

Mixed clitics in Udi: the lexical verb as interface

Berthold Crysmann (CNRS) and Ana R. Luís (U Coimbra)

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 markers in Udi (Harris, 2002) present a particularly challenging case of *mixed clitics* (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.

- (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)
- (2) met'in-al burim sa yärvägä q'əzəl-le
this.ERG-AND apparently one bag.ABS gold.ABS-3SG
aq'sa
take
'She also apparently takes a BAG OF GOLD.' (Harris, 2002)

On closer inspection, complementary distribution between (1) and (2) is controlled by the verb: verb-internal placement constitutes the default (focus-neutral contexts or 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 *endoclis* approach of Harris (2002) essentially plays down the morphological side of the problem, the *mixed clitic* approach (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.

Analyzing 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 analyze 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 the syntactic distribution of Udi person markers, account for their morphological properties and capture the complementary distribution between focus marking and verb-internal morphological realisation.

2 Data

This section examines the syntactic and morphological properties of Udi clitics, particularly focusing on endoclis and 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, p. 95)

2.1.2 Focused constituent

In focus constructions, 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 iſq'armux mano äizi-**q'un** karkesa?
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 person markers also show properties typically associated with affixes, further blurring the line between morphology and syntax.

Infixation With complex verb stems, Udi person 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 aſ-*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, p. 122)

With simple verbs, the person marker is typically infixed before the final consonant of the root (8-10). 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
king-UN boy-ERG much look₁-3SG-look₂-PRES
met'a-laxo
this.GEN-on
'The prince looks at this for a long time.' (Harris, 2002, p. 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'

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

Table 1: Udi person markers (Harris, 2002, p. 28)

- (10) a. bo-*ne*-x-sa
boil-3SG.INV-see₂-PRES
'(s)he boils [it]'
- b. box-*ne*-sa
box-3SG-PRES
'[it] boils' (Harris, 2000, p. 127)

Trapping In certain tenses (e.g., Future II, Subjunctive I and II, and Imperative), person markers must be realised as suffixes in final position, preempting both infixation (11) and attachment to a focused phrase (12):

- (11) ič turinlek'er-ax nut' tämiz-b-al-le
self's shoe-DAT NEG clean-DO-FUTII-3SG
'She won't clean her shoes.' (Harris, 2002, p. 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, p. 118)

Allomorphy The person 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 ga-PRES
'When does she go to Okt'omber?'
- b. okt'omber-a evaxt'-q'un tay-sa?
Okt'omber-DAT when-3.PL ga-PRES
'When do they go to Okt'omber?' (Harris, 2002, p. 30)

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.

- (14) a. vi günäh gölö-vi
you.SG.GEN sin.ABS many-POSS2SG
'You have many sins.'
- b. udin may-bes köfil-esa
Udi song-POSS1SG like
'I like Udi songs.' (Harris, 2002, p. 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. Thus, this allomorphy must count as a lexical property as well.

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

3 Previous approaches

3.1 Harris (2002)

Harris (2002) advocates for an analysis of Udi person markers as endoclititics, proposing an Optimality-Theoretic (OT) account based on alignment constraints. This analysis treats person 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 person 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 person markers or whether an alternative morphological or hybrid account is needed.

3.2 Spencer & Luis (2012)

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 person markers: (i) Udi person 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 (16), and attachment/cliticisation to a complement of the verb (17).

- (16) a. minal-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.’
 (17) 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 & Cysmann, 2013). Phrasal attachment to the complement, by contrast, is captured using edge inflection (Miller & Halpern, 1993; Cysmann, 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 Miller & Halpern (1993) (and its HPSG implementation in Cysmann 2010), a corresponding marking feature will be percolated down the right edge.

4 Analysis

4.1 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 in 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 (18).

$$(18) \left[\begin{array}{cc} \text{HEAD} & \text{verb} \\ \text{COMPS} & [1] \oplus \langle [f]_{\text{FOC}} + \rangle \oplus [2] \end{array} \right] \mapsto \left[\text{COMPS} \langle [f] \rangle \oplus [1] \oplus [2] \right]$$

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 (19) below.

$$(19) \left[\begin{array}{cc} \text{HEAD} & [h] \text{ verb} \\ \text{COMPS} & [c] \end{array} \right] \mapsto \left[\text{COMPS} \left\langle \begin{array}{c} \text{FOC} + \\ \text{MOD} [L|\text{CAT}|\text{HEAD} [h]] \end{array} \right\rangle \oplus [c] \right]$$

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.

$$(20) \left[\begin{array}{cc} \text{HEAD} & \text{verb} \\ \text{COMPS} & \langle [f]_{\text{FOC}} +, \dots \rangle \end{array} \right] \mapsto \left[\begin{array}{cc} \text{INFL|DLS} & [a] \\ \text{COMPS} & \langle [\text{TRIG|R} [a], \dots] \rangle \end{array} \right]$$

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 (Crysmann & Bonami, 2016; Crysmann, 2021). As detailed in (20), we lexically equate the set of delegated morphosyntactic properties of a word (DLS) with the trigger feature of the focus complement.

$$(21) \left[\begin{array}{cc} \text{HEAD} & \text{verb} \\ \text{COMPS} & \text{list}([f]_{\text{FOC}} -) \end{array} \right] \mapsto [\text{INFL|DLS } \{\}]$$

Conversely, if there is no grammaticalised focus complement, delegation of morphosyntactic properties will be illicit, as enforced by the constraint in (21).

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

4.2 Phrasal affix

To start with, we shall consider phrasal affixation on a focused constituent. As shown by the sample tree in Figure 1, the verb takes a focused complement. According to the lexical interface constraint in (20), the delegated morphosyntactic properties are imposed as a trigger feature on a COMPS valency.

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.

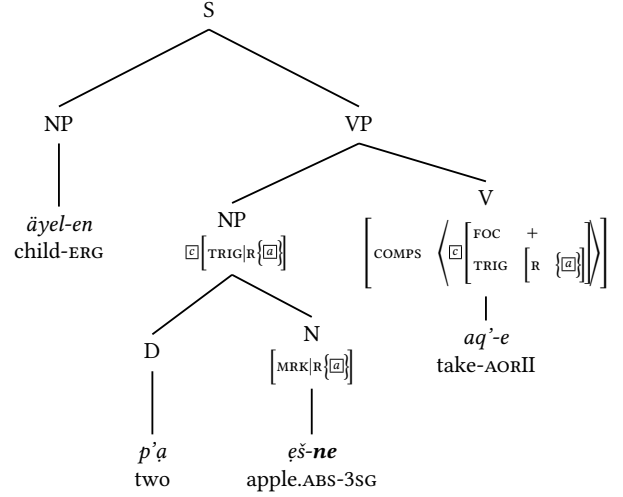


Figure 1: Phrasal affix

4.3 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 (20) are just not met. Moreover, in the absence of a focused complement, (21) disallows delegation, thereby enforcing lexical realisation.

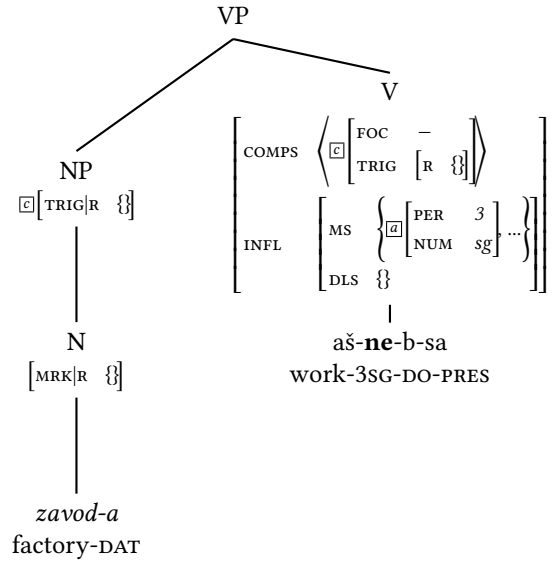


Figure 2: Lexical affix (broad focus)

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 effect exceptional morphotactics, yielding suffixation to the TAM marker instead of infixation.

We suggest that the morphotactically exceptional variants of the agreement realisation rules are conditioned on a particular TAM category, and therefore preempt not only the regular infixed variant, but also the rule of delegation.

As a net result, the morphological feature responsible for delegation DLS will end up to be the empty set in these cases, as shown in Figure 3. Thus, in the case of morphological trapping, the trigger feature will end up sharing an empty set of delegated properties.

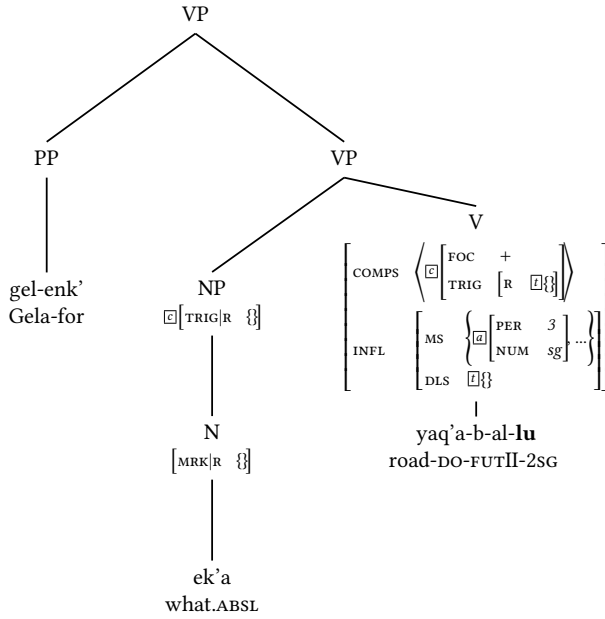


Figure 3: Lexical affix (narrow focus)

The morphology actually plays a rather active role in shaping the lexical phrasal alternation. We shall therefore use Information-based Morphology to provide a full morphological analysis of morphotactics and allomorphy that integrates with the lexical/phrasal alternation characteristic of Udi mixed clitics.

5 Conclusion

In this paper, we observe that Udi person markers show properties of both lexical affixes and phrasal affixes and can therefore be understood in terms of mixed clitics (Spencer & Luis, 2012). We argue that the “endoclitisis” 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.
- Anderson, Stephen R. 2005. *Aspects of the theory of clitics*. Oxford: Oxford University Press.
- Bonami, Olivier. 2015. Periphrasis as collocation. *Morphology* 25. 63–110.
- Bonami, Olivier & Berthold Crysmann. 2013. Morphotactics in an information-based model of realisational morphology. In Stefan Müller (ed.), *Proceedings of the 20th conference on head-driven phrase structure grammar, berlin, 26–29 august*, 27–47.
- Bouma, Gosse, Rob Malouf & Ivan A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* 19(1). 1–65.
- 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 (HPSG 2010)*, 269–287. CSLI Publications.
- 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.
- 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. 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. 593–616.
- Harris, Alice C. 2002. *Endoclititics and the origins of Udi morphosyntax*. Oxford: Oxford UP.
- 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.
- 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.
- Salehi, Ali & Jean-Pierre Koenig. 2023. Argument realization and argument referencing in Sorani 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.
- Spencer, Andrew & Ana R. Luis. 2012. *Clitics: An introduction* Cambridge Textbooks in Linguistics. Cambridge University Press. doi:10.1017/CBO9781139033763.