John Benjamins Publishing Company

CHAPTER 5

Ezafe and the article

Shuan Osman Karim
The Ohio State University

There are a variety of different nominal systems within the Western Iranian languages. Nominals in some of these languages are maximally sensitive to case, number, gender, animacy, ezafe (modification), and definiteness. However, most of these languages only inflect for a subset of these categories. An additional peculiarity of these systems is that some of these languages feature wildly differing systems when inflectional features interact. For instance, Zazaki features a caseless system on indefinite nouns, a bicasual (two case) system on definite nouns, and as many as four cases (following Paul, 1998a) on modified nouns. This paper explores, synchronically and diachronically, the nominal systems observed in Western Iranian languages focusing on the interaction between (in)definiteness and modification by attributive adjectives. This constitutes a near-comprehensive juxtaposition of data that can serve as an essential reference for anyone who seeks to study and understand *Ezafe* and the article in New Western Iranian.

Keywords: ezafe, Iranian, definiteness, nominal morphology, typology

1. Introduction

Within the Western Iranian languages, there are several strategies for marking definiteness and attribution. New Persian, in particular, is famous for differential object marking,[1] or object marking only with definite direct objects and for *Ezafe*, or attribution marker.[2]

---

1. Differential object marking (DOM) is a feature of many Iranian languages. A range of factors influence whether or not a direct object will be marked including definiteness, specificity and animacy. This led Bashir (2008) to coin the term identified object marking (IOM).

2. For an in-depth discussion of the New Persian Ezafe and its syntax and functions see Kahnemuyipour (2014); Samiian (1994); Franco-Rita & Savoia (2012); Haig (2011); Larson & Samiian (2020b, a, etc.)

Standard New Persian does not have a more generalized definite article. This is, perhaps, unsurprising from a diachronic standpoint as the relative pronoun *yat* which was the predecessor of *Ezafe*–at least as Old Persian *haya*[3]–acted as an article in the Old Iranian period (Kent, 1944, 1950); *yat* can be observed through all attested stages of Persian eventually becoming *Ezafe* in New Persian. Absent a definite article, standard Persian innovated a new one recruiting material from other parts of the lexicon (i.e. *-rā < rādiy*). This paper examines various strategies of definiteness and attribution marking in the Western Iranian languages. It differs from other studies of ezafe or definiteness marking because it is primarily descriptive, gathering the data from the Western Iranian languages in one place for the first time. This study focuses narrowly on the interaction of modification and definiteness marking. See Karim (2021) for an examination of how case contributes to these systems. Additionally, I provide some analysis, and through the juxtaposition of these data, I attempt to explain their idiosyncrasies and identify areas where further study is warranted.

1.1   Background on the Iranian languages

The New Iranian languages have traditionally been categorized geographically, Northwestern (e.g. Zazaki), Northeastern (e.g. Yaghnobi), Southwestern (e.g. Persian), and Southeastern (e.g. Pashto). The geographic classification holds for Western Middle Iranian languages (Middle Persian and Parthian) but less so for Eastern Middle Iranian because of languages like Bactrian, which seem to have Eastern and Western features. This categorization does not work for Old Iranian, Old Persian and Avestan, which had not developed the East/West isoglosses that characterize the New Iranian languages. The geographical terminology has come into question due to shared features of non-Persian Iranian languages that seem to transcend the geographical categories. This has led Korn (2016) to propose the classification of Central Iranian; see Korn (2017); Huyse (2017) for more on the classification of these languages. For a more up-to-date family tree incorporating Central Iranian, see Hammarström et al. (2020).

Perhaps without exception, the Old and Middle Iranian languages cannot be understood as the direct ancestors of any of the New Iranian languages. The Old Iranian languages Avestan and Old Persian were probably contemporaries of languages like Proto-Kurdish which have disappeared without attestation in manuscripts or

---

3. Kent (1950) argues for a restricted use of *haya* (< PIE. *sa(s) and *(H)yod Jügel, 2017, 561) as a definite article in Old Persian (Kent, 1950, § 260.IV). This function was not highly grammatical and shown to only have a fraction of the semantic range of the definite article in English.

inscriptions. Therefore, Old Iranian languages are at best approximations of what Proto-Kurdish or Proto-Zaza-Gorani might have been like. For instance, both Avestan and Old Persian show constructions that resemble an early form of *Ezafe* (Jügel, 2017). Among New Iranian languages, the *Ezafe* phenomenon is restricted to Western Iranian. Likewise, New Iranian languages like Kurdish show reflexes of PIE laryngeal consonants which are missing from Old Iranian texts (Kümmel, 2018).

One additional flaw with the geographic designations is the fact that some of the languages that fit genealogically into one category have since migrated to other regions. Most notably, Baluchi fits into the Northwestern group and is spoken in the far east of the Iranian zone (Korn, 2003). As the focus of this inquiry is on the interaction of ezafe and definiteness marking, features of languages histori-cally referred to as Western Iranian, the geographical designation will suffice here. The New Western Iranian languages referenced in this paper are Northwestern: Goranî[4] (Hewramî), Zazaki (Southern and Central), Caspian (Gilaki), Southern Tati (Chali), North Kurdish (Kurmancî), Central Kurdish (Soranî), Southwestern: Persian (Standard and Colloquial), Northern Tat (Şirvan).[5] Data provided will ad-here to the native orthography for languages that use the Latin alphabet,[6] standard transliteration for standard New Persian and Old and Middle Iranian languages and IPA for all others.

2.   Definiteness marking strategies

There are various strategies employed for marking definiteness and attribution in the Western Iranian languages. In these languages, definiteness can not be seen as a binary category consisting of definite and indefinite. Rather, they have a tripartite

---

4.   The term Gorani is used generally by linguists to refer to a group of closely related varieties including Hewramî, Paweyane, Zerdeyane, Gewrecûî, Şebekî, the (classical) Gorani Koiné, etc. The data considered in this study come from MacKenzie (1966) and Holmberg & Odden (2008); the former is from Hewramî (Lihonî) and the latter is Paweyane (the Goranî language from Pawe city). Although these varieties differ on several counts, their use of ezafe and definiteness marking are comparable.

5.   The following abbreviations for languages will be employed when necessary: Soranî (So.), Kurmancî (Ku.), Standard New Persian (NP.), Colloquial New Persian (NP. (col.)), Hewramî (He.), Şirvan Tat (ŞT.), Southern Zazaki (SZ.), Central Zazaki (CZ.) Gilaki (Gi.) and Chali South-ern Tati (CT.).

6.   This includes the use of the Hawar (standard Kurdish) script for all languages spoken in the Kurdish zone, including Hewramî (Lihonî and Paweyane) but not Zazaki, which has its own orthography (although it often appears in the Hawar script as well).

distinction of definite, indefinite and absolute (Thackston, 2006b; Rastorgueva et al., 2012). The categories of definite and indefinite generally behave as they do in other languages, but the absolute has a general sense or is ambiguous. The definite is employed when the referent is either contextually defined or if it has previously been mentioned. The indefinite can have a specific interpretation (i.e. "a certain person") as well as an indefinite interpretation. The absolute form of the noun has a more complex distribution. It can contextually be understood as definite or indefinite/ generic. Compare the Kurmancî sentences *vêrê hirç nîne* "there are no bears here (generic)" and *hirç hat* "The bear came (definite)" (Haig & Öpengin, 2018, 16). All the Western Iranian languages examined here have, minimally, an overt marker for the indefinite. However, the distinction between absolute and definite is not always overtly marked. Several of the languages have an overt definite marker, but others still distinguish definiteness only by the occurrence of case endings (e.g. Zazaki). I outline these strategies in § 2.1 (indefinite marking) and § 2.2 (definite marking).

2.1 The indefinite markers

There are two types of indefinite markers in the Northwestern Iranian languages: YAK (< *aika) and EW[7] (< *aiwa). The languages which have the indefinite article YAK are Soranî (N-*êk*), Kurmancî (N-*ek*) and Central Zazaki (N-*yek*). Standard New Persian (N-*i*), Hewramî (N-*êw(æ)*), Southern Zazaki (N-*ê(n)*), Gilaki (*i* NP), and Chali Tati (*i* NP-(*i*)) have the indefinite article EW. Şirvan Tat (ye NP-*i*) and colloquial New Persian ((*ye*) NP(-*i*)) have developed a hybrid system consisting of both YAK and EW.[8]

The numeral YAK is used in Soranî, Kurmancî, New Persian (standard and colloquial), Hewramî, and Şirvan Tat. The rest of the languages have the numeral EW (e.g. Zazaki (Southern and Central), Gilaki and Chali Tati). The systems of these languages are presented in Table 1. The two Kurdish languages Soranî and Kurmancî have YAK-type systems; the languages Southern Zazaki, Gilaki, and Chali Tati have EW-type systems and New Persian (standard and colloquial), Hewramî and Şirvan Tat have YAK/EW-type systems, or a system with the numeral YAK and indefinite article EW; Şirvan Tat and colloquial New Persian have a bound form of yak which accompanies the indefinite article EW. The opposite, a system with a

---

7. I use small caps here to signify the numeral and article variants YAK and EW regardless of their realization in the individual languages. This is in contrast with YAK-type, a system with both the numeral and article from *aika, ew-type, both the numeral and article from *aiwa and YAK/ EW-type, with the numeral from *aika and the article from *aiwa.

8. In Şirvan Tat, the indefinite construction without the preposed element exists but seems to be marginal (Suleymanov, 2020); in colloquial New Persian, the distribution is more complicated.

numeral ᴇᴡ and an indefinite article ʏᴀᴋ, is only found in Central Zazaki (Werner, 2018). This is odd if it is assumed that both the article ᴇᴡ and the number ᴇᴡ are the inherited forms in Zazaki. In contact with Kurmancî, Central Zazaki borrowed the article ʏᴀᴋ (Werner, 2018). One might expect the number to be borrowed from neighboring Kurmancî and for the number to be recruited as an indefinite article, assuming that it is easier to borrow words than morphemes. However, the inherited numeral ᴇᴡ has been retained but with the form *yew*.[9]

Table 1. The relationship between the numeral "one" and the indefinite article

		So.	Ku.	ST.	NP. (col.)	He.	NP	SZ.	Gi.	CT.	CZ.
YAK	Article	✓	✓	✓	(✓)						✓
	Numeral	✓	✓	✓	✓	✓	✓				
EW	Article			✓	✓	✓	✓	✓	✓	✓	
	Numeral							✓	✓	✓	✓

Unfortunately, the attested Old Iranian languages do not aid in the identification of what is innovation and what is retention. In Old Persian, cardinal numerals are not well attested with the exception of *aiva-* in the phrase *aivam paruvnām X* "one of many Xs" (Kent, 1950, § 204, § 250). In all other instances, the cardinal number "one" is expressed not phonetically but by the orthographic numeral. For this reason, it is unclear if there was an *aika variant. Such a variant does not occur in Avestan either. Because of this lack of clear attestation, it is impossible to say for certain if the *aika variant existed in the Old Iranian period and if it did whether or not there was a semantic distinction.

Both *aika and *aiwa are reconstructed for Proto-Indo-European, and both are attested on the Aryan branch of the Indo-Iranian language family represented by Sanskrit *eka* "one" and *eva* "only," "even," etc.

In Middle West Iranian, the picture is closer to what can be observed today. Parthian (Northwestern) does not show any reflex of the numeral ʏᴀᴋ (Durkin-Meisterernst, 2004, 373), and it always uses ᴇᴡ as the cardinal number "one" (Skjærvø, 2009, 211). This ᴇᴡ-type system is what is observed in the New Northwestern Iranian languages Southern Zazaki, Gilaki and Chali Tati. In contrast, Middle Persian (Southwestern) has both the numeral ʏᴀᴋ and the indefinite article ᴇᴡ, which is the ʏᴀᴋ/ᴇᴡ-type system as observed in the New Southwestern Iranian languages standard New Persian (southwestern) and Hewramî, which is typically thought of as Northwestern (Durkin-Meisterernst, 2004, 100, 373). In

---

9. Other Zazaki varieties have *jew* "one". Based on available data (Paul, 1998b; Werner, 2018), whether or not the *y : j* correspondence is a regular one is unclear. If it is irregular, one possible explanation could be contamination from Kurmancî *yek*.

fact, the function of the Middle Persian indefinite marker *-ew* is to mark a specific indefinite noun just as in New Persian (Skjærvø, 2009, 211, 205).

It is not currently possible to decide whether the existence of a split YAK/EW-type system in New Persian, Hewramî and Şirvan Tat (and Middle Persian) have preserved inherited reflexes of the two lexemes. Another possibility is that the YAK variant was borrowed into these languages from a related language between Old and Middle Iranian. In Kurdish, both the numeral and the indefinite article are from YAK. This fact taken alone suggests that the YAK variant is reconstructable for Proto-Kurdish. The distribution of systems suggests that the YAK-type is the innovative form as it only occurs in Kurdish (see § 4.2 for more discussion). The generally accepted sequence of events was that the numeral became the article and not the reverse. Assuming that in pre-Proto-Kurdish there was a system that better reflected Old Iranian with both EW and YAK variants, the numeral YAK may have become an indefinite article filling the gap created by the loss/repurposing of the indefinite article EW as the indefinite ezafe (which I suggest as a possibility in § 3.1.1). The fact that the only language to feature the numeral EW and the article YAK (Central Zazaki) did so through direct morphological borrowing of the article YAK into a system that preserved the numeral EW suggests that the numeral can be recruited as an indefinite article but an article it is unlikely to break away and become a numeral.

2.2   The definite markers

There is not a clear etymological source for many of the various definite markers in the Northwestern Iranian languages. This is likely due to the fact that there is no record of most of these languages before the modern era. Several definiteness marking strategies are employed: (1) the marker RĀ (New Persian (standard and colloquial) and Şirvan Tat); (2) case marking (Zazaki and Chali Tati); (3) the article *-(h)e* (New Persian (colloquial); and (4) the article *-eke* (Soranî and Hewramî). The historical origins of these particles is not a trivial matter. Their etyma are particularly important to understand the idiosyncrasies of the way they combine with attributes.

2.2.1   *The marker* RĀ
The marker RĀ is employed in several languages to mark a definite direct object. It has a clear etymological path (in Persian) from Old Persian *rādiy* "on account of)" to Middle Persian *ray* "for/because" on to its various forms in New Iranian languages (Korn, 2017, 613). Additionally, in languages which employ the marker RĀ and an overt definite article (e.g. colloquial New Persian), they can occur together (see Example (1)). Due to the fact that this form operates independently from definiteness and ezafe marking, it will not be addressed further in this paper.

(1) *ketab-æ-ro   xund-æm*
     book-DEF-DEF.DO[10] read.PST-1SG
     NP. (col.) "I read the book."

2.2.2   *Case marking alone*

Another strategy is to use differential case marking to mark definiteness. For in-
stance, Zazaki employs oblique to mark definite direct objects and not indefinite/
generic ones (Paul, 1998b, 21). Compare Examples (2) and (3) where the direct
object surfaces as *mırıçıkı* with no oblique marker when used in a generic sense and
mırıçıkerı with the feminine singular oblique marker *-er(i)* when definite.

(2) *seni   kes   mırıçıkı tepş-en-o*
     when person bird   catch.PRS-IPFV-3SG.M
     SZ. "when one catches a bird"
                                    (Berz & Malmîsanij (1988), 113.22 apud Paul (1998b))

(3) *mırıçık-erı   xo   mıyan=dı   nımn-en-ê*
     bird-F.SG.OBL self middle=LOC hide.PRS-IPFV-3PL
     SZ. "(they) hide the bird among themselves"
                                    (Berz & Malmîsanij (1988), 198.4 apud Paul (1998b))

Some Kurmancî dialects also show the pattern where the presence of case marking
(when applicable) only occurs on definite nouns; see Examples (4) and (5) which
only differ by the oblique marker on the direct object *pirtûkê* which marks it as
definite. In the direct case where there is no distinction, there is ambiguity as to
whether the noun is definite or indefinite. As this is not the pattern in standard
Kurmancî, I will not consider these forms in the Kurmancî data.

(4) *Ez   pirtûk di-xwîn-im.*
     1SG.DIR book   IPFV-read.PRS-1SG
     Ku. "I do read books (generic)"

(5) *Ez   pirtûk-ê   di-xwîn-im.*
     1SG.DIR book-M.SG.OBL IPFV-read.PRS-1SG
     Ku. "I read/am reading the book" (definite)

2.2.3   *Colloquial New Persian -*(h)e

The colloquial New Persian article *-(h)e* is perhaps the most puzzling as it was
not attested before the modern period. Its form is similar to *Ezafe* except that it is
stress bearing. *Ezafe* is thought to have descended from the Old Iranian relative

---

10. All glosses have been altered from their original publications to reflect the features discussed
in this paper and according to Leipzig glossing rules when possible. A full list of abbreviations is
presented in appendix A.

pronoun (Avestan: *yat*) acting as an article. However, this cannot be the etymon of *-(h)e*, which in addition to being stressed, occurs with an epenthetic *h* when it follows a vowel. In contrast, *Ezafe* employs an epenthetic glide *y*. The fact that an *h* is used to break hiatus implies that there is an etymological (i.e. not a purely phonological) source.[11]

A possible etymon for this article is the Old Iranian demonstrative pronoun *ha* (< PIIr. *sa(s)). This would account for the phonology and the semantics.[12] It cannot, however, account for the absence of such a marker in historical texts. In Old Persian, the demonstrative *ha-* and the relative pronoun *ya-* only occur in the univerbated form *ha-ya-* which was employed as both a definite article and relative pronoun (Kent, 1944). This may suggest a hybrid solution; i.e. *Ezafe -e* from Old Iranian *ya-* and the definite article *-(h)é* from Old Iranian *(ha)ya-*.

2.2.4 *Soranî -eke*

The article *-eke* found in Soranî and Hewramî is also problematic. It has been suggested that this definite article has its origin in the diminutive suffix *-ek*.[13] Another possibility is that of a hybrid origin. Haig & Mohammadirad (2019) proposed that the diminutive marker *-ek* might in fact be augmented by the demonstrative cir-

---

11. In Persian, the strategy employed to resolve vowel hiatus is the insertion of a glide *-y-*. This is exemplified by ezafe after vowel final nouns and by the personal endings after vowel final verb stems: *xune-ye* and *miy-a-yæm* not *xune-e* and *miya-æm* (there is a colloquial contracted form of *miy-a-yæm* that is realized as *miy-a-m* but, crucially, not *miy-a-hæm*). There are other examples where the insertion of an *-h-* could be perceived by a non-literate speaker as a hiatus resolution strategy. This is exemplified by the plural ending *-(h)a*. The *h* surfaces only after vowel final nouns: *xune-ha* but *ketab-a*. Of course, any literate speaker knows that the /-h-/ is actually part of the plural morpheme orthographic < HĀ>.

12. A demonstrative source is "the most frequent way in which definite articles evolve" (Heine & Kuteva, 2002, 110), e.g., English: *the* and Syriac: *hu*.

13. The diminutive source theory was proposed as early as Soane (1913). His assertion appears to be based on superficial similarity and a fundamental misunderstanding of the suffix *-eke*, which he does not describe as a definite article at all. Mackenzie (1961) has diminutive markers *-çe, -çke, -ek, -ik, -ke, -eke, -ōk, -ōke, -kele, -le, -leke, -île, -ûle, -ûlke, -łe, -ełe, -ōłe*, and *-ōłke*. Each of these is highly lexicalized in its use, and it is not necessarily clear that these are acting as diminutives synchronically or diachronically. For instance, Mackenzie (1961) gives *pûłeke* "fish scale" and *tûreke* "small bag" as examples of the diminutive suffix *-eke*. However, *pûł* "coin, pittance" from which *pûłeke* is derived already connotes something small; the suffix *-eke* narrows the meaning to a specific kind of small round thing. The example *tûreke* "small bag" (Mackenzie, 1961) is not clearly smaller that a *tûr* "small bag". Öpengin (2016) has *tûreke* "a special type of bag" which conforms more to the definite interpretation. However, he categorizes this *-eke* as separate from the definite article albeit homophonous (Öpengin, 2016, 54). This is justified by the fact that these nouns can take the definite article (e.g. *tûrekeke* "the bag").

cumposition =*e*; compare (Soranî) *ew=e* "that" and *ew ktêw=e* "that book". One merit of this explanation is that in addition to the definite article *-eke*, the article *-e* occurs without the *-ek*. The demonstrative clitic *-e* is known as a marker of definiteness in the demonstrative construction (perhaps akin to double determination in Greek). However, it was not described as an article independent of the demonstrative in any Soranî grammar (e.g. Thackston (2006b); McCarus (1956); Mackenzie (1961)) until Öpengin (2016), a study of the Mukrî dialect (cf. (6)).

(6) *kuř-e   name=yî   bird*
     boy-DEF[14] letter=3SG take.PST
     Mukrî: "The boy took the letter."   (Öpengin, 2016, 60)

The plausibility of the diminutive origin has largely gone unchallenged in the literature. This is in part because there seem to be similar constructions in other Iranian languages; e.g. Koroshi Baluchi also has a definite article *-ak/-ok* which bears an affinity to the diminutive marker (Nourzaei et al., 2015). Additionally, there is a precedent for a diminutive becoming a referential marker similar to a definite article (Pakendorf & Krivoshapkina, 2014). One potential problem with using the Èven data from Pakendorf & Krivoshapkina (2014) is that the diminutive marker in Èven is actually an evaluative marker implying some kind of specification (though this is often diminutive). Moreover, "[t]hey are optional, and the speaker's intentions play a large role in whether or not to use them (Pakendorf & Krivoshapkina, 2014, 328). The features indicative of this cline of development are not necessarily features of the Soranî definite article.

　　There are several other problems with/open questions about the argument for a diminutive origin of the Soranî definite article *-eke*. Of these concerns, there are several directly relevant to this inquiry: (1) Why does this article condition the definite form of *Ezafe* (see § 3.1.2)? (2) The Indo-Iranian diminutive marker *-(a)ka-* attached to a variety of bases to create adjectives of appurtenance, diminutives and to impart no discernible meaning (Whitney, 1993, § 1222). Soranî *-eke* is a definite article and has a more clitic like attachment; compare *kuř-eke* "the boy", *kuř-e baş-eke* "the good boy", *kuř-e baş-e drêj-eke* "the good tall boy," *kuř-e şwan-eke* "the shepherd's boy," etc. How can the shift from derivational affix to inflectional clitic be explained? (5) In Hewramî, where both the definite article *-ækǽ* and the demonstrative clitic *-ǽ* exist alongside case marking, both particles have different declensions (MacKenzie, 1966); compare masculine singular oblique *-ækǽy* and *-i=ǽ*. If *-ækǽ* is composed of *-æk* and *=ǽ*, why does the latter occur outside of inflection?

---

14. Öpengin (2016) glosses this as DEM.CL acknowledging its affinity to the postpositional element of the demonstratives.

My intent here is not to necessarily argue against current theories but to express the need for a more thorough investigation of this possibility. Haig (2019) has shown that the diminutive explanation is possible for Soranî. However, he stops short of being able to confirm this cline.[15] Any eventual proposal for the origin of this marker should take into account not just the semantic functions of the definite article but its morpho-syntactic distribution. The way it interacts with ezafe marking may be an important clue to its diachrony.

3. Attribution marking strategies

With perhaps the exception of Baluchi, which has a genitive case and employs an attributive marker on the adjective (not unlike the reverse ezafe[16] discussed below), the Western Iranian languages tend employ some sort of ezafe construction.[17] *Ezafe* is most simply a linker binding a noun to either an attributive adjective or a genitival possessor. In New Persian, Gilaki, Soranî and Hewramî, *Ezafe* is also used to link secondary (de-nominal) prepositions to their complements.

In addition to the three uses of *Ezafe*, ad-attribute, ad-genitive and with prepositions, *Ezafe* makes several morphological distinctions that are specific to each of the languages.

These distinctions are of four types: (1) interactions between ezafe and definiteness, (2) interactions between ezafe and gender, (3) interactions between ezafe and case and (4) type of modification. The combination of these distinctions has been of

---

15. Haig (2019) states that there is no other likely etymon for the article *-eke*. However, this is not necessarily the case. Another currently unexplored possible etymological source for this marker is the relativizer *ke*. Relationships between relativizers and definite articles are well attested, e.g. English: *the ~ that*, German: *der, das*, etc. "(definite article)" ~ *der, das*, etc. "(relative pronoun)". Perhaps the best evidence for the shift from relative to definite article is Iranian-internal. In Old Persian, the relative pronoun *haya* was also used as a definite article (Kent, 1950, § 260.IV). The ambiguity between relative pronoun and definite article can be reanalyzed in either direction. Starting in Classical Greek poetry, the definite article of Classical Greek came to be used as a relative pronoun (van Emde Boas et al., 2019, § 28.31). This likely began with the so-called "polydefinite" construction (i.e. *the man the good* = "the good man"). The phrase *the man the good* could then be reanalyzed as *the man [that] (is) good* with a zero copula.

16. The term reverse ezafe, following Stilo (2018), refers *Ezafe*-like particle that attaches to an attributive adjective that precedes its nominal complement. This particle is sometimes referred to as an attribution marker (e.g. in Suleymanov (2020)).

17. There are varieties of Baluchi which have developed ezafe marking under influence from Persian (cf. Koroshi (Nourzaei et al., 2015)); see Korn (2005) for more on the origin of these particles (gen: p. 108, adj.suff: p. 151).

particular interest to linguists as gender is a feature of the head noun, definiteness is semantically conditioned, case is governed by a verb, preposition or *Ezafe* itself and the type of modifier.[18]

Ezafe canonically falls on the right edge of the head noun and is repeated after each subsequent modifier except for the last (see (7)). In a few of the Northwestern Iranian languages around the Caspian and including Chali Tati and Şirvan Tat (SW), there is a phenomenon known as the reverse ezafe. In this construction, *Ezafe* attaches to an attributive adjective, and the adjective precedes the head noun (see (8)).

(7) *æsp-î   zil-î   syaw-î   xas*
    horse-EZ big-EZ black-EZ good
    He. "(the) good big black horse"   Holmberg & Odden (2008)

(8) *an   råst-ə   gəb-ə*
    that true-REZ word-COP.3SG
    Gi. "It is true? (lit. Is that a true word?)"   (Rastorgueva et al., 2012)

3.1   The interaction of (in)definiteness and ezafe

At the core of this study is the interaction between definiteness marking and *Ezafe*. In this discussion, I do not focus on the distinctions of modifier type (Zazaki and Hewramî) or gender of the head noun (Kurmancî). In order to avoid the complexity of all these factors, examples will include nouns with attributive adjectives.[19]

Gilaki (Rastorgueva et al., 2012), Chali Tati (Yarshater, 1969) and Şirvan Tat (Suleymanov, 2020) employ the so-called reverse ezafe when a noun is modified by an attributive adjective.[20] The reverse ezafe is different from the canonical ezafe in that the Adjectival modifier precedes the head noun and that *Ezafe* attaches to the adjective and not the head noun. Another feature of the reverse ezafe is that it does not have a different form in indefinite, definite and absolute contexts. Compare (9) and (10).

---

18. Samiian (1994) describes *Ezafe* as a case marker with multiple governors. This description only looks at the interaction of case and type of modifier and not definiteness.

19. Case-marked examples are included in order to showcase definiteness. In Chali Tati and Zazaki, definiteness is reliably marked only in the oblique case.

20. Şirvan Tat additionally has an inherited canonical ezafe construction which has been largely replaced by the reverse ezafe. This can be observed in Example (i), where the noun *xuna* house appears as *xune* with its final vowel raised the last vestige of the inherited ezafe marker.

(i) *xune   gilxis=a ataʃ girift-e*
    house.EZ PN=OBL fire   seize-PRF.3SG
    ŞT. "Gilxis's house has caught fire."   (Suleymanov, 2020, Example 335)

(9) *ziyorat-a ǰö*
     visit-REZ place
     ŞT. "a pilgrimage site"   (Suleymanov, 2020, Example 323)

(10) *ye gözäl-ä   imorät=i*
      one beautiful-REZ mansion=INDF
      ŞT. "a beautiful mansion"   (Suleymanov, 2020, Example 749)

3.1.1 *Indefinite ezafe*

With the exception of the reverse ezafe languages, where the head noun hosts the indefinite marking but crucially not *Ezafe*, each of the Western Iranian languages have some idiosyncrasies in the way *Ezafe* may combine with an indefinite marker. These idiosyncrasies may have their origins in a historical incompatibility based on the fact that *Ezafe* is the modern descendant of *yat*, which may have acted as a definite article in the Old Iranian period and therefore could not co-occur with the indefinite article *aiwa*.

A variety of strategies are employed to work around this incompatibility. One such strategy is to have no overt ezafe marker on indefinite nouns which are marked for indefiniteness as can be seen in the juxtaposition of (11) and (12). In addition to this construction in New Persian, the same meaning expressed in (12) can occur as *ketab-e besyar bozorg-i* with *Ezafe* on the head noun and the indefinite article attaching at the end of the noun phrase. This suggests that synchronically it is not an incompatibility between indefiniteness and *Ezafe* but rather between the marker *-i* and the marker *-e*.

(11) *xune-ye besyař bozorg*
      house-EZ very   big
      NP. "(a/the) big book"   (Thackston, 1993)

(12) *xune-i   besyar bozorg*
      house-INDF very   big
      NP. "a very big house"   (Thackston, 1993)

Examples (13) and (14) show the forms of *Ezafe* in the Hewramî dialect of Luhon (MacKenzie, 1966). These examples do not show the incompatibility between indefinite and ezafe seen in the New Persian examples. According to Holmberg & Odden (2008), the occurrence of *Ezafe* with the indefinite article is optional; they deem both (15) and (16) acceptable. It is possible that ezafe marking on indefinite nouns is intrusive in Hewramî. However, the examples from Holmberg & Odden (2008) and MacKenzie (1966) are somewhat ambiguous. For instance, the sequence *-ew-i* can be parsed [IND-OBL]. The oblique marker *-i* blocks the appearance of the adjectival ezafe. Without a full context sentence, it is impossible to tell if the optionallity described by Holmberg & Odden (2008) is actually oblique marking. When

I presented *Ezafe* examples to a native speaker from Pawe city, they were deemed ungrammatical (Hishmat Shiany, p.c.).

(13)   *kitêb-î   sîaw*
      book-EZ black
      He. "(a/the) black book"   (MacKenzie, 1966)

(14)   *kitêb-êw-î   sîaw*
      book-INDF-EZ black
      He. "a black book"   (MacKenzie, 1966)

(15)   *mar-êwæ   zil-î   reş*
      snake-INDF big-EZ black
      He. "(a/the) big black snake"   (Holmberg & Odden, 2008)

(16)   *mar-êw-î   zil-î   reş*
      snake-INDF-EZ big-EZ black
      He. "a big black snake"   (Holmberg & Odden, 2008)

Even if the forms are truly optional as claimed by MacKenzie (1966) and Holmberg & Odden (2008), there is a clear difference between nouns specifically marked for indefiniteness and those that are unmarked; for the latter, there is most assuredly no optionallity. More research is necessary to confirm the inherited incompatibility between indefinite marking and ezafe marking. I tentatively place Hewramî in the group which shows the Persian-like incompatibility, although this disagrees with MacKenzie's (1966) description.[21]

    Another strategy for dealing with the incompatibility of ezafe and indefinite is the Kurdish pattern. In Soranî and Kurmancî, there is a distinct difference between *Ezafe* employed in definite and indefinite contexts. Compare the Kurmancî ezafe in (17) and (18) where there is a different form of *Ezafe* when paired with the indefinite article. This form occurs only when attached to the head noun in most dialects. In other words, additional attributes are marked by the absolute ezafe[22] as can be observed in (19). At first glance, the Soranî ezafe does not show the same type of variation (cf. Examples (21) and (22)). It is unclear from these examples that the

---

21. Ultimately, MacKenzie's (1966) analysis may have to be reevaluated in light of the data from Rasekh-Mahand & Naghshbandi (2014), which shows that the oblique is used in topicalization and to show emphasis or contrast. Added to genitive marking, object marking and as a prepositional complement functioning as a core verbal argument, this may help to disambiguate many unclear examples from MacKenzie (1966).

22. I use the term absolute ezafe to refer to *Ezafe* as it would attach to a noun not specifically marked for definiteness or indefiniteness (following Rastorgueva et al.'s (2012) tripartite classification of definiteness, cf. Thackston (2006b)). This should not be confused with the independent form of *Ezafe* used without a head noun variably termed free (Thackston, 2006a), demonstrative (MacKenzie, 1995), absolute (Paul, 1998b), or elliptic ezafe (Paul, 2009).

same type of alternation takes place in Soranî because the absolute ezafe is the same as the variant that is used in indefinite contexts. When juxtaposed with the definite ezafe (20) *-e*, it is clear that there is, in fact, an indefinite ezafe that is allowed on bare nouns and those marked for indefinite. The fact that Kurmancî shows the indefinite variant only in the restricted context of indefinite (and not as a absolute ezafe) and the fact that that form bears a phonetic resemblance to the Persian indefinite article may suggest that Soranî and Kurmancî reanalyzed the indefinite article EW as the indefinite ezafe, and Soranî has generalized it as the absolute ezafe. This possibility is by no means certain, but it is reasonable to assume that if *Ezafe* is sensitive to definiteness then its etymon must have been as well.

(17)   *welat-ê   mezin*
       country-EZ big
       Ku. "(a/the) great country"   (Ekici, 2007)

(18)   *welat-ek-î   mezin*
       country-INDF-EZ big
       Ku. "a great country"

(19)   *şal-ek-î   rind-ê   erzan*
       pants-INDF-EZ nice-EZ cheap
       Ku. "a nice cheap pair of pants"   (Ekici, 2007)

(20)   *ktew-e   gewre-ke*
       book-EZ big-DEF
       So. "the big book"   (Thackston, 2006b)

(21)   *ktêw-î   gewre*
       book-EZ big
       So. "(a/the) big book"   (Thackston, 2006b)

(22)   *ktêw-êk-î   gewre*
       book-INDF-EZ big
       So. "a big book"   (Thackston, 2006b)

Southern Zazaki alone has developed the idiosyncratic strategy of using a specialized ezafe marker which is fundamentally different from the definite and the absolute ezafe. This ezafe is the d-form ezafe which also occurs in a restricted set of contexts on nouns that are not specifically marked for indefiniteness (cf. Paul's (1998b) oblique IIa). These forms can be understood as different albeit related phenomena.[23] Compare *Ezafe* in Examples (23) and (24); when the indefinite

---

23. The distribution of the d-form ezafe is not well-understood. Paul (1998b) and Todd (2002) disagree about the specific conditioning environments and the realization of the d-forms. The d-forms that surface with variants identical to those used with the indefinite (e.g. with the *-do* variant) only seem to occur with inanimate nouns. The conditioning environments for the

article *-(y)ê* occurs, it blocks the absolute ezafe *-o*, and the specialized form *-do* takes its place.

(23) *meřdım-o gırd*
 man-EZ   big
 SZ. "(a/the) big man"   (Paul, 1998b)

(24) *meřdım-ê-do gırd*
 man-INDF-EZ big
 SZ. "a big man"   (Paul, 1998b)

3.1.2 *Definite ezafe*

In addition to the interaction between the indefinite article and *Ezafe*, each of the Western Iranian languages has some idiosyncrasies in the way *Ezafe* may combine with a definite marker. These idiosyncrasies cannot necessarily be understood in terms of historical semantics as *Ezafe* was itself likely a definiteness marker in the Old Iranian period. Additionally, the definite articles in the languages which have them have disputed etymologies (cf. 2.2). For standard New Persian and Kurmancî which lack any definite article, there is no discernible difference between the absolute ezafe and *Ezafe* employed on a definite noun. Zazaki uses case marking alone to signify definiteness (i.e. nouns which are not definite are uninflected). Zazaki masculine singular nouns are only overtly marked in the oblique case. The Zazaki definite (Example (25)) and absolute (Example (26)) *ezafat*[24] only differ in the oblique. This phenomenon points to a complex (possibly phonological)[25] interaction between the case marker and *Ezafe* and not necessarily an incompatibility.

(25) *embaz-ê*   *bin-i*   *vin-en-o*
 friend-EZ.DEF.OBL other-OBL see.PRS-IPFV-3SG
 SZ. "he sees the other friend"
                                    (Berz & Malmîsanij (1988), 47.23 apud Paul (1998b))

(26) *embaz-o bin*   *vin-en-o*
 friend-EZ other see.PRS-IPFV-3SG
 SZ. "he sees other friends"   (adapted)

---

d-forms may, like the indefinite, have originated in an inherent incompatibility between the proto-ezafe and the relevant context. This is the subject of my ongoing research and beyond the narrow scope of this discussion.

24. *Ezafat* is the plural of ezafe.

25. MacKenzie (1966) describes a blocking of the absolute ezafe after the oblique marker *-ê* in the Hewramî dialect of Luhon. The same phonological blocking of *Ezafe* might be reconstructible for early Zazaki.

Another strategy for dealing with the interaction between definiteness and *Ezafe* can be observed in Soranî and Hewramî. These two groups have the *-eke* type definite article. Compare the Soranî Examples (27) and (28) as well as the Hewramî counterparts (29) and (30). On the noun phrases unmarked for definiteness, (27) and (29), the absolute form of *Ezafe* is used. On the noun phrases marked for definiteness, (28) and (30), a specialized definite ezafe is used.

(27) *ktêw-î   gewre*
     book-EZ big
     So. "(a/the) big book"   (Thackston, 2006b)

(28) *ktêw-e   gewre-ke*
     book-EZ big-DEF
     So. "the big book"   (Thackston, 2006b)

(29) *æsp-î   zil*
     horse-EZ big
     He. "(a/the) big horse"   (Holmberg & Odden, 2008)

(30) *æsp-æ   zil-ækæ*
     horse-EZ big-DEF
     He. "the big horse"   (Holmberg & Odden, 2008)

In addition the Soranî-Hewramî type definite marker, there is a definite article *-(h)e* in New Persian (colloquial). Samvelian (2007) shows that the New Persian (colloquial) definite article can, in fact, co-occur with *Ezafe* (31). She also shows an alternative construction (32) where the definite article follows the entire noun phrase neutralizing ezafe marking on each internal modifier.[26] Assuming that both the strategies for combining definite and ezafe shown in Samvelian (2007) exist, there is a parallel in Soranî. Soranî allows the form *ktêw-eke-i gewre*, but with a slightly different meaning than the canonical *ktêw-e gewre-ke*. *ktêw-e gewre-ke* must be read as *the big book* where the book"s size is an intrinsic part of its identity; *ktêw-eke-î* gewre, in contrast, should be read as the book which happens to be big. More work on the New Persian (colloquial) definite article is necessary to decide whether these are parallel constructions. Based on these data, I group colloquial New Persian with the languages that have a separate definite ezafe Hewramî and Soranî (i.e. Ø).[27]

---

26. There is a great deal of diversity in the regional varieties of Persian including additional definite articles, Bandari *-ü*, Shirazi *-u*, etc.

27. Samvelian (2005) has proposed that the lack of ezafe marking on definite nouns as observed in (32) is actually an example of compounding. The merits of this possibility are not discussed here.

(31) *pesær-e-ye   æhmæq*
      boy-DEF-EZ silly
      NP. (col.) "the silly boy"   (Samvelian, 2007)

(32) *lebas qeɾmez bi astin-e*
      dress red   without.sleeves-DEF
      NP. (col.) "the sleeveless red dress"   (Samvelian, 2007)

3.2   Summary of forms

If the colloquial New Persian definite article is, in fact, similar to what is observed in Soranî and Hewramî, there only seems to be one strategy of dealing with the combination of ezafe and definite article. This strategy involves the use of a specialized definite ezafe on the head noun and definite article at the rightmost edge of the noun phrase. This definite ezafe occurs as *-e* in Soranî, *-æ* Hewramî and as *-Ø* in New Persian (colloquial). I present a summary of these forms in Table 2.

Table 2. Strategies for Marking Definiteness and Adjectival Attribution[28]

	"one"	Indf. Art.	Def Art.	Indf. Ez.	Def. Ez.	Abs. Ez.
S. Zazaki	jew	N-ê(n)	N-Ø/-í[case]	N-IND-do Adj	N-o/-ê Adj	N-o Adj
Hewramî	yæk	N-êw	N-ækæ	N-IND-Ø Adj	N-æ Adj-DEF	N-î Adj
Soranî	yek	N-êk/-yek	N-eke	N-IND-î Adj	N-e Adj-DEF	N-î Adj
N Persian (col.)	yek	(ye) N(-i)	N-(h)e/ -(h)æ-(r)o	N-e Adj-IND	N-Ø Adj-DEF	N-e Adj
C. Zazaki	yew	N-yek	N-Ø/-i[case]	N-INDF-o Adj	N-o/-i Adj	N-o Adj
Kurmancî	yek	N-ek	N-Ø	N-IND-î Adj	N-ê Adj	N-ê Adj
N Persian	yek	N-i	N-Ø/-ra	N-IND-Ø Adj	N-e Adj	N-e Adj
Sirvan Int.	yek	ye N-i	N-Ø/-ra	IND Adj-a N-ind	Adj-a N	Adj-a N
Gilaki	í N	i (ta) N-i	a N	IND Adj-ə N-IND)	Adj-ə N	Adj-ə N
Chali Tati	i X	i N	N-Ø/-e [casc]	IND Adj-a N	Adj-a N	Adj-a N

There are two strategies for dealing with the incompatibility of *Ezafe* and the indefinite article: (1) ezafe-blocking (New Persian and Hewramî), where *Ezafe* cannot co-occur on a noun which hosts the indefinite article; and (2) the use of a

---

28. In addition to the forms in this table, New Persian has an indefinite construction with the absolute ezafe and a phrase final indefinite marker (N-e Adj-IND) as seen in colloquial New Persian. Additionally, Soranî, Hewramî and colloquial New Persian each have the option of stacking the definite article and ezafe on the head noun instead of using the definite ezafe albeit with different semantics (see § 3.1.2; Soranî/Hewramî: N-DEF-i Adj New Persian (col): N-DEF-e Adj). These forms have been omitted from the chart so only comparable forms are presented.

specialized indefinite ezafe (Soranî, Kurmancî and Southern Zazaki). If the (1) strategy is conceptualized as zero marking, then both strategies can be categorized according to which ezafe allomorphs are syncretic the definite, indefinite or absolute. Hewramî and Southern Zazaki have three separate ezafe markers (counting the zero-marked indefinite). Standard New Persian and Kurmancî employ the same marker for the definite and absolute *ezafat* but have a unique marker for indefinite. Soranî, colloquial New Persian and Central Zazaki employ the same marker for indefinite and absolute *ezafat* but have a unique definite ezafe. Gilaki, Chali Tati and Şirvan Tat have the reverse ezafe, one invariable form for all three (see Table 3 for a visual summary).

Table 3. Ezafe Syncretism

DEF.EZ = IND.EZ = ABS.EZ	:	Gilaki; Chali Tati; Şirvan Tat
DEF.EZ = ABS.EZ ≠ IND.EZ	:	Kurmancî; New Persian
IND.EZ = ABS.EZ ≠ DEF.EZ	:	Soranî; New Persian (col.); C Zazaki
DEF.EZ ≠ IND.EZ ≠ ABS.EZ	:	S Zazaki; Hewramî

4. Thoughts and conclusions

Ezafe phenomenon has been described as being sensitive to case, number, gender, definiteness and attribution type. The types of systems and patterns created by these sensitivities has not as of yet been systematically explored. There have been discussions of etymology which is fairly certain for the New Persian ezafe *-e* and definite direct object marker *-rā*. However, for many markers, such as Soranî *-eke* and colloquial New Persian *-(h)e* the origins are still debatable. The point of this inquiry is not to settle these issues once and for all but rather to identify some of the patterns that can guide future inquiry into *Ezafe* phenomenon. One of the guiding principles here is that the etymological source of a particular morpheme must be able to explain its various sensitivities. In other words, an ezafe that marks gender must have descended from a gender bearing formative (e.g. a pronoun). Likewise, since some ezafe markers have a definiteness distinction, they must have descended from a unit bearing definiteness (e.g. an article/demonstrative). Most of these languages were not attested in the Middle or Old Iranian periods. As such, everything we know about the history of these languages is a hypothesis that will constantly be tested as more data is considered. In light of the data presented here, I have identified a few issues that require further study.

4.1   The definite ezafe

There are two types of definite *ezafat* or different ezafe allomorphs for definite contexts. One type is what is observed in both Central and Southern Zazaki. These languages feature differential case marking with the oblique maker only on definite nouns occurring in contexts that call for oblique marking. The distinction between definite and indefinite *ezafat* is therefore only morphologically marked in a restricted set of contexts. When the noun is in the direct case, there is no distinction; when the noun is oblique, there is a definiteness distinction. The distinction has been created by the interaction between case marking and *Ezafe* marker. In Kurmancî, by contrast, *Ezafe* ending supersedes oblique case marking yielding a system with no definiteness distinction for *Ezafe*. This assessment of the Kurmancî ezafe is only valid for the standard language. There are in fact several dialects of Kurmancî that make distinctions for case and definiteness (see Haig & Öpengin (2018)). Any exploration of the evolution of ezafic systems must look into all regional language varieties. Furthermore, the distinctions made in this paper may transcend the genealogical relationships between these languages.

The other type of definite ezafe is that observed in Soranî, Hewramî and colloquial New Persian. These systems are characterized by the existence of a separate definite article (So. *-eke*, He. *-ækæ* and NP. (col.) *-(h)e*) and a reduced form of *Ezafe* (So. *-e*, He. *-æ* and NP. (col.) Ø). In each of these languages the definite article shows a clitic like behavior where the definite article attaches to the end of a noun phrase with any number of intervening adjectives N-DEF.EZ Adj(-DEF.EZ Adj)-def. Alternatively, the definite article can attach to the head noun followed by the absolute ezafe N-DEF-ABS.EZ Adj(-ABS.EZ Adj). The similarities of these systems are striking because they represent three different groupings within Western Iranian that have converged on this point. An additional problem is that the current theory with the most traction regarding the origin of the definite article *-eke* is the diminutive origin theory. Is there a comparable possible etymon for *-(h)e*? It is not clear how this etymology can explain the development of a special ezafe for these definite contexts.

4.2   The indefinite ezafe

The general rule when it comes to the indefinite ezafe is that when the head noun is the host of the indefinite article EW, that same noun cannot host an ezafe. This true of Southern Zazaki, which has innovated an indefinite ezafe, and standard New Persian which does not allow an ezafe in those contexts. According to MacKenzie (1966) and Holmberg & Odden (2008), Hewramî optionally allows the coexistence of these forms. If the reason for the idiosyncrasies of the indefinite ezafe in Southern

Zazaki and standard New Persian is indeed the result of a historical incompatibility between the indefinite article and *Ezafe* (as an article), Hewramî seems to be innovating ezafe marking on indefinite nouns. Another possibility is that this optionallity is actually a misinterpretation by linguists of the *-i* in *-êw-i* as [-INDF-EZ]. The correct interpretation would be *-êw-i-Ø* as [IND-OBL-EZ] the blocking of *Ezafe* after the oblique suffix being a well attested feature of Hewramî (MacKenzie, 1966, § 9.i). Unknown at the time of MacKenzie (1966) is the fact that when a Hewramî noun is topicalized or emphasized it takes oblique marking (Rasekh Mahand & Naghshbandi, 2014). Although colloquial New Persian has an indefinite article EW, it has developed a clitic like placement where it attaches to the end of a noun phrase not necessarily falling on the head noun. Therefore, there is no interaction or conflict between indefinite marking and ezafe marking. Additionally, in the three languages which feature the reverse ezafe, Şirvan Tat, Gilaki and Chali Tati, there is no interaction or conflict between indefinite marking and ezafe marking.

Soranî, Kurmancî and Central Zazaki all have the postposed definite article YAK that must attach to the head noun. Although their placement is akin to the placement of EW in standard New Persian, Hewramî and Southern Zazaki, *Ezafe* and the indefinite article can cooccur; there is no incompatibility. In Central Zazaki, it is clear that the indefinite article YAK is a recent morphological borrowing from Kurmancî. There are, therefore, no diachronic idiosyncrasies in their combination. For Soranî and Kurmancî, it is tempting to reconstruct a numeral YAK and indefinite article YAK as there is no obvious reflex of *aiwa in either language. If the idiosyncrasies of *Ezafe* in standard New Persian, Southern Zazaki and at least partially in Hewramî are in fact due to a historical incompatibility between indefinite and ezafe, then there would be a similar system regardless of the etymon of the indefinite article. The cooccurrence of both *Ezafe* and indefinite article suggests that the form of the indefinite article in Soranî and Kurmancî are more recent innovations albeit not as recent as in Central Zazaki. This innovative form can cooccur with *Ezafe* because definiteness was no longer conveyed by *Ezafe* at the time of its recruitment. Another possibility is that the former indefinite article EW could cooccur with YAK (not dissimilar to Şirvan Tat: *ye* N-*i*), and the indefinite article EW was reinterpreted as the indefinite ezafe.

4.3   Resolution of the historical questions

Ultimately, the origins of the indefinite, definite and absolute *ezafat* are deeply entangled. Soranî and Kurmancî have the same marker for the indefinite ezafe, but Kurmancî shows syncretism between definite and absolute while Soranî shows syncretism between indefinite and absolute. Is the Soranî indefinite ezafe an innovative form that is generalized to absolute contexts or is the Kurmancî definite ezafe

the innovation? Finding the origin of one of these will likely inform conclusions about all of them.

There is rich variation in the ways that attribution and definiteness marking interact within the Western Iranian languages. This paper presents a list of some of the ways these interactions take place. However, it is not an exhaustive list. There seems to be quite a bit of variation within each language. Many of the dialects that contain this variation are under-studied in general, and more work is necessary to fully explore these phenomena. The Iranian languages feature a rich mosaic of nominal modification strategies, each holding a piece of their etymological sources. This study represents the first step on the journey to understanding that mosaic.

A.   Abbreviations

YAK	The reflexes of PIIr. *aika
So.	The Soranî dialect of Central Kurdish
ŞT.	Şirvan Tat (Northern) (Suleymanov, 2020)
NP. (col.)	Colloquial New Persian (Samvelian, 2007, 2005)
NP.	Standard New Persian (Thackston, 1993)
Gi.	Gilaki (Rastorgueva et al., 2012)
DEF	Definite
ABS	Absolute
REZ	Reverse-ezafe
IPFV	Imperfective
1	First-person
M	Masculine gender
SG	Singular
PRS	Present
DIR	Direct case
DO	Direct object
COP	Copula
DOM	Differential object marking
EW	The reflexes of PIIr. *aiwa
Ku.	The Kurmancî dialect of Northern Kurdish
CT.	Chali Tati (Southern) (Yarshater, 1969)
He.	The Gorani varieties Lihonî (MacKenzie, 1966) and Paweyane (Holmberg & Odden, 2008)
SZ.	Southern Zazaki (Paul, 1998b)
CZ.	Central Zazaki (Werner, 2018)
INDF	Indefinite
EZ	Ezafe
DEM	Demonstrative

PRF	Perfect
3	Third-Person
F	Feminine gender
PL	Plural
PST	Past
OBL	Oblique case
LOC	Locative
CL	Clitic
IOM	Identified object marking

References

Bashir, Elena. 2008. Some Transitional Features of Eastern Balochi: An Areal and Diachronic Perspective. In Carina Jahani, Agnes Korn & Paul Titus (eds.), *The Baloch and others: Linguistic, historical and socio-political perspectives on pluralism in Balochistan*, 46–82. Wiesbaden: Reichert.

Berz, Koyo & Malmîsanij. 1988. *Na xumxum a …* Uppsala: Jina Nû Yayınları.

Durkin-Meisterernst, Desmond. 2004. *Dictionary of Manichaean Middle Persian and Parthian*. Turnhout, Belgium: Brepols.

Ekici, Deniz. 2007. *Kurmanji Kurdish Reader*. Dunwoody Press.

Franco-Rita, Ludovico & Manzini-Leonardo Savoia. 2012. Some call it article, some ezafe: "Linkers" as "agreement". *Linguistic Review* 2. 1–54.

Haig, Geoffrey. 2011. Linker, relativizer, nominalizer, tense-particle On the Ezafe in West Iranian. In Foong Ha Yap, Karen Grunow-Hårsta & Janick Wrona (eds.), *Nominalization in asian languages: Diachronic and typological perspectives* [*typological studies in language 96*], 363–390. Benjamins. https://doi.org/10.1075/tsl.96.13hai

Haig, Geoffrey. 2019. Debonding of inflectional morphology in Kurdish and beyond. In Songül Gündogdu, Ergin Opengin, Erik Anonby & Geoffrey Haig (eds.), *Current issuses in kurdish linguistics*, 117–144. Bamberg: University of Bamberg Press.

Haig, Geoffrey & Masoud Mohammadirad. 2019. Definiteness in Central Kurdish: sources and outcomes. Presented at the 4th International Conference on Kurdish Linguistics, University of Rouen, Paris-France.

Haig, Geoffrey & Ergin Öpengin. 2018. Kurmanji Kurdish in Turkey: structure, varieties, and status. In Christiane Bulut (ed.), *Linguistic minorities in turkey and turkic speaking minorities of the peripheries*, 157–230. Wiesbaden: Harrassowitz. https://doi.org/10.2307/j.ctvckq4v1.12

Hammarström, Harald, Robert Forkel, Martin Haspelmath & Sebastian Bank. 2020. Glottolog 4.2.1. https://doi.org/10.5281/zenodo.3754591. https://glottolog.org/

Heine, Bernd & Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge; New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511613463

Holmberg, Anders & David Odden. 2008. The Noun Phrase in Hawrami*. In Vida Samiian, Donald Stilo & Simin Karimi (eds.), *Aspects of iranian linguistics*, 129–152. Newcastle upon Tyne: Cambridge Scholars Publishing.

Huyse, Philip. 2017. The dialectology of Iranian. In Jared S. Klein, Brian D. Joseph & Matthias Fritz (eds.), *Handbook of comparative and historical indo-european linguistics band 1*, chap. 37, 599–608. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110261288-037

Jügel, Thomas. 2017. The syntax of Iranian. In Jared S. Klein, Brian D. Joseph & Matthias Andreas Fritz (eds.), *Handbook of comparative and historical indo-european linguistics band 1*, chap. 35, 549–566. Berlin: de Gruyter Mouton. https://doi.org/10.1515/9783110261288-035

Kahnemuyipour, Arsalan. 2014. Revisiting the Persian Ezafe construction: A roll-up movement analysis. *Lingua* 150. 1–24. https://doi.org/10.1016/j.lingua.2014.07.012

Karim, Shuan Osman. 2021. The synchrony and diachrony of New Western Iranian nominal morphosyntax: The Ohio State University dissertation.

Kent, Roland G. 1944. The Old Persian Relative and Article. *language* 20(1). 1–10. https://doi.org/10.2307/410376

Kent, Roland G. 1950. *Old Persian: Grammar Texts Lexicon*. New Haven, Connecticut: American Oriental Society.

Korn, Agnes. 2003. Balochi and the Concept of North-Western Iranian. In Carina Jahani & Agnes Korn (eds.), *The baloch and their neighbours: Ethnic and linguistic contact in balochistan in historical and modern times*, 49–60. Wiesbaden: Reichert.

Korn, Agnes. 2005. *Towards a Historical Grammar of Balochi: studies in Balochi historical phonology and vocabulary*. Wiesbaden: Reichert.

Korn, Agnes. 2016. A partial tree of Central Iranian. *Indogermanische Forschungen* 121(1). 401–434. https://doi.org/10.1515/if-2016-0021

Korn, Agnes. 2017. The evolution of Iranian. In Jared S. Klein, Brian D. Joseph & Matthias Fritz (eds.), *Handbook of comparative and historical indo-european linguistics band 1*, chap. 38, 608–624. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110261288-038

Kümmel, Martin. 2018. The survival of laryngeals in Iranian. In Lucien Beek, Alwin Kloekhorst, Guus Kroonen, Michaël Peyrot & Tijmen Pronk (eds.), *Farnah: Indo-iranian and indo-european studies in honor of sasha lubotsky*, 162–172. Ann Arbor, New York: Beech Stave Press.

Larson, Richard & Vida Samiian. 2020a. Ezafe, PP and the Nature of Nominalization. *Natural Language & Linguistic Theory* 39. 157–213. https://doi.org/10.1007/s11049-020-09471-1

Larson, Richard & Vida Samiian. 2020b. The Ezafe Construction Revisited. In Richard K. Larson, Sadigheh Moradi & Vida Samiian (eds.), *Advances in iranian linguistics*, 173–236. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.351.10lar

Mackenzie, David N. 1961. *Kurdish dialect, studies 1–2. studies 1–2*. Oxford: Oxford University Press.

MacKenzie, David N. 1966. *The dialect of Awroman (Hawraman-i Luhon): Grammatical sketch, texts, and vocabulary*. Kobënhavn: Kommissionaer : Munksgaard.

MacKenzie, David N. 1995. Proceedings of the second European Conference of Iranian Studies. In Bert G. Fragner (ed.), *Proceedings of the second european conference of iranian studies (held in bamberg, 30th september to 4th october 1991, by the societas iranologica europaea)*, 401–414. Istituto italiano per il Medio ed Estremo Oriente.

McCarus, Ernest. 1956. Descriptive analysis of the Kurdish of Sulaimaniya, Iraq, University of Michigan dissertation.

Nourzaei, Maryam, Carina Jahani, Erik Anonby & Abbas Ali Ahangar. 2015. *Koroshi A Corpus-based Grammatical Description*. Uppsala: Uppsala University. https://doi.org/10.1017/CBO9781107415324.004

Öpengin, Ergin. 2016. *The Mukri Variety of Central Kurdish*. Wiesbaden: Reichert. https://doi.org/10.29091/9783954906314

Pakendorf, Brigitte & Ija V. Krivoshapkina. 2014. Èven nominal evaluatives and the marking of definiteness. *Linguistic Typology* 18(2). 289–331. https://doi.org/10.1515/lingty-2014-0012

Paul, Ludwig. 1998a. The position of Zazaki among West Iranian languages. *Proceedings of the Third European Conference of Iranian Studies: held in Cambridge, 11th to 15th, 1995: Part 1 Old and Middle Iranian Studies (Beiträge zur Iranistik 17)* 163–177.

Paul, Ludwig. 1998b. *Zazaki : Grammatik und Versuch einer Dialektologie*. Wiesbaden: Reichert.

Paul, Ludwig. 2009. Zazaki. In Gernot. Windfuhr (ed.), *The iranian languages*, 545–586. London: Routledge.

Rasekh Mahand, Mohammad & Zaniar Naghshbandi. 2014. The effect of discourse factors on case system in Hawrami. *Language Related Research* 4(4). 87–109.

Rastorgueva, V. S., Aza Alimovna Kerimova, Akhmed Kerimovich Mamedzade, L. A. Pireĭko, D. I. Ėdel man & Ronald M. Lockwood. 2012. *The Gilaki language*. Uppsala: Uppsala University.

Samiian, Vida. 1994. The Ezafe Construction: Some Implications for the Theory of X-bar Syntax. In Mehdi Marashi (ed.), *Persian studies in north america*, 17–41. Bethesda, Md.: Iranbooks.

Samvelian, Pollet. 2005. When morphology does better than Syntax: The Ezafe construction in Persian. (*unpublished*).

Samvelian, Pollet. 2007. A (phrasal) affix analysis of the Persian Ezafe. *Journal of Linguistics* 43(3). 605–645. https://doi.org/10.1017/S0022226707004781

Skjærvø, Prods Oktor. 2009. Middle West Iranian. In Gernot Windfuhr (ed.), *The iranian languages*, chap. 4, 196–278. London: Routledge.

Soane, E B. 1913. *Grammar of the Kurmanji or Kurdish Language*. London: Luzac & Co.

Stilo, Donald. 2018. The Caspian region and South Azerbaijan: Caspian and Tatic. In Geoffrey Haig & Geoffrey Khan (eds.), *The languages and linguistics of western asia: An areal perspective*, 659–829. Berlin, Boston: Walter de Gruyter. https://doi.org/10.1515/9783110421682-019

Suleymanov, Murad. 2020. *A Grammar of Şirvan Tat (Beiträge zur Iranistik 46)*. Wiesbaden: Reichert. https://doi.org/10.29091/9783752005721

Thackston, Wheeler M. 1993. *An introduction to Persian*. Bethesda, Md.: Iranbooks.

Thackston, Wheeler M. 2006a. Kurmanji Kurdish: A Reference Grammar with Selected Readings. https://sites.fas.harvard.edu/{~}iranian/Kurmanji

Thackston, Wheeler M. 2006b. *Sorani Kurdish: A Reference Grammar with Selected Readings*. Cambridge, MA: Harvard University. https://sites.fas.harvard.edu/{~}Iranian/Sorani

Todd, Terry Lynn. 2002. *A Grammar of Dimili Also Known as Zaza*. Stockholm: Iremet Förlag.

van Emde Boas, Evert, Albert Rijksbaron, Luuk Huitink & Mathieu de Bakker. 2019. *The Cambridge grammar of Classical Greek*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781139027052

Werner, Brigitte. 2018. Forms and Meanings of the Ezafe in Zazaki. In Saloumeh Gholami (ed.), *Endangered iranian languages*, 71–91. Wiesbaden: Reichert.

Whitney, William Dwight. 1993. *Sanskrit Grammar*. Cambridge, Mass: Harvard University Press 2nd edn.

Yarshater, Ehsan. 1969. *A Grammar of the Southern Tati Dialects*. The Hague: Mouton.