elements of the periphrastic construction, i.e., it attaches to the past participle. In the past perfect however, they attach within the auxiliary, in second position (see Example (12) *vs.* Example (14)). Thus, the two constructions seem to be different in nature. Instead of simplifying the analysis, assuming a periphrastic construction for the perfect tense forms would require stipulating separate rules for the two perfect tenses, the present perfect following the enclitic placement rule of VP second position enclisis, while the past perfect would follow the verb internal endoclitic placement rules.
2. The periphrastic analysis does correctly account for most perfect tense forms. However, it does not explain the reversed order of the MPMs and PE3, respectively PE2, for the specific cases due to person placement (see Examples (18) and (19)). Another compact operation would be needed to prevent the MPM to attach to the participle in (18) or to the auxiliary stem in (19) directly. But if we consider the forms as periphrastic, compact would not simply operate over morphs but over words or word-like elements, such as the participle in (18). Yet, this is problematic in an approach where we assume autonomous morphology (Aronoff, 1994).

(18) [V xward-û]=[AUX n]=**î**
[V eat.STEM-*û*]=[AUX AUX.P3PL]=**MPM.P3SG**
"He has eaten them." (PRESENT PERFECT)

(19) [V xward] [AUX bû=n]=**î**
[V eat.STEM] [AUX AUX.STEM =AUX.P3PL=**MPM.P3SG**]
"He had eaten them". (PAST PERFECT)

3. A lexicalist approach to morphosyntax entails that morphology and syntax do not submit to rules of the same nature. They are necessarily distinct modules of language. It is hence difficult to even explain the status of the STEM2 within the past perfect constructions. Elements treated within syntax do not undergo the same rules as morphs, they have to be words or word-like elements such as clitics. A bare STEM, such as *xward* in the past perfect tenses, makes for an extremely unlikely syntactic element.[7]
4. Moreover, when perfect tense forms bear a negative polarity feature, MPMs are inserted in internal second position,[8] and not within the alleged auxiliary. This fact alone strongly favours a synthetic analysis.

(20) na=**man**=xward-û=a
NEG=**MPM.1PL**= eat.STEM-PART=P3SG
"We have not eaten it."

(21) na=**man**=xward-bû=in
NEG=**MPM.1PL**=eat.STEM-BÛ=P3SG
"We had not eaten them."

---

are typologically highly unlikely to be anything but periphrastic. Since the auxiliary *bûn* in the past tense seems distinguishable among the elements constituting the past perfect, positing a periphrastic past perfect seems quite reasonable.

[7] One might suggest the existence of apocoped infinitives in Sorani, like those existing in Persian that indeed consist of bare stems (STEM2) (Lazard *et al.,* 2006). Yet this analysis would again run into difficulties due to the different behaviour of the MPMs in the present and past perfect. One would still expect the MPMs to attach to either type of participle.

[8] Except, of course, for the specific placement of third person singular and first person singular markers.

5. Finally, in the past perfect, the STEM and *–bû* are never separated by any morph, even though MPMs may occur immediately before STEM2 and after *–bû.* Thus no morph ever inserts between the presumed predicative element and the auxiliary. This systematic closeness between the two elements does not fit into an account where they would be part of two syntactically independent elements.

5. Completing the formal account of Sorani Kurdish MPMs

5.1 Completing the inventory of compact operations

From the evidence of the preceding section, we can deduce that Sorani perfect tenses are better analysed as synthetic forms. Nevertheless, it appears that these synthetic forms exhibit internal structure, which explains MPM placement properties.

We follow Bonami and Samvelian (2008) in resorting to the morphological compact operation that creates morph-sequences opaque even to the otherwise endoclitic MPMs of Sorani Kurdish. These compact operations apply to STEM2-PE2 or STEM2-PE3 sequences, as shown by Bonami and Samvelian (2008), but also affect the STEM2–*û* and STEM2–*bû* sequences of the perfect tenses, thus creating internal structure.

This approach allows for a homogeneous treatment of all perfect tenses, which given the data seems much more satisfying.

$$\texttt{compact}\left(\begin{bmatrix}\text{MORPHS}\left\langle \textit{xward, û}\right\rangle \\ \text{FEATURES}\begin{bmatrix}\text{ASPECT} + \textit{perfect}\\ \text{TENSE } \textit{present}\end{bmatrix}\end{bmatrix}\right)\longrightarrow\begin{bmatrix}\text{MORPHS}\left\langle \textit{xwardû}\right\rangle \\ \text{FEATURES}\begin{bmatrix}\text{ASPECT} + \textit{perfect}\\ \text{TENSE } \textit{present}\end{bmatrix}\end{bmatrix}$$

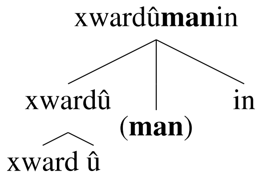
$$\texttt{compact}\left(\begin{bmatrix}\text{MORPHS}\left\langle \textit{xward, bû}\right\rangle \\ \text{FEATURES}\begin{bmatrix}\text{ASPECT} + \textit{perfect}\\ \text{TENSE } \textit{past}\end{bmatrix}\end{bmatrix}\right)\longrightarrow\begin{bmatrix}\text{MORPHS}\left\langle \textit{xwardbû}\right\rangle \\ \text{FEATURES}\begin{bmatrix}\text{ASPECT} + \textit{perfect}\\ \text{TENSE } \textit{past}\end{bmatrix}\end{bmatrix}$$

When confronted with compact units, an MPM cannot make out any internal structure and must thus attach on the outside. The formal representation of the complete verb forms corresponding to above quoted examples (12) to (15) are given below with their corresponding tree structure representations.[9] These trees show the extent of morphological structure created by the compact operations for each type of verb form. The structure available as frontiers for endoclitic insertion corresponds to the first nodes under the tree roots.

(22) xward-û=**man**=in
eat-PART=**P1PL**=P3PL
"We have eaten them." (PRESENT PERFECT)

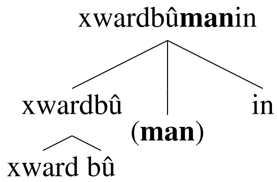
[9] The MPMs are typeset in bold print, as well as noted between brackets so as to express the fact that they insert in the final morph list as endoclitics.

$$\begin{bmatrix} \text{MORPHS } \langle \textit{xwardû} \rangle \oplus \langle \textbf{\textit{man}} \rangle \oplus \langle \textit{in} \rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} + \textit{perfect} \\ \text{TENSE } \textit{present} \end{bmatrix} \end{bmatrix}$$



(23) xward-bû=**man**=in
eat-PASTPERF=**P1PL**=P3PL
"We had eaten them." (PAST PERFECT)

$$\begin{bmatrix} \text{MORPHS } \langle \textit{xwardbû} \rangle \oplus \langle \textbf{\textit{man}} \rangle \oplus \langle \textit{in} \rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} + \textit{perfect} \\ \text{TENSE } \textit{past} \end{bmatrix} \end{bmatrix}$$

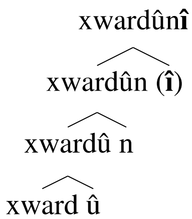


Whenever one of the specific rules applies, the corresponding compact operations described above are activated on top of the perfect compact operations.

(24) xward-û-n=**î**
eat.STEM-PART-P3PL=**P3SG**
"He has eaten them."

$$\texttt{compact}\left(\begin{bmatrix} \text{MORPHS} \left\langle \textit{xwardû, (i)n} \right\rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} + \textit{perfect} \\ \text{TENSE } \textit{present} \\ \text{POLARITY } \textit{positive} \end{bmatrix} \\ \text{SUBJECT} \begin{bmatrix} \text{PERSON } \textit{3} \\ \text{NUMBER } \textit{sing} \end{bmatrix} \end{bmatrix}\right) \longrightarrow \begin{bmatrix} \text{MORPHS} \left\langle \textit{xwardûn} \right\rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} + \textit{perfect} \\ \text{TENSE } \textit{present} \\ \text{POLARITY } \textit{positive} \end{bmatrix} \\ \text{SUBJECT} \begin{bmatrix} \text{PERSON } \textit{3} \\ \text{NUMBER } \textit{sing} \end{bmatrix} \end{bmatrix}$$

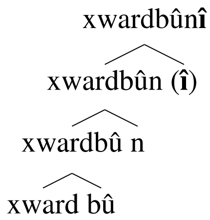
$$\begin{bmatrix} \text{MORPHS } \langle \textit{xwardûn} \rangle \oplus \langle \textbf{\textit{î}} \rangle \\ \text{FEATURES} \begin{bmatrix} \text{ASPECT} + \textit{perfect} \\ \text{TENSE } \textit{present} \\ \text{POLARITY } \textit{positive} \end{bmatrix} \\ \text{SUBJECT} \begin{bmatrix} \text{PERSON } \textit{3} \\ \text{NUMBER } \textit{sing} \end{bmatrix} \end{bmatrix}$$



(25) xward-bû-n=**î**
eat.STEM-BÛ- P3PL=**P3SG**
"He had eaten them."

$$\texttt{compact}\left(\left[\begin{array}{ll}\text{MORPHS} & \left\langle \textit{xwardbû, (i)n} \right\rangle \\ \text{FEATURES} & \left[\begin{array}{l}\text{ASPECT} + \textit{perfect} \\ \text{TENSE } \textit{past} \\ \text{POLARITY } \textit{positive}\end{array}\right] \\ \text{SUBJECT} & \left[\begin{array}{l}\text{PERSON } \textit{3} \\ \text{NUMBER } \textit{sing}\end{array}\right]\end{array}\right]\right) \longrightarrow \left[\begin{array}{ll}\text{MORPHS} & \left\langle \textit{xwardbûn} \right\rangle \\ \text{FEATURES} & \left[\begin{array}{l}\text{ASPECT} + \textit{perfect} \\ \text{TENSE } \textit{past} \\ \text{POLARITY } \textit{positive}\end{array}\right] \\ \text{SUBJECT} & \left[\begin{array}{l}\text{PERSON } \textit{3} \\ \text{NUMBER } \textit{sing}\end{array}\right]\end{array}\right]$$

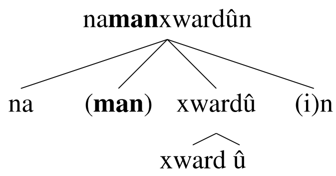
$$\left[\begin{array}{ll}\text{MORPHS} & \langle \textit{xwardbûn} \rangle \oplus \boldsymbol{\langle \hat{\imath} \rangle} \\ \text{FEATURES} & \left[\begin{array}{l}\text{ASPECT} + \textit{perfect} \\ \text{TENSE } \textit{past} \\ \text{POLARITY } \textit{positive}\end{array}\right] \\ \text{SUBJECT} & \left[\begin{array}{l}\text{PERSON } \textit{3} \\ \text{NUMBER } \textit{sing}\end{array}\right]\end{array}\right]$$



Contrarily to the clustering occurring between the prefixes, the stem and the PE2, respectively PE3 person markers, the perfect tenses do not cluster by default with the prefixes. Thus formalisation of perfect tense negative verb forms, where the MPM occurs between the negative prefix and the stem, for example, runs as follows.

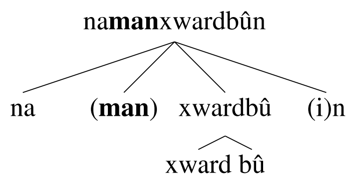
(26) na=**man**=xward-û=n
NEG= **P1PL**= eat.STEM-PART= P3PL
"We have not eaten them."

$$\left[\begin{array}{ll}\text{MORPHS} & \langle \textit{na} \rangle \oplus \boldsymbol{\langle \textit{man} \rangle} \oplus \langle \textit{xwardû} \rangle \oplus \langle \textit{(i)n} \rangle \\ \text{FEATURES} & \left[\begin{array}{l}\text{ASPECT} + \textit{perfect} \\ \text{TENSE } \textit{present} \\ \text{POLARITY } \textit{negative}\end{array}\right]\end{array}\right]$$



(27) na=**man**=xward-bû=n
NEG= **P1PL**= eat.STEM-BÛ= **P3PL**
"We had not eaten them."

$$\left[\begin{array}{ll}\text{MORPHS} & \langle \textit{na} \rangle \oplus \boldsymbol{\langle \textit{man} \rangle} \oplus \langle \textit{xwardbû} \rangle \oplus \langle \textit{(i)n} \rangle \\ \text{FEATURES} & \left[\begin{array}{l}\text{ASPECT} + \textit{perfect} \\ \text{TENSE } \textit{past} \\ \text{POLARITY } \textit{negative}\end{array}\right]\end{array}\right]$$



Perfect clusters containing object affixes, however, do cluster with the prefixes, just as their non-perfect counterparts.[10] This clustering is simply triggered by the operations occurring when constraints for *b)* or *c)* are met (cf. Section 4).

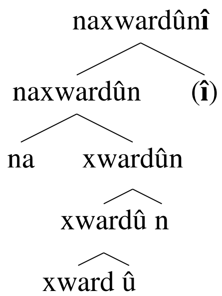
$$\texttt{compact}\left(\begin{bmatrix}\text{MORPHS}\ \left\langle \textit{na, xwardû, (i)n}\right\rangle \\ \text{FEATURES}\ \begin{bmatrix}\text{ASPECT} + \textit{perfect}\\ \text{TENSE}\ \textit{present}\\ \text{POLARITY}\ \textit{negative}\end{bmatrix}\\ \text{SUBJECT}\ \begin{bmatrix}\text{PERSON}\ \mathit{3}\\ \text{NUMBER}\ \textit{sing}\end{bmatrix}\end{bmatrix}\right) \longrightarrow \begin{bmatrix}\text{MORPHS}\ \left\langle \textit{naxwardûn}\right\rangle \\ \text{FEATURES}\ \begin{bmatrix}\text{ASPECT} + \textit{perfect}\\ \text{TENSE}\ \textit{present}\\ \text{POLARITY}\ \textit{negative}\end{bmatrix}\\ \text{SUBJECT}\ \begin{bmatrix}\text{PERSON}\ \mathit{3}\\ \text{NUMBER}\ \textit{sing}\end{bmatrix}\end{bmatrix}$$

$$\texttt{compact}\left(\begin{bmatrix}\text{MORPHS}\ \left\langle \textit{na, xwardbû, (i)n}\right\rangle \\ \text{FEATURES}\ \begin{bmatrix}\text{ASPECT} + \textit{perfect}\\ \text{TENSE}\ \textit{past}\\ \text{POLARITY}\ \textit{negative}\end{bmatrix}\\ \text{SUBJECT}\ \begin{bmatrix}\text{PERSON}\ \mathit{3}\\ \text{NUMBER}\ \textit{sing}\end{bmatrix}\end{bmatrix}\right) \longrightarrow \begin{bmatrix}\text{MORPHS}\ \left\langle \textit{naxwardbûn}\right\rangle \\ \text{FEATURES}\ \begin{bmatrix}\text{ASPECT} + \textit{perfect}\\ \text{TENSE}\ \textit{past}\\ \text{POLARITY}\ \textit{negative}\end{bmatrix}\\ \text{SUBJECT}\ \begin{bmatrix}\text{PERSON}\ \mathit{3}\\ \text{NUMBER}\ \textit{sing}\end{bmatrix}\end{bmatrix}$$

Thus the formal representation for negative perfect forms with morphological stems containing object affixes is as follows. 316

(28) na-xward-û-n=**î**
NEG- eat.STEM-PART- P3PL=**P3SG**
"He has not eaten them."

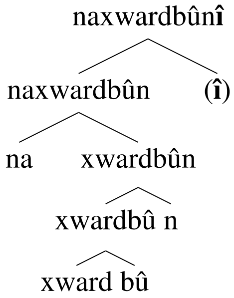
$$\begin{bmatrix}\text{MORPHS}\ \langle \textit{naxwardûn}\rangle \oplus \langle \textbf{î}\rangle \\ \text{FEATURES}\ \begin{bmatrix}\text{ASPECT} + \textit{perfect}\\ \text{TENSE}\ \textit{present}\\ \text{POLARITY}\ \textit{negative}\end{bmatrix}\\ \text{SUBJECT}\ \begin{bmatrix}\text{PERSON}\ \mathit{3}\\ \text{NUMBER}\ \textit{sing}\end{bmatrix}\end{bmatrix}$$



(29) na-xward-bû-n=**î**
NEG- eat.STEM-PART-P3PL=**P3SG**
"He had not eaten them."

[10] The examples given here are for the third person singular subject rule *b).* The same kind of clustering would of course also occur with the first person singular subject or object *vs.* plural person marker, rule *c).*

MORPHS <*naxwardbûn*> ⊕ **<*î*>**
FEATURES [ASPECT + *perfect*, TENSE *past*, POLARITY *negative*]
SUBJECT [PERSON *3*, NUMBER *sing*]



5.2. Summary and respective ordering of the compact operations

The observations and analyses presented above indicate that Sorani clitic MPMs present the following placement properties.

1. If VP is not verb initial, the MPM is an enclitic attaching in second position within the sentence (Samvelian, 2007).
2. If a VP is verb initial, the MPM attaches in verb-internal second position, according to the following constraints.
   a) By default, the MPM attaches after the first element of the morph list, i.e. either the first prefix, or, if there are no prefixes, after the stem.
   b) For the perfect tenses, the stem and the perfect marker (be it present perfect –*û* or past perfect –*bû*) are clustered into one element in the morph-list. They are called morphological stems by Bonami and Samvelian (2008). This constraint is formally encoded by the compact operation A below.
   c) The third singular person marker (be it standard or MPM) always follows the object marker. This constraint is formally encoded by the compact operation B below.
   d) Be it a subject or object marker, the first singular person marker always precedes any plural marker. This constraint is also formally handled by the compact operation B below.
   e) Whenever a verb form contains prefixes and undergoes the compact operation described in *c)* or *d)* above, the prefixes are also clustered to the stem and the person marker during compact operation B.

The two* compact *operations

A. STEM + *û* → < STEM*û* >
STEM + *bû* → < STEM*bû* >

After applying compact A, <STEM*û* >, respectively <STEM*û* >, are considered morphological stems. In particular they count as STEM in compact operation B below.

B. STEM + PE → <STEM-PE >
da + STEM + PE → <*da*STEM-PE >
bi + STEM + PE → <*bi*STEM-PE >
nâ + STEM + PE → <*nâ*STEM-PE >
na + STEM + PE → <*na*STEM-PE >
na +*da* + STEM + PE → <*nada*STEM-PE >
bi + STEM + PE → <*bi*STEM-PE >

Ordering of the placement constraints

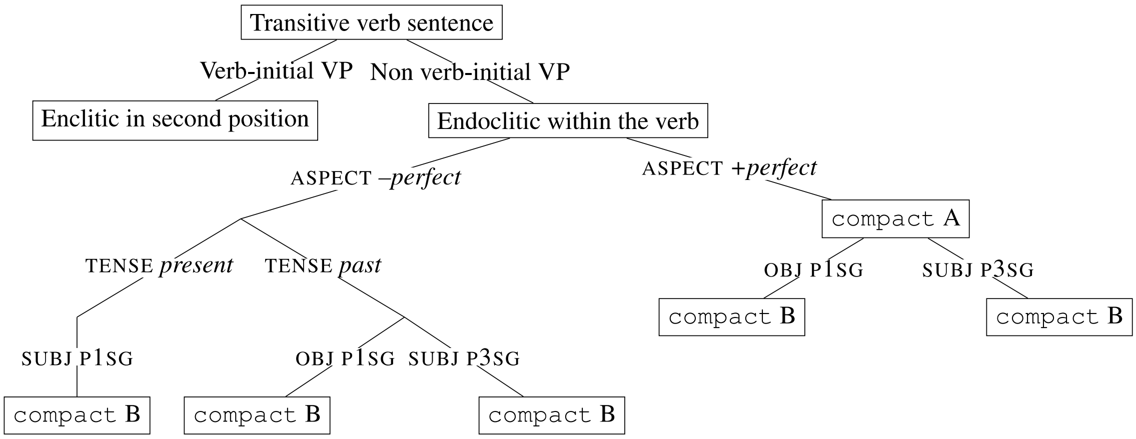
Figure 1 below gives a summary representation the possible combinations of the four compact operations depending on the criteria described above.

Given a sentence containing a transitive verb, the VP can be either verb initial or not verb initial.

If the VP is not verb-initial, the clitic placement follows the pattern in the left first level node: it attaches as an enclitic in second position within the VP.

If the VP is verb initial, the clitic is an endoclitic within the verb and follows the endoclitic placement constraints on the right part of the tree.

If any of the contraints indicated on the tree branches are met, the verb will undergo one of the corresponding nodes' compact operations.



For example, within a verb-initial VP, a past perfect verb having positive polarity, a third person singular subject and a third person plural object, such as *xwardbûinî* 'he will eat them', will follow the following path:

- The initial unclustered morphs are the following: *xward, –bû–, –in–, =î.*
- The verb *xwardbûinî* is part of verb-initial VP. Thus the first path to the right applies.
- It has perfect aspect, thus compact A applies.
  The first clustering outputs *xwardbû, –in–, =î.*
- It has a third person singular subject, thus compact B applies.
  The second clustering outputs *xwardbûin.* To this output the endoclitic *=î* attaches as *xwardbûin=î.*

Note that a past perfect verb having positive polarity does not contain any prefixes. Thus compact B does only affect the stem and the third person singular object affix.

Within a verb-initial VP, a present verb with negative polarity and a first person singular subject and third person plural object, such as *nâxwardimyân* 'he will eat them', will follow the following path:

- The initial unclustered morphs are the following: *nâ–, xward, –im–, =yân.*
- The verb *nâxwardimyân* is part of verb-initial VP. Thus the first path to the right applies.
- It has no perfect aspect, thus compact A does not apply. We follow the second path to the right.
- It has a first person singular subject, thus compact B applies. Compact B here clusters the negative prefix *nâ–*, the stem and the first person singular subject affix
  –im–. The endoclitic *=yân* attaches in the thus new available second position.

The clustered output thus contains *nâxwardim.* Thus the endoclitic *=î* attaches to it as *nâxwardim=yân.*

6. Concluding remarks

We have presented an enhanced formal analysis of Sorani Kurdish endoclitic placement. This analysis extends the analysis and formalisation given by Bonami and Samvelian (2008) by taking into account an additional constraint that had been missed and by including present and past perfect verb forms that had not been studied so far.

In our formalisation, the compact operation used by Bonami and Samvelian (2008) is not limited to clusters containing the negative prefix and/or third person singular subject markers. It also systematically applies to first person singular person marking and to *û/bû* elements, i.e., to clusters containing the perfect and past perfect markers.

Our analysis allows for a non-periphrastic and homogeneous treatment of all Sorani perfect tenses that correctly accounts for the intricate MPM placement phenomena left unexplained until now.

We have shown that a structured representation is needed to account for the clitic placement of MPMs and the morphology-syntax interface in Sorani. Thus, the properties of Sorani MPMs formalised in this paper also argue in favour of partial internal morphological structure.

References

Anderson, Stephen R. 1992. *A-morphous Morphology.* Cambridge, UK: Cambridge University Press.

Anderson, Stephen R. 2005. *Aspects of the Theory of Clitics.* Oxford Studies in Theoretical Linguistics. Oxford University Press, Oxford.

Aronoff, Mark. 1994. *Morphology by Itself.* MIT Press.

Baerman, Matthew. 2007. Morphological reversals. *Journal of Linguistics,* 43, 33–61.

Baker, Mark C. 1988. *Incorporation.* Chicago, USA: University of Chicago Press.

Beard, Robert. 1995. *Lexeme-Morpheme Base Morphology.* Albany, N.Y.: State University of New York Press.

Bedir Khan, Emir Djeladet and Roger Lescot. 1970. *Grammaire kurde (Dialecte Kurmandji).* Paris, France: Librairie d'Amérique et d'Orient.

Beidar, Paul. 1926. *Grammaire kurde.* Paris, France: Librairie Orientaliste P. Garthner.

Blau, Joyce. 1980. *Manuel de kurde, dialecte sorani: grammaire, textes de lecture, vocabulaire kurde-français et français-kurde.* Paris, France: C. Klincksieck.

Blau, Joyce. 2000. *Méthode de kurde sorani.* Paris, France: L'Harmattan.

Bonami, Olivier and Pollet Samvelian. 2008. Sorani Kurdish person markers and the typology of agreement. In *13th International Morphology Meeting,* Vienna, Austria.

Bresnan, Joan. 1995. The lexical integrity principle: Evidence from Bantu. *Natural Language and Language Theory.*

Brown, Dunstan, Marina Chumakina, Greville G. Corbett, Gergana Popova, and Andrew Spencer. 2012. Defining 'periphrasis': key notions. To be published.

Brown, Dunstan and Andrew Hippisley. 2012. *Network Morphology: A Defaults-based Theory of Word Structure.* Cambridge University Press.

Corbett, Greville G. and Norman Fraser. 1993. Network Morphology: a DATR account of Russian nominal inflection. *Journal of Linguistics,* 29, 113–142.

Crysmann, Berthold. 2002. *Constraint-Based Coanalysis: Portuguese Cliticisation and Morphology-Syntax Interaction in HPSG.* Number 15 in Saarbrücken Dissertations in Computational Linguistics and Language Technology. Saarbrücken: Deutsches Forschungszentrum für Künstliche Intelligenz and Universität des Saarlandes.

Crysmann, Berthold. 2006. Floating affixes in Polish. In S. Müller, Ed., *Proceedings of the 13th International Conference on Head-Driven Phrase Structure Grammar,* p. 123–139, Stanford: CSLI Publications.

Di Sciullo, Anna-Maria and Edwin Williams. 1987. *On the Definition of Word.* Cambridge, MA, USA: MIT Press.

Garzoni, Maurizio. 1787. *Grammatica e Vocabulario della Lingua Kurda.* Rome, Italy: Sacra Congregazione di Propaganda Fide.

Haig, Geoffrey. 2010. *Alignment change in Iranian languages: a Construction Grammar approach.* De Gruyter.

Harris, Alice. 2002. *Endoclitics and the Origins of Udi Morphosyntax.* Oxford, UK: Oxford University Press.

Justi, Ferdinand. 1880. *Kurdische Grammatik.* St. Pertersburg, Russian Empire: Commissionäre der Kaiserlichen Akademie der Wissenschaften, Eggers & C.U.J. Glasunow.

Kupść, Anna and Jesse Tseng. 2005. A new HPSG approach to Polish auxiliary constructions. In S. Müller, ed., *The Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar, Department of Informatics, University of Lisbon, August 23-24, 2005,* pages 253–273. Stanford: CSLI Publications.

Lazard, Gilbert, Yann Richard, Rokhsareh Hechmati, and Pollet Samvelian. 2006. *Grammaire du persan contemporain.* Teheran, Iran: Institut Français de Recherche en Iran & Farhang Moaser Edition.

Lieber, Rochelle. 1992. *Deconstructing Morphology: Word Formation in Syntactic Theory.* Chicago, USA: University of Chicago Press.

Lieber, Rochelle. 2006. Syntax of words. In *The Encyclopedia of Language and Linguistics,* p. 405–408. Elsevier.

MacKenzie, David N. 1961. *Kurdish dialect studies.* London Oriental Series. London, UK: Oxford Univ. Press.

McCarus, Ernest N. 1958. *A Kurdish Grammar: descriptive analysis of the Kurdish of Sulaimaniya, Iraq.* PhD thesis, American Council of Learned Societies, New-York, USA.

Samvelian, Pollet. 2007. What Sorani Kurdish absolute prepositions tell us about cliticization. In F. Hoyt, N. Seifert and A. Teodorescu, Eds., *Texas Linguistics Society IX: The Morphosyntaxe of Understudied Languages,* Stanford, USA: CSLI online publications.

Soane, E.B. 1913. *Kurdish Grammar (Grammar of the Kurmandj or Kurdish Dialect).* Luzac & co.

Stump, G. T. 2001. *Inflectional Morphology. A Theory of Paradigm Structure.* Cambridge, UK: Cambridge University Press.

Thackston, W.M. 2006. Sorani Kurdish: A reference grammar with selected readings. Published online.

Wahby, Towfiq. 1956. *The Kurdish Grammar,* volume 1. Baghdad, Iraq: Librairie d'Amérique et d'Orient.

Walther, Géraldine. 2011. A derivational account for Sorani Kurdish passives. Communication at the 4th International Conference on Iranian Linguistics (ICIL4). June 17th-19th 2011. Uppsala, Sweden.

Zwicky, Arnold M. 1977. On Clitics. Indiana University Linguistics Club, 1977.

Zwicky, Arnold M. 1985. How to describe inflection. In M. Niepokuj, M. V. Clay, V. Nikiforidou and D. Feder, Eds., *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistic Society,* p. 372–386, Berkeley,USA: Berkeley Linguistic Society.