

Mixed clitics in Udi: The lexical verb as interface

Berthold Crysmann 

Université Paris Cité, Laboratoire de linguistique formelle, CNRS

Ana Luís 

Universidade de Coimbra

Proceedings of the 32nd International Conference on
Head-Driven Phrase Structure Grammar

Centro de Linguística da Universidade de Lisboa, Lisbon, Portugal

Stefan Müller, Shûichi Yatabe (Editors)

2025

Berlin: Berlin Universities Publishing

pages 5–25



Keywords: HPSG, morphosyntax, clitics, phrasal affixes

Crysmann, Berthold & Ana Luís. 2025. Mixed clitics in Udi: The lexical verb as interface. In Stefan Müller & Shûichi Yatabe (eds.), *Proceedings of the 32nd International Conference on Head-Driven Phrase Structure Grammar, Centro de Linguística da Universidade de Lisboa, Lisbon, Portugal*, 5–25. Berlin: Berlin Universities Publishing. DOI: 10.83067/hpsg.2025.1.  CC-BY 4.0

Abstract

In this paper, we investigate the hybrid properties of subject agreement markers in the Udi language, which exhibit properties of both affixes and clitics. Through detailed analysis of their syntactic and morphological behaviour, including phenomena such as infixation, suffixation and edge inflection, we show that the realisation of these markers is not only sensitive to information structure (focus), as argued by Harris (1996, 2002), but also to the internal morphological structure of the verb. Building on previous work on Sorani Kurdish (Salehi & Koenig 2023), our HPSG analysis treats this hybrid status as an instance of delegated inflection, thereby enabling a modular account of the alternation between lexical and phrasal affixation.

1 Introduction

Clitics constitute a classical problem at the morphology-syntax interface, sharing properties with morphologically bound affixes, while enjoying some degree of syntactic mobility. Subject agreement markers¹ in Udi (Harris 2002) present a particularly challenging case of *mixed clitics* (Luís 2009, Spencer & Luis 2012): when realised as part of the verb, they provide strong evidence in favour of lexical affixes, including infixation (1). However, when realised externally, their attachment properties show the typical behaviour of post-lexical clitics (2), attaching to the edge of a focused constituent.²

- (1) äyel-en a-**t'u**-k'-e k'uč'an-ax
 child-ERG see₁-3SG.INV-see₂-AORII puppy-DAT
 'The child saw the puppy.' (Harris 2002: 28)
- (2) met'in-al burim sa yärvägä q'əzəl-**le** aq'sa
 this.ERG-AND apparently one bag.ABS gold.ABS-3SG take
 'She also apparently takes a BAG OF GOLD.' (Harris 2002: 44)

On closer inspection, complementary distribution between (1) and (2) is controlled by the verb: verb-internal placement constitutes the default (broad focus or

[†]We would like to thank the audience at HPSG 2025, in particular Bruno Estigarribia, Jakob Maché and Manfred Sailer, for comments and discussion, as well as the anonymous reviewers for the conference and the proceedings. The research reported here has partially benefited from a public grant overseen by the French National Research Agency (ANR), as part of the program "Investissements d'Avenir" (reference: ANR-10-LABX-0083), and from CELGA-ILTEC through public funding by the Fundação para a Ciência e a Tecnologia. It contributes to the IdEx Université Paris Cité (reference: ANR-18-IDEX-0001).

¹These markers are also referred to as "Personalzeichen" (Schulze 1982) or person markers (PM; Harris 2002). See Table 1 for full paradigms. For ease of reference, we use bold face for these markers.

²With respect to the characterisation of phrasal hosts, we essentially follow the empirical generalisation by Harris (1996, 2002), who suggests that focus is the notion that best characterises the phrasal hosts, including focused constituents, wh-phrases and negation. Harris (2002: Section 6.8) also provides a detailed discussion that attachment cannot be simply understood in prosodic terms, neither with phrasal nor with lexical hosts.

predicate focus); likewise, suffixal attachment to the verb is triggered by specific TAM properties (e.g. FutureII), preempting focus-driven placement. Furthermore, the verb class decides on the set of allomorphs: the inverse set (1) or the general set (2). This allomorphy applies to internal and external realisations alike.

Currently, no formal account addresses both sides of the coin: while the *endocclisis* approach of Harris (2002) essentially plays down the morphological side of the problem, the *mixed clitic* approach (Luís 2009, Spencer & Luis 2012) does not make explicit the interplay between morphology and syntax. Here, we provide this missing link, suggesting that the lexical verb assumes a pivotal role between internal and external realisation.

Analysing Sorani Kurdish mixed clitics, Salehi & Koenig (2023) combine a morphological approach to verb-internal realisation with edge inflection (Miller 1992, Halpern 1995) verb-externally. They argue that the verb may delegate, via a selectional requirement, realisation of part of its morphosyntactic properties to a prominent dependent, which is the least oblique NP complement in their case. Building on this, we shall analyse the Udi focus position left-adjacent to the verb as such a prominent dependent. To capture this positional restriction in HPSG, we assume that Udi focused dependents are initial elements of the complements list. For arguments, this means placement at the front of the list, whereas for modifiers, we adopt type shifting of an adjunct to first complement (Bouma et al. 2001). Thus, we can lexically control for the syntactic distribution of Udi subject agreement markers, account for their morphological properties and capture the complementary distribution between focus marking and verb-internal morphological realisation.

2 Data

Udi is a Northeast Caucasian (Lezgian) language. It is head-final, with a basic SOV order in main clauses, and allows null subjects. Word order is generally stable, though it may vary under the influence of focus. Udi displays ergative-absolutive alignment: the absolutive case marks the subject of intransitive verbs and the object of transitive verbs, while the ergative case marks the subject of transitive verbs. Grammatical relations are primarily encoded through a rich system of suffixal case marking; the language distinguishes ergative, absolutive, genitive, dative, and a large set of local and spatial cases, including locative, allative, and ablative (Schulze 1982, Harris 2002).

This section examines the syntactic and morphological properties of Udi clitics, critically reassessing the proposal that they should be analysed as endoclititics (Harris 2002), and focusing on their implications for the syntax-morphology interface.

2.1 Syntactic distribution

2.1.1 “Promiscuous” attachment

One of the defining features of Udi clitics is their ability to attach to morphological and syntactic hosts. They may be found suffixed to the verb (3a), but they may also

attach to the main predicate of a clause, even when the predicate is not a verb (3b) or they may attach to focused constituents (3c,d).

- (3) a. xorag box-al-**zu**
 food.ABS boil-FUT-1SG
 ‘I will cook some food.’
 b. beš mässäb š:el-**le**
 our religion.ABS good-3SG
 ‘Our religion is good.’
 c. äyel-en p’a eş-**ne** aq’-e
 child-ERG two apple.ABS-3SG take-AORII
 ‘The child took **two apples**.’
 d. irähät-en mya-**ne** bist’a cil-l-ux
 peasant-ERG here-3SG sow.PRES seed-OBL-DAT
 ‘The peasant sows seeds **here**.’ Harris (2002: 95)

2.1.2 Focused constituent

As proposed by Harris (1996, 2002), Udi clitics attach to the element in focus, which can be preverbal negation (4), wh-phrases (5) or narrow focus structures (6). The focused constituent is systematically found left-adjacent to the verb (Harris 1996, 2002). This distribution suggests a strong correlation between syntactic prominence and clitic placement.

- (4) zu k’inigax te-**z** bese
 I book.DAT not-1SG requested
 ‘I didn’t ask for a book.’ (Harris 1996: 212)
 (5) a. me išq’armux mano äizi-**q’un** karxesa?
 this man.PL.ABS which village.DAT-3PL.D live
 ‘Which village do these men live in?’ (Harris 1996: 210)
 b. xinärmux ma-**q’un** taisa?
 girl.PL.ABS where-3PL.D go
 ‘Where are these girls going?’ (Harris 1996: 210)
 (6) me xinären tägsa k’inigiyo-laxo-**ne** fikirbesa
 this girl.ERG only book.PL.DAT-about-3SG thinks
 ‘Does this girl only think about books?’ (Harris 1996: 208)

2.2 Lexical properties

Udi subject agreement markers also show properties typically associated with affixes, further blurring the line between morphology and syntax.

2.2.1 Infixation

With complex verb stems, Udi subject agreement markers appear as infixes between the lexical root and the light verb, glossed as DO, CAUS, LV, in (7) below. This placement is evidently highly sensitive to the morphological structure of the verb.

- (7) a. zavod-a aš-**ne**-b-sa
 factory-DAT work-3SG-DO-PRES
 ‘She works in a factory.’
 b. me pasčay-en eč-es-**ne**-st’a ... kul
 this king-ERG bring-INF-3SG-CAUS.PRES earth.ABS
 cip-es-*ne*-st’a pak-i
 spread-INF-3SG-CAUS.PRES garden-DAT
 ‘This king has earth brought ...; he has it spread in the garden.’
 c. zer-ev-**ne**-k’-sa
 decorate-CAUS-3SG-LV-PRES
 ‘She arranges [the house].’ (Harris 2002: 122)

With simple verbs, the subject agreement marker is typically infixed before the final consonant of the root, cf. (8). The root is regularly followed by affixes, which means that even here placement needs access to morphological structure.

- (8) pasčay-un yar-en gölö be-**ne**-y-sa met’a-laxo
 king-UN boy-ERG much look₁-3SG-look₂-PRES this.GEN-on
 ‘The prince looks at this for a long time.’ (Harris 2002: 125)

A class of verbs displays an interesting alternation between infixation into the root or suffixation to it. As shown below, we find infixation with the transitive variant (9a/10a), yet suffixation with the inchoative variant (9b/10b).

- (9) a. a-**t’u**-k’-sa
 see₁-3SG.INV-see₂-PRES
 ‘(s)he sees [it]’
 b. ak’-**ne**-sa
 see-3SG-PRES
 ‘[it] is visible’ (Harris 2000: 127)
 (10) a. bo-**ne**-x-sa
 boil₁-3SG.INV-boil₂-PRES
 ‘(s)he boils [it]’
 b. box-**ne**-sa
 box-3SG-PRES
 ‘[it] boils’ (Harris 2000: 127)

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

Table 1: Udi subject agreement markers (Harris 2002: 28)

2.2.2 Trapping

In certain tenses (e.g., Future II, Subjunctive I and II, and Imperative)³, subject agreement markers must be realised as right-most suffixes on the verb, preempting both infixation(11) and attachment to a focused phrase (12):

- (11) ič turinlek'er-ax nut' tāmiz-b-al-**le**
self's shoe-DAT NEG clean-DO-FUTII-3SG
'She won't clean her shoes.' (Harris 2002: 96)
- (12) gel-enk' ek'a yaq'a-b-al-**lu**
Gela-for what.ABSL road-DO-FUTII-2SG
'What will you send to Gela.' (Harris 2002: 118)

2.2.3 Allomorphy

The subject agreement markers exhibit systematic allomorphic variation, distinguishing between four sets of markers (cf. Table 1), including the question set in (13), the possession set in (14), and the inversion set in (15).

Question marking only affects a single form in the paradigm (3SG), and is arguably conditioned syntactically by sentence mood.

- (13) a. okt'omber-a evaxt'-**a** tay-sa?
Okt'omber-DAT when-3.SG go-PRES
'When does she go to Okt'omber?'
b. okt'omber-a evaxt'-**q'un** tay-sa?
Okt'omber-DAT when-3.PL go-PRES
'When do they go to Okt'omber?' (Harris 2002: 30)

³We follow the nomenclature of TAM used by Harris (2002) who in turn builds on the analysis in Schulze (1982), without, however, adopting his nomenclature. Note that Subjunctive I/II are distinct from the Subjunctive. The Subjunctive I/II both build on the future stem, with the Subjunctive I corresponding to the Optative in Schulze (1982). The (Particle) Subjunctive, by contrast, uses the marker -q'a (see Section 2.3.1) and builds on the preterite (=aorist) stem. It corresponds to the adhortative mood in Schulze (1982).

Choice of the possessive set, by contrast, is found in two environments, namely the possessive predicative construction (14a) and with specific verbs, such as *köfil* ‘like’, that select for the possessive marker. At least the latter case must be considered lexical in nature. I.e., even when attached to a non-verbal host, as in (14b), choice of the possessive variant of the marker depends on the verb.

- (14) a. vi günäh gölö-**vi**
 you.SG.GEN sin.ABS many-POSS2SG
 ‘You have many sins.’
 b. udin məy-**bes** köfil-esa
 Udi song-POSS1SG like-PRES
 ‘I like Udi songs.’ (Harris 2002: 29)

The third allomorphic alternation concerns the inverse set. According to Harris (1984), this set is found with a class of verbs that is characterised by its particular case assignment properties, i.e. alternate dative case subjects. However, as shown by the contrasts in (15), choice of the inverse set depends entirely on the verb’s class membership, i.e. case marking potential, not the actual subject case being assigned: i.e., irrespective of whether the subject is in the dative, as the subject pronoun *za* in (15a), or in the ergative, as the subject pronoun *zu* in (15b), the subject agreement marker on the verb is taken from the inverse set, surfacing as *-za* on the verb. Thus, this allomorphy must count as a lexical property as well, depending ultimately on the membership of the verb in a lexically defined class exhibiting exceptional case assignment potential.

- (15) a. za a-**za**-k’-sa šel läzätt’u pak.
 me.DAT see₁-1SG.INV-see₂-PRES good pretty garden.ABS
 ‘I see a good, pretty garden.’
 b. zu a-**za**-k’-sa šel läzätt’u pak.
 I.ERG see₁-1SG.INV-see₂-PRES good pretty garden.ABS
 ‘I see a good, pretty garden.’ (Harris 1984: 247)

Crucially, this allomorphy is observed whether the markers are attached to verbs, predicates or focused constituents, showing that the selection of allomorphic forms is driven by a combination of syntactic and morphological constraints.

2.3 Interaction with other clitics

2.3.1 Subjunctive *-q’a*

The TAM marker *-q’a* (the adhortative in Schulze 1982) combines with the aorist I suffix to express the subjunctive.⁴ Although it encodes verbal properties, Harris

⁴In the examples, intervening markers are set in italics, to distinguish them from the person markers, which are set in bold face throughout.

(2002) argues that *-q'a* is a postlexically attached clitic, mainly because of its promiscuous attachment to a variety of hosts: not only verbs, but also nouns, adverbs, adjectives and pronouns, as shown in (16). Since the subject agreement marker appears always immediately to the right of *-q'a*, Harris (2002) takes this distributional fact to support the clitic status of the subject agreement marker as well.

- (16) a. c'oro-b-en buɣ-buɣ če-r-i-*q'a-n*
 drain-DO-HORT vapor.ABS out-R-AOR-I-SUBJV-3SG
 'Let us drain it, that the vapor may go out.'
- b. murada-*q'a-q'un* ba-p-i
 goal-SUBJV-3PL fill-LV-AOR-I
 'May they attain their goal.'
- c. or-*q'a-n* bak-i?
 because-SUBJV-3SG become-AOR-I
 'How could that happen?'
- d. šono šelgär žal-*q'a-n-p-i* č'oč'o-*q'a-n* bak-i
 it.ABS well hot-SUBJV-3SG-SAY-AOR-I red-SUBJV-3SG become-AOR-I
 'It should be well heated, [it] should become red [i.e. brown].'
- e. šuxo-*q'a-va* q'i-b-i
 who.ABL-SUBJV-INV2SG fear-DO-AOR-I
 'Who should you be afraid of?' (Harris 2002: 103–104)

A closer look, however, shows that the behaviour of *-q'a* does not provide independent support for Criterion F. The crucial point is that the environments in which the subject agreement marker appears are exactly the environments in which *-q'a* appears (modulo the trapping cases, which are conditioned on a different TAM category). When the two markers co-occur, their distribution is fully parallel: if *-q'a* attaches to the verb (whether infixes or suffixes), the subject agreement marker also appears on the verb; if *-q'a* is hosted by a focused constituent, as in (16b) or (16e), the subject agreement marker likewise surfaces on that constituent.

The only safe conclusion regarding Criterion F that one can draw from the placement of *-q'a* between the PM and the host is that both markers probably enjoy the same status: if *-q'a* is a clitic, then the PM must be a clitic too. Conversely, if the PM is an affix, then *-q'a* must be an affix too. Given that the distribution of *-q'a* and the PM is so similar, we cannot tell from Criterion F alone what their status is. However, both the subject agreement marker and the adhortative marker *-q'a* express properties of the verb's paradigm, namely agreement and mood, respectively. The most natural conclusion therefore is that the subject agreement marker and *-q'a* have the same hybrid profile: they may be morphologically integrated into the verbal host, but they can also be attached to the edge of a phrasal host.

2.3.2 Discourse marker *-al*

The discourse marker *-al* ‘and/too’ attaches to an exceptionally wide range of hosts, combining with verbs, nouns, adverbs, adjectives, pronouns and even postpositions. While the exact function is sometimes difficult to pin down (Schulze 1982: 261), two main uses appear common, as indicated by the translation: use as a marker of sentential coordination (=‘and’; cf. (17a,b,d)) and the discourse marker (=‘too/also’; cf. (17c)).

Sentential coordinating in Udi is asyndetic (Schulze 1982). The marker *-al*, however, can be used to optionally mark such coordinations, as in (17a,b,d). When used in this function, the marker appears in second position, i.e. as a suffix/infix on the first constituent of the clause.

In (17c), by contrast, *-al* is a discourse marker on the subject, roughly equivalent with English ‘too’. In this function, it seems that *-al* attaches to the right edge of the phrase it modifies.

Regardless of these minor differences, the properties this marker expresses and the distribution of the marker suggest status of a phrasal affix.

- (17) a. pirog-ax čī-ne-č-er-i la-*al*-**le**-x-i
 pie-DAT out-3SG-CARRY-R-AOR_I put₁-AND-3SG-put₂-AOR
 ‘She took the pie out and put [it under her arm].’
 b. ... paščay-un yar-en-*al* zap-i lay-**ne**-sča
 ... king-GEN boy-ERG-AND pull-PTCPL up-3SG-CARRY.PRES
 ‘... and the king’s son, pulling, brings [them] up.’
 c. ek’a-te giorgi-n-en a-**t’u**-k’e muzeum-a,
 what-REL Giorgi-OBL-ERG see₁-INV3SG-see₂-AOR_{II} museum-DAT
 zu-*al* beš a-**za**-k’e
 I.ERG-AND before see₁-INV1SG-see₂-AOR_{II}
 ‘What Giorgi saw in the museum, I too had seen earlier.’
 d. šuk’ala-*al* te-t’u bak-o p-es te ...
 no one-AND NEG-INV3SG can-FUT say-INF that ...
 ‘No one can say that ...’ (Harris 2002: 102–103)

When *-al* and a subject agreement marker co-occur on the same host, the *-al* always intervenes between the host and the subject agreement marker, both when the host is a verb, as in (17), and when the subject agreement marker appears away from the verb, as in (18).

- (18) abaz-in-*al*-**le** aš-b-esa
 [coin]-INST-AND-3SG business-DO-PRES
 ‘And he works for an abaza [twenty kopeks].’ (Harris 2002: 103)

Harris (2002) therefore takes the bound discourse marker *-al* to be a postlexical clitic. Since the subject agreement markers surface peripheral to *-al*, they would

be postlexical clitics as well, according to Criterion F. However, that status appears incompatible with the broad evidence for the morphological status of the subject agreement markers we have seen above.

Fortunately, phrasal affixation offers a perspective on the syntactic distribution of the marker *-al* which is fully compatible with morphological attachment. Let us briefly consider the classical case of phrasal affixation, namely the English possessive: it (i) expresses a property of a possessive phrase, and (ii) is realised on the word at the right edge of that phrase, but (iii) is subject to haplology with morphological marking of plural and verb agreement.

We argue that the behaviour of the bound conjunction *-al* is highly similar to this classical case of phrasal affixation: *-al* (i) expresses a phrasal property on (ii) the right edge of the left phrase, but (iii) displays host-specific morphotactics, i.e., infixation with verbal hosts. We shall show, as part of our formal analysis below, how this perspective of *-al* as a phrasal affix integrates neatly with an account of the subject agreement markers as hybrid lexical and phrasal affixes.

3 Previous approaches

3.1 Harris (2002)

Harris (2002) advocates for an analysis of Udi subject agreement markers as endoclititics, proposing an Optimality-Theoretic (OT) account based on alignment constraints. This analysis treats subject agreement markers as syntactic clitics that align to specific morphological positions within the word. However, this approach faces several challenges. First, the assumption of endoclitisis is highly marked cross-linguistically, as it involves both infixation — a rare and dispreferred morphological process — and a violation of lexical integrity (Lapointe 1981). Second, the syntactic status of Udi subject agreement markers remains ambiguous, as they do not behave strictly like syntactic clitics nor entirely as affixes. Finally, Harris' alignment-based analysis lacks modularity, as it necessitates full access to both syntactic and morphological structures, a requirement that complicates the theoretical separation between syntax and morphology. These issues raise concerns about whether an alignment-based approach adequately accounts for the distribution and behaviour of Udi subject agreement markers or whether an alternative morphological or hybrid account is needed.

3.2 Mixed clitics (Luís 2009, Spencer & Luis 2012)

Luís (2009), Spencer & Luis (2012) introduce the concept of *mixed clitics*, proposing that clitics can function as lexical affixes when morphologically integrated into the verb, and as phrasal affixes when attached to a separate constituent at the phrasal level. This dual behaviour aligns with the distribution of Udi subject agreement markers: (i) Udi subject agreement markers exhibit complementary distribution between verb-internal (infixal) and external (phrasal) realisation, indicating a complex syntax-morphology interaction; (ii) trapping effects in specific TAM contexts enforce fixed

positioning, overriding infixation or external attachment; and (iii) allomorphy applies to both lexical and phrasal variants, suggesting that their selection is controlled within the lexical entry of the verb rather than purely by syntax.

3.3 Salehi & Koenig (2023)

In a recent paper, Salehi & Koenig (2023) discuss the attachment properties of mobile person markers in Sorani Kurdish, which display an alternation between lexical affixation to or even into the verb (19), and attachment/cliticisation to a complement of the verb (20).

- (19) a. *minał-ekan hermê-kan beş e-ke-n*
 kids-DEF.PL pear-DEF.PL share IPFV-do.PRS-3PL
 ‘The kids share the pears’
 b. *(ême) e-man-kêşan-n*
 we IPFV-1PL:A-pull.PST-3PL:P
 ‘We were pulling them.’ (Salehi & Koenig 2023)
- (20) *heřmê-ek-an-yan beş kird*
 pear-DEF-PL-3PL:A share do
 ‘They shared the pears.’ (Mohammadirad 2020)

Salehi & Koenig (2023) propose what is essentially a mixed clitic analysis, assuming a morphological treatment of verb-internal realisation (Bonami & Crysmann 2013). Phrasal attachment to the complement, by contrast, is captured using edge inflection (Miller & Halpern 1993, Crysmann 2010). In order to capture the complementary distribution between lexical and phrasal affixation and to relate the phrasal affix to the verb whose agreement properties it expresses, they assume that the morphosyntactic property of agreement is delegated by selecting, via COMPS, a phrasal complement with agreement information as a trigger feature. Following the edge inflection theory of phrasal affixation by Miller & Halpern (1993) (and its HPSG implementation in Crysmann 2010), a corresponding marking feature will be percolated down the right edge.

4 Analysis

In this section, we lay out an analysis that captures the hybrid status of the Udi subject agreement marker within HPSG, a constraint-based lexicalist framework. The empirical challenge extends beyond the subject agreement marker itself, given the interaction we observed between the subject agreement marker and *-q’a* and *-al*.

Both the subject agreement marker and the subjunctive marker *-q’a* encode genuinely verbal properties and are typically realised on the verb, where they show tight morphological integration. In focused contexts, both markers surface instead as phrasal affixes, which we understand in terms of edge inflection in the sense of Miller

& Halpern (1993): verbal features are delegated to a focused phrase to be realised at the left edge of a focused constituent.

The conjunction marker *-al*, by contrast, expresses a phrasal property rather than a property of the verbal paradigm. It is percolated to the left edge of the marked phrase and realised either as a suffix or as an infix on the leftmost element of that phrase, in a manner similar to classical cases such as Latin *-que*.

Against this background, the goal of the analysis is to formalise, within HPSG, how verbal inflectional properties (such as subject agreement and the subjunctive) can alternate between lexical affixation on the verb and edge realisation on a focused constituent, and how phrasal markers like *-al* likewise follow a pattern of edge realisation.

4.1 Phrasal affixation in HPSG

Within HPSG, the treatment of phrasal affixes, i.e., bound markers that morphologically integrate with a host word that in turn is determined syntactically, follows quite closely previous proposals within GPSG: as pioneered by Zwicky (1987) and later refined by Miller & Halpern (1993), phrasal affixes, such as the English possessive, are analysed as edge inflection, whereby a featural property of a phrase, e.g. the possessive pre-nominal NP, is inflected for possessive not on the head word, but rather on the rightmost word within that phrase.

(21) [a friend of mine]'s lemur

(22) [Anyone who cheats at roulette]'s hands will be hacked off.

Since the phrasal property is expressed as inflection of a terminal, i.e. lexical element, we expect interaction with other morphological processes: for the English possessive, this is illustrated by the haplology between possessive *'s* and the exponent *-s* on plural nouns or third singular verbs.

(23) the card players' hats

(24) Anyone who snores' mouth will be taped shut.

The GPSG approach, as formalised in Miller & Halpern (1993), uses pairs of features: a trigger feature, that is assigned to the phrase that bears the property, and a corresponding marking feature that serves to percolate the property down to a terminal element at the edge, where it will be expressed by morphology. The distinction between trigger and marking features was motivated by the observation that these features may have different linearisation properties: e.g., the *POSS* trigger feature is assigned to a left NP daughter of NP, while its corresponding *PM* marking feature is found on every daughter at the right edge of the possessive NP, i.e. exclusively on right daughters. Furthermore, the two types of features help distinguish between assignment of phrasal properties and the more technical aspect of edge percolation for means of expression.

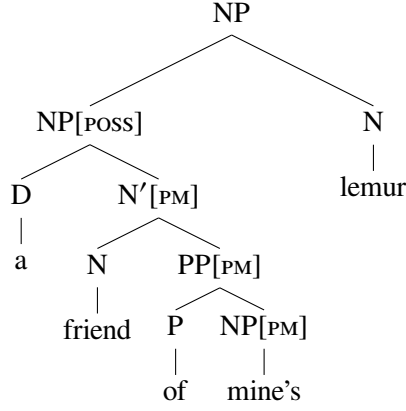


Figure 1: Phrasal affixation as edge inflection

In HPSG, edge feature percolation was first introduced by Tseng (2002) and later expanded in Kupść & Tseng (2005), permitting percolation of both trigger and marking features. In their analysis of Sorani Kurdish mixed clitics, Salehi & Koenig (2023) adopt the more restrictive version formulated in Cysmann (2010), a version which stays very close to the original formulation in Miller & Halpern (1993). Essentially, trigger feature (TRIGL and TRIGR) are asserted by phrase structure rule, differentiating between features that percolate along the left vs. the right edge, using the corresponding marking features (MRKL and MRKR). Percolation to a terminal element is done by the following constraints:

$$(25) \text{ phrase} \rightarrow \left[\begin{array}{l} \text{TRIG} \left[\begin{array}{l} \text{L} \left[\overline{tl} \right] \\ \text{R} \left[\overline{tr} \right] \end{array} \right] \\ \text{MRK} \left[\begin{array}{l} \text{L} \left[\overline{ml} \right] \\ \text{R} \left[\overline{mr} \right] \end{array} \right] \\ \text{DTRS} \left\langle \left[\text{MRK} \left[\begin{array}{l} \text{L} \left[\overline{tl} \oplus \overline{ml} \right] \\ \text{R} \left[\langle \rangle \right] \end{array} \right] \right], \dots \right\rangle \wedge \left\langle \dots, \left[\text{MRK} \left[\begin{array}{l} \text{L} \left[\langle \rangle \right] \\ \text{R} \left[\overline{tr} \oplus \overline{mr} \right] \end{array} \right] \right] \right\rangle \end{array} \right]$$

4.2 Focus position

In our analysis of Udi mixed clitics, we shall build on the recent proposal by Salehi & Koenig (2023) regarding mixed clitics in Sorani.

The main difference between phrasal attachment in Udi and Sorani lies with the properties that characterise the target or host for phrasal affixation. In Sorani, the host phrase corresponds to the least oblique non-subject complement. In Udi, by contrast, the host is defined in terms of its (grammaticalised) information-structural status, which happens to have very strict surface syntactic consequences, i.e. left-adjacency to the verb.

We therefore suggest to capture the strict ordering of the focused constituent by promoting it to the initial position on COMPS. For complements, this will be achieved by permutation of the COMPS list, as stated by the lexical rule in (26).

$$(26) \left[\begin{array}{cc} \text{HEAD} & \text{verb} \\ \text{COMPS} & \boxed{1} \oplus \langle \boxed{f} [\text{FOC} \quad +] \rangle \oplus \boxed{2} \end{array} \right] \mapsto \left[\begin{array}{cc} \text{COMPS} & \langle \boxed{f} \rangle \oplus \boxed{1} \oplus \boxed{2} \end{array} \right]$$

Since the verb-adjacent focus position is not reserved for complements alone, but equally available to modifiers, we suggest that focused modifiers undergo type shifting. To this end, we propose an adjunct-to-complements lexical rule (cf. Bouma et al. 2001), as given in (27) below.

$$(27) \left[\begin{array}{cc} \text{HEAD} & \boxed{h} \text{ verb} \\ \text{COMPS} & \boxed{c} \end{array} \right] \mapsto \left[\begin{array}{cc} \text{COMPS} & \left\langle \begin{array}{c} \text{FOC} \quad + \\ \text{MOD} \quad \left[\text{L|CAT|HEAD} \quad \boxed{h} \right] \end{array} \right\rangle \oplus \boxed{c} \end{array} \right]$$

Representing grammaticalised focus on COMPS not only captures its strict surface-syntactic position, but it also makes the focused constituent available as the target for external realisation (=delegated inflection) of Udi subject agreement.

$$(28) \left[\begin{array}{cc} \text{HEAD} & \text{verb} \\ \text{COMPS} & \langle [\text{FOC} \quad +], \dots \rangle \end{array} \right] \rightarrow \left[\begin{array}{cc} \text{INFL|DLS} & \boxed{a} \\ \text{COMPS} & \langle [\text{TRIG|R} \quad \{\boxed{a}, \dots\}], \dots \rangle \end{array} \right]$$

To effect delegation from inflectional morphology to a syntactic complement, we shall make use of the DLS feature proposed by Aguila-Multner & Crysmann (2025) to address inflectional periphrasis within Information-based Morphology (=IbM; Crysmann & Bonami 2016, Crysmann 2021): they translate the PFM theory of periphrasis by Bonami (2015) to IbM, proposing that individual rules specify, via DEL, the morphosyntactic properties they delegate. Delegated properties from all rules are assembled (=unioned) on the word's DLS feature, which serves as the primary morpho-syntactic interface. Furthermore, the global set of delegated properties is made visible to individual rules.

As detailed by the implicational constraint in (28), we lexically equate the set of delegated morphosyntactic properties of a word (DLS) with the trigger feature of the focus complement.

Conversely, if there is no grammaticalised focus complement, delegation of morphosyntactic properties will be illicit, as enforced by the implicational constraint in (29).

$$(29) \left[\begin{array}{cc} \text{HEAD} & \text{verb} \\ \text{COMPS} & \text{list}([\text{FOC} \quad -]) \end{array} \right] \rightarrow [\text{INFL|DLS} \quad \{\}]$$

In the following two subsections we shall detail how the current approach will capture the different placement properties.

4.3 Phrasal affix

To start with, we shall consider phrasal affixation on a focused constituent. As shown by the sample tree in Figure 2, the verb takes a focused complement. According to the lexical interface constraint in (28), the delegated morphosyntactic properties are imposed as a trigger feature on a COMPS valency.

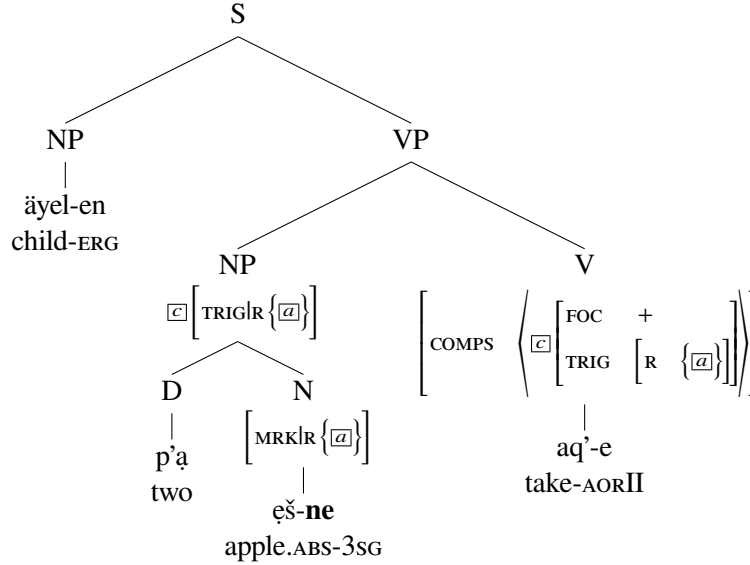


Figure 2: Phrasal affix (cf. example (3c))

The (right edge) trigger feature on the focused complement gives rise to a corresponding marker feature by which the phrasal inflectional property will percolate down along the right edge to a lexical element. This will capture both the phrasal distribution of the agreement marker and the broad set of host categories.

4.4 Lexical affix

Lexical affixation can arise in one of two situations: either in sentences with broad focus or in case of morphological trapping. We shall discuss both cases in turn.

Starting with broad focus or predicate focus, we see that the conditions for delegation to phrasal affixation (28) are just not met. Moreover, in the absence of a focused complement, (29) disallows delegation, thereby enforcing lexical realisation.

The other case to be considered is morphological trapping, which occurs with specific TAM categories. As observed above, such trapping is independent of the availability of a narrow focus phrasal host. What is more, the TAM categories that cause trapping also display exceptional morphotactics, preceding the agreement marker rather than following it.

We suggest that these particular TAM categories preempt the rule of delegation, by requiring that the word's global DLS feature be the empty set. Indeed the usefulness

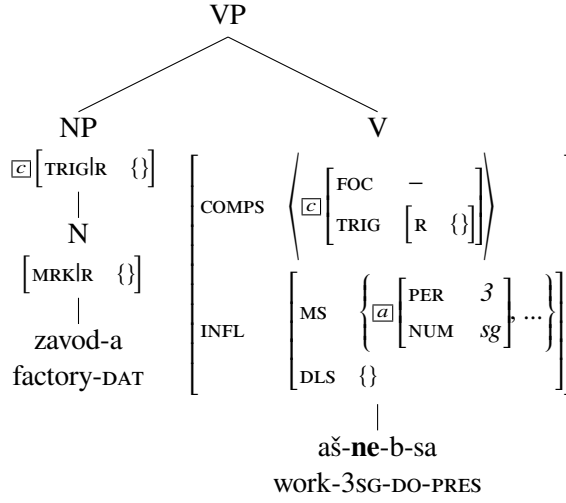


Figure 3: Lexical affix (broad focus; cf. example (7a))

of having individual rules impose global constraints has already been demonstrated for French (Aguila-Multner & Crysmann 2025). As a net result, the morphological feature responsible for delegation **DLS** will end up being the empty set in these cases, as shown in Figure 4. Thus, in the case of morphological trapping, the trigger feature will end up sharing an empty set of delegated properties.

4.5 Interaction

In the remainder of this paper, we shall briefly illustrate how the hybrid agreement marker interacts with other hybrid and phrasal affixes, namely the hybrid lexical/phrasal adhortative marker *-q'a*, as well as the marker *-al*, which we take to be a phrasal affix.

4.5.1 Subjunctive *-q'a*

The first marker we discuss is the subjunctive/adhortative marker. As detailed in Section 2.3.1 above, this marker, when present, always appears adjacent to the subject agreement marker.

Given the similarity of their distribution across verbal and non-verbal hosts, we shall treat adhortative *-q'a* as a hybrid lexical/phrasal affix as well. If realised on the verb, whose properties it expresses, realisation is morphotactically fully integrated, including infixation.

As shown in Figure 5, the subjunctive/adhortative marker “originates” as part of the verb’s inflectional paradigm, but its realisation may get delegated (via **DLS**) to a focused constituent, on a par with agreement marking.

Subject agreement and subjunctive marking are assigned to a right **TRIG** feature on the focused complement, percolating down to its right edge via the corresponding **MRK**

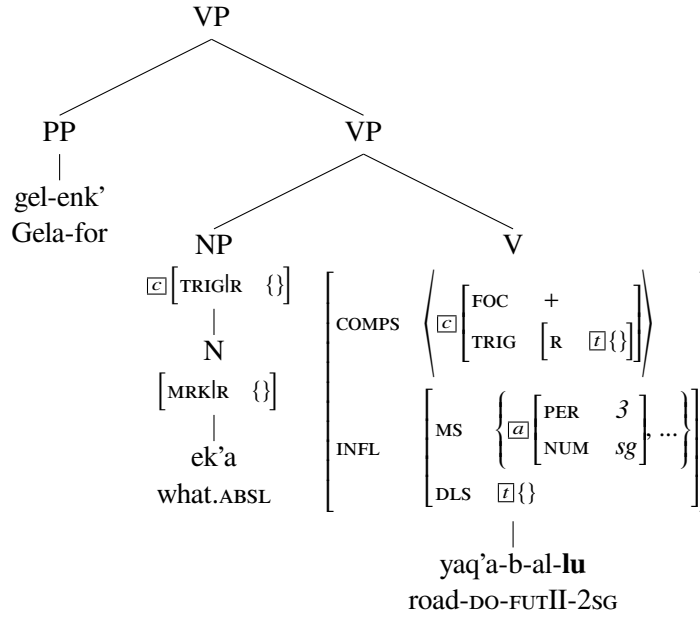


Figure 4: Lexical affix (narrow focus; cf. example (12))

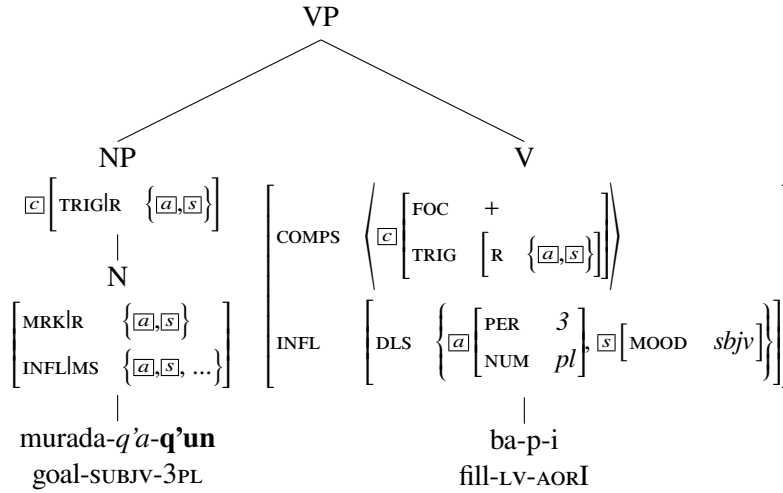


Figure 5: Subjunctive (narrow focus; cf. example (16b))

feature. On the verb, the value of the MRK feature is amalgamated with that word's morphosyntactic property set, making the MRK features available for morphological expression.

4.5.2 Marker *-al*

Turning now to the marker *-al*, recall from Section 2.3.2 that this marker has all the hallmarks of a phrasal affix: it expresses a property of a phrasal constituent, and its host word is selected by syntax as well. We assume that the marker attaches to the right of the first constituent, as in e.g. (17b). This can be achieved by requiring the left daughter of the conjunct to have a corresponding element (e.g. *c-cnj*) on the TRIGLR.⁵ Via MRKLR, it will percolate down along that constituent's right edge to the host word.

Once the host has been determined, exponence ensues in accordance with the morphotactic constraints of the host word's morphology, yielding the two following outcomes: (i) infixation and suffixation for verbs; (ii) suffixation for all other hosts.

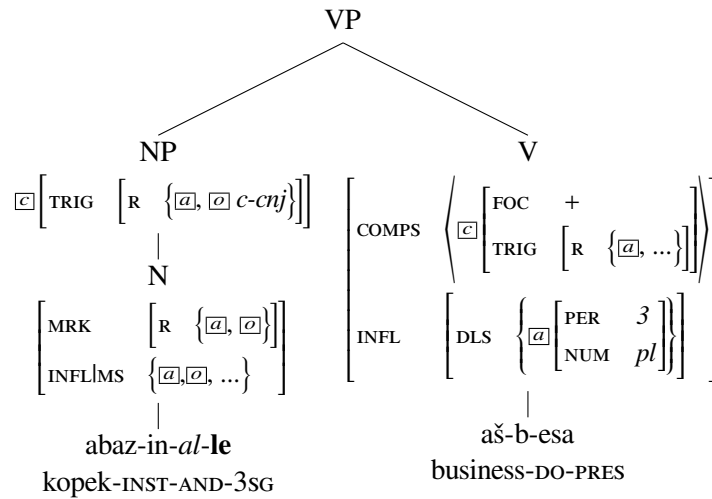


Figure 6: Coordinating conjunction and subject agreement marker (narrow focus; cf. ex. (18))

Given that the marker *-al* is a phrasal affix, but agreement is hybrid, three different situations can arise: if (i) the first constituent of the conjunct happens to be a focused constituent, as in (18), both agreement and coordination marking will end up on the right edge of the focused constituent. If (ii) the focused constituent is not the leftmost constituent, as in (17d), *-al* will attach to the first constituent (*šuk'ala* 'no one') and the agreement markers to the focused constituent (negation *te*). Finally, if (iii) the verb itself is conjunct-initial, as in (17a), both markers will be found on the verb. Since edge features are amalgamated with the host word's morphosyntactic property set, realisational morphology will have full control over exponence and morphotactics, including infixation.

⁵For attachment to the first word, one would specify a TRIGL feature on that daughter.

5 Conclusion

In this paper, we observe that Udi subject agreement markers show properties of both lexical affixes and phrasal affixes and can therefore be understood in terms of mixed clitics (Luís 2009, Spencer & Luis 2012). We argue that the “endocclisis” approaches by Harris (2002) and Anderson (2005), by contrast, downplay the affixal properties and fail to capture allomorphic variation. Most importantly, they require simultaneous access to syntactic and internal morphological structure, an assumption that greatly weakens Lexical Integrity (Lapointe 1981).

Building on a previous proposals for Sorani Kurdish (Salehi & Koenig 2023), we propose a formal approach to mixed clitics that establishes a narrow interface connecting lexical and phrasal affixation and thus maintains a modular view of morphology and syntax. Finally, using recent developments within IbM, the present approach connects mixed clitics to other cases of morphosyntactic delegation, such as periphrasis (Bonami 2015, Aguila-Multner & Crysmann 2025).

References

- Aguila-Multner, Gabrielle & Berthold Crysmann. 2025. The morphology of trapping and climbing in French causatives. In Stefan Müller & Rui Chaves (eds.), *Proceedings of the 31st International Conference on Head-Driven Phrase Structure Grammar, Palacky University, Olomouc, Czechia*. Frankfurt/Main: University Library. DOI: 10.21248/hpsg.2024.1.
- Anderson, Stephen R. 2005. *Aspects of the theory of clitics* (Oxford Studies in Theoretical Linguistics). Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199279906.001.0001.
- Bonami, Olivier. 2015. Periphrasis as collocation. *Morphology* 25. 63–110. DOI: 10.1007/s11525-015-9254-3.
- 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. DOI: 10.21248/hpsg.2013.2.
- Bouma, Gosse, Rob Malouf & Ivan A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* 19(1). 1–65. DOI: 10.1023/A:1006473306778.
- Crysmann, Berthold. 2010. Discontinuous negation in Hausa. In Stefan Müller (ed.), *Proceedings of the workshop on Morphology and Formal Grammar at the 17th International Conference on Head-Driven Phrase Structure Grammar*, 269–287. CSLI Publications. DOI: 10.21248/hpsg.2010.15.
- Crysmann, Berthold. 2021. Morphology. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-driven Phrase Structure Grammar: The handbook* (Empirically Oriented Theoretical Morphology and Syntax 9), chap. 21, 947–999. Berlin: Language Science Press. DOI: 10.5281/zenodo.13644963.

- Crysmann, Berthold & Olivier Bonami. 2016. Variable morphotactics in Information-based Morphology. *Journal of Linguistics* 52(2). 311–374. DOI: 10.1017/S0022226715000018.
- Halpern, Aaron. 1995. *On the placement and morphology of clitics* (Dissertations in Linguistics). Stanford: CSLI Publications.
- Harris, Alice C. 1984. Case marking, verb agreement, and inversion in Udi. In David M. Perlmutter & Carol G. Rosen (eds.), *Studies in Relational Grammar* 2, 243–258. Chicago: University of Chicago Press.
- Harris, Alice C. 1996. Focus in Udi. In Howard I. Aronson (ed.), *NSL 8: Linguistic Studies in the Non-Slavic Languages of the Commonwealth of Independent States and the Baltic Republics*, 201–220. Chicago: Chicago Linguistic Society.
- Harris, Alice C. 2000. Where in the word is the Udi clitic? *Language* 76(3). 593–616. DOI: 10.2307/417136.
- Harris, Alice C. 2002. *Endoclititics and the origins of Udi morphosyntax*. Oxford: Oxford UP. DOI: 10.1093/oso/9780199246335.001.0001.
- Kupść, Anna & Jesse Tseng. 2005. A new HPSG approach to Polish auxiliary constructions. In Stefan Müller (ed.), *Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar, Department of Informatics, University of Lisbon, August 23-24, 2005*, 253–273. Stanford: CSLI Publications. DOI: 10.21248/hpsg.2005.14.
- Lapointe, Steven. 1981. A lexical analysis of the English auxiliary verb system. In Teun Hoekstra, Harry van der Hulst & Michael Moortgat (eds.), *Lexical grammar*, 215–254. Dordrecht: Foris.
- Luís, Ana R. 2009. Patterns of clitic placement: evidence from ‘mixed’ clitic systems. In Patience Epps & Alexandre Arkhipov (eds.), *New challenges in typology: transcending the borders and refining the distinctions*, 11–34. Berlin, New York: Mouton De Gruyter. DOI: 10.1515/9783110219067.1.11.
- Miller, Philip. 1992. *Clitics and constituents in phrase structure grammar* (Outstanding Dissertations in Linguistics). New York: Garland.
- Miller, Philip & Aaron Halpern. 1993. English possessives and the syntax of morphological features. In *FLSM III: Papers from the Third Annual Meeting of the Formal Linguistics Society of Mid-America (1992)*, Northwestern University. IULC Publications.
- Mohammadirad, Masoud. 2020. *Pronominal clitics in Western Iranian languages: description, mapping, and typological implications*. Paris: Université Sorbonne Nouvelle (Paris 3). (Doctoral dissertation).
- 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*, 100–120. Frankfurt/Main: University Library. DOI: 10.21248/hpsg.2023.6.
- Schulze, Wolfgang. 1982. *Die Sprache der Uden in Nord-Azerbajdžan*. Wiesbaden: Harrassowitz.

- Spencer, Andrew & Ana R. Luis. 2012. *Clitics: An introduction* (Cambridge Textbooks in Linguistics). Cambridge University Press. DOI: 10.1017/CBO9781139033763.
- Tseng, Jesse L. 2002. Edge features and French liaison. In Jong-Bok Kim & Stephen Wechsler (eds.), *Proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar*, 313–333. Stanford, CA: CSLI Publications. DOI: 10.21248/hpsg.2002.16.
- Zwicky, Arnold. 1987. Suppressing the Zs. *Journal of Linguistics* 23(1). 133–148. DOI: 10.1017/S0022226700011063.