CrossMark

Association for
Iranian Studies
انجمن ایران‌پژوهی

ARTICLE

Diachronic Development of the K-suffixes: Evidence from Classical New Persian, Contemporary Written Persian, and Contemporary Spoken Persian

Maryam Nourzaei

Department of Linguistics and Philology, Uppsala University and Linguistics,
Otto-Friedrich-Universität, Bamberg
Email: maryam.nourzaei@lingfil.uu.se

Abstract

This paper aims to investigate the usage and frequency of what we refer to as K-suffixes in Classical New Persian of the ninth to thirteenth centuries, Contemporary Written Persian of the late nineteenth to mid-twentieth centuries, and Contemporary Spoken Persian. It shows that K-suffixes are most likely to be the reflexes of earlier evaluative morphemes, traditionally called "diminutives," and are characterized by a high degree of multifunctionality. While evaluative functions continue to dominate in the Classical New Persian works, they have largely been lost in contemporary spoken Persian, and the suffix is now systematically used to express definiteness. The development of the K-suffix as a definiteness marker in contemporary colloquial Persian appears to be innovative, and is mainly dependent on genre, speaker, and speech situation.

Data for Classical New Persian is taken from critical editions of works from the ninth to thirteenth centuries. The data for Contemporary Written Persian comes from comprehensive books of fiction from the late nineteenth to mid-twentieth centuries, and for Contemporary Spoken Persian from an extensive corpus of spoken Persian narratives and a questionnaire answered by fifteen speakers. The results suggest that evaluative morphology can develop into definiteness marking, with the development passing through a stage of combination with a deictic marker.

This paper concludes that the development of definiteness marking can proceed down a new pathway that is different from the one normally assumed for demonstrative-based definite marking, though the endpoint may be similar. The study contributes the second detailed documentation of this process for any Iranian language, and one of the few well-documented cases of a non-demonstrative origin of definiteness marking worldwide.

Keywords: Classical New Persian; Contemporary Written Persian; Contemporary Spoken Persian; diminutive; evaluative; definiteness marking; grammaticalization

1. Introduction

Persian is a term for a collection of closely related western Iranian varieties. It is spoken in Iran, Afghanistan, and Tajikistan, and serves as an official language in these counties. This paper deals with the K-suffix in Classical New Persian of the ninth to thirteenth centuries (CNP), Contemporary Written Persian of the late nineteenth to mid-twentieth centuries (CWP), and Contemporary Spoken Persian (Tehran variety) in Iran (CSP).

In all CNP written works, a suffix of the form -*ak/ek/ag/* is attested, primarily occurring with nouns but also with adjectives and adverbs. It has traditionally been classified as

"diminutive" and presumably is cognate with several formatives containing a velar plosive [k], or a reflex thereof, in other Iranian languages (Balochi, Kurdish, and Lori) and Indo-Aryan. However, in CWP and CSP texts, a suffix of the form *-e* (K-suffix) is attested mostly with singular nouns. The status of this suffix in CWP is largely similar to that of the K-suffix in CNP, but in CSP it is clearly associated with definiteness. The original function of these suffixes is yet to be established with certainty, but available accounts from both CNP and CWP suggest a high degree of multifunctionality of this suffix. There is often a semantic component of "less than expected size," but more frequently we find an evaluative component expressing the speaker's empathy, familiarity, endearment, and respect, or conversely, disdain with respect to the diminutive-marked noun.

Such evaluative connotations are widely attested cross-linguistically[1] and in other Iranian languages such as Balochi, Old Shirazi, and Lari.[2] Given the salience of the evaluative components (and the lack of any reference to "size" in many contexts, see below), I follow Pakendorf and Krivoshapkina in referring to the function of this morphology as evaluative rather than diminutive.[3]

The paper concentrates on what we term the definitizing function of the K-suffix in Persian. It can be demonstrated that, at least in CSP, the K-suffixes are associated with definiteness in a manner approximately comparable to the better-known definite articles of the languages of Europe, e.g., English and Swedish. However, it is still highly dependent on the speaker, genre, and setting.

For that reason, almost all previous studies on the development of definiteness marking assume a demonstrative as its origin (see Section 6). The Persian definiteness marker has considerable implications for our understanding of definiteness systems and their emergence more generally. Looking at the function of the K-suffix in different phases of Persian (CNP, CWP, and CSP), a well-documented New Western Iranian language with available recorded material from its earlier stages, it can be stated with some certainty that the definiteness marker is not related to a demonstrative element.

To the best of my knowledge, there is no previous detailed study of the K-suffix from a diachronic perspective in Persian. The data for this work is taken from extensive corpora of the language phases under study. I complement the quantitative data with a qualitative approach, which demonstrates the various functions with authentic examples and appropriate references to context. I also refer to the results of a questionnaire-based survey with CSP, which is based on the questionnaire used for Kurdish, Balochi, Shirazi, and Lori (see Section 6.2).[4]

One of the most exciting aspects of the data is the high degree of inter-speaker/writer and inter-text variability, particularly in the CWP and CSP corpora. The definiteness function of the K-suffix in CSP is systematically documented for very few texts, typically only for folktales and biographical tales. This is very similar to the results from the questionnaires, which show a high degree of non-conformity and non-systematicity in the definiteness usage across the speakers.

Contrary to the Shirazi data,[5] the grammaticalization development in CSP appears to be fairly sensitive to speech contexts, typically genre rather than linguistic context.[6] Given that

---

[1] Dressler and Barbaresi, *Morphopragmatics*; Jurafsky, "Universal Tendencies"; Steriopolo, "Form and Function"; Pakendorf and Krivoshapkina, "Ėven Nominal Evaluatives"; Ponsonnet, "A Preliminary Typology."

[2] Nourzaei, "Definiteness Marking"; Nourzaei, "History of the Suffix *-ū* in Shirazi."

[3] Pakendorf and Krivoshapkina, "Ėven Nominal Evaluatives."

[4] Haig, "Optional Definiteness"; Nourzaei, "Definiteness Marking"; Nourzaei, "History of the Suffix *-ū* in Shirazi"; Nourzaei and Haig, "An Overview of Definiteness Marking."

[5] The present work is not a comparative study; however, I do refer to some features of the K-suffix in other Iranian languages as well. Examples from these languages can be found in Nourzaei, "Definiteness Marking," "History of the Suffix *-ū* in Shirazi," and works in preparation by Nourzaei and Haig, and Haig et al.

[6] Nourzaei, "Definiteness Marking."

the usage of evaluative morphology is, by definition, primarily determined by interactional context, this finding is not surprising.

This paper is organized as follows: first, it deals with definiteness and types of definiteness contexts and provides an overview of the Persian language and data. Then it covers previous studies of the K-suffix in Persian and demonstrates the multifunctionality of the K-suffix. The evaluative function of K-suffixes in CNP and CWP is then presented, after which K-suffixes functioning as definiteness markers in CSP are illustrated. Data is presented from an extensive text corpus and questionnaire data, and a suggestion is made regarding the original K-suffix in CNP, CWP, and CSP. Finally, the findings are discussed in light of a new grammaticalization pathway from evaluative to definiteness marker.

1.1 Definiteness

Definiteness will be understood here as a property of a noun phrase that is derived from its information status in a given linguistic context. It is thus a contextual property of referring expressions rather than an inherent property of nouns. A number of different approaches to definiteness have been pursued in the literature, including a philosophical approach invoking uniqueness,[7] and a discourse-pragmatic approach.[8] I follow Lyon in considering the primary component of definiteness to be the notion of identifiability.[9] A noun phrase is considered definite if the speaker assumes that its referent is uniquely identifiable by the addressee. Languages differ cross-linguistically in the extent to which, and means by which, they systematically indicate definiteness in morphosyntax. In English, French, or Arabic, definiteness is marked fairly consistently using items generally referred to as "articles." Other languages may mark definiteness by affixes, clitics, word-order properties, or various combinations of these strategies; alternatively, they may have no regular means for indicating definiteness. A noun phrase may have definite status by virtue of several possible contextual factors, which we broadly characterize as follows:[10]

Types of definiteness contexts	
Anaphoric definiteness	The referent has an antecedent in the preceding textual context: *A man and a woman entered.* **The man** *sat down.*
Bridging definiteness	The referent has not been previously mentioned in the discourse context, but its existence can be inferred from associated expressions.[11] *We bought a new car but* **the brakes** *were faulty.*
Proper nouns	The noun is conventionally associated with a specific entity: *Sweden, Angela Merkel, Mount Kilimanjaro.*
Possessed nouns	The noun is accompanied by a grammatical possessor, often syntactically fulfilling the determiner function: *my house, their child, Henry's birthday.*
Deictically modified nouns	The noun is accompanied by a demonstrative element: *this article, that place.*

---

[7] See Russell, "On Denoting"; and Neale, *Descriptions*.

[8] See Christophersen, *The Articles*; Schwarz, *Indirekte Anaphern in Texten*; Clark, "Bridging"; overviews in Abbott, "Definiteness and Indefiniteness"; and von Heusinger, "Definiteness."

[9] Lyons, *Definiteness*.

[10] Following Abbott, "Definiteness and Indefiniteness"; Lyons, *Definiteness*; and Becker, "Articles in the World's Languages."

[11] Lyons, *Definiteness*, 272.

Unique referents	Entities which are assumed to be uniquely identifiable by all members of a given speech community, hence requiring no preceding or inferable mention: *the sun, the river* (in a given community), *the president*.
Situational definiteness	Identifiability is achieved through the immediate speech context, possibly aided by additional gestures and adverbial expressions: **the man** *over there* (pointing).

In contrast to the seven definiteness contexts outlined above, nouns may be indefinite, (either specific or non-specific), or have generic or sortal reference. The correct analysis of generics is beyond the scope of this paper.[12]

2. The Persian Language

Persian belongs to the Western Iranian branch of the Iranian languages, which in turn belong to the Indo-Iranian branch of Indo-European. Persian is the only Iranian language that has documents available from the Old Persian of the Achaemenids, the Middle Persian of the Sassanids, to New Persian (since the eighth century). Different delimitations of the phases in the development of New Persian have been presented by Iranian scholars. For instance, Lazard introduces the following phases: Early New Persian for the language of the tenth to eleventh centuries, and Classical New Persian for the New Persian of the twelfth to nineteenth centuries, with the twelfth century as a transitional period.[13] I find these classifications to be a bit too complicated for the present study. For the sake of brevity, I use Classical New Persian (CNP) of the ninth to thirteenth centuries, Contemporary Written Persian (CWP) of the late nineteenth to mid-twentieth centuries, and Contemporary Spoken Persian (CSP) in the present paper.

Modern Persian is a verb-final language that shows the same alignment system in the past and non-past tenses by not having a morphological case system. Persian is mainly spoken in Iran, Afghanistan, and Tajikistan, and is considered a language of education in these countries. The area where Persian is spoken is highly diverse linguistically. Contact languages include four different language families and different genera: Indo-European (Indo-Aryan and Iranian), Dravidian, Turkic, and Semitic.

Data for CNP is taken from critical editions of works from the ninth to thirteenth centuries (see Table 1), data for CWP come from books of fiction from the late nineteenth to mid-twentieth century (see Table 2), and CSP from an extensive corpus of spoken Iranian Persian narrative and a questionnaire answered by fifteen speakers from Tehran (see Section 5). Fig. 1 presents the location of the data for Contemporary Spoken Persian.

I will briefly comment on other functions of the K-suffix (viz., derivational) than evaluative, before we begin our journey into the K-suffixes in the Persian language.

Derivations with the suffix *-ka- are well attested in Old Indo-Iranic (especially in Old Indo-Aryan). Edgerton offers a detailed survey in two papers with the same title, published in the consecutive issues 2–3 of volume 31 of the *Journal of the American Oriental Society*.[14] He identifies the core semantics of *-ka- for Proto-Indo-Iranic by comparing the Vedic, Sanskrit, and Avestan evidence:[15] "1) the formation of nouns of likeness or adjectiv[e]s of characteristic; 2) the diminutiv[e] and (perhaps) pejorativ[e] formations, 3) occasional formations with *2 ka* [i.e., adjectives of appurtenance or relationship],[16] mainly pronominal adjectiv[e]s, and 4)

---

[12] See Becker, "Articles in the World's Languages."

[13] Lazard, *A Grammar of Contemporary Persian*, 24. See also Windfuhr and Perry, "Persian and Tajik," for a similar delimitation.

[14] Edgerton, "The K-suffixes of Indo-Iranian."

[15] Ibid., 310.

[16] Ibid., 98.

Table 1. List of the critical editions from which data has been extracted.

Title	Date (c.)	Content	Author
Qorʾān-e Qods	9th–10th	Quran translation	Anonymous
Al-abniye ʿan haqāʾiq al-adviye	11th	Traditional and ancient medicine	Asadi Tusi
Tārikh-e Sistān	11th	History of Sistan	Anonymous
Tārikh-e Beyhaqi 1–3	11th	Bayhaqi's history	Abu al-Fazl Beyhaqi
Qābusnāme	11th	Instruction	Amīr ʿOnsor al-Maʿāli Keykāvūs Ebn-e Eskandar Ebn-e Qabūs
Safar name-ye Nāser Khosrow	11th	Travelogue	Abu Moʿin Hamid al-Din Nāser Ebn-e Khosrow Qobādiyāni
Khᵛān al-ekhvān	11th	Religious issues	Abu Moʿin Hamid al-Din Nāser Ebn-e Khosrow Qobādiyāni
Nowruznāme	12th	Nowruz celebration and rituals	ʿOmar Khayyām
Dārābnāme 1–2	12th	A story about the Kiyanid king Darab	Abū Tāher Ebn-e Hasan Ebn-e ʿAli Ebn-e Musā Tarsusi
Bakhtiyārnāme	12th	The story of the ancient proverb	Anonymous
Rowzat al□qul	13th	Ancient animal tales and anecdotes	Mohammad Ebn-e Ghāzī Maltivī
Marzbānnāme	13th	Ancient animal tales and anecdotes	Saʿd al-Din Varāvini

the primary formations from verbal bases, apparently inclining towards the meaning of verbal adjectives or nouns of agent."

The K-suffix *-ak* in Persian largely reflects Edgerton's classification. Iranian traditional grammarians already report a similar classification.[17]

In the CNP works under study, the evaluative semantics of K-suffixes are more predominant than other functions (derivational) including adjective<adverb N<adjective. Note that the K-suffix *-ak* is more productive as a word-creation suffix in CWP and CSP than in CNP, probably because of a national need for creation of words.

In the following example, the adjective *narm* "*soft*" has changed into the adverb *narmak,* "softly, slowly."

Ex. (1)
   ma-rā   **narm-ak**   *āvāz*   *dād*
   PN.1SG-OBJ soft-EV   sing   give.PST.3SG
   "she called me slowly"[18]

3. The K-suffixes in CNP: Initial Observations[19]

Data for analyzing the K-suffixes in CNP comes from critical editions of works from the ninth to thirteenth centuries. Table 1 provides a list of these works.

---

[17] Including Natel Khanlari, *Dastur Zabān-e Fārsi*; Kasravi, *Kāfnāme*; and Kalbasi, *Sakht-e eshteqāqi-ye vāzhe dar Fārsi-ye Emruz.*

[18] *Tārikh-e Sistān*, 63.

[19] Note that the K-suffix has also been attested in the poetry genre, including *Shāhnāme*. Since I already have a large body of prose material at my disposal for studying this suffix, I have not commented on its use in poetry.

Table 2. List of the books from which data has been extracted.[20]

Title	Date	Content	Author
Ketāb-e Ahmad	1890	Didactic novel	ʿAbd al-Rahim Tālebof
Tamsilāt[21]	1874	Play collection	Mirzā Fathʿali Ākhundzade
Siyāhatnāme-ye Ebrāhim Beyg, 1–3	1895	Novel/fictional travelogue	Zeyn al-ʿĀbedin Marāghei
Charand o parand	1907	Short story collection	ʿAli Akbar Dehkhodā
Zibā	1927	Novel	Mohammad Hejāzi
Zende be gur	1930	Short story collection	Sādeq Hedāyat
Se qatre khun	1932	Short story collection	Sādeq Hedāyat
Chamedān	1934	Short story collection	Bozorge ʿAlavi
ʿAlaviye khānom	1934	Short story collection	Sādeq Hedāyat
Hājī āqā	1945	Novel	Sādeq Hedāyat
Nasim	1960	Novel	Mohammad Hejāzi
Buf-e kur	1936	Novel	Sādeq Hedāyat
Sag-e velgard	1942	Short story collection	Sādeq Hedāyat
Cheshmhāyash	1952	Novel	Bozorg ʿAlavi

Across CNP texts, a nominal suffix is found with the forms *-ak/ek/ag.*[22] These are likely to be reflexes of the K-suffix *-ag* in Middle Persian,[23] e.g., *pus-ag* "boy" and CNP *pesar-ak* "boy."

The K-suffix has been attested with nouns, e.g., *pesar-ak* "boy," *darvīš-ak* "dervish," adjectives, e.g., *ǰavān-ak* "young," *saqīr-ak*[24] "small," *andak* "little," and adverbs, *ānak* "now."[25]

Ex. (2)

ammā   kas=ī-rā   ke   ranǰ kam resad   ū-rā   gūšt=e
but   person=IND-OBJ CLM pain little arrive.NPST.3SG PN.3SG-OBJ meat=EZ
gūsāle=e xord-ak   beh-tar   bovād
calf=EZ   small-EV   good-COMP   be.NPST.3SG
"but a person who has less pain, it is better for him/her [to eat] meat of a young calf"[26]

Traditionally this suffix is referred to as a "diminutive." Investigation of the K-suffix in CNP has largely been ignored. However, its existence has been reported by scholars. For Early New Judeo-Persian, Paul reports that "-*ak* functions as diminutive, or it appears without

---

[20] The date here refers to the first edition of the book.

[21] This book is a translation from Azerbaijani Turkish into Persian by Mirzā Jaʿfar Qarājedaghi.

[22] I have not found the suffix *-ag* in my data. However, Sadeghi, "Pasvandha-ye Tahbibi-ye Farsi," reports a few items with the K-suffix *-ag* instead of *-ak,* for instance, *farzandag* "child," *xordag* "little," and *Sahlagī* "?" He also mentions that in another manuscript of *Qorʾān-e Qods* "son" is attested with the K-suffix *-ag,* as in *pusag,* which is similar to *pusag* in Middle Persian. In addition, Khatamipoor, "yā-ye maʿrefeh," based on three manuscripts (titled *hezār hekāyate sūfīyān,* from the thirteenth century), reports the *-ī* suffix including *ak* and considers the *-ī* suffix to be a definiteness marker.

[23] See Durkin-Meisterernst, *Grammatik des Westmitteliranischen*, 253; and Nourzaei and Jügel, "On the Function of -*ag* Suffix in MP," for a detailed discussion of the K-suffix *-ag* in Middle Persian.

[24] The word *saqīr* is an Arabic word meaning small.

[25] See Ciancaglini, "Outcomes of the Indo-Iranian Suffix *-ka in Old Persian and Avestan," for the attestation of this suffix in Old Persian.

[26] *Al-abniye*, 287.

Figure 1. Location of the data for Contemporary Spoken Persian.

semantic modification, e.g., *kanīzak*, 'girl', *xāharak*, 'sister', *mardumakan* and *šamšērak* 'sword'."[27] Gindin, in an unpublished study on Early New Judeo-Persian, mentions *-ak* as a diminutive suffix, such as in "*jūyz-ak*" – a diminutive of *jūy* "river."[28]

Qarib and colleagues introduce the suffixes *-ak/īk/*, *-čeh/*, *-žeh/zeh*, *-īk*, *-ū*, *-ek*, and *-e* as diminutive suffixes; however, they maintain that it covers other semantics, e.g., respect, endearment, and pejorative.[29] Similarly, Ahmadi Givi and Anvari mention *-ak*, *-ū*, *-e*, as a diminutive.[30] Khayyampur reports that *-ak*, *-čeh*, and *-ū* are used as diminutive suffixes, among others.[31] Natel Khanlari[32] considers the suffix *-če* to be diminutive and the suffix *-ak* to be *šebāht* "a likeness suffix."[33]

3.1 Evaluative and Diminutive Usage in CNP

The most frequent usage of the K-suffix is to express evaluative or diminutive semantics, and it is even compatible with indefinite contexts. The term "diminutive" implies the descriptive content "smaller than normally expected," and this is evident in some usages of K-suffixes. However, even in these contexts, an evaluative connotation is often discernible and, for the sake of brevity, following Nourzaei[34] I gloss the suffix with EV, as the most general indication of function, regardless of actual context.

---

[27] Paul, *A Grammar of Early Judaeo-Persian*, 63.
[28] Gindin, "The Early Judeo-Persian *Tafsīr*s of Ezekiel."
[29] Qarib et al., *Dastur-e Fārsi*, 46.
[30] Ahmadi Givi and Anvari, *Dastur-e zabān-e Fārsi 1*, 77.
[31] Khayyampur, *Zabān-e Fārsi*, 34.
[32] Natel Khanlari, *Dastur Zabān-e Fārsi*, 165–67.
[33] "اسمی که به داشتن آن صفت مخصوص است."
[34] Nourzaei, "Definiteness Marking."

In example (3) the K-suffix gives a description of the physical size of the branch, *šāx-ak=ī* "a small branch." Note that the K-suffix is compatible with the indefiniteness context.

Ex. (3) The K-suffix with small size

šāx-ak=ī	az	īn	toxm-hā	bar	ǰast
branch-EV=IND	from	PROX	seed-PL	PREV	spring.PST.3SG

"a **small branch** grew up from these seeds"[35]

Similarly, in example (4), the K-suffix provides a description of the physical size of the deer's fawn. Note that the K-suffix follows a distal demonstrative *ān* "that."

Ex. (4) The K-suffix with small size

be-dān	ke	ān	baxšāyeš	ke	bar	ān	āhū=ye
be[36]-know.NPST.2SG	CLM	PROX	forgiveness	CLM	to	PROX	deer=EZ

mādeh	kard-ī	va	ān	**bačeg-ak**	be=dū	bāz	dād-ī
female	do.PST-2SG	and	DIST	child-EV	to=PC.3SG	again	give.PST-3SG

"Know that the mercy that you have shown to that female deer and that **small child** returned to her [...]"[37]

In example (5), the K-suffix provides a description of a small amount of water.

Ex. (5) The K-suffix with a small amount

be-dīn	dīeh	češme=īst	ke	az	sang	bīrūn
with-PROX	village	spring=COP.NPST.3SG	CLM	from	stone	out

mī-ā-yad						
IMP-come.NPST-3SG						

āb-ak=ī	andak	va	rāh=ī	dawr=e	ǰū	borīd-eh
water-EV=IND	little	and	path=IND	around=EZ	stream	cut.PST-3SG

"There is a spring to this village that comes out of a stone with **little water** and they have paved (lit. cut) a long stream from it"[38]

In example (6), the K-suffix adds a flavor of sorrow on the part of the speaker regarding the Hendu male slave, rather than a description of the physical size of the male slave.

Ex. (6) The K-suffix conveys a flavor of sorrow

man	va	barādar=m	va	**ǧolām-ak=ī**	hendū	ke	bā
PN.1SG	and	brother=PC.1SG	and	male.slave-EV=IND	Hendu	CLM	with

mā	būd	vāred	šod-īm				
PN.1PL	be.PST.3SG	enter	become.PST-1PL				

"I and my brother and **a poor Hendu male slave** who was with us arrived (lit. entered) to [xarzawīl]"[39]

Similar to example (6), example (7) adds a flavor of sorrow on the part of the speaker regarding the deer's mother, who was following the hunter when she repeatedly fell down, rather than a description of the physical size of the deer's mother. Note that the K-suffix follows a proximal demonstrative *īn* "this."

---

[35] *Nowruznāme*, 67.

[36] I have followed Lenepveu-Hotz, Agnés "Evolution of the Subjunctive in New Persian (10th–20th): From disappearance to reappearance", Linguistic, Folia Linguistica, 2018, and glossed *be-* as be at this stage of Persian.

[37] *Tārikh-e Beyhaqi 1*, 250.

[38] *Safarnāme-ye Nāser Khosrow*, 57.

[39] Ibid., 6.

Ex. (7) The K-suffix conveys a flavor of sorrow

bāz	*gašt-am*		*va*	*do*	*se*	*bār*	*ham=čenīn*	*mī-oftād*
PREV	turn.PST-1SG		and	two	three	time	EMPH=PROX	IMP-fall.PST.3SG
va	***īn***	***bīčāreg-ak***	*mī-ām-ad*					
and	PROX	poor-EV	IMP-come.PST-3SG					

"I returned and [I saw that the female deer] two or three times it fell down and **this poor one** still was coming"[40]

The evaluative component is more obvious in the following examples. In example (8), Joseph's father refers to his son with a K-suffix, although the son is grown up. This is obviously a signal of endearment and affection on the part of the speaker towards the son, rather than a description of his physical size. Note that the K-suffix has been attested with vocative and non-vocative contexts.

Ex. (8) The K-suffix with endearment

goft	*yā*	***pesar-ak**=e*	*man*	*ġeseh*	*ma-kon*	*xāb=e*
say.PST.3SG	VOC	son-EV=EZ	PN.1SG	story	IMP-do.NPST.2SG	dream=EZ
to						
PN.2SG						

"He said, O my **lovely son**, do not tell your dream"[41]

Similar to example (8), in the following passage, a dialogue between God and the prophet Noah, Noah refers to his son with a K-suffix, although the son is grown up. Again, this is obviously a signal of endearment and affection on the part of the speaker towards the son, rather than a description of his physical size.

Ex. (9) The K-suffix with endearment

***pesar-ak**=e*	*man*	*az*	*ahl=e*	*man*	*ast*
son-EV=EZ	PN.1SG	from	group=EZ	PN.1SG	COP.NPST.3SG

"my **lovely son** is from my group"[42]

The K-suffix occurs here with an "admiration and respect" connotation. The K-suffix on "Hasan" demonstrates respect towards Hasan, who was an important and influential figure in the Ghaznavid state, rather than a description of his physical size.

Ex. (10) The K-suffix with respect

hasan-ak	*ġarmatī*	*ast*	*vay-rā*	*bar*	*dār*	*bāyad*	*kard*
hasan-EV	Qarmati	COP.NPST.3SG	PN-OBJ	on	wood	must	do.NPST.3SG

"**Hasanak** is Qarmati,[43] he must be executed"[44]

Similar to example (10), the K-suffix in example (11) displays admiration and respect towards Abul Abulqāsem-e Hakīm, rather than a description of his physical size.

Ex. (11) The K-suffix with respect

***Abulqāsem**=e*	***Hakīm-ak***	*ke*	*nadīm*	*Amir*	*Yusef*	*bud*	*mard=ī*
Abulqāsem=EZ	Hakīm-EV	CLM	friend	Amir	Yusef	be.PST.3SG	man=IND

---

[40] *Tārikh-e Beyhaqi 1*, 250.
[41] *Qorʾān-e Qods*, 143.
[42] Ibid., 136.
[43] This term refers to a branch of Islam whose adherents believe in seven Imams.
[44] *Tārikh-e Beyhaqi 1*, 229.

momattˀ	va	bekār	āmadeh	ham	xedmat=e	kas=ī	na-kard
educated	and	skilled		ADD	servant=EZ	person=IND	NEG-do.PST.3SG

va	karīm	bud
and	generous	be.PST.3SG

"**Abulqāsem-e Hakimak**, who was a friend of Amir Yusuf, he was an educated and skilled [man], he was not at the service of anyone, and he was generous"[45]

K-suffixes also occur with pejorative connotations. This can be seen in vocative contexts such as in example (13). The following passage is taken from a dispute between the king and a dervish. Here the K-suffix reflects the king's anger and disapproval of the dervish in the given context.

Ex. (12) The K-suffix with disapproval

īn	darvīš	**mard-ak**=ī	nādān	va	kūhparvar	ast
PROX	dervish	man-EV=IND	ignorance	and	cave man	COP.NPST.3SG

"This dervish is an ignoramus and a caveman"[46]

This can be observed in vocative contexts, as in example (13), where it is taken from a dispute between Halāl and the holy man. Here the K-suffix reflects the king's anger and disapproval of the holy man in the given context.

Ex. (13) The K-suffix with disapproval

halāl	tond	šod	va	goft	ay	**darvīš-ak=e**	nābekār
Halāl	sharp	become.PST.3SG	and	say.PST.3SG	VOC	dervish-EV=EZ	evil

to	čehā	mī-gū-ī
PN.2SG	what	IMP-say.NPST-2SG

"Halāl got angry and said: 'what are you saying, evil dervish?'"[47]

Finally, we should point out that certain words typically indicating both human and non-human referents seem to include the K-suffix as part of the word stem. The suffix lacks any apparent separate semantic content.

   In sum, the K-suffixes of CNP are widely attested with some kind of evaluative semantics, but also as lexicalized and semantically empty elements, and are presumably remnants of the high-frequency evaluative usage associated with certain words. We assume that the multifunctionality of the K-suffix is reasonably representative of earlier stages of Persian and is also compatible with what is known about K-suffixes in earlier stages of other New Western Iranian languages such as Shirazi, Lari, and Balochi. However, in the three phases of Persian (CNP, CWP, and CSP) being studied here, the functionality and frequency of K-suffixes have diverged quite considerably. In particular, in specific genres of CSP, the K-suffix *-e/he* exhibits a regular marking of definiteness in anaphoric and bridging contexts (see Section 6).

   I begin with an outline of K-suffixes in CNP, before focusing on the usage of the K-suffix in CWP (Section 5) and CSP (Section 6) and presenting frequency data from the corpora (Section 7).

3.2 Analysis of the K-suffix in CNP

The K-suffix attaches to nouns, adjectives, and adverbs. The following passage shows the K-suffix with an adjective:

---

[45] *Tārikh-e Beyhaqi 2*, 404.
[46] *Dārābnāme 1*, 419.
[47] Ibid., 417.

Ex. (14) The K-suffix with an adjective
ammā kas=ī-rā   ke   ranǰ Kam   res-ad   ū-rā   gūšt=e
but   person=IND-OBJ CLM pain Little arrive.NPST-3SG PN.3SG-OBJ meat=EZ
gūsāle=e **xord-ak** *beh-tar   Bovād*
calf=EZ   small-EV   good-COMP   be.NPST.3SG
"but a person who has less pain, it is better for him/her [to eat] meat of a young calf"[48]

The K-suffix in CNP has a variety of functions, with no obvious structural constraints. However, there is one type of context that demonstrates a different reading than the normal multifunctional semantics of the K-suffix (see Sections 3.4 and 3.5).

The K-suffix in CNP is compatible with indefinite contexts, as in examples (15) and (16).

Ex. (15) The K-suffix with indefiniteness status
*āvard-e=and   ke   **morǵ-ak=ī**   būd*
bring.PST-PP=COP.PST.3SG CLM bird-EV=IND   become.PST.3SG
"they said that there was a little bird"[49]

Ex. (16) The K-suffix with indefiniteness status
agar **kanīz-ak=ī**   *šāhd   rū   ār-ī*
if   female.servant-EV=IND happy face bring-2SG
"if you bring a happy-faced female servant"[50]

Examples (17) and (18) show that the K-suffix is compatible with proper nouns, for example, the personal names *Hasan-ak* "Hasan," *Mahmūd-ak* "Mahmud," *gandom-ak* "Gandom," *xayr-ak* "Xayrak," *mār-ak ebne allsalāt* "Marak ebne allsalāt," and *sarbāt-ak* "Sarbātak." Note that proper nouns such as these, where the stem and this suffix can be clearly distinguished, are very rare in the manuscripts. The lack of such examples in these works is probably indicative of the strongly interactional nature of the K-suffix in CNP.[51]

Ex. (17) The K-suffix with proper nouns
hasan-ak *albateh   hīč pāsox   na-dād*
hasan-EV   of course   no   answer   NEG-give.PST.3SG
"of course, **Hasan** did not answer"[52]

Ex. (18) The K-suffix with proper nouns
behāzerī=e   kadxodā va   dabīr=aš   **mahmūd-ak** *va   dīgar vakīl-ān*
presence=EZ   khan   and secretary=PC.3SG   Mahmūd-EV   and other lawyer-PL
"in the presence of khan and his secretary, **Mahmūd** and other lawyers [...]"[53]

---

[48] *Al-abniye*, 287.

[49] *Marzbānnāme*, 500.

[50] Ibid., 40.

[51] Ciancaglini, "Outcomes of the Indo-Iranian Suffix *-ka- in Old Persian and Avestan," 94, notes that *-ka- in Old Persian frequently occurs with proper nouns, ethnonyms and toponyms. In Avestan (as well as other ancient Indo-European languages), words with this suffix are often linked to informal registers, occurring in "imprecatory, pejorative, or affective and familiar contexts," ibid., 95. The same observation has been attested for Modern Iranian languages; see, e.g., Nourzaei, "Definiteness Marking."

[52] *Tārikh-e Beyhaqi 1*, 234.

[53] *Tārikh-e Beyhaqi 2*, 747.

We should point out that there are certain words, typically proper names, which seem to include the K-suffix as part of the word stem, i.e., *sīyāmak* "Siyamak," *bābak* "Babak," and *āl barmak* "Albarmak."[54]

Ex. (19) The K-suffix with proper nouns
*pas   az   bar   oftād-an=e   **āl barmak***
after   from   PREV   fall.PST-INF=EZ   Albarmak
"after the collapse of the **Albarmak** [dynasty]"[55]

Ex. (20) The K-suffix with proper nouns
*čon   nobat   be   **sīyāmak**   resīd*
when   turn   to   Siyamak   arrive.PST.3SG
"when it comes to Siyamak's turn"[56]

As with proper nouns, the K-suffix is compatible with place names, for example, "*čenāša*," "*koškak*," and "*ġūzak*," as in the following example:

Ex. (21) The K-suffix with place names
*vazīr   bar   rāh=e   bež=e   **ġūzak**   raft*
vizir   to   way=EZ   hill=EZ   ġūzak   go.PST.3SG
"the vizier set out towards *ġūzak hill*"[57]

Note that it is not at all obvious what semantic content the K-suffixes have in these contexts; they appear to be relatively vacuous. In contrast to the proper nouns, this type of nouns has a high frequency across the critical editions of works, with *Tārikh-e Sistān* being an example.

In CNP, there is no constraint against combining the K-suffix with the plural suffix (see Sections 5 and 6 on this point in CWP and CSP). The following examples illustrate a K-suffix with evaluative sense followed by a plural marker "*-ān.*"

Ex. (22) The K-suffix with plural noun
*besyār   **gabr-ak-ān**   mosalmān   gašt-and*
very   Gabr-EV-PL   Muslim   turn.PST-3PL
"a lot of Gabrs converted to Islam"[58]

Ex. (23) The K-suffix with plural noun
*'bdūs   bāz   gašt   espas   ānke   **kanīz-ak-ān**   bā.vay*
sad   PREV   return.PST.3SG   after   DIST.CLM   servant-EV-PL   with.PN.3SG
bīy-ārāmīd-e   būd-an
be-sleep.PST-PP   COP.PST-3PL
"he sadly returned after the female servants had slept with him"[59]

There is no restriction with the K-suffix in relation to the possessed nouns (see Section 6 on this issue).

---

[54] Diachronically both the proper nouns *sīyāmak* "Siyamak" and *bābak* "Babak" are derived from a noun plus the K-suffix, but at this stage of the language the K-suffix has become an integral part of the stem, as opposed to the proper noun *Mahmūdak*, which consists of Mahmūd + ak.

[55] *Tārikh-e Beyhaqi 1*, 241.

[56] *Dārābnāme 1*, 282.

[57] *Tārikh-e Beyhaqi 2*, 630.

[58] *Tārikh-e Sistān*, 91.

[59] *Tārikh-e Beyhaqi 1*, 281.

Ex. (24) The K-suffix with person-marking clitics

dar	ān	mīyān	az	**kanīz-ak=īš**		xašam	āmad
in	PROX	moment	from	female.servant-EV=PC.3SG		anger	come.PST.3SG

ān	angoštarī	bexašam	bar	vy	zad
PROX	ring	angerly	to	her	hit.PST.3SG

"at this moment, he got angry at his female servant and threw the ring at her in anger"[60]

Ex. (25) The K-suffix with possessed nouns

pesar-ak=e	man	az	ahl=e	man	ast
son-EV=EZ	PN.1SG	from	group=EZ	PN.1SG	COP.NPST.3SG

"my **lovely son** is from my group"[61]

To sum up, the K-suffix in CNP texts has various functions,[62] and is not subject to structural constraints such as obtain for CWP and CSP (see Sections 5 and 6). However, we find singular nouns, often accompanied by proximal/distal demonstratives, taking a K-suffix with no apparent connection to small size or any particular evaluative notion. Such examples are very rare and would require a larger corpus to study. However, in Old Shirazi these functions of the K-suffix predominate.[63]

Before demonstrating the use of K-suffixes as signals of proximity and familiarity/recognition, it would be helpful to outline indefiniteness and definiteness strategies in CNP.

3.3 Indefiniteness and Definiteness Strategies in CNP

In CNP, discourse-new,[64] specific, singular NPs are overtly marked for indefiniteness with an enclitic=ī on the nouns *dōst=ī* "a friend" and *zan=ī* "a woman," as in the following examples. This pattern has been attested in Middle Persian[65] and Old Shirazi.[66] Definite NPs, on the other hand, are generally considered to lack any consistent marker of definiteness and are left unmarked.

Ex. (26)

az	**dost=ī**	šenīd-am
from	friend=IND	hear.PST-1SG

"I heard from a friend"[67]

Ex. (27)

zan=ī	būd	dīvāneh
woman=IND	be.PST.3SG	crazy

"it was a crazy woman"[68]

Once introduced, a referent has the status of definite (anaphoric definite). The two most common strategies for indicating definiteness across CNP (ignoring anaphoric pronouns and zero anaphora) are either combining the noun with a demonstrative pronoun, preferably the distal demonstrative *-ān*, or using the bare form of the noun with

---

[60] *Nowruznāme*, 29.

[61] *Qorʾān-e Qods*, 136.

[62] Similar functions are attested for the Balochi of Sistan; see Nourzaei, "Definiteness Marking."

[63] Nourzaei, "History of the Suffix *-ū* in Shirazi"; Firoozbakhsh, "The Former Dialect of Šīrāz."

[64] The term "discourse-new" is here defined as the first mention of a noun in the discourse.

[65] Nourzaei and Jügel, "On the Function of *-ag* Suffix in MP"; Josephson, "Definiteness and Deixis in Middle Persian."

[66] Nourzaei, "History of the Suffix *-ū* in Shirazi"; Firoozbakhsh, "The Former Dialect of Šīrāz."

[67] *Nowruznāme*, 24.

[68] Ibid.

no additional marking.[69] The following passages (taken from *Dārābname*) demonstrate these two possibilities. A garden is introduced as a singular indefinite in example (28):

Ex. (28)
dar	*bīrūn=e*	*šahr*	***bāġ=ī***	*Būd*
in	outside=EZ	city	garden=IND	be.PST.3SG

"[Enalhayāt] he had a garden outside of the city, (lit. there was a garden for him)"[70]

The second mention (anaphoric definite) takes the distal demonstrative *ān* "that" in combination with the noun *ān bāġ-rā,* "that garden":

Ex. (29)
ān	***bāġ-rā***	*nešāt ābād*	*mī-gū-yand*
DIST	garden-OBJ	Neshāt Abad	IMP-say.NPST-3PL

"they call that garden Neshāt Abad"[71]

After this introductory sequence, there are several lines of intervening text with distal demonstratives referring to the garden before it is mentioned again as a bare noun *bāġ,* "the garden":

Ex. (30)
ayām=e	*bahār*	*būd*	*va*	*Enalhayāt*	*dar*	***bāġ***	*būd*
time=EZ	spring	be.PST.3SG	and	Enalhayāt	in	garden	be.PST.3SG

"it was spring, Enalhayāt was in the garden"[72]

Similar examples with bare nouns can be found in comparable contexts in all works. A similar system has been noted for other Iranian languages such as Vasfi,[73] Balochi,[74] and Kurdish.[75]

In sum, I can conclude that, although discourse-new, singular nouns are consistently marked throughout CNP, the marking of definiteness is not consistent. The two strategies most commonly mentioned are the use of the demonstrative plus noun, or the bare form of the noun.

3.4 K-suffixes as Signals of Proximity

The K-suffixes occur in what I will refer to as contexts of proximity. By this I mean contexts in which the referent is an item within the immediate perceptual range of the interlocutors, and will therefore often be accompanied by a proximate demonstrative. Thus, we have a combination of a proximal demonstrative and a noun carrying a K-suffix, as in example (31).

---

[69] For the same pattern in Middle Persian, see Nourzaei and Jügel, "On the Function of *-ag* Suffix in MP." Josephson, "Definiteness and Deixis in Middle Persian," 27–28, gives examples of the following sequences of first mention and continuation: (a) bare noun – bare noun; (b) noun=*ē(w)* – bare noun; (c) bare noun – *ān* noun; (d) noun=*ē(w)* – *ān* noun.

[70] *Dārābnāme 1*, 40.

[71] Ibid.

[72] Ibid.

[73] Stilo, *A Grammar of Vafsi*, mentions that, "The adnominal proximal demonstrative *in* 'this' tends to have a much higher frequency in extended speech in Vafsi than we might expect. While we see that this bleaching is a tendency in Vafsi, it is clearly not fully grammaticalized, and occurs much less commonly than the definiteness strategy [null marking]."

[74] Nourzaei, "Definiteness Marking."

[75] Haig, "Optional Definiteness"; Haig et al., "Definiteness Markings in Kurdish."

Ex. (31)

ǰavān=ī	*be*	*tāmāxara/e*[76]	*ve-rā*	*goft*	*ey*	*šayx*
youth=IND	with	ridiculed	PN.3SG-OBJ	say.PST.3SG	VOC	old.man

īn	***kamān-ak***	*bečand*	*xarīd-e-ī*			
PROX	hunchback-EV	how much	buy.PST-PP=COP.NPST.2SG			

"a youth ridiculed him, saying: 'O old man, how much have you bought [for] this hunchback?'"[77]

Note that this example lacks any obvious physical size connotations. Instead, it seems to be dependent on a deictic concept of proximity. This is one of most prevalent functions of the K-suffix *-ō* in Old Shirazi.[78]

3.5 K-suffixes as Signals of Recognition and Familiarity

The only evidence of a familiarity/recognitional reading of the K-suffixes occurs in some works under a relatively tightly constrained set of conditions, and only with the singular nouns discussed in examples (32) and (33).

The following passage is taken from an account in *Nowruznāme*.[79] In line 3 of the story, the boy has been introduced for the first time with *pesar=ī* "a boy," and the writer refers to the same referent, "boy," with a proximal demonstrative plus a K-suffix. Among the spectators, the king is pointing to the boy. He says "bring that boy to me," in line 5 of the story, which refers to the same referent again with a demonstrative pronoun plus a K-suffix (when the king commands his ministers to bring that boy to the palace). Interestingly enough, at the end of the same line, he refers to him with a K-suffix without a demonstrative pronoun. In the rest of this account, the writer refers to him either with a bare noun *pesa-rā* "the boy" or a distal demonstrative pronoun plus null form *īn pesar/ān pesar* "this boy/that boy." This passage demonstrates that the K-suffix does not convey the physical size of the boy, but instead illustrates a familiarity/recognitional notion of the reference.

Ex. (32) K-suffix with a familiarity/recognitional reading

čon	*be*	*dar=e*	*darvāz=e*	*šahr*	*resīd*		*češm=aš*
when	to	door=EZ	gate=EZ	city	arrived.PST.3SG		eye=PC.3SG

dar mīyān=e	*nazārgī-yān*	*bar*	***pesar=ī***	*oftād*	*čerk-īn*	*ǰāmeh*
among=EZ	spectator-PL	to	boy=IND	fall.PST.3SG	dirty-ATRR	clothing

beqadr	*davāzdah*	*sāleh*	*ammā*	*saxt*	*nīkū*	*rū*	*va*	*torfeh*	*va*
about	twelve	year	but	very	handsome	face	and	charming	and

zībā	*būd*	*tamām*	*xelqat*	*motadel*	*qāmat*	*enān*	*bāz*	*kešīd*
pretty	COP.PST.3SG	perfect	creation	middle	stature	bridle	PREV	pull.PST.3SG

va	*goft*	***īn***	***pesar-ak-rā***	*pīš=e*	*man*	*ār-īd*
and	say.PST.3SG	PROX	son-EV-OBJ	front=EZ	PN.1SG	bring.NPST-2PL

čūn	*bīy-āvar-dand*	*goft*	***ay***	***pesar***	*to*	*če*
when	be-bring.PST-3PL	say.PST.3SG	VOC	son	PN.2SG	what

kas=ī	*va*	*pedar*	*kī=st*	*goft*	*pedar*
person=COP.NPST.3SG	and	father	who=COP.NPST.3SG	say.PST.3SG	father

na-dār-am	*va laykan*	*mādar=am*	*be*	*folān*	*mohalat*
NEG-have.NPST-1SG	but	mother=PC.1SG	in	such and such	area

---

[76] Moʿin, *Farhang-e Farsi*, 1137.

[77] *Qābusnāme*, 58.

[78] Nourzaei, "History of the Suffix *-ū* in Shirazi."

[79] *Nowruznāme*, 74–77.

nešīn-ad	goft	če	pīšhe	mī-āmūz-ī		goft
sit.NPST-3SG	say.PST.3SG	what	skill	IMP.learn.NPST-2SG		say.PST.3SG
Qorān	hefz	mī-kon-am	farmūd		tā	**ān**
Quran	memorized	IMP-do.NPST-1SG	command.PST.3SG		that	DIST
pesar-ak- rā	be	sarā	bord-and	čon	soltān	forūd āmad
boy.EV-OBJ	to	palace	take.PST.3PL	because	sultan	down come.PST.3SG
pesar-ak-rā	pīš	xānd				
boy.EV-OBJ	front	call.PST.3SG				

"Sultan Mahmud, having arrived at the door of the city gate, among the spectators, he saw (lit. his eyes fell to) a boy, in dirty clothes, about 12 years old, but, very handsome and charming and pretty, with a perfect disposition and of moderate stature. He pulled on the bridle and said, 'bring this boy to me'; when they brought [him], he said, 'O boy, who are you and who is your father?'; he said, 'I do not have a father, but my mother is living in such and such an area.' He said, 'what skill are you learning?' He said, 'I am memorizing the Quran'; he commanded that the boy be brought to the palace; when the sultan got off his horse, he called the boy [...]"[80]

In the works, I only found one particular case of this. In line 1 the doctor is introduced in the discourse for the first time without the K-suffix *tabīb=ī* "a physician," and in line 5 the writer refers to the same referent with a K-suffix *tabīb-ak* "the doctor." In the rest of the story, the same referent appears without the K-suffix, *tabīb* "the physician." Such passages demonstrate that the K-suffix does not express any physical notion about the physician. Instead, it conveys familiarity/recognition.

Ex. (33) K-suffix with a familiarity reading

tabīb=ī	az	sāmānī-yān-rā	selat	nīkū	dād	**tabīb-ak**
physician =IND	from	Samanī-PL-OBJ	gift	good	give. PAST.3SG	physician-EV
čūb band va	talī	āvard		va	goft	īn	pā=ye
wooden stick and	band	bring.PST.3SG		and	say.PST.3SG	PROX	foot=PC.3SG
be-škast		har	rūz	**tabīb-rā**	mī-pors-īd		
be-break.PST.3SG		every	day	physician-OBJ	IMP-ask.PST.3SG		

"He gave a rich reward to a physician from Samānīyan; the physician brought a wooden stick and band and said 'his leg was broken'; he asked the physician every day."[81]

Note that we do not have sufficient examples of this type to draw any significant conclusion. In the later stages of Persian, for instance in *Golestān Saʿdī* and *Totināme*, we cannot find these types of passages. It would be interesting to closely examine this suffix from the fourteenth to the early nineteenth centuries to see which evaluative notions are more predominant.

Summary
The corpus data for CNP demonstrate that the K-suffix has evaluative semantics that account for most of its usage. It is compatible with indefiniteness contexts, and there are no structural constraints (see CWP and CSP on this issue). It somewhat resembles a sporadic remnant of a now defunct morphology that appears to have been incorporated into some items without any discernible change in meaning; see examples (19) and (21).

In CNP, however, we find nouns accompanied by demonstratives and nouns taking a K-suffix, with no clear connotation of small size, little amount, or clear evaluative content. These passages provide some evidence of how evaluative markers might have evolved towards definiteness marking. One of the most recent cross-linguistic studies on diminutives demonstrates that diminutives also convey meanings of endearment, familiarity, and

---

[80] Ibid.
[81] *Tārikh-e Beyhaqi 2*, 495.

proximity.[82] In the case of the proximity and recognitional contexts shown in examples (32) and (33), the concept of familiarity is reduced to physical proximity and shared common ground. Thus, it is not unreasonable to see an evaluative suffix becoming associated with proximity in a non-evaluative sense. We have already observed the concepts of proximity and shared common ground in the K-suffix in Balochi,[83] and it is the most prominent function of the K-suffix *-ō* in Old Shirazi Persian,[84] although in both Sistani Balochi and Old Shirazi, evaluative usage prevails overall. The suggestion here is that the proximate and shared-knowledge usage may have provided a bridging context for the transition from evaluative meaning to definiteness marking.

4. The K-suffix in Contemporary Written Persian: Initial Observations

Data for Contemporary Written Persian are taken from books written in colloquial Persian published from the late nineteenth to mid-twentieth centuries. Table 2 gives an overview of these books.

So far, I have given a detailed discussion of the nature of the K-suffix *-ak* in CNP (see Section 3). Across the works, we only found one form of the K-suffix, namely, *-ak*. However, in the CWP books we found four varied forms of the K-suffix (see Section 7 for a discussion of their origin):

(a) a continuation of the K-suffix *-ak* in CNP as an evaluative notion, e.g., *Hammad-ak*, "Ahmad," *dīb-ak* "demon,"[85] and *hamūm-ak* "bathroom."

(b) the existence of new K-suffixes, e.g., *īk*, in *zan-īk-e*, "woman," *ū*, in *yār-ū* "friend,"[86] *-ī*, in *Hasan-ī* "Hasan," and *-e* in *pesar-e*, "boy," which are mostly found in colloquial and informal written texts with mostly singular nouns.[87] I assume the *-ī* suffix to be a short form of the *-īk* suffix in *Hasan-ī* "Hasan."[88] Determining whether or not they derive from the same origin is not the main point of this paper; what is important is that they display similar (evaluative) semantics.

To the best of my knowledge, Qarib and colleagues[89] and Anvari[90] present the K-suffix *-e*, including *-ū* and *-ak* and *-če*, as a diminutive marker in their studies. However, a definiteness effect associated with the K-suffix *-e* in Modern Persian has already been mentioned by

---

[82] Ponsonnet, "A Preliminary Typology," section 2.

[83] Nourzaei, "Definiteness Marking."

[84] Nourzaei, "History of the Suffix *-ū* in Shirazi."

[85] *Zende be gur*, 108–9.

[86] The K-suffix *-ū/ūk* is found in other Persian varieties, including Bambi, Kermani, Yazdi, e.g., *pesar-ūk, doxtar-ūk*. It has been reported for the Sangsari dialect as well, Sabbaqiyan, *Barrasi-ye zabān-e sangsari*, 133–45.

[87] I have found forms with such words as *martīke, mardīke, mardake* "man" and *zanīke/zanake* "woman," and once with *pesarē/pesarīke* "boy." I am uncertain of the origin of *-īk*; it is an evaluative suffix. Cross-linguistically, it is possible to have more than one diminutive suffix on words, such as in Slavic languages; for Russian, see Volek, *Emotive Signs*. We find the same nouns with two evaluative suffixes in Balochi: *mard-ak-ok* "man," *ǰan-ak-ok* "woman," *maškečok* "goat skin," where the first K-suffix appears to have been re-analyzed as part of a word stem. It is also attested in Kurdish as *ženek*. Note that these words are not common in CSP, but they can be found in some older speakers' daily speech (unpublished Hamedani tale); the standard terms are *mard* and *zan*.

[88] Khatamipoor, "*yā-ye maʿrefeh*," 18, mentions that the K-suffix *-ī* is a definiteness marker in Kashmari dialect. Future corpus-based investigation is needed to ascertain how far this suffix has been grammaticalized as a definiteness marker.

[89] Qarib et al., *Dastur-e Fārsi*, 46.

[90] Ahmadi Givi and Anvari, *Dastur-e zabān-e Fārsi 1*, 7.

various scholars.[91] In the following section, I will discuss the K-suffix -*e* in Contemporary Written (see below section) and Contemporary Spoken Persian in Iran (Section 5).

4.1 K-suffix -e in Contemporary Written Persian

Before we study the status of the K-suffix -*e/he* in CSP, I will give a detailed description of the K-suffix -*e* in CWP. In contrast to the K-suffix -*ak* in CNP (Section 3), the K-suffix -*e* is mostly attested in informal and colloquially written books with a handful of singular nouns.[92] Note that I found three instances of the K-suffix -*e* with the plural marker -*hā* e.g., *čerā mesl e xāle zan-īk-e-hā harf mīzanī* "why are you talking like gossiping women?"[93]

Its semantic domains in CWP are, to a large extent, similar to those in CNP. However, there are some examples of K-suffixes that distinguish CWP from CNP (see Section 4.2).

Analysis of the K-suffix in CWP

As in CNP, the K-suffix in CWP is compatible with indefinite contexts. See example (34).[94]

Ex. (34) The K-suffix with an indefinite context
*īn   hame   māl   va   dolat-rā   **yek   doxtar-e=ye** mast   va*
PROX   all   property and   property-OBJ one   girl-EV=EZ drunk   and
*malang   bar   dār-ad   va   bā   **yek   pesar-e=ye** nāġolā=e ahl=e*
crazy   PREV   take.NPST-3SG   and   with one   boy-EV=EZ   sly=EZ   belonging=EZ
zoleme=ye   aldang   be-xor-and
cruel=EZ   idle   SUBJV-eat.NPST-3PL
"a drunk and crazy girl took this much money and property and spent it together with a sly boy [...]"[95]

Ex. (35) The K-suffix with an indefinite context
*tū=ye   yek=ī   az   īn   otomobīl-hā   **zan-īk-e=ī**   nešast-e*
in=EZ   one=IND   from   PROX   car-PL   woman-EV-EV=IND   sit.PST-PP
būd
COP.PST.3SG
"in one of these automobiles, a woman was sitting"[96]

It has been attested with the proper nouns *ādm-e* and *Havvā-e*, which are signals of the endearment connotations of this suffix.[97] Note that the same writer used *ādm* and *Havvā* without marking them with a K-suffix in his short story titled *Afsāneye Afarīnesh*.

Ex. (36) The K-suffix with proper nouns

sāl-hā	*āmad*		*va*	*sāl-hā*	*raft*	*āš=e*	*pošt=e*	*pā=ye*
year-PL	come.PST.3SG		and	year-PL	go.PST.3SG	soup=EZ	back=EZ	foot=EZ

*ān-hā-rā   ham   sar=e   hafte   **nane   havvā-e   va   bābā   ādam-e***
DIST-PL-OBJ   ADD   head=EZ week   mother Havvā-EV   and   father Adam-EV

xord-e   būd-and
eat.PST-PP   COP.PST-3PL

"Many years passed (lit. years came and years went). By the end of the week, mother Havvā/Eve and father Adam had eaten their farewell soup."[98]

Example (37) is an ambiguous case. The K-suffix could be interpreted as adding a flavor of sorrow/empathy on the part of the speaker regarding the fate of the small, orphaned boy. It could also be interpreted as a recognitional context, when the girl again refers to the boy after several intervening lines.

Ex. (37)

*bābām   češm   na-dāšt   **īn   pesar-e-rā**   be-bīn-ad*
father.PC.1SG   eye   NEG-have.PST.3SG PROX boy-EV-OBJ   SUBJV-see.PST-3SG

zīr=e   korsī   nešast-e   būd-īm   pā=ye   īn   bače   ke
under=EZ   *kursi*   sit.PST-PP   COP.PST-1PL   foot=EZ PROX child CLM

be   korsī   mī-xor-d
to   *korsi*   IMP-eat.NPST-3SG

"my father could not stand to see this boy, (lit. my father does not have eyes to see this boy), we sat under the *korsi*,[99] when the foot of this child touched the *korsi* [...]"[100]

The K-suffix also occurs with pejorative connotations, as in the following examples. This can be observed in vocative contexts. Note that there are two evaluative suffixes on the items in examples (38)–(40).

Ex. (38)

*ay   **pesar-e**=ye   ahmġ*
O   boy-EV=EZ   silly

"O silly boy"[101]

Ex. (39)

zan-īk-e   harf=e   dahan=eš-o   ne-mī-fahm-e
woman-EV-EV   word=EZ   mouth=PC.3SG-OBJ   NEG-IMP.know.NPT-3SG

"[the] woman does not know how to speak"[102]

Ex. (40)

*īn   **mard-īk-e**   ǰonūn   dār-d*
PROX   man-EV-EV   crazy   have.NPST-3SG

"this man is crazy (lit. this man has mania)"[103]

In example (41) the K-suffix occurs in a vocative context:

---

[98] ⁽Alaviye khānom, 76.

[99] A square table covered by a blanket with a brazier beneath it.

[100] *Chamedān*, 80.

[101] *Tamsilāt*, 233.

[102] ⁽Alaviye khānom, 41.

[103] *Chamedān*, 19.

Ex. (41)

šāhzāde	*goft*	*peser-e*	*šanīd-am*	*to*	*xūb*	*so☐bat*
prince	say.PST.3SG	boy-EV	hear.PST-1SG	PN.2SG	good	speak

mī-kon-ī
IMP-do.NPST-2SG
"the prince says, 'O boy, I have heard that you are speaking well'"[104]

Similar to the K-suffix -*ū* in modern Shirazi Persian, I find it in indefiniteness contexts, as in example (42).

Ex. (42)

dīd-an	*ye*	*mart-īk=e*	**ġūzal-ū**	**lāġar-ū**
see.PST-3PL	one	man-EV=EZ	humpbacked-EV	thin-EV

"they saw a humpbacked and skinny man"[105]

Finally, I should point out that certain words, typically indicating place referents, seem to include the K-suffix as part of the word stem, such as in example (43). Note that some compound nouns, such as *Albālū xošk-e* "dry-cheery" in ꜥ*Alaviye khānom*, need further investigation regarding the function of -*e*.[106]

Ex. (43)

se	*nafarī*	*aġlā=šūn-o*	*rū*	*ham*	*rīx-tan*	*ke*
three	person	wisdom.PL=PC.3PL-OBJ	on	add	pour.PST-3PL	CLM

be-r-an	**emām-e**
SUBJV-go.NPST-3PL	NP-

"all three decided to go to Emāme"[107]

In contrast to the K-suffix in CNP, the K-suffixes are not attested with possessed nouns formed with person-marking clitics or copula verbs (see example 24). When a noun and an adjective are combined, the K-suffix is attached to the second constituent of the NP, as in *pesar bozorg-e* "the old brother." See the following example.

Ex. (44)

pesar	**bozorg-e**	*ke*	*dar=e*	*xāne=ye*	*bābāš-rā*	*vāz*	*kard-e*
son	young-EV	CLM	door=EZ	house=EZ	father.PC.3SG-OBJ	open	do.PST-PP

būd
be.PST.3SG
"when the older son has opened the father's door of the house"[108]

Note that in some books written earlier in the period being studied, we find the K-suffix on the first constituent of compound nouns (a noun combined with an adjective) such as *doxtar-e=ye češm sefīd* "impudent girl."[109] It seems that the movement of the K-suffix to the second constituent of the noun phrase occurred in its later stages of grammaticalization.

---

[104] *Siyāhatnāme 1*, 54.
[105] ꜥ*Alaviye khānom*, 112.
[106] Ibid., 80.
[107] Ibid., 106.
[108] Ibid., 75.
[109] *Tamsilāt*, 259.

4.2 Attestation of the K-suffix -e in Non-evaluative Contexts[110]

We have already found some contexts where the K-suffix *-e* does not express a diminutive or evaluative sense. Instead, the item marked with the K-suffix has a referent in the previous clauses or, in some cases, the marked items can refer to common background knowledge.

Before introducing these passages, I will briefly summarize definite and indefinite strategies in CWP. As in CNP (Section 4), discourse-new, specific, singular NPs are overtly marked for indefiniteness across the CWP texts. Definite NPs, on the other hand, are generally considered to lack any consistent signal of definiteness.

Indefinites are marked slightly differently than in CNP (see Section 3.3). The word *ye/yek* "one" preceding the noun (*ye kaftār*, "a hyena") may combine with a suffix=ī (*yek martīke=ī* "a man"). Once introduced, a referent has the status of definite (anaphoric definite). As in CNP, there are two common strategies for indicating definiteness throughout CWP: (a) combining the noun with a demonstrative (*ān doxtar,* "that girl"), (b) using the bare form of the noun with no additional marking (*kaftār* "the hyena").[111]

In the following passage, taken from a story in ʿAlaviye khānom, the word *kaftār* "the hyena" is introduced in the discourse as a singular indefinite.

Ex. (45)

az	tū=ye	qabrestān=e	kohne=ī	**yek**	**kaftār**	bar	mā	mī-gūz-īd
from	in=EZ	graveyard=EZ	old=IND	one	hyena	to	PN.1PL	IMP-fart.PST-3SG

paydā	kard-an
find	do.PST-3PL

"in an old graveyard, they found an arrogant/conceited hyena"[112]

Following the introduction, the second mention (anaphoric definite) takes a bare noun *kaftār.* The writer refers to it several times in the story with a bare noun *kaftār.* He only marks it with the K-suffix *-e* once (on page 127), while in the rest of the story it appears as a bare noun.

Ex. (46)

kaftār-e-ro	bā	dāyereh	va	dombak	vāred=e	kešvar=e	xar	dar
hyena-EV-OBJ	with	tambourine	and	tombak	enter=EZ	country=EZ	donkey	in

čaman	kard
lawn	do.PST.3SG

"[the fox] accompanied the hyena ceremoniously (lit. with tambourine and drum) in the land where donkeys [graze] on the lawn"[113]

It is evident from these passages that the K-suffix does not express an evaluative sense. Still, the K-suffix does not mark the items consistently or systematically. It is hard to find a motivation for the writer to mark the same item with a K-suffix only once, and not in the remaining passages of the story.

Similarly, in the following example, the NP, girl, has been introduced for the first time in the story in a restrictive relative clause, *ān doxtarī ke* "that girl who."

---

[110] Because at this stage the K-suffix *-e* does not systematically appear as a definiteness marker in the texts, I would prefer to keep "EV" as a general term in the glosses.

[111] Note that Meshkat al-Dini, *Dastur-e zabān-e Fārsi*, 148, and Ahmadi Givi and Anvari, *Dastur-e zabān-e Fārsi 1*, 64, consider the first possibility to be a definiteness reading of nouns in Persian.

[112] ʿAlaviye khānom, 121.

[113] Ibid., 127.

Ex. (47)

bā	ham-īn	medād	būd	ke	ǰā=ye	molāɡāt=e	xodam –rā
with	EMPH-PROX	pencil	be.PST.3SG	CLM	place=EZ	meeting=EZ	REFL-OBJ

navešt-am	dād-am	be	**ān**	**doxtar=ī**	**ke**	tāze	bā	ū
write.PST-1SG	give.PST-1SG	to	DIST	girl=IND	CLM	recent	with	PN.3SG

āšenāh	šod-e	būd-am
acquainted	become.PST-PP	COP.PST-1SG

"It was with this pencil that I wrote my meeting place [address] [and] gave it to that girl with whom I have become acquainted recently"[114]

The second mention in line 12 takes the distal demonstrative *ān doxtar*, "that girl." In line 36, the writer again refers to the girl and marks it with the K-suffix, as in the following example.

Ex. (48)

doxtar-e	bekolī	az	yādam	raft-e	būd
girl-EV	totally	from	memory	go.PST-PP	COP.PST.3SG

"I forgot the girl (lit. the girl has gone from my memory)"[115]

In line 38 the writer refers to the girl with a combination of the distal demonstrative and the K-suffix *-e, ān doxtar-e* "that girl."

In the following example, the item *abre* "cloud" marked with the K-suffix *-e* has a referent in the previous context *yek teke abr* "a bit of cloud." Note that it comes with the distal demonstrative. It is also worth noting that throughout the books, there are very few passages where the second mention (anaphoric) is marked with a K-suffix (see CSP on this issue).

Ex. (49)

ye	teke	abr	az	ūn	abr-ā=ye	50×50	metr=e	mokʕab	az
one	piece	cloud	from	DIST	cloud-PL=EZ	50×50	metre=EZ	square	from

pošt=e	kū-hā	padīdār	šod	hamīnke	**ūn**	**abr-e**	šorūʕ
back=EZ	mountain-PL	appear	become.PST.3SG	as soon as	DIST	cloud-EV	start

kard	be	bār-īd-an
do.PST.3SG	to	rain.PST-INF

"a 50×50-square-meter bit of cloud [coming] from those clouds appeared from the back side of the mountains, as soon as that cloud started to rain [...]"[116]

Similarly, in the following example, the item *doxtar-e* "the girl" marked with the K-suffix *-e* has a referent in the previous context *ye yatīm-ī* "an orphan." In the continuation of the story, the same referent appears as a bare noun and PROX+NP. It is notable that, after 17 lines, the writer refers to the girl and marks the referent with a K-suffix *-e*, as *doxtar-e* "the girl."

Ex. (50)

be	hāǰī	xabar	dād	ke	**yatīm=ī**	dar	īn	kočeh
to	Haǰi	news	give.PST.3SG	CLM	orphan=IND	in	PROX	street

hast	gozašt-e	az	īn	ke	**doxtar-e**	az ɡarārī	ke
COP.NPST.3SG	pass.PST-PP	from	PROX	CLM	girl-EV	beside	CLM

šenīd-eh	ast	xošgel	ast
hear.PST-PP	COP.NPST.3SG	beautiful	exist.NPST.3SG

---

[114] *Zende be gur*,12.

[115] Ibid., 14.

[116] *ʕAlaviye khānom*, 115.

"he informed Ḥaǰī [saying] there is an orphan on this street, [...] besides, based on what he heard, the girl is beautiful"[117]

Example (50) is a unique case in the corpus. In the story, *pesar* "the boy" appears as a bare noun. It is marked just once with the K-suffix in combination with the demonstrative when the man points to the boy and says, "he is not a painter, he is reciting a poem for this boy who is sitting in front of the shop." In the rest of the text, the same referent "boy" appears as a bare noun.

Ex. (51)

be-**dīn**	**pesar-e**	šeʿr	mī-band-ad
to-PROX	boy-EV	poem	IMP-close.NPST-3SG

"he is reciting poem[s] for this boy"[118]

The writer similarly marks the item *zan azīz-e* "beloved wife" with a K-suffix, when the woman is pointing to another woman standing close by and says to the man that the beloved wife (lit. dear woman) is over there.

Ex. (52)

zan	azīz-e	ānǰāst
wife	beloved-EV	DIST.COP.NPST.3SG

"the beloved wife is there"[119]

After this, the writer refers back to it either with a bare NP or a combination of demonstrative plus noun.

The following examples, (53) and (54), demonstrate a mutuality reading. Mutuality involves contexts in which the identity of the referent is known by both speakers through their shared world knowledge, even though the referent has not previously been introduced in the linguistic context.

The marked noun dom=e *šotor-e* "the tail of the camel" does not have a referent in the previous clauses. However, the writer still marks it with the K-suffix because it is familiar to both writer and reader via their common cultural background. This usage has been reported for the K-suffix *-ō* in Old Shirazi.

Ex. (53)

tā	ūn	bīy-ā-d	mard	beše	dom=e
till	DIST	SUBJV-come.NPST-3SG	man	SUBJV.become.NPST.3SG	tail=EZ

šotor-e	be	zamīn	mī-res-e
camel-EV	to	earth	IMP-arrive.NPST-3SG

"until that one has become mature (lit. man) the tail of the camel will reach to the ground"[120]

Note that the same expression is not marked with the K-suffix in his other book *Zende be gur*.[121]

---

[117] *Charand o parand*, 112.
[118] *Siyāhatnāme 1*, 163.
[119] *Zende be gur*, 99.
[120] *ʿAlaviye khānom*, 57.
[121] *Zende be gur*, 94.

Ex. (54)

ġese-e	mā	be	sar	resīd	kalāġ-e	be	xūn=aš
story=EZ	PN.1PL	to	end	arrive.PST.3SG	crow-EV	to	home=PC.3SG

na-res-īd
NEG-arrive.PST-3SG

"our story finished (lit. came to the end) [but] the crow did not arrive at its home"[122]

Summary

Across the texts, the K-suffix *-e* of CWP is quite similar to that of CNP, with evaluative connotations accounting for the greatest amount of use. It has been attested in indefiniteness contexts. It shares deictic and recognitional uses with CNP in broader contexts. However, we also encounter some instances in which the K-suffix marks items that have a referent in a previous context and do not convey any evaluative sense. Such examples are rare, but they indicate how an evaluative suffix can develop into a definiteness marker and pave the way towards anaphoric definiteness (for discussion of this as a typical pattern in CSP, see Section 5). In contrast to the K-suffix in CNP (see examples 22–23), the K-suffix *-e* does not occur with plural markers and possessive constructions, typically when the latter are formed with person-marking clitics and enclitic verb copulas.

This observation can be linked to Hawkins's suggestion that each stage of grammaticalization "maintains the usage possibilities of the previous stage and introduces more ambiguity and polysemy, but expands the grammatical environments and the frequency of usage of the definite article."[123]

Finally, what should we call the K-suffix *-e* in CWP?[124] In my view, this is an open question, however, as we can see above and in Section 4.1, the K-suffix *-e* is not yet mature and has not grammaticalized as a definiteness marker as such. It is scattered unsystematically throughout the texts and largely preserves its original evaluative connotations. It is still on the way towards becoming a definiteness marker in Persian, as will be discussed in the next section.

5. Contemporary Spoken Persian

Data for the CSP stem from Persian Language Database (PLD) online corpora,[125] Taghi's corpus,[126] and my new recordings of Tehrani speakers from Tajrish and my field notes.[127] The corpus contains a total of 60,207 words (see Table 3 for an overview). In addition, I use spontaneous speech data from Bamberg-Hamedan joint online data,[128] a variety called Hamedani Persian, and my new recordings. The main speech topics are personal accounts, education, science, and so on.

5.1 Background of Speakers

I do not know the age of the participants for the PLD corpora, as I was informed that the data was recorded from native, educated Tehrani male and female speakers who were born and lived in Tehran. The main speech topics are marriage, women's rights, tales, and free

---

[122] Ibid., 131.

[123] Hawkins, *Efficiency and Complexity in Grammars*, 86.

[124] The definition of a "definite article" is a very controversial issue. Becker, in "Articles in the World's Languages," 86–87, claims that "what definite articles are required to encode are anaphoric, bridging, situationally unique, and established referents"; she emphasizes that the crucial issue is not fully obligatory usage, but rather systematic association with the relevant contexts. Ibid., 36–44.

[125] See http://pldb.ihcs.ac.ir/Default Persian Language Database.

[126] Taghi, *A Typology and Classification of Three Literary Genres*.

[127] Nourzaei, Unpublished texts, recorded between 2018 and 2021.

[128] See https://multicast.aspra.uni-bamberg.de/resources/hambam/.

Table 3. An overview of the corpus.

Stages	Texts	Words	Mean text size	Range text size	N texts>700 words
CNP	12	1,383,426	115,286	14,613–505,808	12
CWP	14	775,664	64,639	26,668–235,010	14
CSP	30	60,207	3,542	610–5,636	9

conversations recorded in (1370/1991) and written down in Persian. I transcribed them for this work. The recorded data is about three hours long.

I use twelve texts published in Taghi.[129] These texts are recorded from two Tehrani speakers aged seventy-three and seventy-five, and written down in Persian. I transcribed them for this study. According to the information supplied by Taghi, both speakers were educated in Islamic schools (*savād maktab*). They were born in Tehran and lived there for their entire lives. The second speaker moved to Sweden at the age of seventy, but traveled back and forth between there and Tehran.

My data consists of recordings of bibliographical tales and accounts (about one hour) told by Tehrani-educated speakers from Tajrish aged between forty and sixty-five years.

Regarding Hamden-Bamberg, the data consists of recordings of male and female Hamdani speakers aged between thirty and seventy years with different backgrounds from 2017 onwards.[130]

For colloquial Tehrani Persian, I complement the quantitative data with qualitative material which illustrates the various functions with authentic examples and appropriate references to context. I also refer to the results of a questionnaire-based survey with Persian speakers based on the English version of the questionnaire used for Kurdish, Balochi, Shirazi and Lori to capture authentic colloquial speech.[131] I have modified the questionnaire slightly by reducing the number of plural NPs due to the incompatibility of the K-suffix with plural nouns.

In the previous section, I gave a detailed discussion of the K-suffix -*e* in CWP. Now I will discuss the status of the K-suffixes -*e/he/ye* in CSP. The K-suffixes -*e/he* have been attested in different varieties of Persian, for instance, Taghi ābād, Esfahani, Hamedani, Yazdi,[132] Najaf ābādi, Qomi, Mashhadi,[133] Birjandi, Qayeni and Neshaburi.[134] Notably, the K-suffix-*e/he* has not been attested in Sistani Persian, which is the variety spoken in Sistan and Balochistan province.[135]

Based on the data available in the Kalbasi,[136] the Taghi[137] and the online Bamberg-Hamedan corpora,[138] and my data, the status of the K-suffix -*e/he* is almost the same across Persian varieties: it is not obligatory but is systematically used in definite contexts. For instance, Hamedani Persian speech is similar to Tehrani Persian; the K-suffix is very sensitive to genre and setting, which means that it is not attested with scientific topics that need a formal setting. The frequency and usage of the K-suffix in anaphoric contexts (particularly its

[129] Taghi, *A Typology and Classification of Three Literary Genres.*

[130] See their online corpus for more details, https://multicast.aspra.uni-bamberg.de/resources/hambam/.

[131] Haig, "Optional Definiteness"; Haig et al., "Definiteness Markings in Kurdish"; Nourzaei, "Definiteness Marking;" Nourzaei, "History of the Suffix -*ū* in Shirazi"; Nourzaei and Haig, "An Overview of Definiteness Marking"; Nourzaei and Haig, *Emerging of Definiteness Markers in New Western Iranian Languages.*

[132] See Kalbasi's data, *Towsife gunehā-ye zabānī-ye īrān.*

[133] See Taghi, *A Typology and Classification of Three Literary Genres.*

[134] Nourzaei, Unpublished texts, recorded between 2020 and 2021.

[135] Nourzaei, Unpublished texts, recorded between 2012 and 2018.

[136] Kalbasi, *Towsife gunehā-ye zabānī-ye īrān.*

[137] Taghi, *A Typology and Classification of Three Literary Genres.*

[138] See https://multicast.aspra.uni-bamberg.de/resources/hambam/.

combination with demonstrative pronouns) diverge in these varieties. Therefore, another study is needed of these varieties using a larger corpus.

In the present study, I will concentrate on the status of the K-suffix *-e-he* in the Tehrani variety of Persian, for which I already have a large corpus at my disposal. Data for this section was taken from a large contemporary spoken online corpus, Persian Language Database (PLD), published texts of Tehrani Persian in Taghi's corpus[139] and my recordings of Persian speakers from Tajrish.

Before discussing the nature of the K-suffix, I will give an overview of the system of discourse-new nouns in this phase of Persian.

The system of discourse-new nouns, specific nouns for the singular, and plural nouns is the same as in CWP: the word *ye/yek* "one" precedes the noun, which may combine with a suffix *=ī/e*[140] on the noun to give an indefinite, singular, specific meaning, as in *ye olāġ=ī* "a donkey" and *ye šīr* "a lion."[141]

Ex. (55)

mī-bīn-an	*ye*	*olāġ=ī*	*gandom*	*bār=eš*	*hast*
IMP-see.NPST-3PL	one	donkey=IND	wheat	load=PC	COP.NPST.3SG

"they see a donkey is loading wheat"[142]

Similar to CNP and CWP, the most common strategy in CSP for marking a referent with a definite status is to use bare nouns or a combination of nouns plus demonstratives. However, in some genres, typically in folktales and biographical tales, a new strategy has emerged that marks the definite nouns with the K-suffix *-e/he* systematically, but not obligatorily, in anaphoric contexts. In the next section, I will illustrate this usage of the K-suffix.

5.2 K-suffixes as Definiteness Markers

The common form of the K-suffix in Contemporary Spoken Persian is *-e/he* (when a word ends with a vowel), for instance *kūze/kūze-he* "the jug," *bābā/bābā-he* "the father." These suffixes have generally not been attested in standard Persian.[143] In contrast to CWP, in CSP K-suffixes are not attested with evaluative or diminutive semantics or in indefinite contexts (see Section 4). In the following subsection I will discuss the K-suffix in CSP.

Anaphoric Definiteness

In CSP, singular nouns that are anaphorically definite take a K-suffix, when the relevant structural conditions obtain. The following examples (56 and 57) illustrate K-suffixes in anaphoric definite contexts, with both human and non-human nouns.

Ex. (56)   Anaphoric definite with a human noun

mī-bīn-an	*ye*	*pīrmard=ī*	*mesle*	*ye*	*ǰūǰe*	*rū=ye*	*zamīn*
IMP-say.NPST-3PL	one	old man=IND	Like	one	chick	on=EZ	ground

---

[139] Taghi, *A Typology and Classification of Three Literary Genres*.

[140] Taghi's corpus, *A Typology and Classification of Three Literary Genres*, 96, is the only one where the speaker introduces a new participant in the discourse with *yek* and *e*, for instance *ye pīrezan-e būde* "there was an old lady." I listened to the sound file of one text together with the author of the book. I can hear a short, unstressed *-e*. It might be another form of indefiniteness marker that so far has not been reported. This is a topic in need of further investigation with more examples of this construction.

[141] Note that in Taghi's corpus, *A Typology and Classification of Three Literary Genres*, 290, the discourse-new nouns appear as bare nouns, as in *mīre barāš **kor-e asp** mīxare ke sareš be īn kor-e asp-e garm beše*, "he buys a foal for him in order to be busy with this foal."

[142] Persian Language Database (PLD).

[143] I have found one instance of the suffix in formal text, with the word *pesar*, as **pesar-e** " داشت حرف میزد که پسره زد زیر چهارپایه " in a novel titled *Khun-khorde*, 133.

nešast-e   mīy-ā-d   šīr-e   bālā=ye sar=e   **pīrmard-e**
sit.PST-PP.COP.NPST.3SG   IMP-come.NPST-3SG   lion-DEF   up=EZ   head=EZ old man-DEF
"they saw an old man sitting on the ground like a chick, the lion came to the old man"[144]

Ex. (57)   Anaphoric definite with animate, non-human noun
ye   vaxt   ye   gorbe=ī   būd-e   bā   **ye   šīr=ī**   šīr-**e**
one   time   one cat=IND   be.PST-PP.COP.NPST.3SG   with one   lion=IND   lion-DEF
be   gorbe   mī-g-e
to   cat   IMP-say.NPST-3SG
"once upon a time, there was a cat with a lion […] **the lion** says to the cat"[145]

Ex. (58)   Anaphoric definite with animate, non-human noun
mī-r-e   mī-res-e   be **ye   gāv=ī**   gav-**e**
IMP-go.NPST-3SG   IMP-arrive.NPST-3SG   to one   caw=IND   caw-DEF
mī-bīn-e   ke   ye   šīr   dār-e   az   dūr mīy-ā-d
IMP-see.NPST-3SG   CLM one lion have.NPST-3SG   from far IMP-come.NPST-3SG
"it goes until [it] arrives at a cow, the cow sees that from a far distance a lion is coming"[146]

Ex. (59)   Anaphoric definite with inanimate nouns
sīb-hā-ro   mī-zāšt   dāxel-e **ye   sabad   sabad-e**   az
apple-PL-OBJ   IMP-put.PST.3SG   into=EZ one   basket basket-DEF from
dast=am   oftād
hand=PC.1SG   fall.PST.3SG
"he put the apples into a basket […] the basket fell down from my hand"[147]

Similar to Shirazi Persian, the K-suffix in CSP does not appear in combination with a demonstrative pronoun in anaphoric contexts, as in the following example:

Ex. (60)
mī-res-e   be ye   gāv=ī   va   mī-bīn-e   **In   gāv**
IMP-arrive.NPST-3SG   to one caw=IND   and IMP-see.NPST-3SG   PROX   cow
bast-e   zamīn
bind.PST-PP   ground
"he arrives at a cow, and he sees this cow bound to the ground"[148]

However, in Taghi's data, there are a few anaphoric contexts with a combination of a demonstrative pronoun plus a K-suffix, as in example (61). I have found a combination of the K-suffix with demonstrative pronouns in anaphoric contexts outside of the storyline when the storyteller explains the situation to the audience.[149]

Ex. (61)
hīčī   ǰavūn-e   xāst   be-r-e   mosāferat barā=ye teǰārat
any way youth-DEF   want.PST.3SG SUBJV-go.NPST-3SG   traveling for=EZ   trading
az   zan=eš   porsīd   čī   mī-xād   az   **īn**
from   wife=PC.3SG ask.PST.3SG what IMP-want-NPST-3SG   from PROX
doxtar-e ham porsīd
girl-DEF   ADD ask.PST.3SG

---

[144] PLD.
[145] Ibid.
[146] Ibid.
[147] Nourzaei, Unpublished texts, recorded between 2018 and 2021.
[148] PLD.
[149] See Taghi, *A Typology and Classification of Three Literary Genres*, 97.

"anyway, the youth wanted to go on a trading journey, he asked his wife what she wants, this girl [he] asked as well"[150]

A combination of the K-suffix with a demonstrative pronoun is common in other Persian varieties such as Hamedani in example (62), and in the Qomi variety of Persian.[151]

Ex. (62)

māl=e	ye	moʿalem=ī	bud	**īn**	**moʿalem-e**	taqrīban
belong=EZ	one	teacher=IND	be.COP.PST.3SG	PROX	teacher-DEF	almost

mī-šod	barādar=e	u	moʿalem=e	kelās	panǰom=e	man
IMP-become.PST.3SG	brother=EZ	PROX	teacher=EZ	class	five=EZ	PN.1SG

"it belonged to a teacher, that teacher, you know he was the brother of my grade five teacher"[152]

The appearance of double marking of definite forms is unexpected in the traditional scenario of developing definiteness marking from a demonstrative, and these instances certainly call for further investigation. However, the construction is not unexpected on the analysis suggested here, where we assume that the definiteness marking evolved from evaluative marking via the marking of proximity and shared knowledge/familiarity, which is supported by our results here (see Section 3 on CNP) and also has occurred in Balochi and Old Shirazi. If this really is the first developmental stage, then it is not surprising that it is still available here in the speech of older speakers. For Old Shirazi, we have evidence that the K-suffix always occurs with a demonstrative in earlier stages of the language. At its current stage we observe a complete absence of the demonstratives in anaphoric contexts and a tendency not to use them in situational contexts.[153]

These observations support my hypothesis that in earlier stages of the grammaticalization of the K-suffix towards definiteness, it occurred with the demonstratives and used them as supporting items/hooks before becoming a pure definiteness marker. In this respect, CSP is at an earlier stage of grammaticalization of the K-suffixes, and traces of this earlier stage can still be found in the speech of older speakers.

Bridging and the K-suffix

Under the heading of bridging definiteness, we include referents that are identifiable based on their unambiguous link to another previously mentioned referent. Generally, bridging contexts appear either with a bare NP or possessed nouns such as *dar* "the door" and *modīr-e madrase šūn* "the principal of their school," as in examples (63) and (64).

Ex. (63)   Bare nouns for bridging

vaǵtī	resīd	xūne	doxtar=e	koček=eš	**dar-o**
when	arrive.PST.3SG	home	daughter=EZ	small=PC.3SG	door-OBJ

bāz	kard
open	do.PST.3SG

"when he arrived home, his youngest daughter opened the door"[154]

Ex. (64)   Possessed nouns for bridging

barādr=am	tu=ye	ye	madrese	tū=ye	tehran
brother=PC.1SG	in=EZ	one	school	in=EZ	Tehran

---

[150] Ibid., 237.
[151] See ibid., 98, 290, 291.
[152] Hamedani's corpus.
[153] Nourzaei, "History of the Suffix -ū in Shirazi."
[154] Nourzaei, Unpublished texts, recorded between 2018 and 2021.

moalem=e	*modīr=e*	**madrase=šūn**	*āġā=ye*	*Irānī=e*
teacher=COP.NPST.3SG	teacher=EZ	school=PC.3PL	Mr=EZ	Iranī=COP.NPST.3SG

"my brother is a teacher at a school in Tehran, the principal of their school is Mr. Irani"[155]

There are some cases with K-suffixes, such as *doktor-e* "the doctor" in example (65). The doctor had not been mentioned previously in the story, but it is common knowledge that a hospital has a doctor/several doctors.

Ex. (65)   The K-suffix for bridging

raft-īm	*bīmārestan*	*gof*	**doktor-e**	*ūmad*	*goft-am*	*āqā=y*
go.PST-1PL	hospital	say.PST.3SG	doctor-DEF	come.PST.3SG	say.PST.1SG	Mr=EZ

doktor
doctor

"we went to the hospital, someone said, the doctor came, I said, Mr. Doctor […]"[156]

Similarly, the singular NP *dūkūndār-e* "the shopkeeper" marked with the K-suffix is identifiable based on its clear connection with the shop, as it is common knowledge that every shop has a shopkeeper.

Ex. (66)   The K-suffix for bridging

īnvar	*ūnvar*	*nešūnī*	*mī-dan*		*belaxare*	*yek*	*ǰā=ī*
this. direction	that. direction	address	IMP-give.NPST.3PL		finally	one	place=IND

sang-o
stone-OBJ

paydā	*mī-kon-e*	**dūkūndār-e**	*mī-g-e*
finding	IMP-do.NPST-3SG	shopkeeper-DEF	IMP-say.NPST-3SG

"he looks here and there, finally he finds the stone in a place [a shop], the shopkeeper says […]"[157]

Situational Contexts

Based on the data, in situational definiteness contexts, CSP uses two strategies: a combination of demonstrative plus K-suffix or just K-suffix. This is contrary to Koroshi Balochi, which always requires a combination of demonstrative plus a K-suffix.[158] The following passage displays a situational definiteness context in which the demonstrative combines with a K-suffix with *īn māšīn-e* "this car." The car has not been mentioned previously in the story. The driver points to the car and explains to the mechanic that this car transports passengers from Kerman to Tehran.

Ex. (67)   The K-suffix for situational definiteness with K-suffix

īn	**māšīn-e**	*tū*	*masīr-e*	*tehrūn*	*kār*	*mī-kon-e*	*āġā*
PROX	car-DEF	in	way=EZ	Tehran	work	IMP-do.NPST-3SG	sir

"this car works the Tehran line, sir"[159]

Example (68) displays a situational definiteness context in which the speaker does not combine a demonstrative with the K-suffix. The basket was previously introduced in line 3 of the story. In the example below (line 4 of the narrative) the speaker points to the basket and says, "give me this basket."

---

[155] Ibid.
[156] Ibid.
[157] Taghi, *A Typology and Classification of Three Literary Genres*, 238.
[158] See Nourzaei, "Definiteness Marking," examples 35–38.
[159] Nourzaei, Unpublished texts, recorded between 2018 and 2021.

Ex. (68) The K-suffix for situational definiteness without K-suffix

man	*pesar-e-ro*	*goft-am*	**ī̆n**	**sabad-o**	*be-d-e*	*be*
PN.1SG	boy-DEF-OBJ	say.PST-1SG	PROX	basket-OBJ	IMPV-give.NPST-2SG	to

man
PN.1SG

"I said to the boy, give me this basket"[160]

Similar to example (68), example (69) displays a situational definiteness context, where the demonstrative combines a K-suffix with *doxtar-e* "the girl."

Ex. (69)   Situational definiteness without a K-suffix

az	*ūn*	*esrār*	*ke*	*ī̆n*	**ham=ūn**	**doxtar-e**	*ast*	*az*
from	DIST	insistence	CLM	PROX	EMPH=DIST	girl-DEF	COP.NPST-3SG	from

mādar-e   *enkār*
mother-DEF   denial

"he insisted that she was that girl, but his mother denied [it]"[161]

5.3 Structural Constraints on K-suffix with Anaphoric Definiteness in CSP

As previously mentioned, anaphorically definite nouns are marked with a K-suffix in CSP. However, the presence of the K-suffix is systematically inhibited under certain conditions. In the following subsections I will describe the main systematic structural constraints on use of the K-suffix with anaphoric definiteness.

Plural

Nouns marked with a plural marker never take a K-suffix regardless of their definiteness status, as in the following examples.

Ex. (70)   Absence of the K-suffix with plural noun

hame	**pedar-ā**	*hame mardār-ā*	*xod=ešūn-ro*	*motaxases=e elm=e*	*talīm*
all	father-PL	all mother-PL	REFL=PC.3PL-OBJ	expert=EZ science=EZ	teaching

va	*tarbīyat*	*mī-dūn-an*	**bābā-he**	*čīz*	*mī-kon-e*
and	education	IMP-know.NPST-3PL	father-DEF	what	IMP-do.NPST-3SG

mī-g-e
IMP-say.NPST-3SG

"all fathers, all mothers consider themselves experts in education, for example, the father says"[162]

Possessed Nouns

In addition to the independent pronouns, there are person-marking clitics (PC), which are used with all functions of the oblique case, direct and indirect objects, and as possessive pronouns. The K-suffix is systematically absent from possessed nouns formed with a clitic possessive pronoun, e.g., "his cow," "your son," and pronouns, e.g., *baxt-e doxtar-e mā* "the fate of our daughter." However, it appears with other possessed constructions formed with ezafe constructions, e.g., *xūneye pedar-e* "the father's house." This system is similar to Shirazi Persian[163] and is contrary to Koroshi. In Koroshi, the K-suffix does not appear with all types of possessive constructions.[164]

---

[160] Ibid.

[161] Taghi, *A Typology and Classification of Three Literary Genres*, 230.

[162] PLD.

[163] Nourzaei, "History of the Suffix -ū in Shirazi."

[164] Nourzaei, "Definiteness Marking."

Ex. (71)   Absence of the K-suffix with a possessed noun

*mādar=etūn   hazer=e   **pesar=etūn**   masalan*
mother=PC.2SG   ready=COP.NPST.3SG   son=PC.2SG   for example

pesar-e   barā   zan=eš   kār   be-kon-e
son-DEF   For   wife=PC.3SG   work   SUBJV-do.NPST-3SG

"is your mother ready, your son for example, the boy should work for his wife"[165]

Ex. (72)

*bad az   arūsī=šūn   īn   bar   gašt   **xūne=ye**   **pedar-e***
after   wedding=PC.3PL   PROX   PREV   turn.PST.3SG   house=EZ   father-DEF

"after their wedding, the girl (lit. this) returned to the father's house"[166]

Proper Nouns and Titles

Generally, the K-suffix is absent from titles and proper nouns, as in examples (73) and (74).[167] It is notable that, as in Central Kurdish[168] and Koroshi,[169] *king* and *mullah* are considered proper nouns in Persian.[170] In Shirazi data, *mullah* is not considered a proper noun and is marked with a K-suffix *-ū*, e.g., *āxūnd-ū* "the mullah," unlike *pādšāh/pādošāh* "king."[171]

Ex. (73)   Absence of the K-suffix with a proper noun

*to   masalan   bā   **sīyāvoš**   rāh   mī-raft-ī*
PN.2SG   for instance   with   Siyāvosh   way   IMP-go.PST-2SG

"for instance, you walked with Siyāvosh [...]"[172]

Ex. (74)   Absence of the K-suffix with a title

*xub   āġā=e   **doktor**   dar   har   sūrat*
well   Mr=EZ   doctor   in   each   face

"well, Mr. Doctor at any rate [...]"[173]

Note that in fairy tales the K-suffix is attested with a title in *āġā dīv-e* "Mr. Demon."[174]

However, both the titles Mrs./Madam and Mr./Sir are marked with the K-suffix when they are used alone, as in example (75).

Ex. (75)

*xode=e   **āġā-he**   na-mī-tūn-es   xod=eš-o   edāre*
REFL=EZ   Mr-DEF   NEG-IMP-be.able.PST-3SG   REFL=PC.3SG-OBJ   manage

be-kon-e   ye   mored=e   dege=am   dār-īm   ke   masalan
SUBJV-do.NPST-3SG   one   case=EZ   another=ADD   have.NPST-1PL   CLM   for example

*čīze   **xānom-e**   marīz   shod*
you know   Mrs-DEF   sick   become.PST.3SG

"Mr. himself, could not take care of himself, we have another case, you know, Mrs. got sick [...]"[175]

---

165 PLD.

166 Ibid.

167 I was informed by my Tehrani speakers that the combination of the K-suffix with items such as man "*marde*" and woman "*zane*" still conveys a pejorative sense in certain contexts. This confirms that some remnant of an evaluative meaning of this suffix can still be found.

168 Cf. Öpengin, *The Mukri Variety of Central Kurdish*; Mackenzie, *Kurdish Dialect Studies.*

169 Nourzaei, "Definiteness Marking."

170 See also Taghi, *A Typology and Classification of Three Literary Genres*, 229.

171 See Nourzaei, "History of the Suffix *-ū* in Shirazi."

172 PLD.

173 Ibid.

174 Taghi, *A Typology and Classification of Three Literary Genres*, 263.

175 PLD.

Some Nouns

The data show that the K-suffix is always absent with some nouns, especially those express-ing conventionalized locations, such as *xūne* "home," *madrase* "school," *šahr* "city," *maktab* "school," *češmeh* "spring," and *hamūm* "bathroom," as in the following example.[176]

Ex. (76) Absence of the K-suffix with conventionalized locations

tū	*taǰrīš*	*ġadīmā*	*ye*	**hamūm**	*bū*	*mī-raft-īm*	*hamūm*
in	Tajrish	past	one	bathroom	COP.PST.3SG	IMP-go.PST-1PL	bathroom

"in the past, there was a bathroom in Tajrish, we went to the bathroom [...]"[177]

Unique Referents

The data demonstrate that the K-suffix is systematically absent with unique referents: *zamīn* "ground," *āsemūn* "sky," *xoršīd* "sun," *setāre* "star."

Ex. (77) Absence of the K-suffix with unique referents

tā	*sāheb=e*	*xūne*	*ham*	*az*	*tū=ye*	**zamīn**	*mī-r-e*
till	owner=EZ	house	ADD	from	in=EZ	earth	IMP-go.NPST-3SG

donbāl=e	*gāv=eš*
following=EZ	caw=PC.3SG

"until the owner of house went after his cow from within the field"[178]

Some Prepositions

The data demonstrate that the K-suffix is absent in some combinations with prepositions in the corpus data: *sorāġ* "after," *az* "from," *be* "to," *az bālā* "above," as in examples (78)–(81). Note that there is great variation among the speakers.

Ex. (78)

xod=eš	*raft*	**sorāǰ-e**	**pīrezan**
REFL=PC.3SG	go.PST.3SG	after=EZ	old lady

"he went after the lady"[179]

Ex. (79)

goft-an	*yek=ī-ro*	*dād-īm*	**be** **pīrezan**
say.PST-3PL	one=PC.3SG-OBJ	give.PST-1PL	to old.lady

"they said, one of them we gave to the old lady"[180]

Ex. (80)

alān	**be**	**šīr**	*mī-res-īm*
now	to	lion	IMP-arrive.NPST-1PL

"now we will arrive at the lion"[181]

Ex. (81) Absence of the K-suffix

omad	*resīd*	*be*	*yek*	*āb*	*va*	*deraxt=ī*	*doxtar-e*	*az*
come.PST.3SG	arrive.PST.3SG	to	one	water	and	tree=IND	girl-DEF	from

bālā=ye	**deraxt**	*faryād*	*zad*
above=EZ	tree	shout	beat.PST.3SG

"he came [and] arrived at a [body of] water and a tree [...] the girl shouted from above the tree"[182]

---

[176] See also Taghi's data, *A Typology and Classification of Three Literary Genres.* A similar pattern has been reported in Kurdish; see Haig et al., "Definiteness Markings in Kurdish."

[177] Nourzei, Unpublished texts, recorded between 2018 and 2021.

[178] PLD.

[179] Taghi, *A Typology and Classification of Three Literary Genres*, 217.

[180] Ibid., 216.

[181] PLD.

[182] Taghi, *A Typology and Classification of Three Literary Genres*, 214–15.

Particle ham/am

The data show a significant variation across the speakers regarding the absence of the K-suffix before the particle *ham/am*. The same speaker systematically does not apply the K-suffix before this particle, as in the following examples.

Ex. (82) Absence of the K-suffix with particle *ham/am*
gāv ham hamīntor barāye xod=eš   rāh mī-r-e
cow ADD like this for   REFL=PC.3SG way IMP-go.NPST-3SG
"the cow, you know, goes like this by itself"[183]

Ex. (83) Absence of the K-suffix with the particle *ham/am*
šīr am bā gorbe hamīntorī rad mī-šod-an
lion ADD with cat   like this pass IMP-become.PST-3PL
"you know the lion passed [from there] with the cat"[184]

In the same text, example (84), the speaker uses the K-suffix before *ham*, as in *doxtar koulī-ye ham*, and does not apply it to the following clause *doxtar koulī ham*. Such examples certainly need more research.[185]

Ex. (84) Absence of the K-suffix with particle *ham/am*
doxtar **koulī-ye** ham zūzan-e   āxar-o   keš-īd
girl   gypsy-DEF   ADD needle=EZ   last-OBJ   pul.PST-3SG
doxtar **Koulī**   ham dorūġakī har če   az   doxtar-e
girl   Gypsy   ADD liar   whatever from girl-DEF
"you know, the gypsy girl took out the last needle, [...], the gypsy girl, whatever she has heard from the girl falsely [...]"[186]

5.4 Unexpected Absence

I have already discussed the attested constraints of the K-suffix in anaphoric contexts. However, there nevertheless remains a residue of nouns in definiteness contexts that lack the K-suffix. Hence the term "unexpected absence" of K-suffix is used.[187] The number of such unmarked definite NPs varies considerably across different speakers in our corpus (see below), indicating considerable inter-speaker variation.

In the following passage, the lion, as the main character in the tale, appears without marking with the K-suffix in the definite contexts. In both examples, the lion and the girl are the main characters in the story, and after several mentions with a K-suffix, they appear without a K-suffix. See also the NP *gorbe*, "cat" in example (56), which lacks a K-suffix despite the cat being one of the important characters in this tale.

Ex. (85)
šīr harče   dast va pā mī-zan-e
lion Whatever hand and feet IMP-beat.NPST-3SG
"the lion is trying a lot (lit. is beating its hands and feet) [...]"[188]

---

[183] PLD.

[184] Ibid.

[185] To be certain, I have checked some passages with the K-suffix in this type of environment with fifteen native speakers. I found the same variation across the speakers. The same observations hold regarding the prepositions.

[186] Taghi, *A Typology and Classification of Three Literary Genres*, 237.

[187] See also more passages with unexpected absence of K-suffixes in Kalbasi, *Towsife gunehā-ye zabānī-ye īrān*, 227–28, such as the NPs *kūze* "jug" and *zan* "the woman."

[188] Taghi, *A Typology and Classification of Three Literary Genres*, 237.

Ex. (86)

pesar=e	*šāh*	*xaylī*	*xoš=eš*	*omad*	*na-dīd-e*	**ʿāšeġ=e**
son=EZ	king	very	love=PC.3SG	come.PST.3SG	NEG-see.PST-PP	lover=EZ

doxtar	*Šod*
girl	become.PST.3SG

"the king's son became interested in her [and] without seeing [her], he fell in love with the girl"[189]

Ex. (87)

doxtar	*ġese=ye*	*xod=eš-o*	*bar=āš*	*goft*
girl	story=EZ	REFL=PC.3SG-OBJ	for=PC.3SG	say.PST.3SG

"the girl told her life story to her"[190]

Similar to examples (84)–(87), in example (88) the old lady is one of the main characters in the story. After several mentions with a K-suffix, she appears without a K-suffix.

Ex. (88)

pīrezan	*goft*	*negarān*	*nabāš*	*doxtar=am*	*balad=e*
old lady	say.PST.3SG	worry	NEG.be.NPST.2SG	daughter=PC.1SG	guide=COP.NPST.3SG

rām=eš	*kon-e*
calm=PC.3SG	SUBJ.do.NPST-3SG

"the old lady said, do not worry, my daughter knows how to make it calm"[191]

Summary
The K-suffixes in CSP are associated with definiteness contexts, usually anaphoric, and very rarely appear in bridging contexts. They are systematically excluded from indefiniteness contexts and are not associated with obvious evaluative or diminutive semantics. In this sense, we speak of a definiteness function of the K-suffix in CSP, and in this sense CSP is distinct from CWP. However, in CSP definiteness is a necessary but not sufficient condition for the K-suffix. There are still many notionally definite NPs in our corpus that do not take a K-suffix. First of all, we noted certain structural conditions that inhibit the presence of a K-suffix:

(a) Plural marking of the noun,
(b) In combination with clitic pronouns and copula,
(c) When the noun can be construed as a title or proper noun,
(d) after some prepositions,
(e) after a particle "*ham/am*,"
(f) after some nouns,
(g) with demonstrative pronouns.

The extent of the residue of definite but unmarked items varies from speaker to speaker and according to genre and speech situation. In the next section, we explore the quantitative data from our corpus to shed light on the nature of the changes that have occurred in Persian.

6. The Emergence of Definiteness: Evidence from the Corpus and the Questionnaire

While the grammaticalization of definite markers has been a central issue in grammaticalization theory, researchers usually cite cases (the languages of Western Europe) in which

---

[189] Ibid., 254.
[190] Ibid., 215.
[191] Ibid., 216.

the source of the definite article is some form of deictic element (a "D-element" according to Himmelmann[192]), and this has become the primary paradigm for understanding the diachronic development of definiteness marking cross-linguistically. However, in our ongoing survey of Western New Iranian languages, and Persian in particular, the definiteness suffix has an entirely different source construction, as it comes from an evaluative suffix. Thanks to the existence of data from earlier phases of Persian, we can formulate some initial hypotheses regarding the developmental sequence that led to the current situation. We can see here that the definiteness marker in Persian does not originate from a demonstrative source. And in particular, its combination with the demonstrative pronoun rules out a demonstrative origin.

An overview of the corpora for CNP, CWP and CSP is provided in Table 3.

A second source of data is a questionnaire conducted between 2018 and 2021 with fourteen Tehrani speakers, which is discussed below. But first I consider two metrics from narrative corpus: overall frequency of the K-suffix and distribution of the K-suffix across the corpora for these three phases.

6.1 Overall Frequency of K-suffixes

Overall frequency is counted as the number of occurrences of K-suffixes across all texts in the corpus per orthographic word,[193] normalized to a value of frequency per 1,000 words, to enable comparison across texts of different lengths. Consideration must be given to the fact that a value of zero is not particularly significant in a small text, while zero occurrences in a larger text is much more significant. Nine texts have fewer than 700 words overall, and in many of them, the number of K-suffixes is high; I left them out of this calculation. The results for the three phases are demonstrated in Fig. 2. The vertical axis represents mean values and the bars give the data for each corpus.

There are some points of interest here. First, the hypothesis that overall frequency would increase with a shift towards a definiteness function is confirmed. In CSP, the mean value of K-suffixes per 1,000 words is 3.2, sixteen times higher than in CWP (0.2), and just over three times more than in CNP (1.0). However, it is also clear that the higher frequency of K-suffixes in CSP is largely the result of three data outliers, with 10.0, 8.0, and 7.0 K-suffixes per 1,000 words, respectively, more than twice the figure for any other texts having a K-suffix, while eight texts still have no items marked with K-suffixes.[194]

Thus, CSP is not characterized by the consistently high level of K-suffixes that one would expect if the forms were uniformly grammaticalized as definiteness markers in this language. Overall frequency is, at best, a very crude measure of grammaticalization, however.[195] Note that this is the opposite of our Shirazi results, in which the K-suffix can be found across all the texts.

Recall that the qualitative investigation of these three phases demonstrates that in CNP and CWP, K-suffixes are used with evaluative meaning in most instances of use. Given

---

[192] Himmelmann, "Regularity in Irregularity."

[193] Some of the critical editions are already available in Word format on the PLD website, which made it easy to calculate the total number of words. Since some are not yet available in Word format, I estimated the number of words by counting the number of words per forty pages of each book separately. I then divided this total by forty to calculate an average number of words per page, and then multiplied this average by the number of pages in each book.

[194] The same result can be found in the Hamedani corpus. Seven texts do not have a single item marked with a K-suffix, and of the rest of the texts, only two show a higher frequency of the K-suffix, with 6.0 and 5.3, more than twice the figures for any other texts with a K-suffix. These two texts are both biographical tales.

[195] Grammaticalization involves increasing obligatoriness, that is, the grammaticalizing element is required in a particular syntactic configuration, and speakers have correspondingly less choice about whether they use it there or not. In the grammaticalization literature this is generally assumed to correlate with "a rise in frequency through the expansion to new contexts where the element becomes obligatory" (Dahl, *Grammaticalization in the North*, 32).

Figure 2. Overall frequency of K-suffixes per 1,000 words.

that K-suffixes in these phases are not associated with a predictable and commonly recurring function, we would not expect a uniform frequency of use. Indeed, frequency of evaluative usage may simply be a matter of genre.

In CSP, on the other hand, K-suffixes are not associated with evaluative and diminutive semantics, but are associated with definiteness. However, the association is not fully regular because, as previously mentioned, structural conditions inhibit the K-suffix. Some definite nouns also lack the expected K-suffix for reasons that are not fully understood. It is highly restricted with regard to inter-speaker, inter-setting, and inter-genre factors.

The second remark concerns the decrease and increase in frequency exhibited by the K-suffix in Persian. On the one hand, we can see a significant drop in the frequency of the K-suffixes in CWP. This decrease may be due to the fact that their syntactic domain is becoming increasingly restricted, which means they can only appear with a handful of singular nouns in informal and colloquial settings. Their semantic domain (polyfunctional evaluative notions) is becoming bleached, and the suffix is moving towards definiteness.

Recall that we can find no restrictions on the K-suffix in CNP. It can be found in all parts of speech, apart from verbs and pronouns, throughout the texts. I have noticed the same result in our ongoing survey in Shirazi and Balochi.[196] It needs to be checked in Kurdish and Lori as well, which are currently being analyzed.

The third exciting point concerns the massive inter-writer/speaker and inter-genre differences found in CWP and CSP, but not in CNP. We observe that the K-suffixes are attested in all the CNP texts studied. What is significant in CNP is the region from which the author of a work comes. We find that works written in the east of Iran have a higher frequency of K-suffixes than ones written in the north. Indications that the K-suffix is developing towards a definiteness marker (see examples 32–33) are also attested in two works titled *Tārikh-e Beyhaqi* and *Nowruznāme*, the authors of which come from Khorasan. This might be connected to Lazar's observation that New Persian originated from Khorasan in eastern Iran.[197] The variety of Persian spoken in Khorasan was influenced by Semitic language earlier than Persian varieties spoken in the north of Iran.

The data from CSP demonstrates that only specific kinds of texts contain K-suffix marking. The texts with a high frequency of K-suffixes in the CSP corpus comprise three traditional folktales and two biographical tales. We cannot find the K-suffix with topics such as education, science, human rights, or the coronavirus, which require formal style. This suggests that genre is the decisive factor in CSP. Development of the definiteness marking within a specific genre has been reported for the Finnish language.[198]

---

[196] Nourzaei, "History of the Suffix -ū in Shirazi"; Nourzaei, "Definiteness Marking."
[197] Lazard, "The Rise of the New Persian Language."
[198] Laury, *Demonstratives in Interaction.*

In the data from CWP, we also find three outliers. The three highest values (10, 0.8 and 0.7) come from a book titled *Tamsilāt* and two other books titled *Hājī āqā* and *ʿAlaviye khānom*. *Tamsilāt* is a colloquial translation into Persian from Azerbaijani Turkish. The highest values of the K-suffix are connected to the same noun, *mard-ak-e* "man," with evaluative meaning. It is worth noting the attested items marked with a K-suffix are *zan-ak-e* "woman," *pesar-e* "boy," and *doxtar-e* "girl," as well as one instance each of *sawār-e* "rider" and *šohar-e* "husband."

The same writer, Hedāyat, wrote *Hājī āqā* and *ʿAlaviye khānom*. These are short, colloquial Persian stories. Recall that the highest values of the K-suffix belong to the same nouns, *mart-ī-ke* "man" in *Hājī āqā* and *mart-īk-e* and *zan-īk-e* "woman" in *ʿAlaviye khānom*, with evaluative meaning.

Surprisingly, K-suffixes have not been used consistently even by the same writer. For instance, some of the books written by Sādeghe-Hedāyet do not contain a single item marked with a K-suffix (such as *Buf-e kur*, *Sag-e velgard*, *Parvin dokhtar-e sāsān*). Another example is Hejazi's book *Nasim*, in which he does not mark any items with a K-suffix, even though he uses the K-suffix in another book called *Zībā*. The results demonstrate that as soon as a text switches to formal style, the author does not use the K-suffix.

Overall, a handful of items are marked with the K-suffix, e.g., boy, girl, man, woman, and very seldom other items, e.g., cloud, demon, hyena. The high frequency of K-suffixes in these texts is associated with evaluative and diminutive functions, as the most frequent usages. Thus, the outliers in CWP have a different underlying cause than those of CSP, where the high frequency of K-suffixes is associated with definiteness marking.

In contrast to Shirazi, the overall picture suggests a small number of speakers who use an overall higher frequency of K-suffixes in a specific genre and presumably act as innovators in the development towards definiteness usage in CSP.

Summary of the Narrative Corpus

The corpus data, combined with the qualitative analysis of the K-suffixes in these three phases of Persian, demonstrate that in CNP and CWP, the K-suffixes are largely restricted to evaluative contexts in their highest rates of usage. In contrast to the K-suffix in CNP, in CWP the overall frequencies vary considerably according to genre and content. The suffix is limited to a small number of nouns within certain structural constraints. In CNP, however, we already find signs of K-suffixes combining with nouns in recognitional and deictic contexts without any obvious evaluative or diminutive connotations (see examples 32–33). CWP and CSP also share this type of usage. I consider this to be the first stage in co-opting evaluative morphology to serve as a definiteness marker in Persian. I am already observing the same result in our ongoing survey of Shirazi and Balochi. I also found a few examples in CWP where the items marked with a K-suffix have a referent in the discourse without any obvious evaluative connection and are not dependent on immediate interaction. CSP also shares this type of usage, and systematically uses it in anaphoric contexts. I would suggest that this is the second stage of development of definiteness from an evaluative origin.

CSP differs from CNP and CWP in its almost complete lack of evaluative functions. Also, it expands some of its structural constraints regarding the use of K-suffixes (see Section 5.3) and spreads the suffixes to more items in definiteness contexts. But in CSP, especially in folktales and biographical tales, we find that the K-suffix is systematically used in anaphoric definite contexts, and not in texts discussing topics such as education, science, human rights and women's rights, which are associated with formal settings. This result is not surprising, and this is what can be expected of evaluative as opposed to descriptive or inflectional morphology. The use of evaluative morphology is situationally sensitive and can therefore be expected to adapt flexibly to content, formality, speaker style, and so on.[199]

---

[199] See Dressler and Barbaresi, *Morphopragmatics*, for the usage of the diminutive in Italian.

Overall, the data does not show a simple picture of a spreading out from an assumed anaphoric usage, commonly taken as prototypical for definiteness marking as suggested in grammaticalization theory for Persian.[200] In the following section I will examine the results of the questionnaire data.

6.2 Presentation of Questionnaire Data

In addition to the corpus data, I tested data from a questionnaire answered by fourteen speakers. The questionnaire used a set of 102 items built into six "mini-narratives" each representing short episodes of approximately ten sentences. In order to capture authentic colloquial speech, we circulated the English form of the questionnaire among participants and asked them to translate it orally into colloquial Persian. Their narratives were recorded with a mobile phone, and the relevant NPs were coded for presence vs. absence of K-suffixes and a number of other features. The results here are from the initial pilot in colloquial Persian based on fourteen speakers (nine female and five male), all of whom come from Tehran.

Fig. 3 presents the percentage of nouns carrying a K-suffix in the respective contexts: first mention (indefinite), bridging, anaphoric, demonstratives, possessed, personal nouns, unique references, and non-referential/generic (as in negated existential, such as "in those days there were no cars"). When considering the questionnaire data, we find more than half of the nouns in anaphoric contexts do not take K-suffixes. Other nouns in these contexts are bare nouns or were in plural, and such cases are not counted here.

As presented in Figs. 3 and 4, overall and across all speakers, we find massive inter-speaker differences in the marking of anaphoric definiteness. Only three speakers use the K-suffix in bridging contexts. The most common forms in bridging contexts are bare nouns or possessed nouns, as we observe in the corpus data.

Moreover, we find consistent observance of the structural constraint against use of K-suffixes with plural markers, possessed nouns formed with person-marking clitics, and generic nouns, along with a complete absence of K-suffixes in the indefinite. Furthermore, we find a consistent lack of K-suffixes with personal names. On the whole, this is the system that was found with the corpus data, as discussed previously. In the following section I will comment on the origin of the various K-suffixes in light of the present data.

7. Origin of the K-suffixes in Persian

7.1 K-suffix -ak

In general, the K-suffixes developing towards a definiteness marker in our New Western Iranian languages survey appear to be derived from *-ka-, presumably with the diminutive (and perhaps) pejorativ[e] formations. The K-suffix -ak in CNP might derive from Middle Persian -g, Pusar-ag<pesar-ak "boy" and duxtag<doxtar-ak "girl."

The K-suffix -ak is attested in Persian varieties such as Shirazi Persian as an evaluative suffix, alongside the K-suffix -ū used as a definiteness marker.[201]

7.2 K-suffix -e/he[202]

The etymological origin of the K-suffix -e/he is not yet clear to me, and I leave it as an open question. However, I can offer the following two hypotheses:

[200] E.g., Hawkins, *Efficiency and Complexity in Grammars*, 84–86; Heine, "On Polysemy Copying and Grammaticalization," 129–30.

[201] Nourzaei, "History of the Suffix -ū in Shirazi."

[202] We do not have enough older material to be able to identify with certainty the origin of the K-suffix -e in Persian.
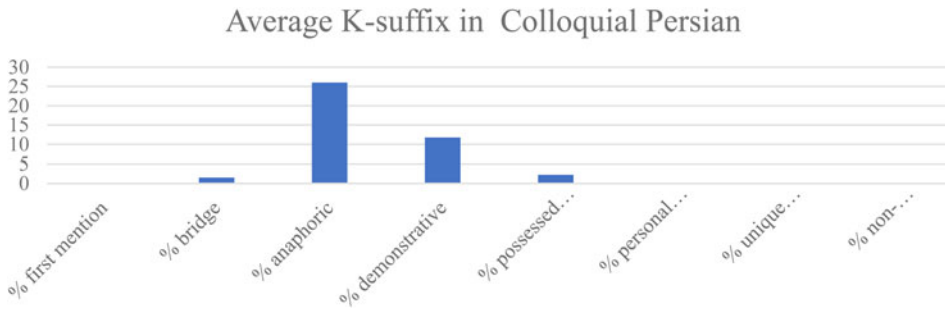
Figure 3. Percentage of K-suffixes, based on questionnaire (fourteen speakers, rounded mean percentages of all speakers' responses).

(1) The K-suffix *-e/he* might be a short form of the *-ak* suffix in CNP. The sound K- has been dropped, and the *a* sound has changed to the *e* sound. This type of sound shift is widespread among Iranian languages such as in *dastag>daste* "handle." In addition, a natural development from Middle Persian to New Persian is the change of Middle Persian *-ag* to *-e*, as is apparent in *setārag>setāre*, and particle *-ag>e kardag-kard-e* as well.

   Across the CWP corpus, however, I found many nouns with a combination of both *-ak* and *-e* suffixes, for instance, *zan-ak-e* "woman," *mard-ak-e* "man," and the following interesting variation of this combination with the same noun "demon." In its first mention in the story, it appears as *yek dīb-**ak**=e sīyā*, "a black demon," and then subsequently as *dīb-**e*** "demon," *dīb-**ak-e*** "demon," and *dīb-**ak*** "demon."[203] If we assume that the K-suffix *-e* is a short form of *-ak*, we should not find both suffixes combined on the same noun. The co-existence of both suffixes *-ak* and *-e* in this scenario seems to be awkward.

   Ex. (89)

yek	***dīb-ak**=e*	*sīyā*	*va*	*kūtol=e*	*dast*	*besīneh*	*ǰol=oš*
one	demon-EV=EZ	black	and	small=IND	hand	on chest	front=PC.3SG

   hazer
   appear

šod	***dīb-e***	*hasan-ī-ro*	*kūl*	*kard*	*dīb-ak-e*
become.PST.3SG	demon-EV	hasan-EV-OBJ	back	do.PST.3SG	demon-EV-EV

ǧeb=eš	*zad-e*	*mesle īnke āb*	*šod*	***dīb-ak-e***
disappeared=PC.3SG	hit.PST-PP	as water	become.PST.3SG	demon-EV-EV

beh=eš	*nešān*	*dād-e*	*būd*	***dīb-ak***
to=PC.3SG	show	give.PST-PP	COP.PST.3SG	demon-EV

   "A black and short demon appeared in front of him with his hands on his chest. The demon put Hasani on his back and disappeared. It was as if he turned to water [sank into the soil]. The demon had shown him [...]"[204]

(2) The K-suffix *-e* might have originated from another source instead of being directly connected to the *-ak* suffix in CNP. However, both of them (*-ak* and *e/he* suffixes) are related originally to the same semantic notions, that is, evaluative (*ke*-suffixes).

An ongoing study by Hashabeiky on Persian (from the sixteenth to eighteenth centuries) shows that only one form of the K-suffix *-ak* with evaluative sense has been written in an informal style, in two of her manuscripts.[205] However, Nadimi Harandi and Atayi Kachooyi provide evidence of the K-suffix *-e* in poetry much earlier (poet, ʿAtar-e

---

[203] *Zende be gur*, 106–7.
[204] Ibid.
[205] Hashabeiky, *A Corpus-Based Description of the New Persian of the 16th–18th Centuries*.
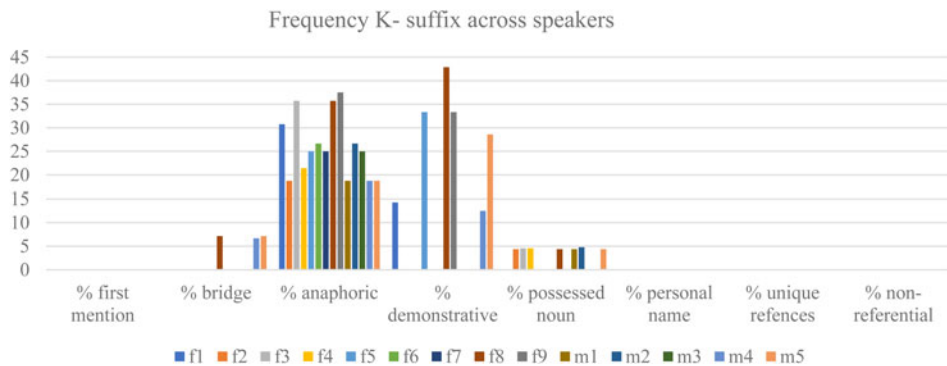
Figure 4. Percentage of the K-suffixes, based on questionnaire (fourteen speakers, rounded mean percentages of individual speakers' responses).

Neshaburi, thirteenth century).[206] This finding suggests that the K-suffix *-e* has been used by Persian speakers (in informal settings) but has not been registered in earlier texts.

Similar to the K-suffix *-ū* in Shirazi Persian, available data with the K-suffix *-e* in Persian shows that this suffix mostly appears with singular nouns and in informal registers. We do not have evidence of its final phonological form. For Shirazi *-ū*, we can trace this suffix back to *-ūk*, used as an evaluative suffix in other Iranian languages such as Bami, Kermani, and Sangsari,[207] while the etymological origin of the Persian *-e* suffix remains a puzzle for the time being.

In this regard, similar to my observation in Shirazi[208] of two K-suffixes *-ak* and *-ū* originally used as evaluative suffixes, I would suggest that there have been different K-suffixes in Persian with an evaluative meaning (*-ak*, *-īk*, *-ūk/*ek*). Whether or not they are related to the same origin is irrelevant here; what matters is that they show similar (evaluative) semantics. These various forms are most probably a matter of Persian dialectal variation, for which we do not have recorded material of the earlier stages. The K-suffix *-e* has been grammaticalized as a definiteness marker, and the *-ak* suffix continued to carry evaluative semantics regardless of genre in written, spoken, formal and informal language settings. However, its evaluative senses, such as endearment when used with proper nouns, have to a large extent been bleached[209] and its pejorative meanings have become colorless.

Note that the short form of the K-suffix *-īk* as *ī* can still be found in Persian speech, such as in *māmī* (my lovely mother) and *xāharī* (my lovely sister), but it is not so frequent. This suffix is very productive as a marker of endearment in other Iranian languages, including Balochi Sistani.[210] Note that in Sistani Persian, the K-suffixes *-ak/ok* are still very productive on proper nouns and reflect endearment and pejorative meanings.

8. Considerations of Sources and Paths of Development

The CNP, CWP and CSP corpora studied here exhibit three different types of development of the K-suffix (the reflexes of cognate and originally evaluative morphemes), which can be interpreted as comprising a scale. In CNP, the most conservative stage in the present study, the K-suffix functions as a polyfunctional evaluative morpheme covering a typical

---

[206] Nadimi Harandi and Atayi Kachooyi, "e-ye maʿrefe dar motune kohan-e Fārsi," 178–79.

[207] Sabbaqiyan, *Barrasi-ye zabān-e sangsari*, 133–45.

[208] Nourzaei, "History of the Suffix *-ū* in Shirazi."

[209] See Sadeghi, "Pasvandha-ye Tahbibi-ye Farsi."

[210] I found the *ī*-suffix on the proper nouns, e.g., *zamzam-ī*, "Zamzam," in my Kholosi data (an Indo-Aryan language spoken in Hormozgan Province of Iran).

array of functions generally associated with diminutives cross-linguistically[211] which are not constrained by definiteness and not subject to structural constraints. However, already at this stage we find some passages with singular nouns in deictic and recognitional contexts. It lies at one end of the scale.

Located in the middle, CWP shows a pre-grammaticalization stage of definiteness marking. The original evaluative meaning of the K-suffix is maintained at its highest usage, but the suffix is subject to structural constraints (i.e., mostly with singular nouns). It shares deictic and recognitional usages of the K-suffix with CNP. The suffix is very immature, and is only sporadically and unsystematically used, even by the same writer, with a handful of nouns.

CSP is found at the other end of the scale. The evaluative usages are not attested, and the suffix is not compatible with indefiniteness contexts. It shares the constraint regarding singular nouns with CWP, but increases in frequency and becomes more closely associated with definiteness contexts. The system does not show a unique spread across the speakers and genres. In the narrative texts investigated, we found a few speakers of CSP who had taken this usage (marking of the NPs with a K-suffix) a step further and now used the K-suffix systematically as a distinct marker of anaphoric definiteness, especially in folktales, biographical genres and informal settings.

This comparison between different stages sheds light on a developmental path from evaluative morpheme to definiteness marker in Persian, as summed up in Table 4. The grammaticalization path is similar to what I already have suggested for other New Western Iranian languages, including Balochi and Shirazi.[212]

These findings suggest that the development of definiteness marking can proceed down a new pathway that is entirely distinct from the one generally presented (demonstrative-based) from a typological perspective. Despite the different pathways, however, the end-points may be fairly similar. Here the starting point is an evaluative marker. In the first stage of the development, evaluative usage is compatible with deictic and recognitional usage, which often occurs with demonstrative pronouns. The latter are anchored to a concrete and interactive speech context involving some form of "attention direction" on the part of the speaker. In the second stage, evaluative usages may disappear entirely/bleach. In contrast, the deictic and recognitional usages are extended to include anaphoric tracking, which would be more independent of setting and not necessarily dependent on immediate interactions. In the final stages, the K-suffix is systematically associated with anaphoric definiteness contexts, although the system continues to co-exist with inherited unmarked definite strategies (bare noun and demonstrative plus noun). Thus, the basic system of definiteness marking with a K-suffix is similar to the more familiar article-based system, of which anaphoric definiteness is generally the core function.

Several differences can also still be discerned, in particular the constraint that prevents definiteness marking in combination with plural marking and possessed nouns formed with a person-marking clitic. In a recent cross-linguistic study on definiteness,[213] Becker found no typological evidence for the compatibility of definiteness markers with plural number (although there is clear evidence for incompatibility between indefiniteness markers and plural number). Thus, the Persian constraints (along with Shirazi and Balochi) remain somewhat of a puzzle, compared to definiteness markers in Lori Bakhtiyari and Central Kurdish based on the same K-suffix, for which no such constraints exist. I leave this as an open question, but assume that the constraint might be due to the following facts: (a) these two suffixes (the plural marker and the K-suffix *-e/he*) are compatible morphologically (since both the plural marker *-hā* and short form of *-e* are new in the language); (b) they are compatible semantically, because the plural marker *-hā* already has a definiteness function, and it does

---

[211] Ponsonnet, "A Preliminary Typology."

[212] Nourzaei, "Definiteness Marking"; Nourzaei, "History of the Suffix *-ū* in Shirazi."

[213] Becker, "Articles in the World's Languages."

Table 4. Overview of grammaticalization path from evaluative to definiteness functions.

	CNP	CWP	CSP
Evaluative usages -possible in non-definite contexts -no structural constraints -frequency depends on context and region	+	+	-
Deictic, recognitional contexts -in combination with demonstratives -singular only	+	+	+
Anaphoric -demonstrative not required - structural constraints -singular only	-	-	+
Bridging	-	-	+

not need to be marked again with another element (*e-he*);[214] and (c) the starting point of an evaluative marker in deictic and recognitional contexts in CNP is singular nouns, which suggests a possible scenario – similar to that of the intrusion of the object marker (*-rā*) into the nominal system with singular nouns, for example in Balochi[215]– where the singular nouns are initially attracted more to the K-suffix than to the plural nouns. I have also noticed a tendency of using the K-suffix with the plural marker in Lori spoken in Fars. This is a topic for future study.

Finally, concerning the development of the definiteness marker in Persian, I would suggest that internal development, for example reducing the case system in Persian, may have favored the emergence of an additional nominal category such as definiteness. So far in the languages in our survey, languages/dialects with a reduced case system exhibit the development of the definiteness marker, for example, Shirazi, Koroshi, Lori, and Central Kurdish. On the other hand, one should not overlook the language contacts (possible earlier Persian contacts with the Semitic languages); see also Haig and Khan.[216] The ongoing project suggests that several New Western Iranian languages have developed some nascent form of definiteness marking based on evaluative morphology.

Due to the extensive documented material from its earlier phases, the Persian case presented here will provide a benchmark for future studies of Iranian languages, and will broaden the database for our understanding of the development of definiteness cross-linguistically.

---

214 Ghomeshi, "Plural Marking."
215 Nourzaei, *Participant Reference in Three Balochi Dialects*, appendix B.
216 Haig and Khan, "Introduction."

She would like to acknowledge Mohammad Rasekhmahand and Geoffrey Haig for sharing the Hamedani Persian spoken corpora with her. She would like to thank Elham Izadifar for helping her with new recordings and testing the questionnaire data on Hamedani speakers. Thanks also to Shokoufeh Taghi for double-checking her published data with their sound files. Any remaining errors are, of course, the author's own responsibility.

Abbreviations

1	first person
2	second person
3	third person
[]	additional information to the text
()	additional information to the gloss
...	incomplete sentence
-	affix boundary
=	clitic boundary
ADD	additive particle
CLM	clause linkage marker
CNP	Classical New Persian
COMP	comparative
COP	copula (present indicative)
CSP	Contemporary Spoken Persian
CWP	Contemporary Written Persian
DEF	definite
DIST	distal
EMPH	emphasis
EV	evaluative
EZ	ezafe particle
IMP	imperfective
IMPV	imperative
IND	individuation clitic
INF	infinitive
NEG	negation
NPST	non-past stem
OBJ	object case
PC	person-marking enclitic (person clitic)
PL	plural
PN	personal pronoun
PP	past participle
PREV	preverb
PROX	proximal deixis
PST	past stem
REFL	reflexive pronoun
SG	singular
VOC	vocative case

Bibliography[217]

Abbott, Barbara. "Definiteness and Indefiniteness." In *Handbook of Pragmatics*, edited by Laurence R. Horn and Gregory Ward, 122–49. Oxford: Blackwell, 2004.

Ahmadi Givi, and Hasan Anvari. *Dastur-e zabān-e Farsī 1*. 4th ed. Tehran: Fātemi, 1390/2011.

Akhundzade, Mirza Fathᶜali. *Tamsilāt*. Tehran: Enteshārat-e Khᵛārazmi, 1874.

ᶜAlavi, Bozorg. *Chamedān*. Tehran: Ketābkhāne-ye Matbaᶜe-ye Dānesh, 1313/1934.

---

[217] The critical editions of the Classical Persian works are arranged according to the names of the editor/s.

ʿAlavi, Bozorg. *Cheshmhāyash*. N.p.: Enqelāb va Adabiyāt, 1363/1984.

Bahmaniyar Ahmad, ed. *Al-abniye ʿan haqāʾiq al-adviye*. Tehran: Tehran University Press, 1346/1967.

Becker, Laura. "Articles in the World's Languages." PhD diss., University of Leipzig, 2018.

Christophersen, Paul. *The Articles: A Study of Their Theory and Use in English*. Copenhagen: Munksgaard, 1939.

Ciancaglini, Claudia. "Outcomes of the Indo-Iranian Suffix *-ka in Old Persian and Avestan." In *Persepolis and Its Settlements: Territorial System and Ideology in the Achaemenid State*, DARIOSH Studies II, edited by Gian Pietro Basello and Adriano Rossi, 91–100. Napoli: L'Orientale, Dipartimento Asia, Africa e Mediterraneo, 2012.

Clark, Herbert H. "Bridging." In *Theoretical Issues in Natural Language Processing*, edited by R. C. Schank and B. L. Nash-Webber, 169–74. New York: Association for Computing Machinery, 1975.

Dabir Siyaqi, Mohammad, ed. *Safarnāme-ye Nāser Khosrow*. N.p.: Enteshārāt-e Zavvār, 1370/1991.

Dahl, Östen. *Grammaticalization in the North: Noun Phrase Morphosyntax in Scandinavian Vernaculars*. Berlin: Language Science Press, 2015.

Dehkhoda, Ali Akbar. *Charand o parand, Kānun-e maʿrefat*. N.p.: 1286/1907.

Dressler, Wolfgang, and Lavinia Barbaresi. *Morphopragmatics: Diminutives and Intensifiers in Italian, German, and Other Languages*. Berlin: Mouton de Gruyter, 1994.

Durkin-Meisterernst, Desmond. *Grammatik des Westmitteliranischen (Parthisch und Mittelpersisch)*. Wien: Verlag der Österreichischen Akademie der Wissenschaft, 2014.

Edgerton, Franklin. "The K-suffixes of Indo-Iranian, Part I: The K-suffixes in the Veda and Avesta." *Journal of the American Oriental Society* 31, no. 2 (1911): 93–150.

Edgerton, Franklin. "The K-suffixes of Indo-Iranian, Part I: The K-suffixes in the Veda and Avesta." *Journal of the American Oriental Society* 31, no. 3 (1911): 296–342.

Firoozbakhsh, Pejman. "The Former Dialect of Šīrāz in the Poetry of Šams Son of Nāṣir of Šīrāz (d. 763 AH/1362 CE)." PhD diss., University of Hamburg, 2019.

Ghomeshi, Jila. "Plural Marking, Indefiniteness, and the Noun Phrase." *Studia Linguistica* 57, no. 2 (2003): 47–74.

Gindin, Thamar E. "The Early Judeo-Persian *Tafsīrs* of Ezekiel. Vol. 3: Grammar." Unpublished manuscript.

Haig, Geoffrey. "Optional Definiteness in Central Kurdish and Balochi: Conceptual and Empirical Issues." Talk at the third workshop on Information Structure in Spoken Language Corpora (ISSlaC3), University of Münster, December 7–8, 2018.

Haig, Geoffrey, and Geoffrey Khan. "Introduction." In *The Languages and Linguistics of Western Asia: An Areal Perspective*, edited by Geoffrey Haig and Geoffrey Khan, 1–29. Berlin: Mouton de Gruyter, 2018.

Haig, Geoffrey, Maryam Nourzaei, and Masud Rad. "Definiteness Markings in Kurdish" (in preparation).

Hashabeiky, Forogh. *A Corpus-Based Description of the New Persian of the 16th–18th Centuries in Three Socio-political Spheres*. Uppsala: Acta Universitatis Upsaliensis, forthcoming.

Hawkins, John. *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press, 2004.

Hedayat, Sadeq. *Parvin dokhtar-e sāsān*. N.p.: Ferdosi, 1309/1931.

Hedayat, Sadeq. *Afsāneye Afarīnesh*. Paris, 1324/1946.

Hedayat, Sadeq. *Hājī āqā*. Tehran: Enteshārāt-e Amir Kabir, 1330/1951.

Hedayat, Sadeq. *Se qatre khun*. Tehran: Sorush, 1333/1932.

Hedayat, Sadeq. *ʿAlaviye khānom va Velengāri*. Tehran: Enteshārāt-e Amir Kabir, 1342/1963.

Hedayat, Sadeq. *Sag-e velgard*. Tehran: Enteshārāt-e Amir Kabir, 1342/1963.

Hedayat, Sadeq. *Zende be gur*. Tehran: Enteshārāt-e Amir Kabir, 1342/1964.

Hedayat, Sadeq. *Buf-e kur*. Isfahan: Enteshārat Sādeq Hedāyat, 1383/1936.

Heine, Bernd. "On Polysemy Copying and Grammaticalization in Language Contact." In *Dynamics of Contact-Induced Language Change*, edited by Claudine Chamoreau and Isabelle Léglise, 125–66. Berlin: Mouton de Gruyter, 2012.

Hejazi, Mohammad. *Nasim*. Tehran: Kebkhāne-ye Ebn-e Sinā, 1346/1960.

Hejazi, Mohammad. *Zibā*. Tehran, 1340/1962.

Himmelmann, Nikolaus. "Regularity in Irregularity: Article Use in Adpositional Phrases." *Linguistic Typology* 2 (1998): 315–53.

Jahani, Carina. "On the Definite Marker in Modern Spoken Persian." Sixth International Conference on Iranian Linguistics (ICIL 6), Ilia State University, Tbilisi, GA, June 23–26, 2015.

Josephson, Judith. "Definiteness and Deixis in Middle Persian." In *The Persian Language in History*, Beiträge zur Iranistik 33, edited by Mauro Maggi and Paola Orsatti, 23–39. Wiesbaden: Reichert, 2011.

Jurafsky, Dan. "Universal Tendencies in the Semantics of the Diminutive." *Language* 72, no. 3 (1996): 533–78.

Kalbasi, Iran. *Sakht-e eshteqāqi-ye vāzhe dar Fārsi-ye Emruz*. Tehran: Pezhūheshgāh-e olum-e ensāni, 1380/2001.

Kalbasi, Iran. *Towsife gunehā-ye zabānī-ye īrān*. Tehran: Pezhūheshgāh-e olum-e ensāni, 1388/2009.

Kasravi Ahmad. *Kāfnāme*. Tehran, 1330/1936.

Khatamipoor, Hamed. "yā-ye maʿrefeh: nokte-I noyāfte dar dastur tārikhi-ye zabān-e Fārsi." *Dastur (Vizhenāme ye name ye Farhangestān)* 9 (1392/2013): 12–19.

Khatib Rahbar, Khalil, ed. *Tārikh-e Beyhaqi 1–3*. Tehran: Nashr-e Mahtāb, 1383/2004.

Khayyampur, Abdolrasul. *Zabān-e Fārsi*. Tabriz: Enteshārāt-e Shafaq, 1344/1965.

Laury, Ritva. *Demonstratives in Interaction: The Emergence of a Definite Article in Finnish.* Amsterdam: John Benjamins Publishing Company, 1997.

Lazard, Gilbert. *A Grammar of Contemporary Persian.* Costa Mesa, CA: Mazda Publishers, 1992.

Lazard, Gilbert. *Grammaire du persan contemporain.* Paris: Klinksieck, 1957.

Lazard, Gilbert. "The Rise of the New Persian Language." In *The Cambridge History of Iran.* Vol. 4, *The Period from the Arab Invasion to the Saljuqs,* edited by R. N. Frye, 595–632. Cambridge: Cambridge University Press, 1975.

Lyons, Christopher. *Definiteness.* Cambridge: Cambridge University Press, 1999.

Mackenzie, David Neil. *Kurdish Dialect Studies.* Vols. 1 and 2. London: Oxford University Press, 1961/62.

Malek al-Shoʾara Bahar, ed. *Tārikh-e Sistān.* Tehran: Ketābkhāne-ye zavvār, 1314/1935.

Maraghei, Zeyn al-ʿĀbedin. *Siyāhatnāme-ye Ebrāhim Beyg.* N.p.: 1347/1895.

Meshkat al-Dini, Mehdi. *Dastur-e zabān-e Fārsi bar pāye-ye Nazariye-ye gashtari.* Mashhad: Dāneshgāh Ferdosi-e Mashhad, 1370/1991.

Minovi, Mojtaba, ed. *Nowruznāme.* Tehran: Ketābkhāne-ye Kāve, 1312/1933.

Moʿin, Mohammad. *Farhang-e Farsi.* Vol. 1. Tehran: Enteshārāt-e Amir Kabir, 1364/1985.

Nadimi Harandi, Mahmood, and Tahmine Atayi Kachooyi. "e-ye maʿrefe dar motune kohan-e Fārsi." *Dastur* 14 (1397/2018): 173–82.

Natel Khanlari, Parviz. *Dastur Zabān-e Fārsi.* Tehran: Enteshārāt-e Bonyād-e Farhang-e Iran, 1351/1972.

Neale, Stephen. *Descriptions.* Cambridge, MA: MIT Press, 1990.

Nourzaei, Maryam. "Definiteness Marking from Evaluative Morphology in Balochi: Internal Variation and Diachronic Pathway." *Iranian* Studies 54, no. 5–6 (2021): 699–735. https://doi.org/10.1080/00210862.2020.1813555.

Nourzaei, Maryam. "History of the Suffix -ū in Shirazi" (in preparation).

Nourzaei, Maryam. *Participant Reference in Three Balochi Dialects: Male and Female Narrations of Folktales and Biographical Tales.* Uppsala: Acta Universitatis Upsaliensis, 2017. https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-314090.

Nourzaei, Maryam. Unpublished texts of colloquial Persian, recorded between 2018 and 2021.

Nourzaei, Maryam. Unpublished texts of Hamedani Persian, recorded in 2021.

Nourzaei, Maryam. Unpublished texts of Neshaburi Persian, recorded between 2020 and 2021.

Nourzaei, Maryam. Unpublished texts of Sistani Persian, recorded between 2012 and 2018.

Nourzaei, Maryam, and Geoffrey Haig. Emerging of Definiteness Markers in New Western Iranian Languages (in preparation).

Nourzaei, Maryam, and Geoffrey Haig. "An Overview of Definiteness Marking in New Western Iranian Languages" (in preparation).

Nourzaei, Maryam, and Thomas Jügel. "On the Function of -ag Suffix in MP" (in preparation).

Nye, Gertrude Elizabeth. "The Phonemes and Morphemes of Modern Persian: A Descriptive Study." PhD diss., University of Michigan, 1955.

Öpengin, Ergin. *The Mukri Variety of Central Kurdish: Grammar, Texts and Lexicon.* Wiesbaden: Harrassowitz, 2016.

Pakendorf, Brigitte, and lja V. Krivoshapkina. "Ėven Nominal Evaluatives and the Marking of Definiteness." *Linguistic Typology* 18, no. 2 (2014): 289–331.

Paul, Ludwig. *A Grammar of Early Judaeo-Persian.* Wiesbaden: Reichert, 2013.

Ponsonnet, Maïa. "A Preliminary Typology of Emotional Connotations in Morphological Diminutives and Augmentatives." In "Morphology and Emotions across the World's Languages," edited by Maïa Ponsonnet and Marine Vuillermet. Special Issue. *Studies in Language* 42, no. 1 (2018): 17–50.

Qarib, Abdolazim, [Mohammad Taqi] Malek al-Shoʾara Bahar, Badiʾozzaman Foruzanfar, Jalal Homayi, and Rashid Yasemi [Panǰ Ostad]. *Dastur-e Fārsi.* 9th ed. Tehran: Enteshārāt-e Ashrafi and Enteshārāt-e Vāzhe, 1370/1991.

Qavim, Ali Akbar, ed. *Khᵛān al-ekhvān.* Tehran: Asātīr, 1384/2005.

Ravaqi, Ali. ed. *Qorʾān-e Qods. Kohantarin bargardān-e Qorʾān ba Fārsi.* Tehran, 1984.

Rowshan, Mohammad, ed. *Bakhtiyārnāme.* Tehran: Enteshārāt-e Bonyād-e Farhang-e Iran, 1348/1969.

Rowshan, Mohammad, ed. *Marzbānnāme.* Tehran: Enteshārat-e Bonyād-e Farhang-e Iran, 1335/1976.

Rowshan, Mohammad, and Abullqasm Jalil Pur, eds. *Rowzat al-ʿoqul.* Tehran: Farhangestān Zabān va adab Fārsi, 1383/2004.

Russell, Bertrand. "On Denoting." *Mind* 14 (1905): 479–93.

Sabbaqiyan, Naser. *Barrasi-ye zabān-e sangsari.* Amol: Shomal-e paydar, 1350/1971.

Sadeghi, Ali Ashraf. "Pasvandha-ye Tahbibi-ye Farsi dar dowre-ye eslami." *Vizhename-ye Farhangestan.* Vol. 13. Tehran, 1397/2018.

Sadeghi, Ali Ashraf, and Qolamreza Arzhang. *Dastur-e zabān-e Fārsi.* Iran: Sāzemān ketabhā-ye darsi, 1355/1976.

Safa, Zabihollah, ed. *Dārābnāme.* Tehran: Bongāh Tarjome va Nashr-e ketāb, 1339.

Samiian, Vida. "Structure of Phrasal Categories in Persian: An X-bar Analysis." PhD diss., University of California, Los Angeles, 1983.

Schwarz, Monika. *Indirekte Anaphern in Texten.* Studien zur domänengebundenen Referenz und Kohärenz im Deutschen. Tübingen: Niemeyer, 2000.

Steriopolo, Olga. "Form and Function of Expressive Morphology: A Case Study of Russian." *Russian Language Journal* 59 (2009): 149–94.

Stilo, Donald L. *A Grammar of Vafsi.* Vol. 1: *Phonology, The Noun Phrase, Verb Tense-Aspect-Mood Paradigms and Their Uses* (in preparation).

Taghi, Shokoufeh. *A Typology and Classification of Three Literary Genres.* Uppsala: Acta Universitatis Upsaliensis, 2016.

Talebof, Abd al-Rahim. *Ketāb-e Ahmad.* Tehran: Enteshrate shabgir, 1336/1977.

Volek, Bronislava. *Emotive Signs in Language and Semantic Functioning of Derived Nouns in Russian.* Amsterdam: John Benjamins Publishing Company, 1987.

von Heusinger, Klaus. "Definiteness." In *Oxford Bibliographies Online: Linguistics*, edited by M. Aronoff. New York: Oxford University Press, 2011. https://doi.org/10.1093/OBO/9780199772810-0063.

Windfuhr, Gernot. *Persian Grammar: History and State of Its Study.* Berlin: Mouton Publishers, 1979.

Windfuhr, Gernot, and John R. Perry. "Persian and Tajik." In *The Iranian Languages*, edited by Gernot Windfuhr, 416–544. London and New York: Routledge, 2009.

Yazdani Khorram, Mehdi. *Khun-khorde.* N.p.: Nashr-e Cheshme, 1397.

Yusefi, Gholam Hosayn, ed. *Qābusnāme.* Tehran: Bongāh-e tarjome va nashr-e ketāb, 1345/1966.