

De-collective plurals and the development of slot competition between definiteness and plurality: Evidence from Iranic

Shuan Osman Karim

*Lehrstuhl für Vergleichende Sprachwissenschaft, Julius-Maximilians-Universität Würzburg, Campus Hubland Nord,
Oswald-Külpe-Weg 84, Würzburg 97074, Germany*

Received 21 January 2026; revised 5 July 2026; accepted in revised form 5 July 2026;



Abstract

This study explores idiosyncratic slot competition between definiteness and plurality that characterizes languages like Southern Kurdish and Colloquial New Persian. One strategy by which Iranian languages have developed definiteness marking is by the recruitment of diminutive/evaluative morphology. A strategy by which they have developed plural marking is by recruiting collective nouns/morphemes. The relative chronology of these developments dictates whether there is slot-competition or not. If these languages had a plural marker before the recruitment of the definite article, they would feature a definiteness distinction in plural contexts. However, if the definiteness distinction was the earlier innovation, plural nouns do not feature definiteness marking. This competition results from persistent features of their etyma, which can be used as heuristics for identifying the origins of innovative morphology.

© 2026 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Keywords: Diachrony; Typology; Definiteness; Collective; Analogy

1. INTRODUCTION

Across the Iranian languages, there are innovative definiteness marking suffixes. These were perhaps emergent in Middle Iranian (e.g., Middle Persian, see [Nourzaei and Jügel, 2023](#)), and they continue to emerge in New Iranian languages (e.g., Koroshi Balochi, [Nourzaei, 2020](#)). One distinctive feature of these markers is that, in some languages, they are incompatible with plural marking. In other words, there is competition¹ for realization between the morphemes that express definiteness and plurality. This incompatibility is not semantically motivated. As a result, some scholars

E-mail address: shuan.karim@uni-wuerzburg.de

¹ I use the term competition here to mean that when a plural noun appears in definite contexts, the formative marking definiteness is blocked by the appearance of the formative marking plural. The concept of competition between morphemes has been well studied in the context of inferential realizational approaches to morphological theory. My focus here is on the development of these systems, not their synchronic functions. For more on competition in morphological theory, see [Stump \(2001\)](#), [Stump and Finkel \(2015\)](#); [Rainer et al. \(2019\)](#), [Ackerman and Stump \(2004\)](#).

have questioned the status of these formatives as definiteness markers, labeling them idiosyncratically, e.g., specificity marker, deictic marker, etc.

Another source of morphological innovation across the Iranian world is the use of plural affixes from collective nominal and affixal sources. This study looks at the typological variation in definiteness and plurality marking strategies to identify the source of their incompatibility. Ultimately, I show that there is only a diachronic explanation for their incompatibility: the relative chronology of these innovations dictates whether or not definiteness and plurality markers can or cannot co-occur. Furthermore, their co-occurrence is one of many persistent² features that one can use to reconstruct de-collective plural markers and de-evaluative definite markers after their grammaticalization.

In Section 2, I show the forms that have developed innovative definiteness marking and their various sources. I explore innovative plural morphology in Section 3, outlining the system's morphological and nominal sources. In Section 4, I show how the source of the definiteness marker (evaluative/demonstrative) and the plural marker (morphological/nominal) can motivate competition between definiteness and plurality.

2. DEFINITENESS MARKERS FROM EVALUATIVE SOURCES

There are at least four definiteness marking strategies in Iranian languages: (1) definite suffixes, (2) demonstratives, (3) differential case marking (including =RĀ), and (4) no marker. Of these, only the first is relevant to this discussion. There are several definiteness-marking suffixes in Iranian languages, including the Paweyane *-eke* and colloquial New Persian *-(h)e*. These formatives are now thought to come from diminutive/evaluative morphology, as Pakendorf and Krivoshapkina (2014) has proposed for the Siberian language Even. In some Iranian varieties of Balochi, e.g., Koroshi, Sistani, and Coastal, the stages of development from evaluative to definite can be observed as an emergent feature (Nourzaei, 2020).

Nourzaei (2020) follows Pakendorf and Krivoshapkina's (2014) proposal, building upon Haig and Mohammadirad (2019), to classify the Balochi suffix *-ok* on a cline from evaluative to definite. Coastal Balochi features the inherited *evaluative usage*, i.e., it is possible in "non-definite contexts," there are "no structural constraints," and their frequency is "dependent on context, speaker, and content." Sistani Balochi features the inherited *evaluative usage* and has been extended to *deictic and recognitional contexts*, i.e., they occur "in combination with proximal demonstratives." In contrast, Koroshi Balochi features the use of the *k*-form suffix in *deictic and recognitional contexts* and as an *anaphoric definite* but not the inherited evaluative function.³ This nascent definiteness marker does not have the same distribution as the well-established definiteness markers that have demonstrative etyma, "the most frequent way in which definite articles evolve" (Heine and Kuteva, 2002, 110).

The evaluative usage can be observed synchronically in Koroshi. However, the Indo-Iranian evaluative suffix *-(a)ka-*, the ultimate etymon of Koroshi Balochi *-ok*, has traditionally been described as diminutive. However, this is a grammatical label assigned to this formative by scholars and does not adequately reflect its distribution; 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). In other words, the diminutive label may be too narrow. Regardless, suffixes that feature the *-KA* formative or are known reflexes of **(a)ka-* are widely attested across the Iranian world. These formatives are challenging to study due to a lack of documentary work on many of these languages and inconsistent terminology. They are referred to differently by different scholars, e.g., definite (Mackenzie, 1961; MacKenzie, 1966; Mahmoudveysi et al., 2012; Mahmoudveysi and Bailey, 2013; Öpengin, 2016, etc.), demarcative (McKinnon, 2011), determinative (Windfuhr, 2012), or deictic (Windfuhr, 1991).

These markers include the K-type markers Emāmzāda Esmā'īlī (Fars): *doft-ak-ō* 'the girls [girl-DEF-PL]' (Windfuhr, 2012), Bušehrī (Fars): *havā-y-akū* 'this weather' (Windfuhr, 2012), Gīonī (Lor): *asp-Ø gap-eka* 'the big horse [horse-DEF.EZ big-DEF]' (McKinnon, 2001), Northern Lori: *-(e)ka* (McKinnon, 2011), Dezfuli and Šuštari (S Lori): *-aka* (McKinnon, 2011), Bakhtiāri (S Lori): *-ekū* (McKinnon, 2011), *-(e)ke* (Anonby et al., 2019, 452), Central Kurdish: *-eke* (Mackenzie, 1961), Southern Kurdish: *-eke*, *-ege* and *-eġe* (Fattah, 2000, 245), Hewramī (Lihon/Text): *-eke* (MacKenzie, 1966; Mohammadirad, 2026), Paweyane: *-eke* (Holmberg and Odden, 2008), Zerdeyane: *-aka* (Mahmoudveysi and Bailey, 2013), and Gewrecūi: *-eke* (Mahmoudveysi et al., 2012). Additionally, these markers

² Here, I use the term persistent in the linguistic sense. Following Hopper (1991), persistence is the tendency for forms to retain features of their etyma.

³ According to Karim (2021, 101), the fact that this usage is present in Koroshi Balochi but not its closest relatives points to the shift from evaluative to definite being a more recent innovation.

include forms without the *k* that are likely to share the same etymon⁴ *-(a)ka- Sīvāndī: -u [M.SG.DEF] and -e [F.SG.DEF] (Windfuhr, 1991), Judeo Isfahānī: -e [SG.DEF] (Windfuhr, 1991), Khuri: -u [SG.DEF] (Windfuhr, 1991), Kermani languages: -u (Borjān, 2017a), the Median dialects (Kašān): a/-e (Borjān, 2012), Kešā'i: -é (Borjān, 2017b), Kumzari: -ō (van der Wal Anonby, 2019, 631), and Colloquial New Persian: -(h)e (Jahānī, 2015; Nourzaei, 2023a). Following Karim (2021), most of these languages share syncretism between the definiteness marker and a diminutive/evaluative suffix.

3. COLLECTIVE PLURALS IN IRANIC

There are many plural-marking strategies in Iranian languages. Among the Western Iranian languages,⁵ there are systems that Karim (2021) has referred to as 'layered' because they feature innovative markers attached to historically inflected forms. For instance, Table 1 shows a system where non-definite nouns show one unmarked form for direct, oblique, and direct object cases regardless of number, while the genitive and locative are distinguished with unique forms sensitive to number.

In contrast, definite nouns are unmarked in the direct case but distinguish singular and plural and the oblique, direct object, genitive, and locative cases with unique markers. This reflects Stilo's (2008) proposal that Iranian languages feature nominal systems built on two axes, the "reduction axis" and the "innovation axis."

Iranian languages have lost a great deal of morphology due to the "Rhythmic Law," where PIE strong stems lost their word-final codas and most of the case system along with it (Sims-Williams, 1982). In the wake of this loss, new material is recruited for case and definiteness marking. The plural markers in the various systems come from inherited and innovative sources. The inherited plural is usually -ān, from the Indo-Iranic genitive plural *-ānām, which lost its word-final coda due to the aforementioned Rhythmic Law. The innovative plural markers come from collective nouns and affixes that have univerbated with the head noun. For Colloquial New Persian, the innovative marker is -(h)ā. Sundermann proposed that -(h)ā is a combination between the abstract suffix -īh and the plural suffix -ā(n) used as a collective marker (Sundermann, 1989). However, Sims-Williams (2004) has the Old Iranian abstract suffix *-iyaθwa, as the etymon. Both etymologies have problems; for instance, the loss of *n* is not a regular sound change. In other words, if the *n* was lost, it must have been lost due to analogical forces like boundary reanalysis, which has never been proposed. The lack of a clear etymon is not an issue for this analysis; what is important is that the suffix -hā does not occur as such until early New Persian, placing its development after the development of the definite suffix. Additionally, its attested development shows characteristics of plurals from collective sources.

For the Sogdian language (Middle Iranian), the innovative plural marker is -t-, from a collective suffix in *ta*, e.g., Avestan: *yazata*- 'divinities.' These forms are declined as if singular (see Yoshida, 2009 for the archaic *aka-ān* ending). In Šebekī (MacKenzie, 1956), Āštīāni (Tati; Windfuhr, 1998), and Sīvāndī (Windfuhr, 1991), the innovative plural markers are from a word meaning 'flock, troop' -geṭ, *ger*, and -gel, respectively. For Southern Kurdish, there are various innovative plural

Table 1
Turkmen Balochi case (Karim, 2021, 86, following Axenov, 2006).

	SG	PL	DEF.SG	DEF.PL
DIR		-Ø		-Ø
OBL			-ā	-ān
OBJ			-ā-rā	-ān-ā
GEN	-ay ⁶	-ān-ī	-ay	-ān-ī
OBL.II(LOC)	-ay-ā	-ān-ī-ā	-ay ⁵ -ā	-ān-ī-ā

⁶ The genitive singular suffix is a contracted form of the oblique singular -ā and the genitive suffix -ī/-ay (Axenov, 2006: 72).

⁴ In some Iranian Varieties, the intervocalic *k voiced to *g*, e.g., Middle Persian -ag. The final *g* was subsequently lost, resulting in New Persian -a/-e. Not enough is known about all the varieties here to say whether or not similar processes are responsible for the other *k*-less articles. See Nourzaei (2023b) for a similar development from *K*-form to -ū in Shirazi.

⁵ Western Iranian is a now-defunct designation containing the groups Southwestern Iranian and Northwestern Iranian. Korn (2019) has proposed that the Iranian languages should be divided into a core Central Iranian, including the old Northwestern designation, and a series of peripheral groups that are not easily sub-categorized. I use the term Western Iranian here because I am examining Southwestern Iranian and Northwestern Iranian, and the phenomena targeted by this study span any proposed subdivisions. In other words, the de-collective plurals and definiteness markers from evaluative morphology are areal features not necessarily genealogical ones.

Table 2
Definiteness and number interactions in Colloquial New Persian.

	INDEF.SG	DEF.SG	DEF/INDEF.PL
DIR	<i>ketāb</i>	<i>ketāb-e</i>	<i>ketāb-ā</i>
ACC	<i>ketāb</i>	<i>ketāb-æ-ro</i>	<i>ketāb-ā(-ro)</i>

suffixes all transparently from the same etymon, *-geʔ*, *-eyʔ*, *-êʔ*. Kurmancî has an indefinite plural marker *-in* from the word *hin* ‘some.’.

For Koroshi Balochi (Nourzaei et al., 2015), the innovative plural marker is *-obār*. According to Haig (2019), there is as of yet no convincing etymology for the formative *-obār*. However, I propose here that this is from a collective noun *bār* ‘load’ and the formative *o* ‘and,’ which is used in forming compounds.⁶ The reason for this proposed etymon of *-obar* is discussed in the following Section 4.

4. THE COMPETITION BETWEEN DEFINITENESS AND PLURALITY

In many Iranic languages, the innovative plural and definiteness markers cannot co-occur. When a noun is both plural and definite, only the plural marker surfaces. For instance, the colloquial New Persian definiteness and plural markers “are in complementary distribution,” as exemplified in (1-b) (Samvelian, 2007, 618).

- (1) a. **ketāb-e-hā*
book-DEF-PL
b. **ketāb-hā-e*
book-PL-DEF
(putatively) ‘the books’ (Samvelian, 2007, 619)

The Colloquial New Persian nominal paradigm has layers similar to the Turkmen Balochi paradigm in Table 1. The Colloquial New Persian paradigm is summarized in Table 2. The indefinite singular⁷ is unmarked. Definite singular is marked with the definite article *-(h)e* and both the definite article and definite direct object marker *-ro* in the accusative. However, in the plural, the expected combined form **ketāb-æ-hā* or **ketāb-ā-he* does not occur.

Likewise, Şebekî and Koroshi (Balochi) have innovative plural markers, *-geʔ* and *-obār*, respectively. These plural markers occur in both definite and indefinite contexts. However, only singular nouns are overtly marked for definiteness, e.g., Koroshi (Balochi): *oştör-ok* [camel-DEF] ‘the camel’ - *oştör* ‘(a) camel’ - *oştör-ay* ‘of (a) camel’ - *oştör-obār* ‘(the) camels,’ etc. (Nourzaei et al., 2015). Eastern Hill Balochi, the Southern Kurdish varieties of Bîcar, Qorwe, Bîlewar, Çemçemal, Mendilî, Sencawî, Serpol, Xalêse, Dînewer, Duşeyx, Ke/hor, Kirmanşa, Kolyay, Qesri Şîrîn, Sehene, and Kehrîz feature innovative plural markers, while maintaining a definiteness distinction in the singular and plural. However, the plural suffix and the definiteness marker are incompatible. When the definiteness marker is employed, the inherited plural marker *-an* is used, e.g., Kirmanşa (SK): *jîn* ‘woman’ *jîn-ege* [woman-DEF] ‘the woman’ *zimsan-eyl* [winter-PL] ‘(the) winters’ *mař-eg-an* [snake-DEF-PL] ‘the snakes.’.

5. DISCUSSION

There are persistent tendencies associated with plural markers descended from collectives. (1) These formatives carry singular inflection, (2) they feature hybrid agreement, (3) they are often incompatible with high-animacy nouns, and (4) they are incompatible with definiteness marking. These four persistent tendencies can be used as heuristics for identifying de-collective plural markers. In the languages with the inherited plural marker from **ānām*, the subsequent development of a de-evaluative definiteness marker adds a definiteness distinction in both the singular and plural. However, the subsequent innovation of a new plural marker from a collective source disrupts the system, generating competition between definiteness and plurality because of their persistent tendencies.

⁶ See Karim (2021, 183) for a formal discussion of the use of ‘and’ as a compounding device.

⁷ In this case, indefinite is probably better described as non-definite or generic. In New Persian, there is a specific indefinite marker *-i* which is used to mean ‘a certain X’ and can co-occur with the definite direct object marker *-ro*.

Table 3
 “Declension of [Sogdian] Heavy Stems” (Yoshida, 2009, 289).

	M.SG ‘day’	PL	F.SG ‘life’	PL	cf. PL ‘gods’
DIR	<i>mēθ</i>	<i>mēθ-t</i>	<i>žwān</i>	<i>žwān-t</i>	<i>βaγ</i>
OBL	<i>mēθ-t̄</i>	<i>mēθ-t-t̄</i>	<i>žwān-t̄</i>	<i>žwān-t-t̄</i>	<i>βaγ-ān</i>

5.1. Collective plurals carry singular inflection

Plural markers from collective sources feature singular morphology. For instance, the Sogdian collective plural *-t* takes the singular case endings of heavy stems, e.g., direct singular *-Ø*, oblique singular *-t̄*, and vocative singular *-ē*; see Table 3. This stands in contrast to the inherited oblique plural marker *-ān*, only used in a few archaic constructions, e.g., *βaγānβaxtam* ‘godliest of gods’ (Yoshida, 2009, 290). Note that the collective morpheme *-t* is something that dates back to a common Proto-Indo-Iranic, as evident from Sanskrit *deva-ta* ‘divinities.’

Just like Sogdian, Koroshi Balochi features singular suffixes on the oblique plural marker, e.g., *janek-obār-ay* ‘daughter[PL-GEN] *pūldā-obār-ā* ‘rich[PL-OBL]’ (Nourzaei et al., 2015, 187–188). Likewise, the New Persian indefinite singular suffix can co-occur with the collective plural marker *-(h)ā*, e.g., *ketāb-hā-i* ‘book[PL-INDEF]’ (Thackston, 1993, 4). This development is easily explained by the fact that the collective is seen as a singular entity. Compare the English phrases *a flock of birds*, *a load of junk*, and *a ton of boxes*. Each of these collectives *flock*, *load*, and *ton* receive the singular indefinite article *a* even though they refer to a plurality.

5.2. Collective plurals feature hybrid agreement

According to Gardelle (2019, 27), Hybrid agreement occurs for collective nouns when the NP that they head both licenses syntactic agreement (the singular, for a head noun in the singular) and overrides plural agreement (the verb or the anaphoric pronoun is in the plural although the subject head noun is in the singular). In the English sentences in (2), Gardelle (2019) shows that the same noun *audience* can license both singular and plural agreement, the former conditioned by the singular noun *audience* and the latter conditioned by the plurality of the referent (multiple people that make up the audience).

- (2) a. *The audience was invited to turn up early and picnic on the grounds* (Gardelle, 2019, 27).
 b. *The audience were in their seats for the first of the two evening performances* (Gardelle, 2019, 27).

There is competition between the grammatical singularity of the collective noun and the discourse-level collective reference, reflected in the Persian sentences in (3). The two sentences are identical except for the verbal agreement. In (3-a), the singular copula agrees with the grammatically singular collective, and in (3-b), the verb agrees with the semantic plurality.

- (3) a. *in ketāb-hā xub-ast*
 DEM.PROX book-PL good-COP.SG
 b. *in ketāb-hā xub-and*
 DEM.PROX book-PL good-COP.PL
 ‘The books are good’ (Thackston, 1993, 6–7)

However, this interpretation of these data stands in contrast with the typical interpretation that the singular agreement is possible because of the animacy of the noun; ‘[a]s a general rule, plurals of inanimate nouns take singular verbs’ (Thackston, 1993, 6). Gardelle (2019, 100) argues that Hybrid agreement “could not be regarded as a defining feature of the class of collective nouns; rather, it is an addition to the system, as an effect of the Animacy Hierarchy.” Following Gardelle’s (2019) conclusion, I propose that this be considered subordinate to the following heuristic of animacy.

5.3. Collective plurals are often incompatible with high-animacy nouns

Here, I assume that collective plurals from nominal sources, including those in Iranic languages, begin as partitive constructions, constructions referring to a part or subset of a definite superset, e.g., measure phrases like *a flock of birds*. Following Rutkowski (2007), the partitive construction undergoes a process of grammaticalization, becoming a pseudopartitive (coined by Selkirk, 1977). Pseudo-partitives are defined as “expressions referring to an amount/quantity of some (indefinite) substance rather than to a part/subset of a (definite) superset” (Koptjevskaja-Tamm, 2001 apud

Rutkowski, 2007). These constructions are identical to partitives in their surface structure. However, they have different underlying syntactic structures (for a syntactic analysis, see Stickney, 2004). These pseudo-partitive constructions become common first with nominals on the low end of the Animacy Hierarchy (4).

- (4) The Animacy Hierarchy — speaker > addressee > 3rd person > kin > human > animate > inanimate (Corbett, 2000, 56)

The animacy sensitivity of the collective plurals can be understood as one persistent aspect of collective plurals. This feature of collective plurals is evident in the restricted use of the inherited plural from the genitive plural suffix **-ānām*. In Sogdian, the *-ān* suffix only occurs with a small set of animate nouns, e.g., *βaγ-ān* ‘gods,’ *pīd-ān* ‘elephants,’ *martaxm-ān* ‘man.’ Likewise, early/classical New Persian showed the inherited *-ān* plural suffix used idiosyncratically with animate nouns. According to Thackston (1993, 58), “[s]ubstantives denoting animate beings, parts of the body that occur in pairs, and a few other nouns (determined by idiomatic usage) for their plurals by suffixing *-ān*.” These include third person pronouns *ān-ān* [DEM.DIST-PL] ‘those/they,’ nouns denoting people *dānešju-yān* [student-PL] ‘students,’ body parts *dast-ān* [hand-PL] ‘hands,’ nouns denoting living things *deraxt-ān* [tree-PL] ‘trees,’ and nouns with the evaluative suffix *bačč-eg-ān* [child-EVAL-PL] ‘children.’ In Koroshi Balochi, the plural suffix *-obār*, which is now widespread, does not occur with third-person singular pronominal forms, e.g., *eš-ān* not **eš-obār*. To my knowledge, no Iranic languages show innovative collective plural markers on first- and second-person pronouns.

In colloquial New Persian, the innovative plural marker *-(h)ā* occurs productively with all nouns and third-person pronouns. At the same time, its use was originally most prominent on low-animacy nouns. This shows a slow whittling away of the persistent qualities of the collective plural as it becomes fully grammaticalized and productive.

5.4. Collective plurals are incompatible with definiteness marking

As discussed in Section 4, collective markers are incompatible with definiteness markers. This competition is not the product of a semantic incompatibility but rather a persistent feature of the plural markers’ pseudo-partitive etyma, which refer to a quantity of some indefinite substance, not to a part/subset of a definite superset. Collective plurals tend to be incompatible with definiteness marking. In the English examples, *load* and *ton* are often used to describe things that are not loaded or heavy. In these cases (5), a simple plural (a) and the collective with the indefinite article are felicitous. In contrast, the collective *ton* with the definite article is infelicitous.

- (5) a. *I carried the boxes.*
 b. *I carried a ton of boxes.*
 c. **I carried the ton of boxes.*

Note that this phenomenon does not hold for things that come in collectives, e.g., *Did you see the flock of birds?* This incompatibility is a feature of many but not all languages with innovative plural markers. Some languages allow for the combination of the definite article and the innovative plural marker. However, there is reason to believe this is a secondary development with a clear grammaticalization path.

In Table 4, I show four groups of languages: (1) languages with a definiteness distinction in the singular and plural from inherited sources, e.g., the evaluative suffix **(a)ka-* and the genitive plural **-ānām*; (2) languages with a definiteness distinction in the singular and plural from the inherited suffixes **(a)ka-* and **-ānām* but with an innovative non-definite plural marker from a collective source; (3) languages where the collective plural ending has been generalized as the only plural marker, displacing the inherited definite-plural marking, and (4) languages with an innovative definite-plural marker built from the productive plural marker and definite suffix. Note the morphemic order in group (4) is [-PL-DEF], running contrary to groups (1) and (2), which show the order [-DEF-PL].⁸ It should be understood that the size of each category in Table 4 does not correspond to the commonality of any particular construction. Every language/variety in group (4) is classified as Southern Kurdish (or Laki), which share a recent common ancestor. Group (3) shows the most diversity regarding sub-groupings of Iranic that feature that system, but they seem to show variable time depths. The distribution of these types can be observed in Fig. 1, where the bias to documented data from Kurdish is apparent.

⁸ Haig (2019, 119) states that the original order DEF-PL seems to violate Bybee’s (1985) Relevance Principle, “morphology with higher relevance will be closer to the stem.” Assuming an idiosyncratic grammaticalization path that created the -DEF-PL *-ek-an* forms, the shift to -PL-DEF *-ê/e* appears to be a correction (See Haig, 2019 for a discussion of these developments).

Table 4
Definiteness and number interaction across Western Iranian.

	Language/variety	DEF.SG	DEF.PL	PL
1	Turkmen Balochi	-Ø/ā (DIR/OBL)	-Ø/ān (DIR/OBL)	-Ø
	Central Kurdish	-eke	ek-an	-an
	Hewramî (Luhon)	-aka/ê	-ek-ā	-a
	Paweyane	-eke/ê (f)	-ek-an	-an
2	Eastern Hill Balochi	-Ø/ā (DIR/OBL)	-Ø/ān (DIR/OBL)	-yal
	Bîcar (SK)	-e(ġe)	-eġan	-ġel
	Qorwe (SK)	-e(ġe)	-eġan	-eyl
	Southern Kurdish 1 ¹⁰	-e(ġe)	-eġan	-ġel
	Southern Kurdish 2 ¹⁰	-ege	-egan	-eyl
	part Kethor (SK)	-eke	-ekan	-gêl
3	Colloquial New Persian	-(h)e		-(h)ā
	Şebekî	-eke		-geġ
	Koroshi (Balochi)	-ok		-obâr/-obâr-ā
	Āštîāni	-Ø/i/a/o (M/F/OBL.M/OBL.F)		-gal/-gal-ān (DIR/OBL)
4	part Duşeyx (SK)	-e(ġe)	-eyle	-eyl
	Heresem (SK)	-e(ge)	-eyle	-eyl
	Îwan (SK)	-e(ġe)	-eyle(ġa)	-eyl
	Bîsitun (SK)	-e(ge)	-êyle(ge)	-eyl
	Kordelî (SK)	-e(ke)	-ele(ke)	-el
	Beyrey (SK)	-e(ge)	-êle(ge)	-êl
	Çîhr (SK)	-e	-yele	-yêl
	Southern Kurdish 3 ¹⁰	-e	-ele	-el

¹⁰ There is great variation within Southern Kurdish. It is not uncommon for people in the same city to speak different dialects. The best description of Southern Kurdish varieties is Fattah (2000), which is fairly comprehensive in that he includes all the variations he encountered. However, the work does not show a complete picture of any variety. For instance, Fattah (2000) states that in Beyrey, some speakers have only the definite article -e, while others have both -e and -ege. Likewise, he states that a great part of Beyrey has the plural marker -êl, while only a small part has the plural marker -el. He then states that the definite plural is either -êle(ge) or -ele(ge). One could assume that the speakers that have only -e would show the definite plural -ele or -êle depending on whether they had the plural markers -êl or -el. However, it is unclear from the text whether each combination exists (two definite markers, two plural markers, and four definite plural markers). What Fattah (2000) gains in thoroughness, he loses in painting a clear picture of any given variety (idiolect). Here, I take some liberties in assuming that the variation observed here is across different speakers in the same cities, towns, and villages and not within the speech of single speakers. Additionally, I assume coherent systems, e.g., a language with the definiteness marker -e and the plural marker -el would have the definite marker -ele (not -êlege) (see Fattah 2000: 245-254). Southern Kurdish1: Bîlewar, Çemçemal., Mendiî, Sencawî, Serpol, Xâlesa; Southern Kurdish 2: Dînewer, Duşeyx, Kethor, Kirmanşa, Kolyay, Qesri Şîrîn, Sehene; Southern Kurdish 3: Hersîn, Myexas, Şêrwan.

5.5. Relative chronology

Group (1) features the inherited definiteness and plural markers. As such, this is the inherited schema. Here, there definiteness and plurality are fully compatible. Both the *k*-form definiteness marker and the plural marker -ān were already part of the language in Proto-Indo-Iranic as evidenced by forms in Sanskrit and Avestan. Group (2) represents the first innovation, recruiting the collective marker as a new non-definite plural. These forms likely began with low-animacy nouns, and none of the languages in this category show the innovative plural marker on pronouns. Group (3) has extended the collective plural to definite contexts, usurping the function of the combination of definite and plural *-egan. Note that some of these languages now use the collective plural with third-person pronouns, e.g., Persian: *inhâ* 'these,' Şebekî: *inegle* 'these.' Others have maintained the inherited plural on the same pronouns, e.g., Koroshi (Balochi): *eşân* 'these.' These instances of climbing the animacy hierarchy (4) can be seen as continued development, as these formatives lose properties of their etyma. Finally, group (4) no longer features competition between definiteness and plurality marking. All of these languages are Southern Kurdish varieties. They clearly show the innovative plural marker on the inside of the definiteness markers -e and -eke/-ege/eġe, although -e is more widespread than the *k*-form examples. Haig (2019) argues against the view that the morpheme order observed in Central Kurdish [DEF-PL]

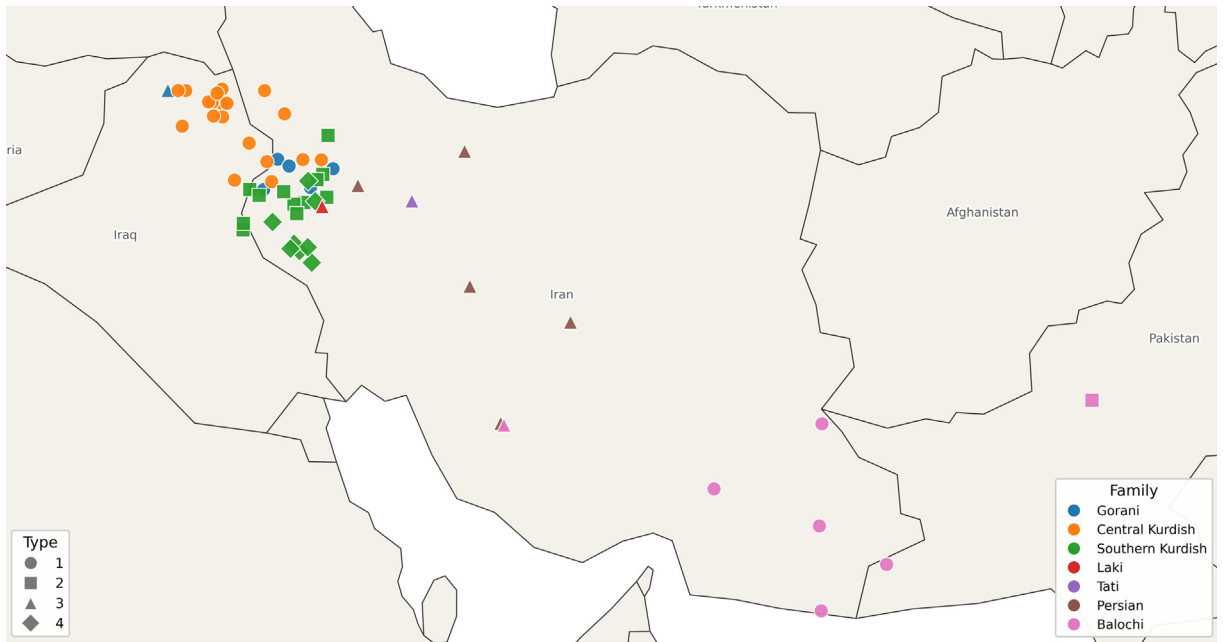


Fig. 1. Distribution of Definiteness Types. Map lines delineate study areas and do not necessarily depict accepted national boundaries.

-ek-an is debonding, a separation of bound morphology contrary to grammaticalization. Assuming that he is correct, it seems that group (4) is a more convincing example of debonding.⁹

The debonding of the definiteness suffix can be explained by the presence of a medial form that provides the analogical exemplar for this development. In the Kurdish-zone languages, demonstrative constructions are bipartite, consisting of a prepositive and post-positive element. For instance, Suleymanî (Central Kurdish) has *em=e* [DEM.PROX = DEM.POST] - *em şt-an=e* [DEM.PROX thing-PL=DEM.POST], where the enclitic demonstrative postposition *=e* occurs at the end of the noun phrase, and the demonstrative *em* is preposed. Likewise, Şebekî (Group 3) has *in=e* [DEM.PROX=DEM.POST] - *in tişt-ge=e* [DEM.PROX thing-PL=DEM.POST], following the same pattern but with the innovative plural marker. In some varieties scattered throughout the Kurdish zone, the demonstrative clitic can be used as a definiteness marker alongside the *k*-form marker.

- (6) N(P)=e :: N-*eke* (both [N.SG.DEF])
N-*ge*=e :: X (both [N.PL.DEF])

This analogy is represented in (6), where all three forms are possible inherited constructions, and X = N-*ge*-*eke*. The definiteness marker from the demonstrative postposition *=e* is the most common way of expressing definiteness in the southern part of Southern Kurdish (Group 4); in six varieties, the *k*-form definiteness suffix has been extended to those same plural forms, effectively repairing the incompatibility between plural and definiteness with the innovative definiteness marker *-a*. This marks the end of persistence in these varieties.

5.6. Āštīānī

In group (3), Āštīānī *-galān* (Windfuhr, 1991), seems to be a combination of the inherited and innovative plural markers. This is likely a recombination of the two suffixes, i.e., the inherited suffix on the collective noun *gal* 'flock/troop,' which was grammatically singular, referring to a semantic plurality. This only occurs in the oblique case and should

⁹ Note that the geographic distribution in Fig. 1 seems to reflect this relative chronology, the conservative Type 1 having the broader distribution. However, the map should not be interpreted in this way. The Balochi points in the far southeast of the map are old transplants from Northwestern Iran. The chronology is based rather on what is known about Old Iranian: **-aka* and **-ān*, and the fact that varieties within each subgroup have innovated different collective markers.

be understood as the innovative plural marker and a case-marking suffix of the same etymology as the plural marker from the rest of these languages.

5.7. Balochi

In Table 4, I have included Turkmen (Axenov, 2006) and Eastern Hill Balochi (Elfenbein, 1980). However, the oblique suffix *-ā* is not clearly descended from the evaluative marker **(a)ka-* (suggested as a possibility by Karim, 2021).

Koroshi (Balochi) *-obār*: The Koroshi (Balochi Nourzaei et al., 2015) plural marker shares the same distributional idiosyncrasies as the plural markers known to be descended from collective nouns, e.g., Şebekî *gaḥ* 'flock.' For instance, (1) the plural suffix *-obār* governs singular agreement; (2) the suffix is incompatible with definiteness marking like the colloquial New Persian *-ha*; (3) the suffix is inflected with singular endings for oblique/object and genitive, and (4) it has not yet been extended to third-person singular pronouns. These four facts lead to the conclusion that the Koroshi Balochi suffix *obār* is also a collective plural. It is likely related to the term *kārobār* 'work (lit. work and a load).' As in many other Iranian languages, the word 'and' acts as a compounding marker, not as a simple conjunction.¹⁰ The resulting *obār*-marked compounds are something akin to a load of a noun or a noun and related things, then plural.¹¹

6. CONCLUSION

The seemingly odd competition between definiteness and plurality in Iranian languages is just another example of persistence, where 'when a form undergoes grammaticization from a lexical to a grammatical function, so long as it is grammatically viable, some traces of its original lexical meanings tend to adhere to it, and details of its lexical history may be reflected in constraints on its grammatical distribution' (Hopper, 1991: 22). When reconstructing the origin of plural markers, one can determine that a particular formative has its origins in a collective noun by several persistent tendencies: (1) Singular inflection, (2) Hybrid agreement, (3) Incompatibility with high-animacy nouns, and (4) Incompatibility with definiteness marking. However, heuristic (2) may be epiphenomenal, i.e., due to the low animacy of the nouns, to which the innovative plural marker attaches (following Gardelle, 2019: 100).¹²

The diminutive/evaluative source of the definiteness markers in Iranian languages has its own set of persistent qualities; see Haig (2019), Nourzaei (2020, 2023a) for more on the development and features of these formatives. The interaction between two aspects of the definiteness and plural markers likely led to their incompatibility: (1) Nouns marked by the diminutive/evaluative formative are evaluable, expressing sorrow, affection, admiration, endearment, small size, small portion, pejoration, or derivation, which naturally differentiates the noun from members of the same class. These nouns cover kinship terms, humans, animate beings, and inanimate objects. They range the animacy hierarchy (4) covering all except the speaker, addressee, and third-person, reflecting an incompatibility with pronominal forms, distributionally similar to definite articles. This contrasts with the innovative collective plural markers that begin on the low end of the animacy hierarchy; they are by nature undifferentiated and incompatible with a specific or definite reading. (2) The relative chronology of these formatives allowed the plural marker to usurp the domain of the combination of the definiteness marker and inherited plural marker. The collective plural marker replaced the older construction, retaining its incompatibility with the definiteness marker as a persistent feature of the two formatives. This incompatibility is not otherwise explicable.

These facts allow us to make several predictions about the diachronic typology. Specific changes occur as collective plurals extend up the animacy hierarchy to more domains. First, the plural marker moves from inanimate to animate nouns, humans, kinship terms, and third-person pronouns. The domains of the definite article in plural contexts get usurped by the innovative plural, first partially, then completely, when it begins to be used with kinship terms. Only then does the innovative plural come to be used with third-person pronouns. Likely, the suffix will not appear on third-person pronouns until it has completely taken over the definite contexts, i.e., a language that has an innovative plural on pronouns must be either a language from Group (3) or (4), showing one innovative plural marker for definite and indefinite contexts or having passed through such a stage. Finally, we can use these heuristics to reconstruct the etyma of

¹⁰ See Karim (2021) for a theoretical explanation of the use of *û* 'and' as a compounding marker.

¹¹ Note that the forms marked with *-obār* can take the singular case-marking suffixes; compare: Singular: *janek-Ø* [woman-NOM], *janek-ā* [woman-OBL], *janek-ay* [woman-GEN]; Plural: *janek-obār-Ø* [woman-PL-NOM], *janek-obār-ā* [woman-PL-OBL], *janek-obār-ay* [woman-PL-GEN], not **janek-obār-ān* [woman-PL-OBL], **janek-obār-āni* [woman-PL-GEN] (Nourzaei et al., 2015, 28).

¹² I have chosen to focus this discussion on the animacy analysis. An anonymous reviewer suggested that evaluative meanings may align with Gundel et al.'s (1993) "referential" category, while collective meanings may correspond to "type identifiable" on their "Givenness Hierarchy." Animacy may be only one of several factors involved. Exploring these relationships further is a promising direction for future research, especially if they prove to be cross-linguistically robust.

innovative plural markers (and definite articles) using the heuristics proposed here (or in Nourzaei, 2020, for definiteness markers). This allows me to confidently say that the Koroshi Balochi plural marker *-obār* is of a collective source, of which *bār* 'load, burden' is the most obvious choice. Ideally, these heuristics can be used in the future as a tool for identifying and reconstructing the sources of some innovative plural markers. Furthermore, this analysis points to an explanation for morphological competition between definiteness and plurality that may give some insight to other types of morphological competition in the future.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Shuan Osman Karim: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Data availability

All data is openly available through published works

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

APPENDIX A. SUPPLEMENTARY MATERIAL

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.lingua.2026.104243>.

REFERENCES

- Ackerman, F., Stump, G.T., 2004. Paradigms and periphrastic expression: a study in realization-based lexicalism. In: Sadler, L., Spencer, A. (Eds.), *Projecting Morphology*. CSLI Publications, Stanford, CA, pp. 111–157.
- Anonby, E., Mohammadirad, M., Sheyholislami, J., 2019. Kordistan Province in the Atlas of the Languages of Iran: research process, language distribution, and language classification. In: Gündogdu, S., Opengin, E., Anonby, E., Haig, G. (Eds.), *Current Issues in Kurdish Linguistics*. University of Bamberg Press, Bamberg, pp. 9–38.
- Axenov, S., 2006. The Balochi language of Turkmenistan a corpus-based grammatical description Ph.D. thesis. Uppsala University, Uppsala.
- Borjian, H., 2012. KASHAN ix. The median dialects of Kashan, <http://www.iranicaonline.org/articles/kashan-ix-the-median-dialects-of-kashan>.
- Borjian, H., 2017a. KERMÁN xvi. Languages, <https://iranicaonline.org/articles/kerman-16-languages>.
- Borjian, H., 2017b. KEŠĀ'Ī dialect, <http://www.iranicaonline.org/articles/keshai-dialect>.
- Bybee, J., 1985. *Morphology: a study of the relation between meaning and form*. Benjamins, Amsterdam.
- Corbett, G., 2000. *Number*. Cambridge University Press, Cambridge.
- Elfenbein, J., 1980. *The Balochi Language: Dialectology and Texts*. Indus Publications, Karachi.
- Fattah, I.K., 2000. Les dialectes kurdes méridionaux: étude linguistique et dialectologique (Iranica 37). Peeters, Leuven.
- Gardelle, L., 2019. *Semantic Plurality English Collective Nouns and Other Ways of Denoting Pluralities of Entities*. John Benjamins, Amsterdam.
- Gundel, J.K., Hedberg, N., Zacharski, R., 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69, 274–307.
- Haig, G., 2019. Debonding of inflectional morphology in Kurdish and beyond. In: Gündogdu, S., Opengin, E., Anonby, E., Haig, G. (Eds.), *Current Issues in Kurdish Linguistics*. University of Bamberg Press, Bamberg, pp. 117–144.
- Haig, G., Mohammadirad, M., 2019. Definiteness in Central Kurdish: sources and outcomes. In: Paper presented at the 4th International Conference on Kurdish Linguistics. University of Rouen, Paris-France.
- Heine, B., Kuteva, T., 2002. *World Lexicon of Grammaticalization*. Cambridge University Press, Cambridge, New York.
- Holmberg, A., Odden, D., 2008. The Noun Phrase in Hawrami*. In: Samiian, V., Stilo, D., Karimi, S. (Eds.), *Aspects of Iranian Linguistics*. Cambridge Scholars Publishing, Newcastle upon Tyne, pp. 129–152.
- Hopper, P.J., 1991. On some principles of grammaticization. In: Traugott, E.C., Heine, B. (Eds.), *Approaches to Grammaticalization*, Vol. I. John Benjamins, Amsterdam, pp. 17–36.

- Jahani, C., 2015. On the definite marker in Modern Spoken Persian. In: Paper presented at the Sixth International Conference on Iranian Linguistics (ICIL6).
- Karim, S.O., 2021. The synchrony and diachrony of New Western Iranian nominal morphosyntax Ph.D. dissertation. The Ohio State University.
- Koptjevskaja-Tamm, M., 2001. "A piece of the cake" and "a cup of tea": partitive and pseudo-partitive nominal constructions in the Circum-Baltic languages. In: Dahl, Ö., Koptjevskaja-Tamm, M. (Eds.), *The Circum-Baltic Languages Typology and Contact, Grammar and Typology, Volume 2*. John Benjamins, Amsterdam, pp. 523–568.
- Korn, A., 2019. Isoglosses and subdivisions of Iranian. *J. Histor. Linguist.* 9, 239–281.
- Mackenzie, D., 1961. *Kurdish dialect studies I*. Oxford University Press, London, New York.
- MacKenzie, D.N., 1956. Bajalani. *Bulletin of the School of Oriental and African Studies, University of London* 18, 418–435.
- MacKenzie, D.N., 1966. *The dialect of Awroman (Hawraman-i Luhon): Grammatical sketch, texts, and vocabulary*. Munksgaard, Kobenhavn, Kommissionær.
- Mahmoudveysi, P., Bailey, D., 2013. *The Gorani language of Zarda, a village of West Iran*. Reichert, Wiesbaden.
- Mahmoudveysi, P., Bailey, D., Paul, L., Haig, G., 2012. *The Gorani Language of Gawraju, a village of West Iran: texts, grammar, and lexicon*. Reichert, Wiesbaden.
- McKinnon, C., 2001. GIONI, <https://iranicaonline.org/articles/gioni-giani>.
- McKinnon, C., 2011. LORI language i. LORI Dialects, <https://www.iranicaonline.org/articles/lori-dialects>.
- Mohammadirad, M., 2026. *A grammar of Hewramî*. (Comprehensive Grammar Library 11). Berlin: Language Science Press. <https://doi.org/10.5281/zenodo.17140766>.
- Nourzaei, M., 2020. Definiteness marking from evaluative morphology in Balochi: Internal variation and diachronic pathway. *Iranian Stud.* 54, 699–735. <https://doi.org/10.1080/00210862.2020.1813555>.
- Nourzaei, M., 2023a. Diachronic development of the k-suffixes: evidence from classical new persian, contemporary written persian, and contemporary spoken persian. *Iran. Stud.* 56, 115–160. <https://doi.org/10.1017/irm.2021.27>.
- Nourzaei, M., 2023b. The history of the k-suffix -ū in shirazi. *J. Roy. Asiatic Soc.* 33, 589–626. <https://doi.org/10.1017/S1356186322000323>.
- Nourzaei, M., Jahani, C., Anonby, E., Ahangar, A.A., 2015. *Koroshi a corpus-based grammatical description*. 13, Uppsala Universitet, Uppsala. doi: 10.1017/CBO9781107415324.004.
- Nourzaei, M., Jügel, T., 2023. On the function of -ag in middle persian: evaluative marker or derivational suffix?. *Stud. Iran.* 52 7–38. <https://doi.org/10.2143/SI.52.1.3294136>.
- Öpengin, E., 2016. *The Mukri Variety of Central Kurdish*. Reichert, Wiesbaden.
- Pakendorf, B., Krivoshapkina, I.V., 2014. Even nominal evaluatives and the marking of definiteness. *Linguistic Typol.* 18, 289–331.
- Rainer, F., Gardani, F., Dressler, W.U., Luschützky, H.C., 2019. *Competition in Inflection and Word-Formation*. Springer, New York. <https://doi.org/10.1007/978-3-030-02550-2>.
- Rutkowski, P., 2007. *The syntactic structure of grammaticalized partitives (pseudo-partitives)*. University of Pennsylvania Working Papers in Linguistics 13.
- Samvelian, P., 2007. A (phrasal) affix analysis of the Persian Ezafe. *J. Linguist.* 43, 605–645.
- Selkirk, E., 1977. Some remarks on noun phrase structure. In: Culicover, P.W., Wasow, T., Akmajian, A. (Eds.), *Formal syntax: Papers from the MSSB-UC Irvine Conference on the Formal Syntax of Natural Language*, Newport Beach, California. Academic Press, New York, pp. 285–316.
- Sims-Williams, N., 1982. The double system of nominal inflexion in Sogdian. *Trans. Philol. Soc.* 80, 67–76.
- Sims-Williams, N., 2004. The Parthian Abstract Suffix -yft. In: Penney, J.H.W. (Ed.), *Indo-European Perspectives*. Oxford University Press, Oxford, *Studies in Honour of Anna Morpurgo Davies*, pp. 539–547.
- Stickney, H., 2004. *The pseudopartitive and its illusory projections* Master's thesis. University of Massachusetts.
- Stilo, D., 2008. Case In Iranian: From Reduction and Loss to Innovation and Renewal. In: Spencer, A., Malchukov, A.L. (Eds.), *The Oxford Handbook of Case*. Oxford University Press, Oxford; New York, pp. 700–715.
- Stump, G., Finkel, R.A., 2015. The complexity of inflectional systems. *Linguist. Vanguard* 1, 101–117. <https://doi.org/10.1515/lingvan-2014-1007>.
- Stump, G.T., 2001. *Inflectional morphology*. Cambridge University Press.
- Sundermann, W., 1989. Westmitteliranische Sprachen. In: Schmitt, R. (Ed.), *Compendium Linguarum Iranicarum*. Reichert, Wiesbaden. chapter 3.1.0, pp. 106–113.
- Thackston, W.M., 1993. *An introduction to Persian*. Iranbooks, Bethesda, Md.
- van der Wal Anonby, C., 2019. Kumzari. In: Khan, G., Haig, G. (Eds.), *The languages and linguistics of western Asia: An areal perspective*. De Gruyter, Berlin, Boston, pp. 625–658.
- Whitney, W.D., 1993. *Sanskrit Grammar*. Harvard University Press, Cambridge, Mass, London.
- Windfuhr, G., 1991. Central dialects. <https://www.iranicaonline.org/articles/central-dialects>.
- Windfuhr, G., 1998. AŠTIANI. https://doi.org/10.1163/2330-4804_EIRO_COM_6022.
- Windfuhr, G.L., 2012. FARS viii. Dialects, <https://iranicaonline.org/articles/fars-viii>.
- Yoshida, Y., 2009. Sogdian. In: Windfuhr, G.L. (Ed.), *The Iranian Languages*. Routledge, New York, NY, pp. 279–335.