Argument realization and argument referencing in Soranî Kurdish

Ali Salehi D

University at Buffalo

Jean-Pierre Koenig

University at Buffalo

Proceedings of the 30th International Conference on Head-Driven Phrase Structure Grammar

> University of Massachusetts Amherst Stefan Müller, Elodie Winckel (Editors) 2023

Frankfurt/Main: University Library

pages 101-121

Keywords: Soranî Kurdish, Argument Indexing, Edge Inflection, Morphology

Salehi, Ali & Jean-Pierre Koenig. 2023. Argument realization and argument referencing in Soranî Kurdish. In Stefan Müller & Elodie Winckel (eds.), *Proceedings* of the 30th International Conference on Head-Driven Phrase Structure Grammar, University of Massachusetts Amherst, 101–121. Frankfurt/Main: University Library. DOI: 10.21248/hpsg.2023.6.

Abstract

Soranî Kurdish can reference up to two arguments morphologically, a subject agreement marker and an incorporated object pronoun. One of the argument referencing morphs is verb-bound and occurs in a fixed position in the verb template (after the stem), while the other is a mobile morph that can occur either verb-internally (in second or last position) or verb externally. Either the subject agreement marker or the object incorporated pronoun can be verb bound or mobile morphs, depending on the tense and presence of an NP complement. Previous literature has analyzed mobile morphs as (VP) endoclitics. We argue that this is not the case as verb-external mobile morphs occur at the end of the last word of the least oblique NP complement and cannot attach to the last word of VP-internal PPs. We provide an edge-feature based analysis of verb-external mobile morphs and show that the same realizational rules account for the exponents of mobile morph features whether they occur verb-internally or verb-externally. We furthermore suggest that the dissociation between paradigm class (verb-bound or mobile morph) and syntactic status (subject or object; agreement marker vs. incorporated pronoun) challenges views that treat morphological structure as isomorphic to syntactic structure.

1 Introduction

Semantic arguments of predicators can be realized syntactically or be morphologically referenced on the head. Haspelmath (2013) uses the term *argument indexing* to cover both possibilities, a usage we follow, although our paper focuses on the second possibility, the use of affixes to index arguments. We use the term morphological argument referencing or argument referencing for short for the morphological indexing of semantic arguments. Sorani's argument referencing poses interesting challenges to the interface between morphology and syntax and has for this reason received quite a bit of attention (Samvelian 2007, Bonami & Samvelian 2008, Walter 2012, Bonami & Crysmann 2013, Crysmann 2021, Akkus et al. 2023). The issue that makes Soranî argument referencing of particular interest is that it lives a double life. Some argument indices are always expounded verb internally after the verb stem as runof-the-mill verbal affixes but some argument indices either occur verb internally (in all but one case in second position within the verb form) or are affixed/cliticized on the last word of a constituent that precedes the verb within the VP. We refer to these two ways of expounding Soranî argument indices as verb-bound and mobile morphs, respectively (Bonami & Samvelian (2008) use the terms Verbal Person Endings and Mobile Person Markers). In previous HPSG literature, mobile morphs have been treated as second-position (endo-)clitics within the VP, in a way that is reminiscent of Pashto endoclitics (Dost 2007). In this paper, we argue mobile morphs are not second position (endo-)clitics. They do not necessarily occur at the end of the first VP constituent; they occur at the end of the least oblique NP complement. We further argue that a proper analysis of Soranî morphological argument referencing requires dissociating the syntactic status of the argument being referenced (which members of the ARG-ST list is being referenced) from the morphological status of the

inflectional feature being expounded (which paradigm the inflectional feature belongs to). We present an analysis of Soranî morphological argument referencing within Information-based Morphology (IbM) (Crysmann & Bonami 2016) and provide a set of realizational rules for mobile morphs that ensure they are second-position/last position affixes either within the verb form or after the last word of the NP complement to which they are affixed (making use of edge features in the latter case, see Miller & Halpern 1993).

2 Morphological argument referencing in Soranî

Soranî, previously described as a split-ergative language (Thackston 2004, McCarus 2009, MacKenzie 1960), can morphologically reference up to two arguments. When referencing the subject argument in the present tense, a *verb-bound* post-stem affix is employed and a subject external NP can co-occur and agree with the verbal affix. Verb-bound morphs in sentences (1)-(2), for example, agree with the external subject NP¹.

- (1) minał-ekan/(ewan) e-řo-n kid-DEF.PL/they IPFV-gO.PRS-3PL
 'The kids/they are leaving.'
- (2) (*êma*) nan e-xo-yn we food IPFV-eat.PRS-1PL 'We are eating food.'

In contrast to what is the case with subject NPs, object NPs are in complementary distribution with argument referencing *mobile morphs*. Thus, in example (3), the object is expressed as a lexical NP, while in example (4) from Thackston (2004), the object is morphologically referenced by the mobile morph -*t* in second position within the verb form (preceding the stem in this case). Importantly, an object NP cannot co-occur with the mobile morph. In other words, while the morphs that reference the subject in examples (1)-(2) are subject agreement markers, morphs that reference the object realize the relevant argument of the verb in the sense of Levin & Rappaport Hovav (2005). The complementary distribution between morphological argument referencing and syntactic argument realization suggests that object argument referencing affixes are incorporated pronouns in the terminology of Bresnan & Mchombo (1987) who discuss a similar distinction in Chicheŵa.

(3) (min) koř-eke e-bîn-im I boy-DEF.SG IPFV-see.PRS-1SG:A 'I (will) see the boy.'

¹All examples in this work are from fieldwork data unless stated otherwise. Soranî speakers were all from Suleymanî in Iraq or Baneh in Iran.

(4) *e-t-bîn-ê* IPFV-2SG:P-see.PRS-3SG:A 'S/he (will) see you.'

In the past tense, the correspondence between syntactic arguments and verbbound vs. mobile morphs switches. Subject agreement markers are now second position mobile morphs while post-stem verb-bound morphs reference and realize the verb's object. Consider examples (5) and (6). The second-position mobile morphs *-man* and *-yan* are now subject agreement markers and can co-occur with a lexical NP while the object markers are now verb-bound morphs that occur after the stem and they cannot co-occur with external NPs.

- (5) (*ême*) *e-man-kêşan-n*we IPFV-1PL:A-pull.PST-3PL:P
 'We were pulling them.'
- (6) (*ewan*) girt-yan-în they catch.PST-3PL:A-1PL:P 'They caught us.'

Although verb internal mobile morphs typically occur in second position, a third singular mobile morph occurs after the verb-bound morph rather than in the usual second position within the verb template, as shown in (7) and (8) (from Thackston 2004).

- (7) *xward-in-î* eat.PST-3PL-3SG 'S/he ate them'
- (8) dît-în-î see.pst-1pL-3sg
 'S/he saw us.' (Thackston 2004)

Table 1 summarizes the phonology and position (PC) of verb-bound and mobile morphs (we omit a table of the Soranî verb template for reasons of space). Note that although this paper deals almost exclusively with the position of mobile and verbbound morphs, the exponents of the two classes of morphs also differ segmentally. The distinction between verb-bound and mobile morphs is thus one of paradigm class.

The distinction between mobile morphs and verb-bound morphs is straightforward when there is no NP preceding the verb as in the examples we provided so far. The presence of external NPs complicates the distribution of mobile morphs. In example (9), the proto-agent argument is referenced by the verb-bound suffix -n in the present tense; however, in the past tense example (10), the proto-agent is referenced by the mobile morph -yan suffixed to the last word of the object NP.

 (9) minat-ekan hermê-kan beş e-ke-n kids-DEF.PL pear-DEF.PL share IPFV-do.PRS-3PL
 'The kids share the pears'

	Verb Bound	PC	Mobile Morph	PC
1sg	-m	9	-(i)m	2nd
2sg	î(t)	9	(i)t	2nd
3sg	$-\hat{e}(t)/a(t)/\mathcal{O}$	9	-î	LAST
1pl	-în	9	-man	2nd
2pl	-(i)n	9	-tan	2nd
3pl	-(i)n	9	-yan	2nd

Table 1: Morphological argument marking and position class in Soranî

 (10) heřmê-ek-an-yan beş kird pear-DEF-PL-3PL:A share do
 'They shared the pears.' (Mohammadirad 2020)

The mobile morph in sentence (10) is suffixed to the direct object NP. The position of mobile morphs in sentences that contain complex predicates shows that mobile morphs can be suffixed to a complement NP even if it is not a direct object or does not correspond to a semantic argument. Mobile morphs simply attach to the least oblique NP complement.² Complex predicates in Soranî can be formed by combining a light verb and a so-called deverbal, something that is widely assumed to be a nominal. A mobile morph can correspond to the proto-patient argument and be suffixed to the deverbal (11) with the proto-agent expounded through a verb-bound morph on the light verb in the present tense or can be a subject agreement affix (12) in the past tense while the proto-patient is expounded morphologically on the light verb of the complex predicate. The deverbal in both sentences behaves as an NP.

- (11) (ewan) bang-man e-ke-n they call-1pl:p IPFV-do.prs-3pl:A 'They (will call) are calling us.'
- (12) (*ême*) bang-man kird-in we call-1PL:A do.PST-3PL:P 'We called them.'

Importantly, when a complex predicate also combines with a direct object external NP, the subject agreement marker (in the past tense) is now affixed to the last word of that object NP. This phenomenon can be observed in example (13), where the subject agreement is suffixed to the object complement of the complex predicate *heřmekan* 'the pears.' The subject marker can still be affixed to the last word of the constituent headed by the deverbal even when the verb semantically selects for a

²B. Crysmann (p.c.) asked whether mobile morphs attach to NP complements or to second least oblique NP syntactic argument (the second NP member of the ARG-ST list). Examples that involve displaced NP arguments are critical to decide between these two alternatives. Unfortunately, the data we have been able to gather up to this point is not entirely conclusive. We continue to talk about least oblique NP complement for convenience's sake.

proto-patient argument, provided the object NP modifies the deverbal, as illustrated in example (14). The complex predicate in this example is semantically dyadic, corresponding to the English verb *look*. However, the realization of the second semantic argument, the entity being looked at, modifies the deverbal of the complex predicate. This modification is indicated by the presence of an ezafe morph (EZF) on the deverbal. As the constituent headed by the deverbal is now the only NP within the VP, the subject agreement marker is suffixed to the last word of the modifier of the deverbal. Overall, the contrast between examples such as (11)-(12) or (14), on the one hand, and (13), on the other, shows that verb external mobile morphs are suffixed to the last word of the *least oblique* complement NP, the deverbal when there is no other NP complement, the object NP, when there is one.

- (13) heřmê-ek-an-yan beş kird
 pear-DEF-PL-3PL:A share do
 'They shared the pears.' (Mohammadirad 2020)
- (14) [*seyr-î wêne-kan-yan*] *kird* look-ezf photo-DEF.PL-3PL:A do.PST 'They looked at the photos.'

While complex predicates in Soranî are typically semantically dyadic, some complex predicates are semantically monadic and still exhibit a subject agreement marking pattern typically associated with dyadic predicates as examples (15) and (16) illustrate. In both cases, the complex predicate is semantically monadic, yet subject agreement is marked by a mobile morph on the deverbal: the subject agreement marker is suffixed to the right edge of the deverbal NP complement just as for verbs that have an object NP complement.

- (15) *pyase-man kird* walk-1pl do.pst 'We walked.'
- (16) *yarî-yan kird* play-3PL do.PST 'They played.'

The contrast between the verb form-internal subject agreement marker in (17) and the verb form-external subject agreement markers in (15) and (16) shows that it is the presence of an external NP complement, rather than semantic dyadicity, that governs the occurrence of the mobile morph outside of the verb form.

(17) (ewan)/dar-ek-an kewt-in they/tree-DEF-PL fall.PST-3PL:s 'They/the trees fell.'

Stepping back, the dissociation between the morphological and syntactic status of Soranî argument referencing this section discusses poses significant challenges to linguistic theories that assume an isomorphism between the linear order of morphs and constituency or grammatical function (e.g., some version of the Mirror Principle proposed by Baker 1985). Verb-bound and mobile morphs do not consistently correspond to subject agreement markers or the realization of the proto-patient argument, *contra* what one would expect if morphology mirrored syntactic structure.

We have up to now provided examples that suggest that verb external mobile morphs are suffixed to the least oblique NP complement. This contrasts with previous HPSG literature that proposed that mobile morphs are second position clitics (within the VP) when they occur verb externally. We turn to examples that support our generalization and invalidate previous descriptions of Soranî mobile morphs as endoclitics. As PP complements play a critical role in comparing the two descriptions of the distribution of mobile morphs, we first discuss general properties of Soranî adpositions. Traditionally, Soranî adpositions (prepositions) have been grouped into two classes, simplex and compounded adpositions. The inflectional features and morphological attachment of these adpositions have been described extensively (MacKenzie 1960, Thackston 2004, McCarus 2009). Critical for our purposes, each class of adpositions can be further divided into simple and absolute adpositions, as proposed first in MacKenzie (1960). Basically, the pronominal complement of simple adpositions are separate words while the pronominal argument of absolute adpositions is referenced as a morph suffixed to those adpositions.

Most relevant to the issue of the status of the verb-external occurrence of mobile morphs is the fact that the position of mobile morphs seems to depend on whether the adposition is semantically potent. When the adposition is not semantically potent, the mobile morph may be suffixed to the last word of the phrase that contains the adposition, as seen in (18), where the adposition *be* functions as a case marker. By contrast, the same adposition is semantically potent (encoding the notion of goal of motion) in (19) and the mobile morph is now on the last word of the object NP, not on the last word of the PP. The prepositions in (19)-(21) are similarly semantically potent and the mobile morph does not occur on the last word of the phrases that contain these semantically potent adpositions; rather, the mobile morph occurs verb internally, as if the verb had no complement.

- (18) *be Ali-m gut* to Ali-1sg say.pst 'I told Ali.'
- (19) *bo qotabxane minał-ekan-man nard* to school kid-DEF.PL-1PL:A send.PST 'We sent the kids to school'
- (20) *le Kurdistan bar-man kird-in* from Kurdistan load-1PL:A do.PST-3PL:P 'We loaded them from Kurdistan.'

(21) *berew dorge-ke mele-yan e-kird* toward island-DEF.SG swim-3PL IPFV-dO.PST 'They were swimming toward the island.'

Data from conjunction reduction (22) and VP-anaphors (23) show that semantically potent PPs are VP internal in Soranî. The mobile morph agreement marker for the subject *Ali* is suffixed to the last word of the direct object NP in the first conjunct in (22), but occurs verb internally in the second conjunct because the complement of the complex predicate is a PP, not an NP. The fact the combination of the PP and the complex predicate (but to the exclusion of the subject) can be combined with the combination of a direct object and a complex predicate shows that the PP is indeed VP internal. The same conclusion can be drawn from (23). The anaphor in the second conjunct is a VP anaphor, as the two conjuncts have different subjects and its antecedent is *leget mindatekan qisey kird* suggesting that the PP *leget mindatekan* 'with the kids' is VP internal.

- (22) Ali mindał-ekan- \hat{i}_{MM} timaşa kird \hat{u} leget mîwan-ekan qise- y_{MM} kird Ali kid-DEF.PL-3SG watch do.PST and with guest-DEF.PL talk-3SG do.PST 'Ali watched the kids and talked to the guests.'
- (23) Ali leget mindat-ekan qise- y_{MM} kird \hat{u} min-iş herwa Ali with kid-DEF.PL talk-3sG:A do.PST and I-too so 'Ali talked with the kids and so did I.'

To conclude, the contrast between (18)-(19) and (20)-(21) indicates that mobile morphs do not always attach to the last word of the first complement of the VP when they occur verb-externally, as had been previously claimed (Samvelian 2007, Bonami & Samvelian 2008). Rather, they are suffixed to the least oblique NP complement, if there is one, or to a phrase that contains a semantically inert preposition (as in (18)). We analyze semantically inert prepositions as markers: the head of *be Alim* 'to Ali' in (18) is thus an NP, in line with our hypothesis that verb external mobile morphs suffix to the last word of the least oblique NP complement.

We round up our description of Soranî morphological argument referencing with two constructions, the applicative and possessor raising, that add an NP argument to the ARG-ST list and can affect argument referencing. Absolute adpositions can function as applicative markers and increase the number of direct syntactic arguments of the verb as shown in (24)-(25) and discussed in Karim & Salehi (2022). The number of direct arguments increases from 1 to 2 or 2 to 3 when $l\hat{e}$ is added. As the recipient argument in (24) is now expressed as a direct syntactic argument, it is referenced via a verb-bound suffix and the subject agreement marker becomes a mobile morph. Similarly, the addition of $l\hat{e}$ in (25) allows the proto-patient to be referenced by a mobile morph.

(24) *pirsyar-man lê kird-in* question-1pL:A ABS.P do.PST-3PL:P 'We asked them a question.' (25) *lê-t-e-ç-ê* ABS.P-2SG:P-IPFV-go-3SG 'S/he looks like you.'

More generally, adpositions can head a PP that realize a recipient argument or they can combine with a verb as absolute adpositions. In the latter case, the member of the ARG-ST that corresponds to the recipient of a typical semantically tryadic verb is an object NP (when the verb is in the active voice; we omit discussion of passives for reasons of space) and that object NP can be referenced morphologically when it bears a pronominal index. The contrast between these two alternative ways of expressing a recipient is illustrated in (26) vs. (27) and (28).

- (26) *mamosta-ke bo ême kitêb-eke-y hênaw* teacher-DEF.sg for we book-DEF.sg-3sg:A bring.pst 'The teacher brought a book for us.'
- (27) *mamosta-ke kitêb-eke-y bo hênaw-în* teacher-DEF.SG book-DEF.SG-3SG.A ABS.P bring-1PL 'The teacher brought the book for us.'
- (28) *kitêb-ek-it bo e-hên-im* book-DEF.SG-2SG:BEN for IPFV-bring.PST-1SG:A 'I will bring a book for you.'

All Soranî verbs with three NP dependents must include an absolute adposition that functions as an applicative marker, except for the verb *dan* 'give', as shown in (29), where the pronominal recipient is referenced by a verb-bound morph or as in (30) where it is referenced by a mobile morph.

- (29) *kitêb-ek-an-im da-n-ê* book-DEF-PL-1sG:A give.PST-3PL:R-ABS.P 'I gave them the books.'
- (30) (min) sêw-êk-it e-de-m-ê
 (1sg) apple-IND-2sg:R IPFV-give.PRS-1sg.A-ABS.P
 'I will give you an apple' (Mackenzie 1961: Sul.2)

The verb *dan* 'give' is also exceptional in that when the postposition $-\hat{e}$ is present, all three arguments of the verb can be referenced morphologically, as shown in (31).

(31) da=m-in-in-ê give.pst=1sg.A-3pl.o-3pl.R-Appl
'I gave them to them.' (Karim & Salehi 2022)

We now turn to the possessor raising construction, which like applicative uses of absolute adpositions, adds an NP to the ARG-ST list and thus affects argument referencing. Soranî simple verb stems, complex predicates and applicative constructions all allow for possessor raising when the verb is in the past tense under certain conditions. In such cases, the pronominal possessor of an NP is realized as a verb-bound morph after the verb stem and the subject agreement marker is a mobile morph suffixed to the least oblique member of the ARG-ST. Examples are provided in (32)-(34). While possessor raising can be observed with a variety of verbs, it is subject to certain constraints. For example, the verbs *birdin* 'to take' and *peřandin* 'to fly/to jump' allow possessor raising only for certain person/number combinations and some verbs do not allow possessor raising at all. Compare (32) to (34): the possessor of the book in sentence (32) is referenced on the verb with a verb-bound morph but the possessor of the kids in sentence (34) is referenced on the direct object itself, despite the main verb being identical in both sentences. The exact conditions under which possessor raising occurs are outside the scope of this paper.

- (32) *ktêb-ek-it bird-im* book-DEF.SG-2SG:A take.PST-1SG:POSS 'You took away my book.'
- (33) *xew-it lê-peřand-im* sleep-2sg:A APPL-jump-1sg:Poss 'You ruined my sleep.'
- (34) *minal-ek-m-yan bird* kid-DEF.PL-1sG:POSS-3PL:A take.PST 'they took my kids'

3 An IbM account of Soranî argument referencing

Having described how morphological argument referencing works in Soranî, we turn to our model of these facts. To model Soranî morphological argument referencing, we make use of Information-based Morphology (Crysmann 2021, Crysmann & Bonami 2016). In Information-based Morphology (IbM), an INFL feature that records all relevant information for inflection is included in the representation of each word, as shown in (35).

(35) word $\Rightarrow \begin{bmatrix} MPH \ list(mph) \\ MUD \ set(msp) \\ MR \ \begin{bmatrix} MUD \ set(msp) \\ MS \ set(msp) \\ MPH \ list(mph) \end{bmatrix} \end{bmatrix}$

The INFL feature includes three distinct features. The first feature, MPH, is a partially ordered list of the exponents that make up a word and corresponds to the "output" of the set of realizational rules the word instantiates. This list is critical in establishing the linear order of exponents that make up the phonological shape of the word. The fact that the list is partially ordered allows for some degree of flexibility in the ordering of morphs within the word when a language allows for it. The second feature, RR, or set of realizational rules, is responsible for mapping morphosyntactic properties to exponents. Each realizational rule is made up of three properties. MUD (Morphosyntactic features Under Discussion) includes features that a rule expounds while MPH is the set of exponents of the MUD features. Finally, MS is the entire set of morphosyntactic features a word expounds and provides a context for the application of each realizational rule when appropriate. Each member of MPH includes both phonological information (recorded in the value of PH) along with position in the word template (recorded in the value of PC). The final feature of INFL is the MS feature, which, as just mentioned, includes all the morphosyntactic features a word expounds, including the form of the stem on which the word is based. In order to ensure that the word's morphosyntactic property set is present on each realizational rule and provides a possible context for the application of rules, structure sharing is employed between the MS value of each realizational rule and the MS value of the INFL feature.

To guarantee the application of realizational rules and the generation of wellformed words, Crysmann (2021) proposes a general principle for the well-formedness of words. This principle, stated in (36), ensures that the properties expounded by any rule are consistent with the word's morphosyntactic features and that the exponents for each rule appear in the list of morphs of the word: each morphosyntactic feature is expounded in one and only one rule and each exponent of a rule appears in the word form.

$$(36) \quad word \Rightarrow \left[INFL \begin{bmatrix} MPH & \underline{e_1} \bigcirc \dots \bigcirc \underline{e_n} \\ MPH & \underline{e_1} \\ MR & \left\{ \begin{bmatrix} MPH & \underline{e_1} \\ MUD & \overline{m_1} \\ MS & 0 \end{bmatrix}, \dots \begin{bmatrix} MPH & \underline{e_n} \\ MUD & \overline{m_n} \\ MS & 0 \end{bmatrix} \right\} \end{bmatrix} \right]$$

To model Sorani's mobile morphs one important feature needs to be added since the position of mobile morphs in the verb template is not absolute but relative: mobiles morphs (except for third person) occur after the first morph, irrespective of the position in the verb template of that first morph. Crysmann (2021) proposes an analysis of mobile morphs that makes use of a "pivot feature" 1st-Pc to model their relative position within the verb template. This pivot feature is included in all members of the list of morphs (as shown in (37)) and can thus be referenced by rules that realize mobile morphs verb-internally such as the one he proposes in (39).

$$(37) \quad word \Rightarrow \left[\text{INFL} \left[MPH \left\langle \begin{bmatrix} PC & \boxed{1} \\ 1ST-PC & \boxed{1} \\ STM-PC & \boxed{S} \end{bmatrix}, \begin{bmatrix} 1ST-PC & \boxed{1} \\ STM-PC & \boxed{S} \end{bmatrix}, \dots, \begin{bmatrix} 1ST-PC & \boxed{1} \\ STM-PC & \boxed{S} \end{bmatrix} \right\rangle \right] \right]$$

The rule in (39) is intended to model the mobile morph *-man* in (38): it requires *-man* to occur one position after the first position in the verb template, i.e. in second position.

(38) *e-man-xward-in* IPFV-1PL:A-eat.PST-3PL:P 'We didn't eat them.'

$$(39) \begin{bmatrix} MUD \left\{ \begin{bmatrix} PERS & 3\\ NUM & pl \end{bmatrix} \right\} \\ MPH \left\langle \begin{bmatrix} PH & \langle man \rangle \\ 1sT-PC & 1 \\ PC & 1 + 1 \end{bmatrix} \right\rangle \end{bmatrix}$$

The analysis presented in Crysmann (2021) is meant to cover verb-internal exponence of mobile morphs. In the rest of this section, we provide a more general account of Soranî argument referencing, an account that covers both verb-bound morphs and verb-external as well as verb-internal occurrence of mobile morphs. To provide a comprehensive analysis of Soranî morphological argument referencing, we need to first distinguish between two sorts of argument referencing features or arg-ref whose informational structure is provided in (40). Vb-arg-ref and mm-arg-ref (for verb-bound and mobile morph argument referencing features) are subsorts of arg-ref (see Bonami & Crysmann (2013) for a similar distinction between verbal person ending and mobile person markers morphosyntactic feature sorts). These two subsorts of arg-ref encode the fact that verb-bound and mobile morphs constitute two distinct paradigms. Realizational rules can then distinguish between the position of the exponents of these two subsorts (second morph or post-stem morph) as well as their phonology, except in the case of the first person singular form where the segmental make-up of the morph is the same for the two subsorts of argument referencing features (but their position remains distinct). The realizational rule for the first person affix will target the *arg-ref* sort rather than either of its subsort, ensuring that the PH, but not the position class, of first person affixes is shared between the relevant verb-bound and mobile morphs.

(40)
$$\begin{bmatrix} arg\text{-}ref \\ ARG\text{-}IND nom\text{-}ind \end{bmatrix}$$

Having introduced the feature that are targeted by argument referencing realizational rules, we now turn to how argument referencing is achieved. Conceptually, Soranî argument referencing involves three distinct sets of constraints. The first set relates indices of syntactic arguments (members of ARG-ST) with morphosyntactic features (members of MS or the scrapbook feature MM-FEAT, see below), what we call *argument referencing features* for ease of reference. This set of constraints ensures that argument referencing features are assigned the right sort, *vb-arg-ref* or *mm-argref*, respectively, as well as the appropriate person and number. The rule we provided in (39) from Crysmann (2021) is thus amended as follows so that it only applies to argument referencing features of subsort *mm-arg-ref*.

(41)
$$\begin{pmatrix} \text{MUD} \left\{ \begin{bmatrix} mm\text{-}arg\text{-}ref \\ \text{ARG-IND} \begin{bmatrix} \text{PERS} & 3 \\ \text{NUM} & pl \end{bmatrix} \right\} \\ \text{MPH} \left\langle \begin{bmatrix} \text{PH} & \langle \text{man} \rangle \\ 1\text{ST-PC} & \boxed{1} \\ \text{PC} & \boxed{1} + 1 \end{bmatrix} \right\rangle$$

The second set of constraints—which is only relevant to verb-external mobile morphs—ensures that a *mm-arg-ref* feature is present on the last word of an NP complement if there is one, rather than being expounded on the verb. The third set of constraints are the realizational rules themselves, for both verb-bound and mobile morphs. We tackle each set of constraints in turn.

The first step in modeling Soranî argument referencing involves relating indices of syntactic arguments with argument referencing features. As discussed extensively in previous literature as well as in Section 2, subjects of present tense verbs are referenced by verb-bound affixes whether the verb is intransitive or transitive. In other words, whether the verb's ARG-st includes one or two NPs, the argument referencing feature for the subject is of sort *vb-arg-ref*. The constraint in (42) models this generalization as follows. The ARG-st includes an NP corresponding to the verb's subject followed by a (possibly empty) *list* of other syntactic arguments. The NP argument's index is structure-shared with the value of the ARG-IND feature of *vb-arg-ref* in the morphosyntactic property set. This structure-sharing ensures that the exponent of the argument referencing feature agrees with the subject NP. As the argument referencing feature (42) introduces is of sort *vb-arg-ref*, subject agreement affixes will be verb-bound morphs.

(42)
$$\begin{pmatrix} \text{HEAD}|\text{TNS } pres \\ \text{ARG-ST } \langle ^{\text{NP}} \underline{1} \rangle \oplus list \end{pmatrix} \Rightarrow \begin{bmatrix} \text{INFL } \begin{bmatrix} \text{MS } \left\{ \begin{bmatrix} vb\text{-}arg\text{-}ref \\ \text{ARG-IND } \underline{1} \end{bmatrix} \right\} \uplus set \end{bmatrix} \end{bmatrix}$$

When the verb is in the past tense, the argument referencing feature of subjects is not uniformly of sort *vb-arg-ref*. It is of sort *vb-arg-ref* if the verb only has one NP syntactic argument (is intransitive); it is of sort *mm-arg-ref* if it has two or more NP syntactic arguments. The constraint in (43) ensures that the subject argument referencing feature of past tense verbs with only one NP in their ARG-ST list is of sort *vb-arg-ref*; the constraint in (44) ensures that the subject argument referencing feature of past tense verbs with two or more NPs in their ARG-ST list is of sort *mm-arg-ref*. Note that this last constraint makes use of the MM-FEAT feature whose value is a set of *mm-arg-ref* features. As we discuss shortly, this feature—a sort of scrapbook feature—proves useful when modeling the alternation between verb-internal and verb-external exponence of mobile morph features.

$$(43) \quad \begin{bmatrix} \mathsf{HEAD}|\mathsf{TNS} \ past\\ \mathsf{ARG-ST} \ \langle \mathsf{NP}_{\boxed{1}} \rangle \oplus \mathit{list}([\mathsf{CAT} \neg \mathit{noun}]) \end{bmatrix} \Rightarrow \begin{bmatrix} \mathsf{INFL} \left[\mathsf{MS} \left\{ \begin{bmatrix} \mathit{vb-arg-ref}\\ \mathsf{ARG-IND} \end{array} \right] \right\} \oplus \mathit{set} \end{bmatrix} \\ (44) \quad \begin{bmatrix} \mathsf{HEAD}|\mathsf{TNS} \ past\\ \mathsf{ARG-ST} \ \langle \mathsf{NP}_{\boxed{1}}, \mathsf{NP} \rangle \oplus \mathit{list} \end{bmatrix} \Rightarrow \begin{bmatrix} \mathsf{INFL} \left[\mathsf{MM-FEAT} \left\{ \begin{bmatrix} \mathit{mm-arg-ref}\\ \mathsf{ARG-IND} \end{array} \right] \right\} \end{bmatrix} \end{bmatrix}$$

The sorts of argument referencing features of pronominal direct objects is the mirror image of the sorts of subject argument referencing features. Object argument referencing features are of sort *mm-arg-ref* when the verb is in the present tense and of sort *vb-arg-ref* when the verb is in the past tense, as shown in the constraints in (45) and (46), respectively. Note that while there was no restriction as to the sort of *synsem* targeted by the constraints introducing argument referencing features for subjects (Soranî is a *pro*-drop language), object argument referencing features introduced by the constraints in (45) and (46) require the relevant ARG-ST member to be of sort *pron-aff*, i.e. to be a non-canonical *synsem* that is not structure-shared with valence lists, following the work of Miller & Sag (1997).

$$(45) \begin{bmatrix} |\text{HEAD}|\text{TNS } pres \\ |\text{ARG-ST} \rangle & |\text{NP, NP}^{pron-aff}|_{2} \end{pmatrix} \oplus list \end{bmatrix} \Rightarrow \begin{bmatrix} \text{INFL} \begin{bmatrix} \text{MM-FEAT} \left\{ \begin{bmatrix} mm-arg-ref \\ |\text{ARG-IND} | 2 \end{bmatrix} \right\} \end{bmatrix} \end{bmatrix}$$
$$(46) \begin{bmatrix} \text{HEAD}|\text{TNS } past \\ |\text{ARG-ST} \rangle & |\text{NP, NP}^{pron-aff}|_{2} \end{pmatrix} \oplus list \end{bmatrix} \Rightarrow \begin{bmatrix} \text{INFL} \begin{bmatrix} \text{MS } \left\{ \begin{bmatrix} vb-arg-ref \\ |\text{ARG-IND} | 2 \end{bmatrix} \right\} & |\text{JSET}|_{2} \end{bmatrix}$$

The set of constraints we just provided ensures that an argument referencing feature of the appropriate sort is part of the inflectional structure of the verb and is co-indexed with the relevant ARG-ST member. In most languages, all that would then be needed would be realizational rules that expound those argument referencing features verb internally. But, Soranî's mobile morphs can occur within the verb form (mostly, in second position in the verb template) or be suffixed at the right edge of the first NP complement. So, we need to account for the fact that inflectional features of sort *mm-arg-ref* are alternatively expounded verb-internally and verb-externally. A second set of constraints is needed to model this alternate exponence of *mm-arg-ref* features. As we want the same constraints relating ARG-ST members and inflectional features to apply whether a mm-arg-ref feature is expounded verb-internally or verbexternally, the scrapbook feature MM-FEAT we introduced in constraints (44) and (45) comes in handy, as this feature helps keep the introduction of an argument referencing feature independent of whether that feature is expounded verb-internally or verb-externally. The constraint in (47) says that when a verb has no NP complement, the mobile morph feature set (the value of MM-FEAT) is part of the verb's MS (and will thus be expounded within the verb form, as per the Well-formedness constraint in (36)).³ The constraint in (47) suffices to ensure the verb-internal realization of mobile morph features when no NP complement is present.

(47)
$$\begin{bmatrix} \text{INFL} & [\text{MM-FEAT} \boxed{1} \\ \text{COMPS} & list([\text{CAT} \neg noun]) \end{bmatrix} \Rightarrow [\text{INFL} & [\text{MS} \boxed{1} \uplus set] \end{bmatrix}$$

³We assume here that the constraint applies to *in situ* NP complements and therefore references NP members of the COMPS list. As mentioned above, the data we have been able to gather so far is not entirely probative and further fieldwork is needed. If mobile morphs can be suffixed to displaced syntactic arguments, the constraint would reference ARG-ST members. Nothing substantial hinges on this issue.

The constraint in (48) introduces via structure-sharing the mobile morph feature set (the value of MM-FEAT) on the least oblique NP complement as an edge feature rather than on the MS of the verb, thus guaranteeing the feature will not be expounded verb internally. Standard approaches to edge features further guarantee that the mobile morph argument referencing feature set is part of the representation of the last word of the least oblique NP complement (the NP whose *synsem* is tagged [2] in (48)).

$$(48) \quad \begin{bmatrix} \text{INFL} & [\text{MM-FEAT} \ \boxed{1}] \\ \text{COMPS} & \langle \boxed{2} \text{NP} \rangle \oplus \boxed{3} \end{bmatrix} \Rightarrow \begin{bmatrix} \text{COMPS} & \langle \boxed{2} \begin{bmatrix} \text{EDGE} \end{bmatrix} \text{RIGHT} & \boxed{1} \uplus \text{set} \end{bmatrix} \rangle \oplus \boxed{3} \end{bmatrix}$$

Although many different approaches to edge features would fit the bill, we follow the approach developed in Miller & Halpern 1993 and Halpern (1995) (see also Tseng 2003) and its HPSG implementation in Crysmann (2010). In that approach, edge features "percolate down" from the root of the phrase they are introduced on to the leftmost/rightmost word they are expounded on: in our case, from the root of the least oblique NP to the last word of that NP. These approaches distinguish between edge trigger features, launching an edge inflection dependency, and edge marking features. Percolation of feature values is achieved through the Edge Feature Principle as follows: "The right (left) MARK feature of the right (left) daughter is the concatenation of the right (left) MARK and TRIG features of the mother." (Crysmann 2010: 278), as stated semi-formally in (49) (adapted from Crysmann 2010). The principle in (49) together with the constraint in (48) guarantees that the mobile morph is realized at the right edge of the noun phrase that corresponds to the first NP member of the comps list.

(49)
$$\begin{bmatrix} ss & edge & MARK | RIGHT & 2\\ TRIG | RIGHT & 1 \end{bmatrix} \\ DTRS & \left< [ss|edge|MARK | RIGHT & 1] \\ \end{bmatrix} \right> \oplus list \end{bmatrix}$$

The second set of constraints we have laid out makes sure that the mobile morph argument referencing feature is either part of the verb's set of inflectional features (member of its MS) or is part of the set of right edge features of the least oblique NP and percolates down to the last word of that NP. Before one can apply the third set of constraints (the realizational rules themselves for either verb-bound or mobile morphs), we need to make sure mobile morph argument referencing features on the set of right edge features are part of the MS of the last word of the least oblique NP complement. To that end, we posit the word-to-word construction in (50).

$$(50) \begin{bmatrix} wd-to-wd-infl\\ INFL & \left[Ms \left\{ \boxed{3}, \left[mm-stem-lid \\ STEM & \boxed{2} \right] \right\} \right] \\ EDGE & \left[MARK|RIGHT \left\{ \boxed{3} \right\} \right] \\ DTRS & \left\langle \left[word \\ PH & \boxed{2} \right] \right\rangle \end{bmatrix}$$

A few comments on this construction are in order. First, since mobile morphs can attach to uninflected as well as inflected last words of NPs, the argument referencing feature is only part of the value of the MS attribute of the mother, as the daughter may already be a fully inflected word. Second, the construction creates a new stem whose stem phonology is identical to that of the phonology of the word to which the mobile morph is suffixed. Third, the extended word created by the construction contains in its inflectional feature set (the value of its MS) the right edge feature ([1] in (50)), i.e., the mobile morph argument referencing feature as well as the new stem feature ([6]in (50)). That's it! Nothing else is needed aside from the *wd-to-wd-infl* construction to properly expound mobile morph argument referencing features verb-externally. The same rules used for expounding the features verb-internally apply when they are expounded verb-externally. So, the realizational rule in (41) applies whether *-man* is realized verb-externally as in (12) or verb-internally as in (38). The *wd-to-wd-infl* construction provided in (50) directly accounts for the expounding of the *mm-arg-ref* in the same position class as when it is expounded verb-internally: the mobile morph follows the first morph of the extended word just as it does when it follows the first morph of an inflected verb.

Interestingly, the introduction via (50) of the mobile morph argument referencing feature onto the MS of last word of the least oblique NP complement also accounts for cases where the mobile morph is not the second morph in a word. Until now, we have focused on cases where the mobile morph is the second morph in an inflected word. But this is not always the case, as (53) and (54) show (to be compared to the more usual pattern in (51) and (52)). As noted by Samvelian (2007) and others, third singular mobile morphs occur *after* verb-bound morphs; in fact, they occur in the final position in the verb template, i.e. as the last suffix of the verb.

- (51) *dît-man-Ø* see.pst-1pl:A-3sG:P 'We saw him/her.'
- (52) *dît-man-in* see.pst-1pl:A-3pl:P 'We saw them.'
- (53) *dît-în-î* see.pst-1pl:p-3sg:A 'He saw us.'
- (54) *xward-in-î* eat.PST-3PL:P-3SG:A 'He ate them.'

The realizational rule in (55) models such verb-final realization of mobile morphs. As the position class is again relative rather than absolute (last morph in the inflected word), we make use of the feature LAST-PC to ensure $-\hat{i}$ is the last morph. This rule also applies when the mobile morph occurs verb-externally: as the MS of the mother node of any construct based on the construction in (50) contains only two inflectional

features, the stem feature and the mobile morph argument referencing feature, the last position (for third singular mobile morphs) and second position (for all other mobile morphs) are the same position.

(55)
$$\begin{bmatrix} MUD \left\{ \begin{bmatrix} mm-arg-ref \\ ARG-IND \begin{bmatrix} PERS & 3 \\ NUM & sg \end{bmatrix} \right\} \end{bmatrix}$$
$$\begin{bmatrix} MPH \left\langle \begin{bmatrix} PH & \langle \hat{1} \rangle \\ LAST-PC & 1 \\ PC & 1 \end{bmatrix} \right\rangle$$

We have now discussed three sets of constraints required for modeling morphological argument referencing in Soranî. The first set of constraints identifies indices of NP members of ARG-ST to argument referencing inflectional features and assign the appropriate paradigm sort to these inflectional features. The second set of constraints ensures the mobile morph argument referencing feature is either part of the verb's MS (when the mobile morph occurs verb-internally) or on the MS of the last word of the least oblique NP (when the mobile morph occurs verb-externally). The third set of constraints are the realizational rules themselves that provide segmental and positional information for combinations of person and number that are appropriate for the verb-bound and mobile morph paradigms. The last aspect of Soranî argument referencing to cover is the neutralization of paradigm sort for the phonological, but not positional, exponence of first person indices, and the variable morphotactics of mobile morphs (last position for exponents of third singular indices, second position in all other cases). In both cases, our model makes use of the dissociation between phonology and morphotactics proposed in Crysmann & Bonami (2016) and the notion of Online Type Construction first proposed in Koenig & Jurafsky (1994) and Koenig (1999), as Figure 1 illustrates.

The MORPHOTACTICS dimension specifies positional information about exponents of argument referencing features. Verb-bound morphs always occur in position 9 in the verb template; mobile morphs occur in last or second position depending on whether the argument referencing feature is third singular or not. The PHON dimension specifies segmental information about exponents of argument referencing features. The segmental exponence of first person singular is the same for verb-bound and mobile morphs and the rule *1sg-phon* therefore mentions the supersort *arg-ref*. It applies whether the argument referencing feature is of subsort *vb-arg-ref* or *mm-arg-ref*. Exponents of other person/number combinations differ segmentally between verb-bound and mobile morphs, as illustrated in (41) above for the first plural mobile morph *-man*.

4 Final remarks

At first glance, the morphological argument referencing system of Soranî appears to resemble the second position (endo-)clitic system of Pashto and this is indeed how it has been analyzed in previous HPSG work. We showed in this paper that Soranî's



Figure 1: A partial hierarchy of argument referencing realizational rules for Soranî argument referencing

verb-external argument referencing morphs are not second position clitics. Rather, they function as edge affixes on the least oblique NP complement. Mobile morphs are, we claim, best modeled as second position inflectional affixes (leaving aside the case of the third person singular morph), using the relative morph placement analysis proposed by Crysmann (2021: 983), as well as the approach to edge inflection outlined in Miller & Halpern (1993) and Crysmann (2010).

More broadly, Soranî morphological argument referencing demonstrates a unique dissociation between syntax and morphology, which has implications for the architecture of grammars. Specifically, both subject and object argument referencing morphs can be either *verb-bound* or *mobile*, and both *verb-bound* and *mobile* morphs can correspond to either agreement markers or so-called incorporated pronouns. Such dissociations between syntactic status and morphological status are difficult to reconcile with theories that assume an isomorphism between morphological linear order and constituency/grammatical function, such as the mirror principle proposed by Baker (1985), or the view that morphological expounding feeds off functional syntactic terminals as proposed by Embick (2015) (but see Akkus et al. (2023) for an analysis of Soranî argument referencing within an approach of the kind advocated in Embick (2015) that may overcome the difficulties the Soranî data presents). In frameworks that maintain a distinction between inflectional features (members of the Ms set) and syntactic information (order on the ARG-ST list or presence on both the ARG-ST list and valence lists), such as HPSG, such dissociations are easily modeled.

References

- Akkus, Faruk, David Embick & Mohammed Salih. 2023. Case and the syntax of argument indexation an analysis of Soranî Kurdish. University of Massachusetts Amherst and University of Pennsylvania.
- Baker, Mark. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16(3). 373–415.
- Bonami, Olivier & Berthold Crysmann. 2013. Morphotactics in an information-based model of realisational morphology. In Stefan Müller (ed.), *Proceedings of the* 20th International Conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin, 27–47. Stanford, CA: CSLI Publications.
- Bonami, Olivier & Pollet Samvelian. 2008. Soranî Kurdish person markers and the typology of agreement. Talk presented at the 13th International Morphology Meeting, Vienna.
- Bresnan, Joan & Sam A. Mchombo. 1987. Topic, pronoun, and agreement in Chicheŵa. *Language* 63(4). 741–782.
- Crysmann, Berthold. 2010. Discontinuous negation in Hausa. In Stefan Müller (ed.), Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar, Université Paris Diderot, Paris 7, France, 269–287. Stanford, CA: CSLI Publications.

- Crysmann, Berthold & Olivier Bonami. 2016. Variable morphotactics and Information-based Morphology. *Journal of Linguistics* 52(2). 311–374.
- Crysmann, Bertholt. 2021. Morphology. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head Driven Phrase Structure Grammar*, 1st edn., chap. 21, 219–244. Berlin: Language Science Press. DOI: 10.5281/ zenodo.5543318.
- Dost, Ascander. 2007. A domain-based approach to second position clitics in Pashto. In Frederick Hoyt, Nikki Seifert, Alexandra Teodorescu & Jessica White (eds.), *Texas Linguistic Society IX: the morphosyntax of underrepresented languages*, 89– 110. Stanford, CA: CSLI Publications.
- Embick, David. 2015. *The morpheme: a theoretical introduction* (Interface Explorations 31). Berlin: de Gruyter Mouton.
- Halpern, Aaron. 1995. *On the placement and morphology of clitics* (Dissertation in Linguistics). Stanford, CA: CSLI Publications.
- Haspelmath, Martin. 2013. Argument indexing: a conceptual framework for the syntactic status of bound forms. In Dik Bakker & Martin Haspelmath (eds.), *Languages across boundaries: studies in memory of Anna Siewierska*, 197–227. Berlin: De Gruyter Mouton.
- Karim, Shuan & Ali Salehi. 2022. An applicative analysis of soranî "absolute prepositions". In Sara Pacchiarotti & Fernando Zuniga (eds.), *Applicative Morhphology: Neglected Syntactic and Non-syntactic Functions*, 1st edn., chap. 10, 263–298.
 Berlin, Boston: De Gruyter Mouton. DOI: 10.1515/9783110777949.
- Koenig, Jean-Pierre. 1999. *Lexical relations* (Stanford Monographs in Linguistics). Stanford: CSLI Publications.
- Koenig, Jean-Pierre & Dan Jurafsky. 1994. Type underspecification and on-line type construction in the lexicon. In.
- Levin, Beth & Malka Rappaport Hovav. 2005. *Argument realization* (Research Surveys in Linguistics). Cambridge, UK: Cambridge University Press.
- MacKenzie, D N. 1960. The Origins of Kurdish. Oxford University Press. 2-16.
- Mackenzie, David N. 1961. Kurdish Dialect Studies. Oxford: Oxford University Press.
- McCarus, Ernest N. 2009. Kurdish. In Gernot Windfuhr (ed.), *The iranian languages*, 1st edn., chap. 10, 587–633. New York: Routledge.
- Miller, Philip & Aaron Halpern. 1993. English possessives and the syntax of morphological features. In *Proceedings of the 3rd meeting of the formal linguistics society of Mid-America*, 219–234. Evanston, Northwestern University.
- Miller, Philip & Ivan Sag. 1997. French clitic movement without clitics or movement. *Natural Language and Linguistic Theory* 15(3). 573–639.
- Mohammadirad, Masoud. 2020. Pronominal clitics in Western Iranian languages: Description, mapping, and typological implications. Paris: Université Sorbonne Nouvelle Paris 3. (Doctoral dissertation).
- Samvelian, P. 2007. What Sorani Kurdish absolute prepositions tell us about cliticization. In *Texas linguistic society ix: the morphosyntax of underrepresented languages*, 263–283.

- Thackston, Wheeler M. 2004. Sorani Kurdish: a reference grammar with selective readings. Cambridge, MA: Harvard University.
- Tseng, Jesse L. 2003. Edge features and French liaison. In Jong-Bok Kim & Stephen Wechsler (eds.), *On-line proceedings of HPSG 2002*, 313–333. Stanford, CA: CSLI Publications.
- Walter, Géraldine. 2012. Fitting into morphological structure: accounting for Sorani Kurdish endoclitics. In Angela Ralli, Geert Booij, Sergio Scalise & Athanasios Karasimos (eds.), On-line proceedings of the eighth Mediterranean morphology meeting, 299–321.