

The structure of the Kurdish noun phrase

Jan Strunk

December 15, 2003

1 Introduction

Kurdish is an Indo-Iranian language spoken in Armenia, Azerbaijan, Iran, Iraq, Syria, and Turkey (cf. The Ethnologue). It is divided into several dialects some of which are sometimes regarded as distinct languages. This paper proposes an analysis of the Kurdish noun phrase within the framework of Lexical Functional Grammar. Specifically, I will be looking at the noun phrase in the northern dialect of Kurdish called Kurmancî. The data I use in this paper are taken from interviews and recordings which were carried out at the University of Bochum, Germany, with a native speaker of the dialect of Muş, Turkey. As a second source, I cite sentences from Wurzel (1997), which is a Kurmancî introductory text book for German speakers.

2 General overview

The general linear structure of the nominal phrase in Kurmancî is as shown in figure (1).

(1) DEM Num/Qnt N Prn(OBL) AP , NP(OBL) PP

A noun phrase minimally consists of the head noun¹, cf. example (2).

(2) *gul sor=e.*
flower red=be.2/3.SG.PRES
'The flower is red.'

Other elements and phrases can appear either to the left or to the right of the head noun. Section 3 will deal with elements appearing to the left of the head noun. Section 4 discusses rightward modification which has interesting properties in Kurmancî.

3 Pre-nominal elements

There are two pre-nominal positions in Kurmancî. Demonstratives always occur leftmost in a nominal projection; cf. the examples in (3). I will assume in this paper that they represent the category determiner and that they head the nominal functional projection DP. One reason

¹There are some exceptions that will be covered later.

Following the practice in Bresnan (2001), I assume that regular personal pronouns, too, are of the category D. It still has to be determined whether pronouns can also act as a determiner to an embedded noun phrase. If they cannot, their PRED feature has to be non-optional in contrast to that of demonstratives, cf. the lexical entry for *wi* above.

4 Post-nominal modification

All possessor phrases (DPs), adjective phrases (APs), and prepositional phrases (PPs) that modify a noun have to follow it². Possessor phrases always have to be in the oblique case. They usually directly follow the head noun. While this position is obligatory for pronominal DPs, heavier possessor phrases can also be preceded by an AP. PPs can solely occur at the right edge of an NP. I have only ever encountered noun phrases with a single PP adjunct, so it is not clear to me whether multiple PPs can modify a single noun. In contrast to English, demonstratives and possessor phrases do not exclude each other, so that example (6) is perfectly grammatical. Determiners and possessors are thus neither positionally nor functionally in complementary distribution.

- (6) *wa jun-ên bir=ên wi*
 DEM.PL wife-EZ.PL brother=EZ.PL 3.SG.OBL
 'those her sisters-in-law'

A very interesting phenomenon in Kurmancî (and also in the other Kurdish dialects) is the so-called *ezafe* (or *izafet*) which I gloss as EZ in this paper³. It is a type of linker morpheme⁴ that has to appear between a modified noun and a post-nominal modifier: an AP, a DP, or even a PP. This morpheme agrees with the head noun in gender and number (the feminine singular is -a, masculine singular -e, and plural is -en). Consider the following examples:

- (7) a. *mal-a* [_{DP} *mîn*]
 house-EZ.FEM.SG 1.SG.OBL
 'my house'
- b. *mal-a* [_{AP} *sor*]
 house-EZ.FEM.SG red
 'the red house'
- c. *mal-a* [_{NP} *vi mer-e*]
 house-EZ.FEM.SG DEM.SG.OBL man-SG.OBL
 'this man's house'
- d. *mal-a* [_{PP} *sere çiya*]
 house-EZ.FEM.SG on mountain
 'the house on the mountain'

²It is not clear yet whether adjuncts and complements behave differently in any significant way. Further work should examine this possibility.

³This phenomenon also occurs in Farsi, cf. e.g. Kahnemuyipour (2000) for a recent work on this subject. A general description of the *ezafe* in different Kurdish dialects can be found in McKenzie (1961).

⁴In the remainder of this paper, I will argue that the *ezafe* marker is a suffix when it appears on the head noun and a clitic when it appears between two modifying phrases.

The corresponding NPs in (8) where the head noun is not marked with the ezafe are completely ungrammatical.

- (8) a. * *mal* [_{DP} *mın*]
house 1.SG.OBL
‘my house’
- b. * *mal* [_{AP} *sor*]
house red
‘the red house’
- c. * *mal* [_{NP} *vi* *mer-e*]
house DEM.SG.OBL man-SG.OBL
‘this man’s house’
- d. * *mal* [_{PP} *sere çiya*]
house on mountain
‘the house on the mountain’

Ezafe marking is, however, not restricted to the head noun. Every element that is followed by another modifier of the head noun shows an ezafe marker agreeing with the head noun.

- (9) a. *mal-a* [_{DP} *mın*]=*a* [_{AP} *sor*]
house-EZ.FEM.SG my=EZ.FEM.SG red
‘my red house’
- b. *mal-a* [_{AP} *sor*]=*a* [_{NP} *vi*]
house-EZ.FEM.SG red=EZ.FEM.SG DEM.SG.OBL
mer-e
man-SG.OBL
‘this man’s red house’

Basically, all post-nominal modifiers are marked with the ezafe morpheme⁵; only the right-most one does not carry the ezafe marking because no further modifier follows and accordingly no linker is needed.

So far, I have glossed the ezafe morpheme as a suffix on the head noun but as a clitic on the modifiers. I have followed Kurdish orthographic practice which treats the ezafe as part of the head noun but the ezafe morphemes between two modifiers as independent particles. Before I attempt an initial analysis, I would like to look at the constituency of the ezafe. Does it form a constituent with the preceding element or with the following or does it indeed lie in between the preceding and the following elements? If the ezafe really is a suffix on the head noun, then we obviously have to consider it as forming a (sublexical) constituent with the modified noun if we want to respect the principle of lexical integrity. However, other arguments lead to the conclusion that the ezafe (at least the one which occurs between two modifiers) is part neither of the preceding nor of the following constituent but really lies between and links the two of them. First, if we want to consider it as a suffix on a preceding

⁵There are some forms of apposition where ezafe marking does not occur. I will not deal with these constructions in this paper.

modifier we would have to regard it as a kind of phrasal affix (i.e. clitic) comparable perhaps to the English 's genitive marker because it always appears on the right edge of the preceding modifier regardless of what word class the preceding word belongs to. In (9a) e.g. it follows a pronoun, in (9b) an adjective and in principle it can follow any word which can be the right edge of a phrase which modifies the head noun. Second, as can be seen in the examples in (10), the possessor in a possessive construction can be modified itself but still the *ezafe* following the whole possessive has to agree with the head noun because it links it to another adjunct, an AP in this case. It is easier to model these agreement facts if we assume that the *ezafe* is not embedded inside an adjunct or other grammatical function.

- (10) a. *mal-a* *bıra=yê* *man=a* *piçûk*
house-EZ.FEM.SG brother=EZ.MASC.SG 1.SG.OBL=EZ.FEM.SG small
'my brother's small house'
- b. *mal-a* *bıra=yê* *man=ê* *piçûk*
house-EZ.FEM.SG brother=EZ.MASC.SG 1.SG.OBL=EZ.MASC.SG small
'my small brother's house'

(Wurzel 1997:26)

Third, the *ezafe* only ever occurs before a following dependent of the head noun, i.e. its occurrence depends on the following phrase. Moreover, there is a phenomenon called *absolute ezafe* by Wurzel (1997). This is a construction which is comparable to the English one-pronoun, e.g. "a green one", "these big ones", etc. It involves an *ezafe* morpheme and a following adjunct or possessor but no overt head noun, cf. example (11).

- (11) a. *ya* *man*
EZ.FEM.SG 1.SG.OBL
'mine' (feminine)
(Wurzel 1997:25)
- b. *nav-ê* *yê* *mezın*
name-EZ.MASC.SG EZ.MASC.SG big
'the big one's name'
(Wurzel 1997:31)

The last two arguments could be regarded as evidence for an analysis which assumes that the *ezafe* is part of the modifying phrase. However, I maintain that the *ezafe* morpheme is not part of the modifying phrase because it only occurs inside a noun phrase and never when an AP, DP or PP is used predicatively at least not without implying an "invisible" head noun. This e.g. is in contrast to Albanian which has a comparable linker morpheme (cf. Himmelmann 1997:chapter 5). In contrast to the Kurdish *ezafe*, this morpheme only appears with certain classes of adjectives and possessive pronouns and is thus lexically conditioned. Moreover, it also precedes adjectives when they are used predicatively, cf. example (12).

- (12) *vajz-a* *është* *e* *bukur*
 girl-DEF.NOM.SG.FEM be.3.SG.PRES LK.NOM.SG.FEM pretty
 'the girl is pretty'⁶
 (Himmelman 1997:chapter 5, example 23)

I will thus assume the following constituent analysis for the ezafe construction in the rest of this paper.

- (13) [_{NP} [_N X] EZ [_{ADJ/POSS} Y]]

4.1 First analysis

In order to propose a first draft of an LFG analysis, I will assume in this section that contrary to Kurdish orthographic practice, the ezafe on the head noun, too, is a simple clitic and not a suffix. I will propose an alternative, but more complicated analysis in section 4.2.

Based on the discussion above, I assume the following c-structure rules for the Kurmançî NP.

$$\begin{array}{l}
 N' \quad \rightarrow \quad (N) \quad (EZ \quad DP) \quad , \quad (EZ \quad AP) \\
 \quad \quad \quad \uparrow=\downarrow \quad \uparrow=\downarrow \quad (\uparrow \text{ POSS})=\downarrow \quad \uparrow=\downarrow \quad \downarrow \in (\uparrow \text{ ADJ}) \\
 \quad \quad \quad \quad \quad \quad (\downarrow \text{ CASE})=\text{OBL}
 \end{array}$$

$$\begin{array}{l}
 NP \quad \rightarrow \quad (N') \quad (EZ \quad PP) \\
 \quad \quad \quad \uparrow=\downarrow \quad \uparrow=\downarrow \quad \downarrow \in (\uparrow \text{ ADJ})
 \end{array}$$

These rules do not conform to the general x-bar scheme, cf. Bresnan (2001:chapter 6). One reason is that I annotate the ezafe morpheme so that its f-structure is unified with that of the head noun. The second reason is that although the APs and PPs which modify a noun are mostly assumed to be adjuncts (as I do here), they do not show Chomsky adjunction behavior in Kurdish, i.e. the head noun projects to a higher projection when it is modified by these phrases. This has the consequence that one noun can only be modified by at most one (possibly complex) AP and one PP⁷.

The reason I chose to construct such unusual rules is that it seems to be the case that whenever a noun is modified by more than one AP, these APs obligatorily have to be combined into a complex AP using a conjunction such as *û* (and). They cannot each be individually added to a head noun with their own accompanying ezafe marker⁸, cf. example (14).

⁶LK means linker. I have slightly changed the category abbreviations in order to keep them consistent with those used in the rest of this paper.

⁷If multiple PPs should be possible, the corresponding c-structure rule simply has to be changed to be non-projective, i.e. a typical adjunction rule.

⁸If it should turn out that multiple APs linked to the head noun are possible (maybe with different semantics), a more standard adjunction rule could be proposed.

- (14) a. *mal-a min=a kevin û şil*
house-EZ.FEM.SG 1.SG.OBL=EZ.FEM.SG old and humid
'my old, humid house'
(Wurzel 1997:26)
- b. * *mal-a min=a kevin=a şil*
house-EZ.FEM.SG 1.SG.OBL=EZ.FEM.SG old=EZ.FEM.SG humid

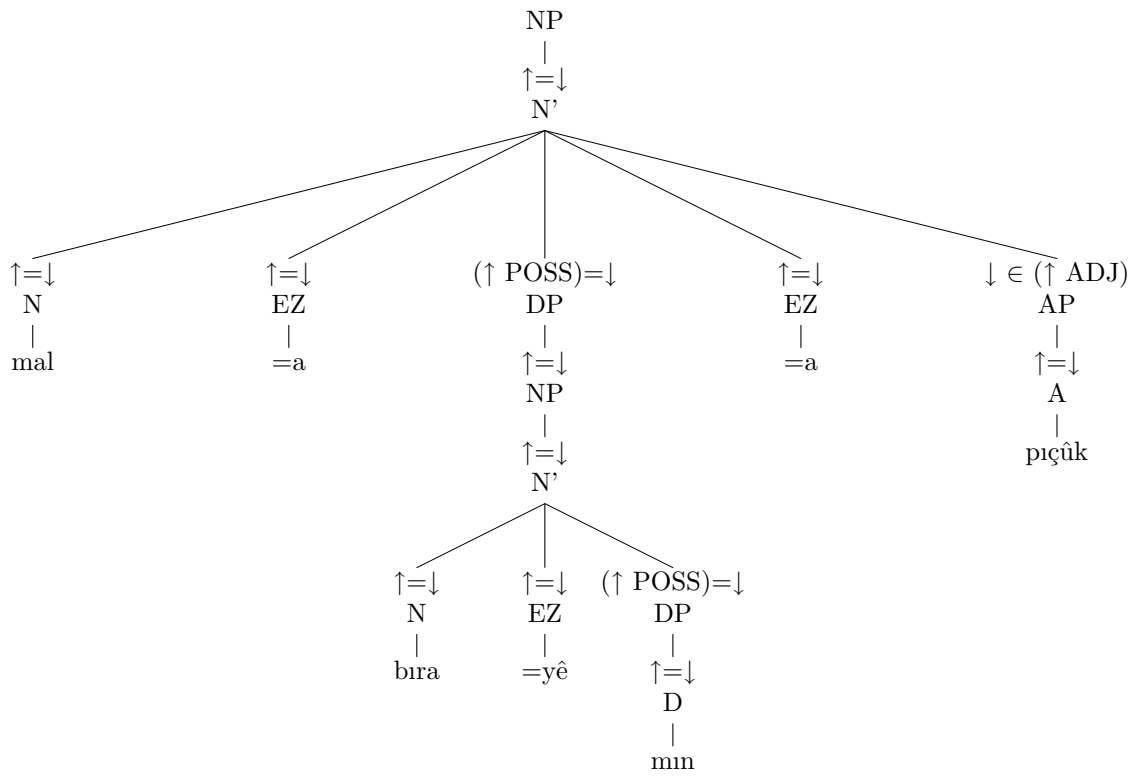
I will now give some lexical entries and use them to analyze three sample noun phrases.

mal	N	(↑ PRED) = 'house' (↑ GEND) = FEM	birâ	N	(↑ PRED) = 'brother' (↑ GEND) = MASC
=(y)a	EZ	((↑ PRED) = 'PRO') (↑ GEND) = FEM (↑ NUM) = SG (↑ DEF) = +	=(y)ê	EZ	((↑ PRED) = 'PRO') (↑ GEND) = MASC (↑ NUM) = SG (↑ DEF) = +
min	D	(↑ PRED) = 'PRO' (↑ PERS) = 1 (↑ NUM) = SG (↑ CASE) = OBL	piçûk	A	(↑ PRED) = 'small'

I will now give c-structures and f-structures for the three examples in (15).

- (15) a. *mala birayê min a piçûk*
'my brother's small house'
- b. *mala birayê min ê piçûk*
'my small brother's house'
- c. *mala yê piçûk*
'the small one's house'

The following tree gives the c-structure for (15a).

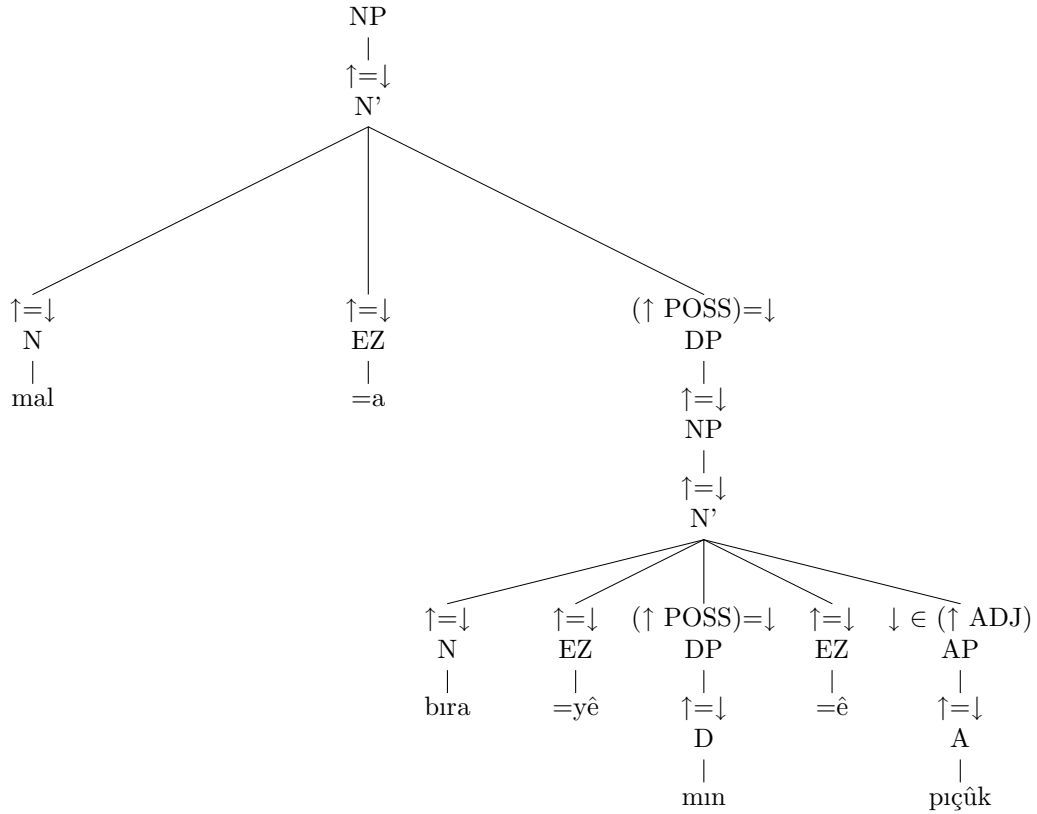


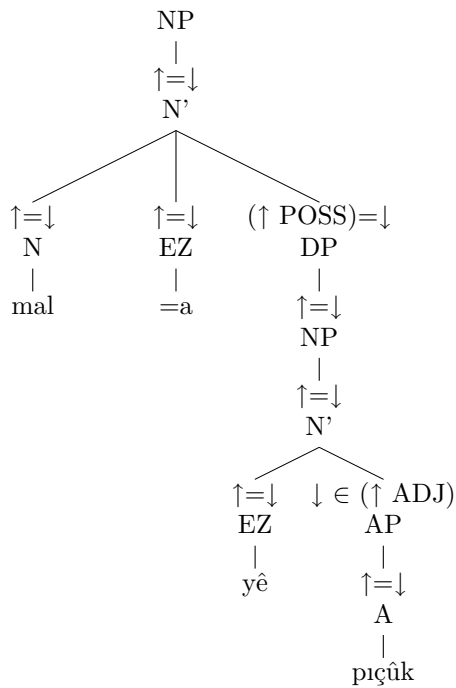
The resulting f-structure is shown in (16).

(16)

[PRED	'house']			
	GEND	FEM				
	NUM	SG				
	DEF	+				
	POSS	[PRED	'brother']
				GEND	MASC	
				NUM	SG	
				DEF	+	
				CASE	OBL	
	POSS	[PRED	'PRO']
PERS			1			
NUM			SG			
ADJ	{	[PRED	'small']	

As can be seen in this example, the form of the *ezafe* restricts which head noun the corresponding adjunct can modify. In this sentence, *piçûk* (small) can only modify the first noun *mal* (house) and not the embedded one, *bira* (brother) although this would be possible structurally. Notice also that the optional PRED of the *ezafe* morpheme is not unified with the f-structure of the head noun precisely because there is an overt head noun and thus already a PRED. The structures corresponding to (15b) are shown below.





(18)

PRED	'house'												
GEND	FEM												
NUM	SG												
DEF	+												
POSS	<table border="1"> <tr> <td>PRED</td> <td>'PRO'</td> </tr> <tr> <td>GEND</td> <td>MASC</td> </tr> <tr> <td>NUM</td> <td>SG</td> </tr> <tr> <td>DEF</td> <td>+</td> </tr> <tr> <td>CASE</td> <td>OBL</td> </tr> <tr> <td>ADJ</td> <td>{ [PRED 'small'] }</td> </tr> </table>	PRED	'PRO'	GEND	MASC	NUM	SG	DEF	+	CASE	OBL	ADJ	{ [PRED 'small'] }
	PRED	'PRO'											
	GEND	MASC											
	NUM	SG											
	DEF	+											
CASE	OBL												
ADJ	{ [PRED 'small'] }												

4.2 Second modified analysis

In the examples, I have glossed the ezafe markers on modifiers as clitics because they show slightly different behavior from the ezafe marker appearing on the head noun. The grade of morphophonological integration of the ezafe marker on the head noun seems to be greater, consider example (19) which shows that the ezafe morpheme can cause stem alternations.

- (19) a. dê (mother) + a (EZ.FEM.SG) → diya
 b. rê (way) + a (EZ.FEM.SG) → riya

These kinds of alternations do not seem to occur when the ezafe morpheme cliticizes onto the right edge of a modifier, cf. example (20).

- (20) *keç-a* *te=ya* *mezın*
 daughter-EZ.FEM.SG 2.SG.OBL=EZ.FEM.SG big
 'your big daughter'

However, further study is needed to clarify this issue. A better argument for considering the ezafe on the head noun as a suffix is the fact that the ezafe in Kurmancî suppresses all other possible inflectional suffixes on the noun. The noun *mal* in (21a) carries a case suffix because the DP it heads is assigned oblique case by the verb. This suffix cannot appear if the noun is modified by a post nominal element as in (21b). However, the whole DP is still in the oblique case because the demonstrative still has to appear in its oblique form *vi*.

- (21) a. *ez* [*DP vi* *mal-e*] *dı-bin-ım.*
 1.SG.NOM DEM.SG.OBL house-SG.OBL DUR-see.PRES-1.SG
 'I see this house.'
- b. *ez* [*DP vi* *mal-a* *sor*]
 1.SG.NOM DEM.SG.OBL house-EZ.FEM.SG red
dı-bin-ım.
 DUR-see.PRES-1.SG
 'I see this red house.'
- c. * *ez* [*DP vi* *mal-e* *sor*]
 1.SG.NOM DEM.SG.OBL house-SG.OBL red
dı-bin-ım.
 DUR-see.PRES-1.SG

The same phenomenon occurs for possessors that are themselves modified. Normally, possessors have to appear in the oblique case, but they cannot bear any case morpheme when they, too, are modified. Instead they have to be marked with the ezafe, cf. (22).

- (22) a. *mal-a* [*NP jin-e*]
 house-EZ.FEM.SG woman-SG.OBL
 'the woman's house'
- b. *mal-a* [*NP jin-a* *min*]
 house-EZ.FEM.SG woman-EZ.FEM.SG 1.SG.OBL
 'my wife's house'
- c. * *mal-a* [*NP jin-e* *min*]
 house-EZ.FEM.SG woman-SG.OBL 1.SG.OBL

These facts suggest that the ezafe morpheme on the head noun is an inflectional affix and not a clitic anymore. In order to better account for this interpretation, I will propose a second, more complicated analysis in this section.

While I propose a revised analysis for the ezafe morpheme on the head noun, I will maintain the initial clitic analysis for the ezafe marker which appears between two modifiers. The main problems which thus arise are how to make sure that the ezafe form of a noun only occurs when the noun is modified and how to avoid that a noun in the ezafe form can be followed by another ezafe clitic.

I propose the following sublexical rule for the ezafe form of a noun:

$$N_{EZ} \rightarrow N_{Stem} -EZ$$

$$\uparrow=\downarrow \quad \uparrow=\downarrow$$

The lexical entry of the ezafe suffix then has the following form. I provide one example for -EZ.FEM.SG.

$$-ya \quad -EZ \quad (\uparrow \text{ GEND}) = \text{FEM}$$

$$(\uparrow \text{ NUM}) = \text{SG}$$

$$(\uparrow \text{ DEF}) = +$$

$$(\uparrow \text{ ADJ}) \vee (\uparrow \text{ POSS})$$

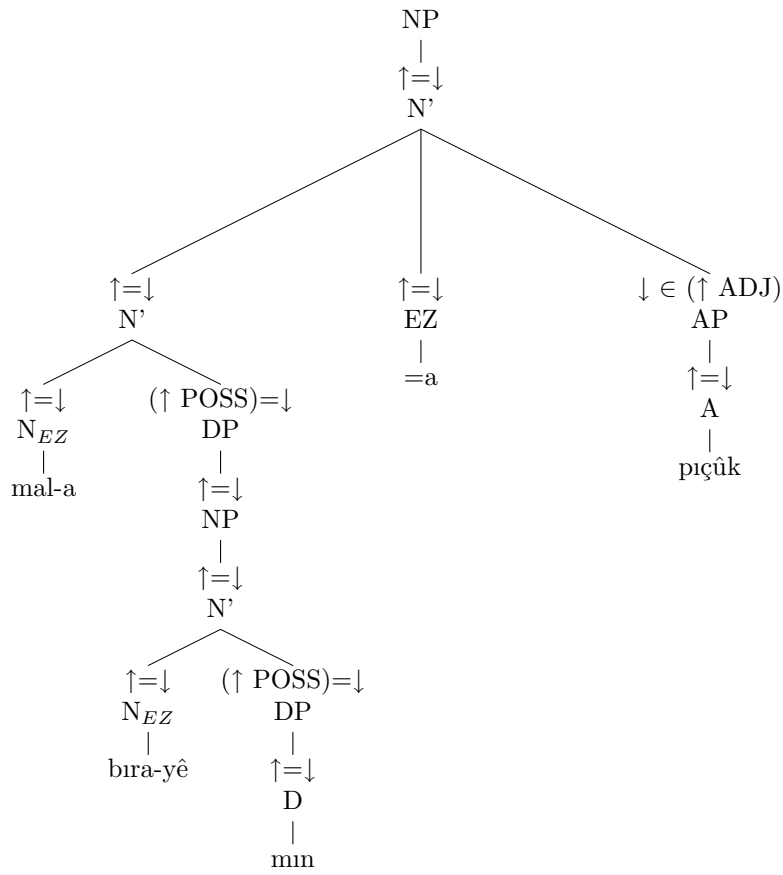
The last line ensures that a noun in the ezafe form always has to be modified by at least one adjunct or a possessor. The problem of avoiding a doubling of the ezafe marking between a head noun and its leftmost modifier is not that straightforward. On the basis of the present analysis, I have to assume that it is a structural property of Kurdish that nouns which carry an ezafe suffix are followed directly by a modifying phrase. An alternative analysis where doubled ezafe marking is excluded because of f-structure principles would require some kind of non-unifiable feature. However, since under the analysis proposed here multiple ezafe markers do indeed unify with the head noun's f-structure if the head noun is followed by more than one modifying phrase, this choice is not viable here. Moreover, I want to avoid devising a special ad hoc feature just for the sake of the Kurdish ezafe construction⁹.

Therefore, I propose the following modified c-structure rules¹⁰ for the Kurmancî NP.

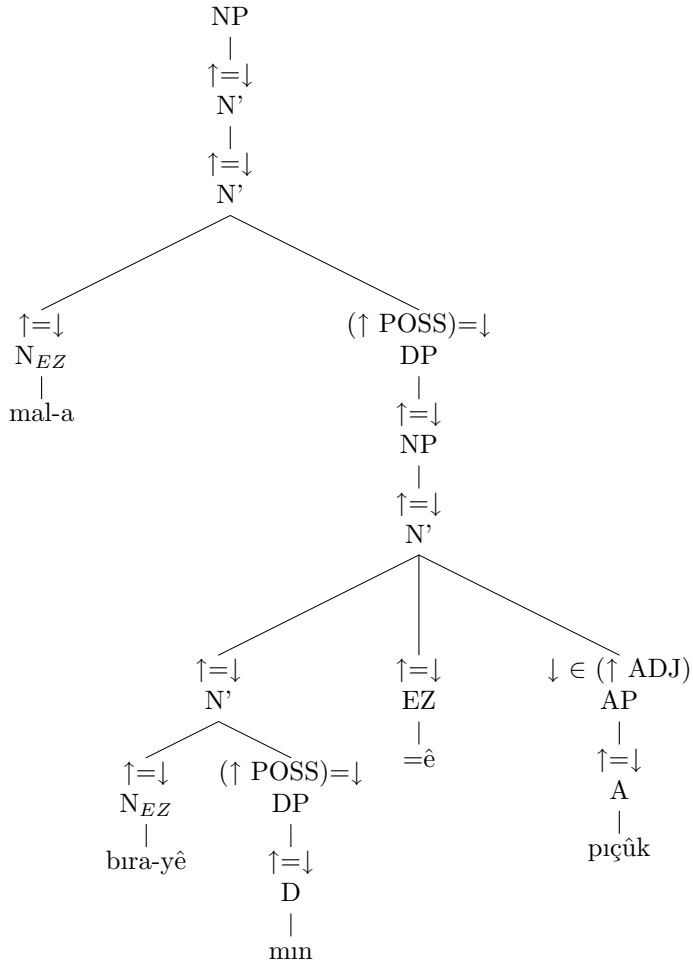
⁹One could however make use of some version of the principle of economy, cf. e.g. Sadler 1997.

¹⁰Rule (d) should also make sure that pronominal possessors cannot occur separated from the head noun by another modifier, e.g. by allowing only DPs as possessor which have non-pronominal binding features.

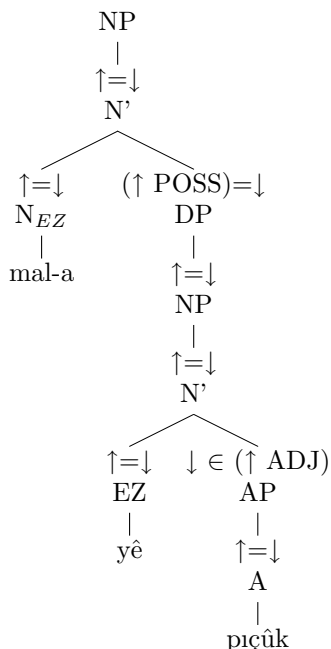
I will now apply these modified rules to the three example NPs from (15). I first show the c-structure for (15a). The corresponding f-structure is the same as that in (16).



The c-structure of example (15b) is the following, while the corresponding f-structure is again the same as (17).



Last but not least, (15c) can also be analyzed with the modified rules. Again the corresponding f-structure is the same as that in (18).



4.3 Interaction between case and ezafe marking

Under both analyzes (4.1 and 4.2), we have to account for the fact that case is no longer marked on the head noun when the head noun is modified. This is straightforward for the analysis from section 4.2. We can simply assume that a noun which carries an ezafe suffix has no case feature by itself. It can however still be assigned case by a verb or a construction such as the possessive construction. We can then make sure that unmodified nouns have to appear in the correct case form, either nominative¹¹ or oblique, by specifying the case of nouns that have no ezafe suffix as OBL in their lexical entries if they carry a case suffix or as NOM if they do not carry case marking. For the initial analysis from section 4.1, the problem is less easy to solve. We cannot simply state e.g. that the ezafe morpheme carries a lexical information such as $\neg(\uparrow \text{CASE})$ because a DP and thus also the noun which heads it can be assigned case by verbs, prepositions, or specific constructions. This case is not expressed on the head noun if it is modified but for example on a demonstrative that acts as its determiner. This is also the reason why I proposed the more complicated second analysis in the first place. I will leave the question open whether the first analysis which assumes that all ezafe markers are clitics can be adapted to provide for the interaction of modification and case marking.

¹¹This case should better be called casus rectus.

5 Further work

Further work on the Kurdish *ezafe* should clarify a number of important questions about the Kurdish noun phrase. It should e.g. be investigated whether it really is the case that a given head noun can be modified by only one AP and only one PP. Furthermore, it is not clear yet whether one can distinguish between adjunct and complement dependents of the noun in Kurmancî and what syntactic differences could be manifestations of this distinction. Some further complicating facts that I have ignored in this paper should also be looked at more closely. First, there is a distinction between definite and indefinite *ezafe* markers in some dialects of Kurdish, while others e.g. the dialect of our informant from Turkey seem to be losing the indefinite *ezafe* altogether. Second, the conditions and rules for noun and NP coordination should be investigated. Example (23) shows that only the last noun in a coordination determines what *ezafe* suffix it appears with even though the whole complex is modified.

- (23) *Mizgin l şûna di-ya xwe l [xwisk û*
 Mizgin instead mother-EZ.FEM.SG REFL on sister and
bir-ê] xwe di-nêr-e.
 brother-EZ.M.SG REFL DUR-watch.PRES-2/3.SG
 'Mizgin cares for her sisters and brother in her mother's stead.'
 (Wurzel 1997:63)

Although it is clear from the preceding context that Mizgin has several younger sisters and only one younger brother, the *ezafe* appearing on the last noun of the coordination is only determined by that noun itself making it masculine singular in example (23). However, example (24) shows that a clitic *ezafe* marker preceding an additional modifier would agree with the whole modified NP and thus appear in the plural form.

- (24) *[Xwisk û bira-k-i] wê yên*
 sister and brother-INDEF-EZ.MASC.SG.INDEF 3.SG.FEM.OBL EZ.PL
piçûk he-ne.
 small exist-PL.
 'She has little sisters and a little brother.'
 (Wurzel 1997:63)

Last but not least, it would be interesting to compare the Kurdish *ezafe* with the *ezafe* in Farsi which seems to show slightly different behavior. There have been several proposals for formal analyzes of the Farsi construction (such as Kahnemuyipour 2000) which could also be compared to the present proposal.

6 Conclusion

Although the analyzes proposed in this paper are only a first attempt at describing the syntax of the Kurdish noun phrase in the formal syntactic theory LFG, they already show how difficult but also how fascinating this task can be. It is my impression that "linking constructions" such as the Kurdish *ezafe* have not received enough attention by theories

of formal syntax. The phenomenon itself is not so rare. Related constructions appear in a number of different languages and language families, e.g. in Tagalog (cf. Himmelmann 1997). It should thus be a fruitful task to see whether a formal theory such as LFG can provide elegant analyzes for these phenomena or whether a wholly new construction type would have to be integrated into existing theories of syntax.

7 REFERENCES

- Bresnan, Joan (2001): *Lexical-Functional Syntax*. Blackwell Textbooks in Linguistics 16. Blackwell Publishers, Malden, Massachusetts.
- Himmelmann, Nikolaus P. (1997): *Deiktikon, Artikel, Nominalphrase: zur Emergenz syntaktischer Struktur*. Max Niemeyer Verlag, Tübingen.
- Kahnemuyipour, Arsalan (2000): *Persian Ezafe Construction Revisited: Evidence for Modifier Phrase*. In: *Proceedings of the 2000 annual conference of the Canadian Linguistic Association*, pp. 173-185.
- MacKenzie, D. N. (1961): *Kurdish Dialect Studies I*. Oxford University Press, London.
- Sadler, Louisa (1997): *English Auxiliaries as Tense Inflections*.
Online: citeseer.nj.nec.com/sadler97english.html.
- Sadler, Louisa (1998): *Welsh NPs without Head Movement*. In: *Proceedings of the LFG98 Conference, University of Queensland, Brisbane*. CSLI publications.
Online: <http://www-csli.stanford.edu/publications/>
- Wurzel, Petra (1997): *Rojbaş. Einführung in die kurdische Sprache*. Sprachenreihe Reichert. Dr. Ludwig Reichert Verlag, Wiesbaden.