# ESSAYS ON TYPOLOGY OF IRANIAN LANGUAGES

*Edited by Alireza Korangy, Behrooz Mahmoodi-Bakhtiari*

Alireza Korangy, Behrooz Mahmoodi-Bakhtiari (Eds.)
**Essays on Typology of Iranian Languages**

# Trends in Linguistics
# Studies and Monographs

# Volume 328

# Essays on Typology of Iranian Languages

———

Edited by
Alireza Korangy
Behrooz Mahmoodi-Bakhtiari

**DE GRUYTER**
MOUTON

Alireza Korangy dedicates this book to two people. He dedicates it to his "sunrise, sunshine, and sunset", his four year old daughter Iran Ghazal Korangy: "my little linguist" ; and to the memory of Professor Lazard whose dedication to Iranian linguistics has not known and will not know an equal.

Behrooz Mahmoodi-Bakhtiari dedicates this book to the memory of Professor Lazard (d. 2018), whose very last article appears in this volume.

# Acknowledgments

The editors would like to thank the contributors to this volume whose persistence on excellence is truly the purple patch of the volume as can be seen in the articles herein. We would also like to thank Julie Meiss of Mouton whose professionalism was infinite as was her kindness. We would like to thank Stefan Diezmann for his astute attention to detail in the typesetting and copyediting stages of the volume: truly amazing. Last but certainly not least we thank the editors of TiLSM who were as rigorous as they were forthcoming in their communications and guidance. A very special thanks goes to Hans Henrich Hock and Daniël Van Olmen of TiLSM, the editors responsible for this particular volume.

# Contents

Bernard Comrie
# Iranian languages and linguistic typology

Although I am not an Iranianist, as a typologist I recognize that Iranian languages can make an important contribution to our understanding of cross-linguistic variation. Hitherto, this potential has barely been realized, though there have been a few notable exceptions: thus, the now standard term "Differential Object Marking" and its abbreviation "DOM" were introduced by Bossong (1985) in his treatment of this phenomenon in Iranian languages. The present volume is an important step in redressing this balance.

As the reference to DOM has already shown, case marking is an area where Iranian languages have already made a contribution to linguistic typology, and this extends to other instances of flagging (case marking, use of adpositions) and indexing (pronominal indices on the verb) of core arguments, including the phenomenon of ergativity. The chapter by Bernhard Scheucher succeeds in compressing many aspects of the synchronic and diachronic richness of Iranian, especially New West Iranian[1] languages into a digestible presentation that addresses Iranianists and typologists alike. In particular, it provides further empirical evidence of the "horizontal" or "double-oblique" alignment type, where the same form is used for both Agent and Patient of the transitive verb, a different form for the Single argument of the intransitive argument – see, for instance, the Northern Kurdish examples in his Section 5.1. This alignment pattern was first drawn to the attention of general linguists by Payne (1980), with material from Pamir languages, and Scheucher's contribution shows typologists that the pattern is more widespread in Iranian.

Another nominal category subject to variation across Iranian languages is gender, absent from innovative languages like Persian, but present in more conservative languages like Pashto. Don Stilo's contribution shows that two neighboring, closely related, indeed mostly mutually intelligible Tatic varieties, Kafteji (Kabatei) and Kelasi, nonetheless differ strikingly in this regard. Kafteji has not only retained the grammatical masculine–feminine gender opposition, but has even extended its application to new domains within its verb-agreement system. Kelasi has lost the category completely. In addition to the detailed examination of the structural mechanisms involved, this chapter also points to important general issues in the study of language contact: given that Kafteji and Kelasi speakers are in close contact, might the retention of gender in Kafteji be perhaps a "shibboleth"

---

**1** I will retain the traditional Iranianist use of "New" rather than "Modern".

by which Kafteji speakers assert their identity? Stilo does not unequivocally answer this question in the affirmative in his Section 5.3, but raises a possibility that merits investigation in other similar instances of language contact across the world.

Plurality as a nominal category is retained across the Iranian languages, but the chapter by Hasan Rezai-Baghdidi and Behrooz Mahmoodi-Bakhtiari shows that there is nonetheless considerable typological variation within New West Iranian, both synchronically and diachronically, concerning both nominal marking and its interaction with syntax. While the most widespread markers are of the type *-ān* (from an earlier oblique plural marker) and *-hā* (from an earlier abstract noun marker), there is also *-gal* deriving from a noun 'troop', and even an outlier language, Abuzaydābādi, that sometimes uses a plural prefix *pāk-*.

Rounding out the chapters on nominal categories, Ketevan Gadilia examines expressions of definiteness and indefiniteness in Iranian languages. This includes both the use of definite and especially indefinite suffixes, but also the use of demonstratives as definite markers and the numeral "one" as an indefinite marker – plus, of course, DOM via reflexes of Old Persian *rādiy* as a combination of case marking and definiteness marking.

Turning to the verb, Gilbert Lazard presents a taste of the range of variation in the expression of aspect by contrasting two Iranian verb systems, those of Persian and Pashto. Of particular typological interest is the fact that Persian has overt marking for durative aspect (with the prefix *mi-*), while Pashto has overt marking of perfective aspect (via the prefix *wë-* or its stem-change or stress-shift allomorphs). The clear correlation between formal and semantic markedness in these two contrasting systems would provide an interesting *tertium comparationis* for the investigation of systems where the correlation between formal and semantic marking is less clear, e.g., the Latin perfect (and its Romance reflexes) or the Slavic perfective. This chapter, though concentrating on two Iranian languages, thus opens up new perspectives for the study of aspect more generally from a typological perspective.

Two chapters each deal more generally with an individual language or a group of languages, providing a typological profile. Habib Borjian's chapter on Mazandarani follows the areal typological approach developed for West Iranian by Don Stilo, and shows, with a helpful tabular summary, a range of Mazandarani feature values that locate it relative to a selection of other West Iranian languages, in terms of whether Mazandarani shares or does not share the feature value with that other language. The results point to a particularly close typological affinity of Mazandarani with Gilaki, followed by Aftari and Semnani. An unusual typological feature of some varieties of Mazandarani is the distinction among four verbs "to be" in terms of combinations of equation, existence, containment, animacy, and emphasis.

Thamar Eliam's chapter is concerned with typological features of Iranian language varieties spoken by Jewish communities, i.e., Judeo-Persian (several usually mutually intelligible varieties), Juhuri (aka Judeo-Tat), and the secret jargon Lutera'i. As noted by the author, the term "Judeo-Iranian" delimits a social rather than a linguistic unity. Judeo-Iranian languages can profitably be studied typologically along two axes. One is the influence of New Persian, traditionally stronger on Judeo-Persian than on the local non-Jewish dialects of the same geographical region, which means that Judeo-Persian is here innovative, e.g., lacking gender and case marking of nouns even where these are present in local non-Jewish speech. The other is the presence of the Hebrew(-Aramaic) component in all Judeo-Iranian varieties, distinguishing them lexically from non-Jewish varieties.

Finally, two chapters examine one particular phenomenon in one particular language. Sascha Völlmin's chapter examines the quotative suffix in Gilaki, more specifically the Rasht dialect within the Western Gilaki dialect group; in this variety, the suffix is *-ə*. The chapter is based on the examination of an extensive corpus supplemented where necessary by elicitation. Völlmin shows that while instances of quotative *-ə* have appeared in previous works, they were not correctly identified as such, although the more extensive corpus shows that this is indeed a quotative marker, nearly always present when another's words are quoted (and thus crucially absent when one quotes one's own words). Quoted speech (thought, etc.) in Gilaki is always "direct" speech, lacking for instance the deictic shifts characteristic of "indirect speech" – although the language does have a logophoric form *xu*, literally 'self', to express coreference with the reporter. The richness of the system and the fine judgments that the author reveals lead one to desire equally detailed studies of this much-neglected domain in other oral Iranian languages.

Haig and Adibifar treat the phenomenon of null referential subjects in spoken Persian from a usage- and discourse-based perspective. By analyzing retellings of the Pear Story video by twenty-nine different speakers, they throw light onto possible factors governing variation in the frequency of null referential subjects. They investigate both factors that might be expected to lead to different rates of use (text length; number of new referents introduced in the text; speaker's familiarity with the interviewer) as well as those where expectations from studies of other languages are either absent or inconsistent (gender; age), and conclude that none of these factors yields a statistically significant correlation. This negative result is nonetheless interesting, in that it provides support for the hypothesis of the overall homogeneity of spoken language, in contrast to the heterogeneity of written language. Moreover, the chapter opens the door to enriched study of Persian in a usage-based approach, including in particular the much neglected spoken language.

In summary, this volume acquits itself fully of the task of bridging the gap between Iranian languages and linguistic typology, both advancing our understanding of existing problems and pointing to new areas that require investigation. While I have emphasized more what the typologist can learn from Iranian languages, the volume also shows how new insights into the synchrony and diachrony of Iranian languages can be gained by adopting a typological perspective. Of course, this volume can still only scratch the surface. There are many other areas in Iranian languages that merit study from a typological perspective; to cite just one: the processing complications that arise from the combination of verb-final constituent order with a frequently head-initial noun phrase, as seen particularly clearly in Persian, and the discourse strategies used to resolve potential conflicts (e.g., preposing noun phrases with a postnominal relative clause). More volumes like the present one are needed!

# References

Bossong, Georg. 1985. *Empirische Universalienforschung: differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Gunter Narr.
Payne, John R. 1980. The decay of ergativity in Pamir languages. *Lingua* 51. 147–186.

Bernhard Scheucher
# Ergativity in New West Iranian

## 1  General

Indo-European languages are known for their accusative structure. Yet there are exceptions to this pattern: Various modern Indo-Iranian languages display split ergativity with a tense-aspect split. The past tenses in these languages are constructed ergatively. The ergative past traces back to an Old Indian and Old Iranian periphrastic perfect, which is formed by the past participle in -ta- with passive meaning and the present tense of the copula. In later linguistic stages of the Iranian and Indo-Aryan languages, this construction replaced the synthetic past tenses.

Many New Iranian languages have a split ergative construction. The historic roots of this ergative construction trace back to the Old Iranian linguistic stage, when the old synthetic perfect was replaced by an analytic construction with passive value. In Old Persian this replacement is attested by the renowned *manā kartam* construction ('by me it was done'). The subject of a transitive verb in the new perfect takes the genitive-dative case, whereas in Sanskrit the agent in that construction is marked by the instrumental case. The oblique case, which marks the agent of transitive verbs in the past tenses of Middle Iranian and New Iranian languages, stems from the Old Iranian genitive-dative case. In Hindi the subject of a transitive verb in the past tense is marked by a special agentive or ergative particle that has instrumental etymology (cf. Pirejko 1979).

Most New East Iranian languages have ergative past tenses. The remarkable exception is Ossetic, which is spoken in the Caucasus area, which is famous for the multitude of ergative languages spoken there. Among the West Iranian languages are many languages that have lost the ergative structure of the past tenses. For instance, New Persian, the Iranian language with the greatest number of speakers, is a purely accusative language. However, languages like Kurdish, Zāzākī, Gōrānī, and other smaller languages have retained the ergative construction.

## 2  Basic notions of ergativity

Ergativity is a morphosyntactic property of the ergative languages. In these languages the subject of a transitive verb is marked by a non-nominative case, which is often called ergative or agentive case. The Iranian languages have no

unique ergative case marking in the ergative past tenses. Instead they use the oblique case, which is also used in the accusative present tense and future tense to mark the direct and indirect object. The subject of an intransitive verb, however, is marked by a case, called nominative or absolutive, which is mostly characterized by zero marking. This case is often called absolutive case. The direct object of a transitive verb is also marked by the absolutive case. Transitive verbs in ergative languages often agree with the direct object or have no agreement at all.

Ergative languages are, for example, Basque, the Caucasian languages, Eskimo-Aleut, Tibetan, indigenous American, and Australian languages. There are also some ergative languages attested in the ancient Near East: Sumerian, Hattic, Elamite, Urartian, and Hurrian.

The ergative pattern is quite contrary to that which we know from most Indo-European languages, the Semitic languages, the Uralic languages, the Altaic languages, etc. In the latter languages, which are of the accusative type, the subjects of intransitive verbs and the subjects of transitive verbs form a morphosyntactic unity in the way that both are marked by the same case, namely, the nominative case, which is in most of these languages not characterized by an affix or by any other kind of morphological marking. The direct object of the transitive verb, however, is distinguished from the subject by a different case, which is traditionally called "accusative". The verb of a transitive sentence usually agrees with the subject.

The term "ergative" was coined by Adolf Dirr, who labeled the case of the transitive subject in the Caucasian languages that way (cf. Dirr 1928). Before that it was common to read the ergative constructions, which were then already known from various languages as "passive".

Many languages show ergativity only on the morphological intra-clausal level. Some languages, however, have ergative features on the syntactical level.

The term "intra-clausal ergativity", also called "morphological ergativity" or "surface ergativity", indicates that within a single clause S (i.e., the intransitive subject) and O (i.e., the direct object) are marked the same way, and A (i.e., the transitive subject or agent) is marked in a different way.

There are various means of distinguishing A and O: cases (Basque), particles (Tongan from the Austronesian language family), adpositions, constituent order, and pronominal cross-referencing on the verb (Abaza, Abkhaz). Many languages combine these strategies, e.g., Georgian and Circassian use cases and pronominal cross-referencing affixes (cf. Dixon 1994: 39).

The term "inter-clausal ergativity", also called "syntactic ergativity" or "deep ergativity", indicates that there are syntactic constraints on clause combinations, or on the omission of coreferential NPs in clause combinations, which treat S

and O in the same way and A differently. For instance, the Australian language Dyirbal is a language with inter-clausal ergativity (cf. Dixon 1994: 143).

The studies of Matras (1992–93, 1997) show that Northern Kurdish has intra-clausal ergativity, but not inter-clausal ergativity.

Many languages are not entirely ergative and follow the accusative pattern in parts of their system. In the Iranian languages we have a tense/aspect split. The present tense and the future tense are constructed accusatively and the past tenses show an ergative system. According to Dixon (1994: 99), this pattern is common among ergative languages with tense/aspect split:

> Many languages can, of course, have nominative-accusative marking in all aspects and tenses, and others have absolutive-ergative marking unimpeded by aspect or tense [...]. But if a split is conditioned by tense or aspect, the ergative marking is always found either in past tense or in perfective aspect.
>
> There are several other kinds of split ergativity. Splits can also be conditioned by the semantic nature of the verb: In some languages the S is marked like A (by the ergative case) when the meaning of the (intransitive!) verb implies that the S exerts control over the action, and it is marked like O (by the absolutive case) when the meaning of the verb implies that the S is affected by the action (cf. Dixon 1994: 70).

Another kind of split is determined by the semantic nature of the NP. According to a nominal hierarchy, a first person pronoun is to be expected more than any other part of speech to operate as A rather than as O. Second person agreement is next in the hierarchy, then demonstratives and third person pronouns, and at last proper names and common nouns (cf. Dixon 1994: 83; DeLancey 1981: 628). In Balōčī, an Iranian language with the already mentioned tense/aspect split, first and second person pronouns are not marked by the oblique case when they act as A in the ergative domain (past tense), whereas third person pronouns, demonstrative pronouns, common nouns, and proper names in A function are marked by the oblique case in the ergative domain. As a consequence in Balōčī, there is a nominal hierarchy split in addition to the tense/aspect split (cf. Farrell 1990: 67):[1]

---

**1** 

| | | | |
|---|---|---|---|
| 1SG | First Person Singular | N | Noun |
| DIR | Direct Object | | |
| SG | Singular | IMPF | Imperfective |
| PST | Past | PL | Plural |
| 3SG | Third Person Singular | M | Masculine |
| OBL | Oblique Case | PP | Past Participle |
| V | Verb | ACC | Accusative |
| O | Object | PRON | Pronoun |
| PRES | Present | 3PL | Third Person plural |

(1)  *mɘn*      *bɘcɪk*       *ja*
     1SG.DIR   boy.DIR.SG   hit.PST.3SG
     A          O             V(=O)
     'I hit the boy'.

(2)  *tɘo*      *bɘcɪk*       *ja*
     2SG.DIR   boy.DIR.SG   hit.PST.3SG
     A          O             V(=O)
     'You hit the boy'.

(3)  *aya*      *bɘcɪk*       *ja*
     3SG.OBL   boy.DIR.SG   hit.PST.3SG
     A          O             V(=O)
     'He hit the boy'.

# 3  New West Iranian languages

These languages belong to the New West Iranian group, which traces back historically to the Old Iranian languages Old Persian and Median and the Middle Iranian languages Middle Persian and Parthian. New West Iranian is subdivided in two groups, in (a) the Southwest Iranian languages, which are descendants of Old and Middle Persian or genetically related languages, which are not attested, and in (b) the Northwest Iranian languages, which are descendants of Median and Parthian or genetically related languages, which are not attested.

The grouping inside the New West Iranian languages does therefore not rely on the present geographical location of the single languages but on their historical filiation. The following list shows which New West Iranian languages have ergative past tenses and which have accusative past tenses:

(a)  North-Western subgroup
     Kurdish (further subdivision in North, Central, and South Kurdish) (split ergativity)
     Zāzākī (split ergativity)
     Gōrānī (split ergativity)
     Balōčī (split ergativity)
     Māzandarānī (pure accusative system)
     Gīlakī (pure accusative system)
     *Āzarī* dialects (split ergativity)
     Tālešī (split ergativity)
     Semnānī (remnants of ergativity in the inflection of transitive verbs in the past tense)

Dialects of the region Semnān (remnants of ergativity in the inflection of transitive verbs in the past tense)

Central dialects (split ergativity)

(b) South-Western subgroup

Persian and its dialects (including Tajik and Darī) (pure accusative system)

Tātī (pure accusative system)

Lurī (pure accusative system)

Fārs Dialects (split ergativity)

Lārestānī, Baškardī, und Kumzārī (split ergativity)

# 4 Diachronic roots of the ergativity in New West Iranian

Ergativity in the New West Iranian languages emerges from a periphrastic perfective construction with passive value that is already recorded in Old Persian texts. In the western Middle Iranian languages, this construction has replaced the synthetic past tenses of the Old Iranian period. Due to the passive value of the construction, the agent is marked by a non-nominative case, namely, in Old Persian by the genitive-dative case, and in Parthian and Middle Persian by the oblique case.

## 4.1 Old Persian

Old Persian is the only representative of Southwest Iranian in the Old Iranian period. Consequently it is the ancestor of Middle Persian and New Persian. Old Persian is a highly fusional language with a rich case system and a complex tense system. It follows the accusative pattern. See the following example of a transitive verb in the imperfect:

(4) *pasāva    adam   kāram    pārsam      frāišayam*
    thereafter 1SG.N  army.A.SG  Persian. SG   send.IMPF.1SG
    'After that I sent a Persian army' (DB 3.2).

However, the original Iranian synthetic perfect has been replaced by an analytic construction with a passive value.[2] Some scholars have labeled this construction as

---

**2** There exists one single remnant of the old synthetic perfect in Old Persian: *caxriyā* 'has been done'.

*manā-kartam* construction. It consists of the participle perfect passive and optional the present tense form of the copula, which is often omitted in the third person.

The past participle of an intransitive verb has active meaning; its subject is marked by the nominative case:

(5) *hamiçiyā        hagmatā        paraitā*
   insurgent.N.PL.M gather.PP.N.PL.M  advance.PP.N.PL.M
   *Patiš    Vivānam*
   against   Vivana.A.SG.M
   'The insurgents gathered and advanced against Vivana'
   (DB 2.32).

The participle of a transitive verb has passive meaning when there is no logical subject/agent. If there is an agent, it is marked by the genitive-dative case. The patient/logical object is marked by the nominative case (examples from Schmitt 1989a: 80):
   without agent:

(6) */taya Bardiya        ava-jata/*
   that   Smerdis.N.SG.M   kill.PP.N.SG.N
   'that Smerdis has been killed'

without agent, with copula:

(7) */yadi  kāra       Pārsa         pāta         ahati/*
   when   army.N.SG.M Persian.N.SG.M protect.PP.N.SG.M be.PRS.CONJ.3SG
   'when the Persian army has been protected' (DPe 1.22).

with agent:

(8) */ima kartam/*                    *taya*             *manā*
   PRON (3-N/A-SG-n) do (PP-N/A-SG-n)  RELPRON (N/A-SG.n)  PRON (1SG-G/D)
   'that what I have done' (DB 1.27).

The *manā-kartam* construction has traditionally been regarded as a passive form (cf. Pirejko 1979; Jamison 1979a, 1979b). Some, however, for instance Emile Benveniste, took the view that it is a possessive construction because the agent is marked by the genitive-dative case. Benveniste stated that the *manā-kartam* construction is an active perfect, which is expressed by a possessive syntactic structure (in analogy to Latin mihi factum est → habeo factum) (cf. Benveniste 1952: 56). As a matter of

fact, the oblique endings found in various Middle Iranian and New Iranian languages derive diachronically from Old Iranian genitive endings (cf. Pirejko 1979; Jamison 1979a, 1979b). The suffixed pronouns, which denote the ergative agent in Middle and New Iranian, also have possessive meaning. Nevertheless this theory cannot explain the passive meaning of the agentless *manā kartam* construction (cf. Skjaervo 1985: 217).

The passive value of the past tense stem is still visible in ergative New Iranian languages, when a transitive verb in the past tense has no agent, cf. the Northern Kurdish examples:

(9) *ez*             *dîtim*
    PRON (1SG-DIR)   see (PAST-1SG)
    S           V(=S)
    'I was seen'.

But with agent:

(10) *wî*         *ez*         *dîtim*
    PRON (3SG-OBL)  see (PAST-1SG)  pron (1SG-DIR)
    A          O          V(=O)
    'He saw me'.

The passive meaning of the past tense stem is lost in the new Iranian languages with accusative past tenses, cf. the New Persian examples:

(11) *man*      *dīdam*
    PRON (1SG)  see (PAST-1SG)
    S         V(=S)
    'I saw'.

(12) *man*      *to-rā*      *dīdam*
    PRON (1SG)  see (PAST-1SG)  PRON (2SG)-OBJPART
    A         O          V(=A)
    'I saw you'.

## 4.2 Parthian and Middle Persian

In Middle Persian and Parthian, the Old Iranian tense system has been reduced radically. The synthetic past tenses have vanished entirely and have been replaced

by a construction like that found in the Old Persian analytic perfect. In older Middle Persian and Parthian, the past participle has retained its passive value.

### 4.2.1 Parthian

If there is no agent, transitive verbs in the past tense have passive value. If there is an agent, it is marked by the oblique case, the logical object is unmarked (direct case), and the verb agrees with the subject. However, in Manichaean Parthian only the personal pronoun of the first person singular distinguishes between direct and oblique case. In the inscriptions of the Arsakid period, there is an oblique case for the noun in the plural. In the singular, however, there is no distinction between direct case and oblique case (cf. Sundermann 1989b:130).

Examples (from Gippert 1996: 152):

(13) *hawīn*        *abgundām*
     PRON (3Pl)      uncover (PRES-1sG)
     'I uncover them'.

(14) *abgust*       *ahēnd*
     uncover (PP)    be (PST-3PL)
     'they were uncovered'

(15) *man*         *abgust*       *(a)hēnd*
     PRON (1SG-OBL)    uncover (PP)    be (PRES-3PL)
     'I uncovered them'

(16) *az*           *Kāram*
     PRON (1SG-DIR)    do (PRES-1SG)
     'I do'

(17) *man*         *kird*
     PRON (1SG-OBL)    do (PP)
     'I did'

(18) *az*           *vāžam*
     PRON (1SG-DIR)    say (PRES-1SG)
     'I say'

(19)  *man*                *vaxt*
      PRON (1SG-OBL)      say (PP)
      'I said'

### 4.2.2 Middle Persian

In Middle Persian – a Southwest Iranian language – the synthetic past tenses have been replaced by an analytic construction, similar to what we find in Parthian. In early texts the past tenses are formed passively (or ergatively, respectively). In later texts this construction has been activized under the influence of the present tense.

Examples (from Sundermann 1989a: 152):

(20)  *awēšān*        *paymōzam*
      PRON (3PL)     dress (PRES-3PL)
      O              V(=A)
      'I dress them'

(21)  *paymōxt*    *hēnd*
      dress (PP)   be (PRES-3PL)
      'they got dressed'

(22)  *man*                *paymōxt*    *hēnd*
      PRON (1SG-OBL)      dress (PP)   be (PRES-3PL)
      A                   V(=O)
      'I dressed them'

# 5 Strategies of marking the agent in the past tenses

Two strategies to mark the ergative agent in the past tenses are employed in New West Iranian languages that have preserved the old ergative construction: Strategy one is to mark the agent by the oblique case; strategy two is to mark it by an agentive suffix, i.e., a suffixed pronoun, which also has possessive meaning and is used to indicate direct and indirect objecthood in the present tense. These two ways of expression can co-occur in languages that have retained both case inflection and suffixed pronouns.

## 5.1 Agent is marked by oblique case

The arguments of the verb are marked by case inflection, i.e., the nouns are marked by case endings, independent pronouns by suppletion. In the present tense and the future tense, the subjects of intransitive as well as transitive verbs are marked by direct case, namely, they are not characterized by any case affixes; direct objects are marked by oblique case. The ergativity of transitive sentences in the past tenses is expressed by the fact that the oblique case marks the subject of the transitive verb, and the direct case marks the direct object. Northern Kurdish and Zāzākī employ only that kind of agent marking because they do not have suffixed pronouns. Other West Iranian languages using the oblique case to mark the ergative agent are the following: several Central Kurdish dialects Hawrāmī and related dialects of Gōrānī, Tālešī, Āzarī dialects, the Central dialects spoken around Tafreš, and Balōčī.

Examples from Northern Kurdish:

(23) *ew*　　　　　　*min*　　　　　　*dibîne*
PRON (3SG-DIR)　PRON (1SG-OBL)　see (PRES-3SG)
A　　　　　　　　O　　　　　　　　V(=A)
'He sees me'.

(24) *wî*　　　　　　*ez*　　　　　　*dîtim*
PRON (3SG-OBL)　PRON (1SG-DIR)　see (PAST-1SG)
A　　　　　　　　O　　　　　　　　V(=O)
'He saw me'.

(25) *keçik-ek*　　　　*hêspan*　　　　*dibîne*
girl (DIR-SG)-INDET　horse (OBL-PL)　see (PRES-3SG)
A　　　　　　　　　　O　　　　　　　　V(=A)
'A girl sees the horses'.

(26) *keçik-êk-ê*　　　　*hesp*　　　　*dîtin*
girl (SG)-INDET-OBL　horse (DIR-PL)　see (PAST-3PL)
A　　　　　　　　　　O　　　　　　　　V(=O)
'A girl saw the horses' (example from Blau 1989b: 331).

(27) *tu*　　　　　　*min*　　　　　　*nas dikî*
PRON (2SG-DIR)　PRON (1SG-OBL)　know (PRES-2SG)
A　　　　　　　　O　　　　　　　　V(=O)
'You recognize me'.

(28) *te*                   *ez*                    *nas kirim*
    PRON (2SG-OBL)   PRON (1SG-OBL)   recognize (PAST-1SG)
    A                         O                          V(=O)
    'You recognized me' (example from Blau 1989b: 331).

(29) *ez*                   *te*                    *tînim*
    PRON (1SG-DIR)   PRON (2SG-OBL)   pick up (PRES-1SG)
    A                         O                          V(=A)
    'I pick you up'.

(30) *min*                  *Tu*                    *anîyî*
    PRON (1SSG-OBL)  PRON (2SG)   pick up (PST-2SG)
    A                         O                  V(=O)
    'I picked you up' (example from Blau 1989b: 331).

## 5.2 Agentive suffix

The other strategy of marking the ergative agent is by agentive suffixes, i.e., suffixed pronouns in agent function. In ergative Iranian languages without case inflection, they are the only way to maintain the ergative pattern. Even when there is an overt nominal agent, there is usually also an agentive suffix resuming the agent inside the sentence. In the ergative past of languages without cases, the use of the agentive suffix is compulsory. These suffixed pronouns occur in nearly all major Iranian languages, except Northern Kurdish and Zāzākī. Their basic function is to express possessive and oblique case relations:

> Enclitic pronouns are found in all dialects except Dimili, Northern Kurdish, and Sangesari; they are confined to *xod* 'own, self' in Mazandarani and Gilaki. They have the general function of oblique cases. This applies even to non-inflecting dialects such as Persian, where the suffixes mark possession, e.g., *ketāb-aš* 'his book', direct object, e.g., *dīd-aš* 'he saw him, it', and indirect object, e.g., *yād-aš āmad* 'he remembered' (literally 'it came him to memory') (Windfuhr 1989c: 259).

In the present tense they express direct or indirect object; in the past tense of ergative languages, however, they indicate the agent.

Here are examples from the ergative past tense in the Sulaimanīya dialect of Central Kurdish:

(31) *pyāw-aka*  *sag-aka*  *-y*          *kušt*
    man-DET       dog-DET    -CLIT (3SG)  kill (PRET-3SG)
    A                    O             AS               V(=O)
    'The man killed the dog'.

(32) *pyāw-aka nān-aka -y dā ba sag-aka*
     man-DET    bread-DET   -CLIT (3SG)  give (PRET-3SG)  to  dog-DET
     A           O              AS          V(=O)              IO
     'The man gave the bread to the dog'.

(33) *pyāw-aka ba sag-aka -y Dā*
     man-DET   to  dog-DET  -CLIT (3SG)  give (PRET-3SG)
     A          IO          AS          V(=O)
     'The man gave it to the dog'.

(34) *kušt -ī*
     kill (PRET-3SG)-  CLIT (3SG)
     V(=O)             AS
     'He killed him/she/it'.

In languages that have abandoned ergativity in the past tenses and generalized the accusative pattern of the present and the future, as for instance Persian, the suffixed pronouns indicate the direct and indirect object also in the past tenses. Nevertheless there are languages that use both case endings and suffixed pronouns in the ergative construction of the past tenses. Essentially these languages are Hawrāmī and related dialects of Gōrānī, some dialects of Central Kurdish as well as the Caspian dialects. In some Balōčī dialects there is also the possibility to indicate the agent in the ergative past by means of suffixed pronouns. However, in Balōčī it is far more common to mark the ergative agent only by case. Consider the example from the Balōčī dialect of Karachi.

> In Semnānī and the dialects of the region Semnān the agentive suffixes are attached permanently to the past tense stem of the verb and thus form a new set of endings for the past tense of transitive verbs (cf. Lecoq 1989a: 308).

Suffixed pronouns indicating the agent were already used in earlier linguistic stages of the West Iranian language history. Already in Old Persian there are examples of suffixed pronouns acting as agent of a passive verb (examples from Pirejko 1979):

(35) *utā -šām Auramazdā naiy ayadiya*
     and  -CLIT (3PL)  Auramazda (N-SG-m)  not  rever (IMPF-PASS-3SG)
     'and Auramazda was not revered by them'

or:

(36)  *utā   Auramazdā           -šām      naiy    ayadiya*
      and   Auramazda (N-SG-m)   -CLIT (3PL)   not     rever (IMPF-PASS-3SG)
      'and Auramazda was not revered by them'

Suffixed pronouns in agent function are also attested in the Middle Persian past tenses (Nyberg 1974: 282):

(37)  *api   -š          Ohurmazd   guft*
            CLIT (3SG)   Ohurmazd   speak (PP)
      *api   Ohurmazd   guft         aš*
            Ohurmazd   speak (PP)   CLIT (3SG)
      'he, Ohurmazd, spoke'

(38)  *api   -mān         nē       ōzat*
            CLIT (1PL)   NEG     kill (PP)
      'was not killed by us' (Nyberg 1974: 124).

(39)  *cē-   šān         Zarērē-c              ōzat*
      why   -CLIT (3PL)   Zarer (OBL-SG)-EMPH   kill (PST-3SG)
      'Why did they kill Zarer of all people?' (Nyberg 1974: 184).

## 5.3  Personal endings express the indirect object or the possessor of the direct object

In several New West Iranian languages, transitive verbs in the past tenses can agree with the indirect object or the possessor of the direct object. This specific variant of the ergative construction is found in dialects of Central Kurdish, in Hawrāmī (possibly also in other Gōrānī dialects), in some Central dialects, in Somġunī (possibly also in other Fārs dialects), and in Lārestānī and Baškardī a pronominal indirect object ("to him", "from him", etc.) or possessor ("my", "your", "his", etc.) can be indicated by the personal endings of a transitive verb governing a third person pronoun or noun as direct object (cf. Windfuhr 1989c: 260).

Hawrāmī examples of verb agreeing with the indirect object (examples from MacKenzie 1966: 53):

(40)  *kiteb-aka           -š        dāne          pana*
      book (DIR-SG)-DET   -CLIT (3SG)   give (PAST-1SG)   to
      O                  AS        V(=IO)
      'He gave me the book'. (literally: 'The book, by him I was given to'.)

(41)    *kiteb-aka*        *-m*        *dāni*        *pana*
        book (DIR-SG)-DET  -CLIT (1SG)  give (PAST-2SG)  to
        O                    AS           V(=IO)
        'I gave you the book'. (literally: 'The book, by me you were given to'.)

(42)    *i*    *zamin-a*        *-tā*       *čana*  *n-asā-yme*
        this   land (DIR-SG)-DET  -CLIT (2PL)  from  NEG-buy (PAST-1PL)
               O                AS              V(=IO)
        'You (PL) did not buy this land from us'. (literally: 'This land, by you we
        were not bought from'.

Sulaimānīya (dialect of Central Kurdish) examples of verb agreeing with the indi-rect object (examples from MacKenzie 1961: 115):

(43)    *awān*     *-ī*      *lē*    *sand-īn*
        PRON (3PL)  -CLIT (3SG)  from  take (PAST-1PL)
        O            AS         V(=IO)
        'He took them from us'. (literally: 'They, by him we were taken from'.)

In this dialect the verb stem can bear two endings indicating as well the direct as the indirect object:

(44)    *lē*     *-y*      *sand-in-īn*
        from    -CLIT (3SG)  take (PAST-3PL-1PL)
                AS         V(=O-IO)
        'He took them from us'.

Hawrāmī examples of verb agreeing with the possessor of the direct object (POSS) (examples from MacKenzie 1966: 53):

(45)    *bāxča-ka*    *-š*      *diayme*
        garden-DET  -CLIT (3SG)  see (PAST-1PL)
        O            AS          V(=POSS)
        'He saw our garden'. (literally: 'The garden, by him we were seen'.)

(46)    *bāxča-ka*    *-m*      *diene*
        Garten-DET  -ENKL (1SG)  sehen (PAST-3PL)
        O            AS          V(=POSS)
        'I saw their garden'. (literally: 'The garden, by me they were seen'.)

Sulaimānīya examples of verb agreeing with the possessor of the direct object (POSS) (examples from MacKenzie 1961: 115):

(47) *bak-ak-ān*    *-ī*       *a-xwārdim*
     child-DET-PL   -CLIT (3SG)   eat (IMPF-1SG)
     O            AS          V(=POSS)
     'It used to eat my children'. ('The children, by him I used to be eaten'.)

(48) *šēt-aka*    *das*   *-ī*       *gazīm*
     maniac-DET   hand   -CLIT (3SG)   bite (PAST-1SG)
     A            O     AS         V(=POSS)
     'The maniac bit my hand'. (literally: 'The maniac, the hand, by him I was eaten'.)

# 6 Verbal compounds

Verbal compounds of the type nominal meaningful part + auxiliary verb carrying the grammatical information are very common in the New Western Iranian languages. In the languages and dialects that maintain split ergativity, transitivity, or intransitivity of the auxiliary verb are crucial in the formation of the past tense of the compound. For instance, in Northern Kurdish verbal compounds with the auxiliary verbs *kirin* 'do' or *dan* 'give' are treated as transitive verbs and follow the ergative pattern in the past tenses albeit the meaning of the entire compound. That is, they are treated as transitive verbs even if they are semantically intransitive. The nominal part of the compound seems to act as direct object. Cf. the Northern Kurdish example *limêj kirin* 'pray' (examples from Bedir Khan and Lescot 1970: 187):

(49) *ez*           *limêj*           *di-k-im*
     PRON (1SG-DIR)   prayer (DIR-SG)   do (PAST-3SG)
     A              O             V(=A)
     'I pray'

(50) *me*           *limêj*           *kir*
     PRON (1SG-OBL)   prayer (DIR-SG)   do (PAST-3SG)
     A              O             V(=O)
     'I prayed'

Zāzākī examples (examples from Paul 1998: 132):

(51) *Mi*            *žī*   *zūrī*       *kerd-ī*
PRON (1SG-OBL)  also  lie (DIR-PL)  do (PAST-PL)
A                      O             V(=O)
'I lied as well'.

(52) *mā-di*          *qisey*        *kerd-ē*
PRON (1PL)-about  story (DIR-PL)  do (PERF-PL)
                         O             V(=O)
'He has spoken about us'.

Balōčī examples (from Farrell 1995: 232):

(53) *kangi-a*        *bal*   *kω*
crow (OBL-SG)  flight  do (PAST-3SG)
A                   O      V(=O)
'The crow flew'.

(54) *mor-a*        *der*   *kω*
ant (OBL-SG)  delay  do (PAST-3SG)
A                 O        V(=O)
'The ant was late'.

In Northern Kurdish semantically transitive compounds like *nas kirin* 'know' in this respect seem to govern two direct objects, whereby the nominal part of the compound has to be regarded as generic and therefore it is not marked by the oblique case in the present tense and the future tense:

(55) *ez*             *te*           *nas*           *di-k-im*
PRON (1SG-DIR)  PRON (2SG-OBL)  cognition (RECT-SG)  do (PRES-1SG)
A                  $O_1$             $O_2$           V(=A)
'I know you'.

(56) *me*             *tu*           *nas*           *kir-î*
PRON (1SG-OBL)  PRON (2SG-DIR)  cognition (DIR-SG)  do (PAST-2SG)
A                  $O_1$             $O_2$           (=$O_1$)
'I knew you'.

# 7 Results

Split ergativity in various New West Iranian languages has its roots in the Old Iranian linguistic stage. In Old Persian, the oldest attested representative of West Iranian, the inherited synthetic perfect has been replaced by an analytic construction, which is composed of the passive past participle and the present tense of the auxiliary verb. The agent, if there is any, is marked by the genitive case. The auxiliary verb agrees with the logical object, which is marked by the nominative case. In Middle Persian and Parthian this construction forms the basis of the past tense. Due to the passive value of the past participle, the subject of a transitive verb in the past tense is marked by the oblique case. The verb agrees with the direct object.

The past tense stem of the New Iranian verb, with which the past tenses are formed, etymologically goes back to the Old Iranian past participle in *-ta-*. In many languages the passive character of the past tense stem, which is the historical source of the ergative construction in the past tenses, has been preserved.

The ergative domain in the New West Iranian languages, which have ergativity, covers the entire past tenses, which are formed with the past tense stem and the perfect stem, which is a derivation of the former. The past continuous tense in Balōčī, which follows an accusative pattern, is not an exception to this schema, since it is not formed with the past tense stem.

As a consequence of the case marking resources in the individual languages, there arise varying specifications and forms of the ergative construction. The main patterns are:

The agent of a transitive verb in the past tense is marked by the oblique case. This implies that there is case inflection, a criterion that is often not met.

The agent in the ergative domain is expressed by a suffixed pronoun indicating the actor of an action referred to by a transitive verb. This pronoun can be attached to any constituent in the clause. A suffixed pronoun is used to express the agent even if there is an overt nominal subject at the beginning of the sentence.

Languages with case inflection often use both principles: they mark pronominal and nominal subjects in the ergative domain with oblique case and at the same time indicate the subject by a suffixed pronoun.

In many New West Iranian languages transitive verbs in the ergative domain agree with the direct object in person and number. (for instance, Zāzākī, Kurdish, and Hawrāmī). In other languages verbal agreement has been reduced to a great extent: the verb occurs only with the ending-less third person singular in the ergative domain (for example, the most Central Dialects, Sivandi, etc.).

In some languages, namely, Semnāni and the dialects of the region Semnān, the suffixed pronouns occur exclusively attached to the past tense stem of transitive verbs in the past tense and that way form a second (transitive) set of past tense endings along with the regular endings of the intransitive past.

In dialects of Central Kurdish, in Hawrāmī and several other languages and dialects, a pronominal indirect object or possessor can be indicated by the personal endings of the transitive verb in the ergative domain.

**Table 1:** Outline of the results:

|  | A → obl.[1] | AS[2] | AS as verbal ending[3] |
|---|---|---|---|
| **Zāzākī** | + | − | − |
| **Northern Kurdish** | + | − | − |
| **Central Kurdish** |  |  |  |
| Dialects of Bingird, Piždar, Arbil, Rewandiz, Xošnāw | + | + | − |
| Mukrī | − | + | − |
| Dialects of Sulaimānīya and Warmāwa | − | + | − |
| **Gōrānī (Hawrāmī)** | + | + | − |
| **Caspian dialects and North-western dialects** |  |  |  |
| Tālešī | + | + | − |
| Āzarī dialects | + | + | − |
| Semnānī | + | − | + |
| Dialects of the region Semnān | + | − | + |
| **Central dialects** |  |  |  |
| Tafreš | + | + | − |
| Northwest | − | + | − |
| Northeast | − | + | − |
| Southwest | − | + | − |
| Southeast | − | + | − |
| Kavir dialects | − | + | − |
| **Dialects in Southwest Iran** |  |  |  |
| Sīvandī | − | + | − |
| Fārs Dialects | − | + | − |
| **Dialects in Southeast Iran** |  |  |  |
| Lārestānī | − | + | − |
| Baškardī | − | + | − |
| Balōčī | + | (+) | − |

**1** The subject of a transitive verb in the ergative domain is marked by the oblique case.
**2** Agentive suffixes: Suffixed pronouns denote the subjects of transitive verbs in the ergative domain.
**3** Agentive suffixes occur exclusively attached to the past tense stem of transitive verbs in the past tense.

# Bibliography

## Books

Bedir-Khan, Djeladet (Emir), and Roger Lescot. 1970. *Grammaire kurde: Dialecte kurmandji*. Paris: Adrien Maisonneuve.

Brandenstein, Wilhelm, and Manfred Mayerhofer. 1964. *Handbuch des Altpersischen*. Wiesbaden: Harrassowitz.

Collett, N. A. 1983 *A grammar, phrase book and vocabulary of Balochi (as spoken in the Sultanate of Oman)*. Abingdon: Burgess and Son.

Dirr, Adolf. 1928. *Einführung in das Studium der kaukasischen Sprachen*. Leipzig: Asia Major.

Dixon, Robert M. W. 1994. *Ergativity*. Cambridge: Cambridge Univeristy Press.

Dorleijn, Margreet. 1996. *The decay of ergativity in Kurmanci: Language internal or contact induced?* Tilburg: Tilburg University Press.

Elfenbein, Josef. 1966. *The Baluchi language: A dialectology with texts*. London: n.p.

Elfenbein, Josef. 1990. *An anthology of classical and modern Balochi literature*. 2 vols. Wiesbaden: Harrassowitz.

Farrell, Tim. 1990. *Basic Balochi: An introductory course*. Naples: Istituto universitario orientale.

Geiger, Wilhelm, and Ernst Kuhn (eds.). 1898–1901. *Grundriss der iranischen Philologie*. Strassburg: J. Trubner.

Lecoq, Pierre. 1979. *Le dialecte de Sivand*. Wiesbaden: Dr Ludwig Reichert Verlag.

MacKenzie, D. N. 1961. *Kurdish dialect studies*. 2 vols. London: Oxford Univeristy Press.

MacKenzie, D. N. 1966. *The dialect of Awroman (Hawrāmān-ī Luhōn)*. Copenhagen: Enjar Munsksgaard.

Nyberg, Henrik Samuel. 1974. *A manual of Pahlavi*. 2 vols. Wiesbaden: Ludwig Reichert Verlag.

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

Plank, Frans (ed.). 1979. *Ergativity*: Towards a Theory of Grammatical Relations. London: Academic Press.

Schmitt, Rüdiger (ed.). 1989. *Compendium Linguarum Iranicarum* [Guide to Iranian Languages]. Wiesbaden: Dr. Ludwig Reichert Verlag.

Wurzel, Petra. 1997. *Rojbaş. Einführung in die kurdische Sprache*. 2 vols. Wiesbaden: Ludwig Reichert Verlag.

## Articles

Asatrian, Garnik S. 1996. Dīm(ī)lī. In Ehsan Yarshater (ed.), *Encyclopaedia Iranica* VII, 405–411. Costa Mesa, CA: Mazda Publishers.

Benveniste, Emile. 1952. La construktion passive du parfait transitif. In *BSL* (Bulletin de la Société de linguistique de Paris) 48 (1952). 52–62.

Blau, Joyce. 1989a. Gurn et Zz. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum* [Guide to Iranian Languages], 336–340. Wiesbaden: Dr. Ludwig Reichert Verlag.

Blau, Joyce. 1989b. Le Kurde. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum* [Guide to Iranian Languages]. 327–335. Wiesbaden: Dr. Ludwig Reichert Verlag.

Bynon, Theodora. 1979. The ergative construction in Kurdish. *BSOAS* (Bulletin of the School of Oriental and African Studies) 42. 211–224.

DeLancey, Scott. 1981. An interpretation of split ergativity and related patterns. *Language* 57 (3). 626–657.

Dixon, Robert M. W. 1979. Ergativity. *Language* 55 (1). 59–138.

Elfenbein, Josef. 1989a. Baluchistan III. In Ehsan Yarshater (ed.), *Encyclopaedia Iranica*, vol. III, 633–644. London and New York: Routledge and Kegan Paul.

Elfenbein, Josef. 1989b. Baloči. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum* [Guide to Iranian Languages]. 350–362. Wiesbaden: Dr. Ludwig Reichert Verlag.

Farrell, Tim. 1995. Fading ergativity: A study of ergatity in Balochi. In T. Bynon, D. Benett & G. Hewitt (eds.), *Subject, voice and ergativity*. 218–241. London: School of Oriental and African Studies.

Geiger, Wilhelm. 1893. Die Passivkonstruktion des Prateritums transitiver Verba im Iranischen. In Ernst Kuhn (ed) *Festgruss an Rudolph v. Roth zum Doktor-Jubiläum*, 1–5. Stuttgart: W. Kohlhammer.

Gippert, Jost. 1996. Die historische Entwicklung der Zaza-Sprache. *Dergi: Ware. Pêseroka Zon u Kulturê Ma: Dımıli-Kırmanc-Zaza*, 10:148–154.

Grierson, Sir George Abraham. 1921. Baloči. In George Abraham Grierson (eds.), *Linguistic survey of India*. 327–362. Calcutta: Office of the Superintendent of Government Printing.

Hopper, Paul J. & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56 (2). 251–299.

Jamison, Stephanie W. 1979a. The case of the agent in Indo-European. *Die Sprache* 25. 129–143.

Jamison, Stephani W 1979b. Remarks on the expression of agency with the passive in Vedic and Indo-European. *KZ* (*Zeitschrift für vergleichende Sprachforschung*) 93. 196–219.

Lecoq, Pierre. 1989a. Les dialectes caspiennes et du nordouest. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum* [Guide to Iranian Languages]. 296–312. Wiesbaden: Dr. Ludwig Reichert Verlag.

Lecoq, Pierre. 1989b. Les dialectes du centre de l'Iran. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum* [Guide to Iranian Languages]. 313–326. Wiesbaden: Dr. Ludwig Reichert Verlag.

Lecoq, Pierre. 1989c. Les dialectes du sud-ouest de l'Iran. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum* [Guide to Iranian Languages]. 341–349. Wiesbaden: Dr. Ludwig Reichert Verlag.

Matras, Yaron. 1992–1993. Ergativity in Kurmanji (Kurdish): Notes on its use and distribution. *Orientalia Suecana* XLI (XLII). 139–154.

Matras, Yaron. 1997. Clause combining, ergativity, and coreferent deletion in Kurmanji. *Studies in Language* 21 (3). 613–653.

Moshkalo, V. V. 1991. In Beludzhskij jazyk (ed.), *Osnovy iranskogo jazykoznanija. Novoiranskie jazyki*, vol. I. 5–99. Moskva: Izdat Nauka.

Pirejko, Lija A. 1979. On the genesis of the ergative construction in Indo-Iranian. In Frans Plank (ed.), *Ergativity*. 481–488. London: John Benjamins.

Schmitt, Rüdiger. 1989a. Altpersisch. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum* [Guide to Iranian Languages]. 56–85. Wiesbaden: Dr. Ludwig Reichert Verlag.

Schmitt, Rüdiger. 1989b. Mitteliranische Sprachen im berblick. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum* [Guide to Iranian Languages]. 95–105. Wiesbaden: Dr. Ludwig Reichert Verlag.

Skjaervo, Prods Oktor. 1985. Remarks on the Old Persian verbal system. *Münchener Studien zur Sprachwissenschaft* 45 (2). 211–227.

Skjaervo, Prods Oktor. 1989. Languages of Southeast Iran: Lārestānī, Kumzārī, Baškardī. In Rüdiger Schmitt (ed.), Compendium Linguarum Iranicarum [Guide to Iranian Languages]. 363–369. Wiesbaden: Dr. Ludwig Reichert Verlag.

Socin, Albert. 1898–1901. Die Sprache der Kurden. In Wilhelm Geiger & Ernst Kuhn (eds.), *Grundriss der Iranischen Philologie*. 249–286. Strassburg: J. Tr'ubner.

Sundermann, Werner. 1989a. Mittelpersisch. In Rüdiger Schmitt (ed.), Compendium Linguarum Iranicarum [Guide to Iranian Languages]. 138–164. Wiesbaden: Dr. Ludwig Reichert Verlag.

Sundermann, Werner. 1989b. Parthisch. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum* [Guide to Iranian Languages]. 114–137. Wiesbaden: Dr. Ludwig Reichert Verlag.

Sundermann, Werner. 1989c. Westmitteliranische Sprachen. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum* [Guide to Iranian Languages]. 106–113. Wiesbaden: Dr. Ludwig Reichert Verlag.

Windfuhr, Gernot L. 1989a. Western Iranian dialects. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum* [Guide to Iranian Languages]. 294–295. Wiesbaden: Dr. Ludwig Reichert Verlag.

Windfuhr, Gernot L. 1989b. New Iranian languages: Overview. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum* [Guide to Iranian Languages]. 246–250. Wiesbaden: Dr. Ludwig Reichert Verlag.

Windfuhr, Gernot L. 1989c. New West Iranian. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum* [Guide to Iranian Languages]. 251–262. Wiesbaden: Dr. Ludwig Reichert Verlag.

Windfuhr, Gernot L. 1996a. Central dialects. In Ehsan Yarshater (Ed.), *Encyclopaedia Iranica*, vol. V, 243–252. Costa Mesa, CA: Mazda Publishers.

Windfuhr, Gernot L. 1996b. Fars VIII. Dialects. In Ehsan Yarshater (ed.), *Encyclopaedia Iranica*, vol. IX, 362–373. New York: Bibliotheca Persica Press.

Gilbert Lazard†*

# Aspect in Iranian two systems: Persian and Pashto

## 1 Introduction

Aspect is a domain where the structural diversity of Iranian languages is conspicuous. I would like to illustrate that diversity by comparing the verb systems of two languages, (New) Persian and Pashto. I shall first summarily recall the main features of the Persian system, then those of the Pashto one, then examine the two systems and present a few remarks on some important points.

## 2 Persian

The Persian verb system consists of a number of oppositions, which may or may not combine with one another.

(1)  Time: present vs. past. Ex. *miravad* 'goes, is going' ~ *miraft* 'went, used to go, was going', *rafte-ast* 'has gone' ~ *rafte bud* 'had gone'.

(2)  Durativeness: non-durative (or aorist) vs. durative. Ex. *raft* 'went' ~ *miraft* 'was going', *rafte*(-*ast*) ~ *mirafte*(-*ast*). A durative verb form presents the action in development, usually before its completion (but see below), while an aorist verb form presents it as completed. In a narration a sequence of aorist forms expresses a succession of actions without any overlapping, while an action expressed by a durative form may be interrupted by another action, as in (1)[1]:

(i)  *kâr     mikardam     ke     u     vâred     shod*
     work     I.was.doing     that     he     entering     became
     'I was working when he came in'.

---

[1] Abbreviations: CL clitic, EZ *ezâfe*, IMP imperative, NEG negation, OPT optative, PERF perfective, PN proper name, SG singular. In Pashto *R* and *D* are retroflexes, ë is a central vowel.

---

(3) Progressiveness: non-progressive vs. progressive. Ex: *miravad* 'goes, is going' ~ *dârad miravad* 'is going', *miraft* 'went, was going' ~ *dâsht miraft* 'was going' (on the notion of Progressive, see Cohen 1989).

(4) Resultativeness: simple vs. perfect. Ex: *miravad* 'goes, is going' ~ *rafte-ast* 'has gone', *miraft* 'went, was going' ~ *rafte bud* 'had gone'.

(5) Mode: indicative vs. subjunctive. Ex: *miravad* 'goes, is going' ~ *beravad* 'may go', *rafte-ast* 'has gone' ~ *rafte bâshad* 'may have gone'.

(6) Mediativity (or evidentiality): simple vs. mediative (or evidential). Ex: *raft* 'went' ~ *rafte(-ast)* 'went, as it appears, as it is said', *miraft* 'went, used to go, was going', ~ *mirafte(-ast)* 'went, used to go, was going, as it appears, as it is said'. In Persian there are two "registers" for the past, i.e., two ways of narrating past events. In Register I, they are stated simply, without any particular qualification. In Register II, they are affected by an abstract reference to the way in which they were known by the speaker, either by hearsay, or by observation of consequences (inference), or perhaps otherwise. Register I includes the forms *raft*, *miraft*, *dâsht miraft*, *rafte bud*; Register II includes the forms *rafte(-ast)*, *mirafte(-ast)*, *dâshte mirafte(-ast)*, *rafte bude(-ast)* (see Lazard 1985, 1996, 2000, 2006).

Time opposition combines with progressiveness and resultativeness. Does it combine also with durativeness? This point is problematic (see below).

Mediativity opposition combines with durativeness, progressiveness, and resultativeness.

Mode opposition combines only with resultativeness, and perhaps with durativeness (see below).

Those combinations are presented in Table 1. There are still two verb forms that I have not mentioned and have not included in the example: the imperative *berow* 'go!' and the future *xâhad raft* 'will go'. They are isolated and do not take part in any opposition.

**Table 1:** Combinations as per recommendation of series editor

|  | Aorist | Durative | Progressive | Resultative |
|---|---|---|---|---|
| Present | — | *miravad* | *dârad miravad* | *rafte-ast* |
| Past | *raft* | *miraft* | *dâsht miraft* | *rafte bud* |
| Mediative | *rafte(-ast)* | *mirafte(--ast)* | *dâshte mirafte(-ast)* | *rafte bude(-ast)* |
| Subjunctive | *beravad* | — | — | *rafte bâshad* |

Several points call for particular remarks:

(i)  The durativeness opposition is akin to an aspect opposition between completed action (aorist) and uncompleted action, which is very common in languages. It is also very common that there is no form expressing the completed action in the present, as in Persian. Although time (present vs. past) and aspect (uncompleted vs. completed) are quite different notions, for the first refers to the *speaker*'s time and the second to the internal time of the *action*, it is difficult to conceive of a completed action in the present, i.e., at the time of speaking. A completed action is naturally represented as a past event, even though a very recent one. In actual fact, in Persian a "past" form like *raft* is used for expressing the immediate past, as in (2):

(2)  *mâshin    âmad*
     car       came
     'Here is a car [coming]'.

Or even the very near future, as in (3):

(3)  *bejomb     mâ   raftim*
     move.IMP   we   went
     'Hurry up, we are going'.

Similar uses are known in other languages. Perhaps, instead of saying that there is no form for the completed action in the present, it would be better to say that the durativeness opposition is not combinable with the time opposition and to analyze forms like *raft* as pure 'aorists', including no time reference: their only content would be the idea of a compact indivisible event.

(ii)  The aorist may also refer to the future in subordination, as in (4) and (5); in these constructions, it alternates with the subjunctive:

(4)  *vaqti / agar  Parviz  âmad   / biâyad, xabar-am  kon*
     when/if     PN     comes    news-CL1SG       do.IMP
     'When / If Parviz comes, tell me'.

(5)  *shâyad    Parviz  âmad / biâyad*
     perhaps  PN      come
     'Perhaps Parviz will come'.

One may wonder whether in sentences like (4) and (5), there is a shade of meaning between the use of *âmad* (aorist) and that of *biâyad* (subjunctive). Should we

think that there is an aspect difference? If such is the case, we would have to make the subjunctive take part in the aspect (durativeness) opposition. The point remains to be investigated.

(iii)   Forms like *miraft*, which I have summarily described as presenting an action as uncompleted, more often than not does express events in progress or habitual. However, differently from "imperfects" in other languages, it may also express a durative completed action, as in (6):

(6)   *tamâm-e moddat-e    mosâferat sohbat   (mi)kard*
       all-EZ    duration-EZ  trip           talking  did
       'He talked/did not stop talking during the whole time of the trip'.

In (6) both *kard* and *mikard* may be used: the meaning remains nearly the same, but *mikard* emphasizes the duration. Both refer to an action that is undeniably completed. The content of the form *mikard* includes both the idea of a completed action and that of a durative one. Thus, on the level of the verb system, such forms as *miraft* should be characterized as able to express either an uncompleted action or a durative (completed) action. This situation can be diachronically explained by referring to the etymology of the prefix *mi-*, originally a particle (*hami*) meaning 'always, unceasingly'.

(iv)   The reader may be surprised to find *rafte-ast* in two different places in (1). I think that that form is indeed to be mentioned in two different places because it does fulfil two different functions in the verb system. On the one hand, it is part of the resultative opposition: it is a present resultative opposed to *miravad* as a resultative and to *rafte bud* as a present. On the other hand, it belongs to Register II of past narration, along with *mirafte(-ast)*, *dâshte mirafte(-ast)*, *rafte bude(-ast)*: it is a non-durative mediative opposed to *raft* as mediative and to *mirafte(-ast)* as non-durative, also opposed to *dâshte mirafte(-ast)* as non-progressive, and to *rafte bude(-ast)* as non-resultative.

# 3 Pashto

The oppositions that form the Pashto verb system are the following:

(i)   Time: Present vs. past. Ex. *taRi* 'ties' ~~ *taRë* 'was tied', *taRëlay dëy* 'has been tied' ~ *taRëlay wu* 'had been tied'. Past, perfect, and pluperfect forms are translated in English in the passive because Pashto is a language with so-called split ergativity, which means that with all verb forms referring to past time, the construction is ergative and the verb form agrees with the object.

(ii)  Aspect: Imperfective vs. perfective. Ex. *taRi* 'is tying' ~ *wëtaRi* 'ties, may tie', *taRë* 'was being tied' ~ *wëtaRë* 'was tied'. The relevant feature is completedness of the action: the perfective form expresses a completed action; for details, see below. Perfective aspect is marked either by prefix *wë-*, or by shifting stress position or by a change of stem, according to the verb class.

(iii)  Resultativeness: Simple vs. perfect. Ex. *taRi* 'is tying' ~ *taRëlay dëy* 'has been tied', *wëtaRi* 'ties / may tie' ~ *taRëlay wi* 'may have been tied'.

(iv)  Injunctiveness: Indicative vs. imperative. Ex. *taRi* 'is tying' ~ *taRa* 'tie!', *wëtaRi* 'may tie' ~ *wëtaRa* 'tie!' (definitely).

(v)  Actuality: Actual vs. unactual. Ex. *taRi* 'is tying' ~ *bë taRi* 'will tie', *wëtaRi* 'may tie' ~ *wë-bë-taRi* 'will tie' (completely).

Most oppositions combine with one another. Aspect opposition combines practically with all other oppositions: it is nearly pervasive. Time opposition combines with aspect, resultativeness, and actuality. Resultative opposition combines with aspect, time, and actuality. Particle *bë*, marker of unactuality, combines practically with any other opposition with various meanings (see below). Imperative combines only with aspect.

I have left out a few other more peripheral forms, optative and the potential phrase, and the hortative particle *di*.

A few remarks need to be made:

(i)  The semantic value of the aspect opposition is best perceivable in combination with the past and the imperative. The perfective past is used in narration: a sequence of such forms describes a succession of completed actions with no overlapping. The imperfective past may express an action in progress, a habitual action, even an intended but not initiated action, as in (7) (Vogel 1994: 123).

(7)  *mënDe me*          *kawële  xo  saxt     bârân sho  no*
     running me          did      But  heavy rain  became then
     *wë-me-në-swây*       *kRây*
     PERF-me-NEG-become.OPT  do.OPT
     'I was going to run, but it rained heavily, I could not do it'.

*Kawël*e is an imperfective past form: here it expresses an only intended action; in another context it might mean "I was running". An uncompleted past form may also refer to an action that took place in the past, but it does not specify whether it was completed or not. On the whole, the perfective past verb form indicates that the

action was indeed completed, while the imperfective past form does not exclude that the action was completed, but it does not imply anything about its completion.

In the imperative, with the perfective form the action is regarded as a compact whole whose completion is in any case to be reached. The imperfective form does not imply that the action must be continued until its full completion; it may leave some freedom to the addressee (Vogel 1994).

In the present, the aspect opposition is more ticklish. "The imperfective present forms refer to action going on at the present time or any action or state at the present time level" (Penzl 1955: 113). On the other hand, for reasons explained above (Section 2), perfective present forms can hardly express actual actions in the present. According to Penzl (1955: 114):

> "the perfective present forms occur in a great variety of syntactical situations. In main clauses they seem to express habitual and probable occurrence; they also occur in questions, and in commands or wishes [ ... They] occur most frequently in dependent (subordinate) clauses [...] It is obvious that in many of their occurrences [they] express a subjunctive mood rather than a perfective aspect".

(ii)  The particle *bë*, which I have labeled as a marker of unactuality, conveys several meanings. With present forms, imperfective or perfective, it expresses futurity. With past forms, imperfective or perfective, it expresses past habit. With the imperfective past it also expresses counterfactual in the apodosis of conditional sentences. With the perfective perfect, it gives a presumptive meaning. With pluperfect, it expresses past counterfactual in an apodosis or injects doubt into the statement.

# 4 Persian and Pashto

The original Iranian (and Indo-Iranian) verb system was similar to the Classical Greek one. The main oppositions were between a so-called present set of forms, with present (proper) and past, an aorist set, and a perfect set, with present and past, plus a series of modes in each set. The opposition between "present" and aorist was aspectual; the perfect forms were resultative. During the evolution of Iranian languages, those oppositions were lost, and new oppositions were created (for Persian, see Lazard 2003).

Aspect oppositions in Persian (durative vs. non-durative) and in Pashto (imperfective vs. perfective) are semantically akin to the original opposition between "present" and aorist. They are also akin to each other, but they are different in several respects:

(i) The opposition in Persian is between an aorist, which presents the action as a compact, indivisible block, and a durative, which expresses its internal duration. In Pashto it is between a perfective, which presents the action as a completed whole, including its endpoint, and an imperfective, which implies nothing concerning the completion of the action.

(ii) Persian aorist and Pashto perfective obviously resemble each other. However, the oppositions are the reverse of each other with respect to markedness. In Persian the marked term is the durative term, which positively indicates duration; in Pashto it is the perfective, which positively implies completion of the action.

(iii) The functional markedness in both languages is reflected in the morphology. In Persian, the simplest form of the whole system is the aorist (*raft*), while the durative forms bear the durative prefix *mi-*. In Pashto, the more simple forms are the imperfective ones, while the perfective has the prefix *wë-* or another conspicuous mark.

Theoreticians of aspect often consider that the main conceptual opposition in the domain of aspect proper is between the action in progress before its completion (phase 1) and the action taken as an indivisible whole, completion included (phase 2), with two minor concepts, namely, the progressive (action in progress at a reference point) and the perfect (resultative state) (see Cohen 1989, also Bertinetto 2006). These notions are at the basis of a number of aspectual systems with many variations. Persian and Pashto are no exceptions. The Pashto system seems to be more "regular", with its pervasive aspect opposition. The Persian one is more variegated, with its progressive and its series of mediative forms. It has probably reached a more advanced stage of evolution. It is likely that the Pashto perfective prefix *wë-* is etymologically identical with the Persian subjunctive marker *be-* and that the latter was formerly an aspect marker. We have seen that the Pashto perfective present forms often assume modal meanings. In the future evolution the Pashto system might possibly follow the same path as the Persian one and develop a modal opposition, while the aspect opposition would be devoid of present forms.

# References

Bertinetto, Pier Marco. 2006. Rigid vs. flexible (A)TAM systems. Paper presented at the symposium Current Topics in Typology, Berlin, 13–14 January.
Cohen, David. 1989. *L'aspect verbal*. Paris: PUF.

Lazard, Gilbert. 1985. L'inférentiel ou passé distancié en person. *Studia Iranica* 14. 27–42.

Lazard, Gilbert. 1996. Le médiatif en persan. In Z. Guentchéva (ed.), *L'énonciation médiatisée*, 21–30. Louvain & Paris: Peeters.

Lazard, Gilbert. 2000. Le médiatif: considérations théoriques et application à l'iranien. In L. Johanson and B. Utas (eds.), *Evidentials: Turkic, Iranian and neighbouring languages*, 209–228. Berlin and New York: Mouton de Gruyter.

Lazard, Gilbert. 2001. *Etudes de linguistique générale: Typologie grammaticale*. Leuven & Paris: Peeters.

Lazard, Gilbert. 2003. Aspect, temps, mode de procès. In J. Lentin and A. Lonnet (eds.), *Mélanges David Cohen*, 357–369. Paris: Maisonneuve et Larose.

Lazard, Gilbert. 2006. *Grammaire du persan contemporain*, 2 edn. Edited in collaboration with Y. Richard, R. Hechmati, and P. Samvelian. Tehran: IFRI (Institut Français de Recherche en Iran).

Penzl, Herbert. 1955. *A grammar of Pashto*. Washington, DC: American Council of Learned Societies.

Vogel, Sylvain. 1994. Oppositions aspectuelles et injonction en pashto. *Bulletin de la Société de linguistique de Paris* 89 (1). 121–159.

Donald Stilo

# Loss vs. expansion of gender in Tatic languages: Kafteji (Kabatei) and Kelāsi

## 1  Introduction

In this chapter, I document a case in which two neighboring, closely related dialects, even though mostly mutually intelligible and in regular socioeconomic contact and interdependence, can still diverge from each other in striking ways. While areal phenomena and contact are important in such situations leading to convergence, *divergence* may also be taking place at the same time due to reasons internal to the specific dialects as well as to external factors that cause them to drift apart, e.g., areal phenomena with different isoglottic patternings encompassing one dialect but not the other; the retention of archaic features in one of these dialects as a factor of a more peripheral location; sociolinguistic attitudes; intermarriage patterns; and language shift, to name a few.

Below I will present an extreme and somewhat unique case of just such divergence on one issue: gender marking. Kafteji and Kelāsi[1] are two neighboring Tatic[2]

---

**1** I have used the traditional transcription symbol *ā* in names (Kelās, Jeyshābād, Māsule, Māsāl, and others) in Section 1 of this chapter only for identification purposes on traditional maps. In articles for a linguist readership, I generally avoid using the symbols *a* and *ā*, traditionally used in Iranian studies, respectively, for (a) a low front vowel *(æ)* and (b) a low back slightly rounded vowel *(ɒ)*, as they mislead the reader into thinking there is a true distinction of length in Persian and (at least) some other languages of Iran. In the six-vowel system of these languages, *ā* is the only vowel in the traditional transcription that shows a supposed length distinction. Not only is this a typologically strange situation, i.e., only one long vowel in the system, but in fact is really only used out of lack of a more acceptable way of distinguishing æ and ɒ. In this chapter, and in my writing in general, I simply prefer to use æ and a for these two vowels, respectively.
**2** The term "Tatic" (Stilo 1981) refers to both Tati languages and all forms of Talyshi, two very closely related groups. The Tatic group also includes two other groups, Rudbāri and Tāleqāni, which I consider to be "Tatoid", that is, languages of the Tatic family that, under the influence of other groups with which they are in heavy contact, have lost all the characteristic morphology of Tatic languages (including gender) and have retained only the (more or less) original Tatic lexical composition.

---

dialects of the Central Tati group (Stilo 1981) that have taken the original two-term masculine/feminine gender distinction of Tatic languages – still quite robust in many members of the group but totally lost to about 40% of the Tati languages and all Talyshi and Tatoid dialects – and have each pushed the pattern in completely opposite extremes. That is, Kafteji has expanded the marking of gender beyond the grammatical domains where gender agreement is usually encoded in these languages and has extended it to additional loci in the language. Kelāsi, its immediate neighbor, on the other hand, having gone in the opposite direction on the issue of gender, has lost all traces of an original masculine–feminine distinction throughout its grammatical system.

## 1.1 Background

Kelāsi and Kafteji, both undocumented until now, are the easternmost members of the Tāromi subgroup (Yarshater 1970) of Tati languages and, due to the drawing of administrative borders, are located just outside the Upper Tārom district of Zanjan province and are instead included within the Rudbār district of Gilan province. The official Persian names for these villages are Kælas (Kalās) and Kæbæte (Kabate), but the native names Kelas/Kelās and Kæfti/Kafti and their names of the inhabitants and their language varieties in both dialects, Kelāsi and Kafteji, are used throughout this chapter.

Kelās and Kafti are located about twelve kilometers from each other in a sparsely populated area with no other villages in the intervening area between them. To the west of the area is the Upper Tārom district with most of the Tāromi dialect cluster consisting of another eleven subdialects (Charazei, Jeyshābādi, Hezār-rudi, Bākoluri, Jamālābādi, Bandargāhi, Siāvarudi, Nowkiāni, Gowjāni, Gandomābi, and Kalāsari; see Map 1 in Section 4.1). The closest genealogical relatives of Kelāsi-Kafteji are the dialects of Jeyshābād (*šava* in Kelāsi) and Charaze, both of which are geographically closer to Kelās than to Kafti. The populations of Jeyshābād and Charaze in ca. 1965 were 141 and 165 inhabitants respectively (Yarshater 1970: 454). Other villages to the west, southwest, and south on this expanse of mountainous plateau are Azerbaijani-speaking.

To the north we encounter the southern limits of Talyshi (also spelled Ţāleshi/Taleshi in Iranian sources), an independent branch of Tatic and the language with the largest number of speakers of any Tatic group, although the southern, central, and northern varieties of Talyshi, due to significant structural and lexical differences and the resulting wide gaps in interintelligibility, should really be considered independent languages (Stilo 2015). Kelās, Kafti, and Upper Tārom are located on the arid southern slopes of the Talesh mountains (the western

extension of the Alborz, west of the Sefidrud and the Rudbār Valley) facing the Qezel Owzān River and the Iranian plateau, in contrast to Māsule and Māsāl, the closest Talyshi-speaking neighbors to Kelāsi-Kafteji, located on the other side of mountain passes on the verdant northern slopes of the mountains facing the Caspian littoral. The summer camps (*yeylaq*) of Kelās and Kafti and those of the latter two Talyshi groups, albeit on opposite sides of the local Talesh mountain passes, are even closer to each other – probably around a half-day's trek.

To the east of Kafti at a driving distance of some twenty-four kilometers in the Rudbār Valley below at a steep drop from the plateau, we find the city of Rudbār and the Rudbāri language group.[3] By contrast to the high plateau, the Rudbār Valley – due to the presence of one of Iran's major rivers (the Sefidrud), its relatively low elevation, its proximity to the Caspian littoral, and lack of physical barriers from it – has a quite humid, mild, subtropical climate known for its rice paddies and olive groves. The deceptively close distance of twenty-four kilometers between Rudbār and Kafti represents *in real time*, at least when I visited it in 1976, a trip of about three and a half hours by jeep due to the precipitous ascent via a dangerously narrow road – most likely an original donkey trail – lacking guard rails and filled with numerous hairpin turns that are too sharp even for a small jeep to make in one turn. Thus, in practical terms, Kafti and the Rudbār Valley are quite removed from each other and it would be difficult to estimate how intense their contact was before the existence of motorized transport.

My language consultants, Mr. Mo'men (Behruz) Montazeri for Kafteji and Mr. Mostafā Rastegār for Kelāsi, informed me that their respective villages even as late as the late 1960s or early 1970s consisted of about 350 inhabitants each, some fifty to sixty families. With the economic upheavals in Iran of the early and mid-1970s, however, almost all the villagers left for larger cities and industrial centers. In the summer of 1976, I found only a handful of people left in Kafti, mostly elderly or under school-age children, and I was obliged to leave to seek speakers elsewhere. My fieldwork sessions in each dialect were conducted in the city of Rasht for Kafteji or, for Kelāsi, in the "Industrial City of Qazvin", a conglomeration of various factories outside Qazvin city with residences and shops for the workers and their families. Given that the populations of these two villages were so small to start with and that most of the population had already migrated to large urban centers in the mid-1970s, the present status of the dialects and the prognosis for their future survival are not very positive – as is most likely also the case for the whole Tāromi dialect cluster for similar reasons.

---

**3** Rudbāri is of a completely different type from Tāromi or other typical Tati languages and has been tentatively classified as "Tatoid" (Stilo 1981).

## 1.2 Kelāsi, Kafteji, and the linguistic environment

These two Tāromi-type Central Tati dialects are closely related and are completely mutually intelligible. When people from these two villages interact, as I was told but did not observe, they each speak their own respective dialects and have no difficulty in understanding each other. It is my guess that, although their dialects are completely intelligible to one another, speakers may in fact compensate by adapting their speech to each other's in conversation. On this front, however, I do not know how they handle the issue of gender, probably the most significant and all-pervasive difference between the two dialects.

# 2 Points of Kafteji and Kelāsi grammar

Now I would like to present a treatment of what I was able to elicit on the points of grammatical gender in Kafteji. I will start by fleshing out, one at a time, the domains of the grammar in which gender plays a role, thus requiring a slightly broader grammatical description than just gender issues, and will also contrast them with the equivalent genderless Kelāsi forms as I progress.

## 2.1 Kafteji and Kelāsi noun morphology

Kafteji nouns formally distinguish two genders (masculine and feminine), two cases (direct and oblique), and two numbers (singular and plural). Gender is distinguished only in the singular; the plural has common forms for both genders. Thus, as Table 1 shows, Type I Kafteji nouns have two case forms for masculine singular (including -Ø for the masculine direct form), two for feminine singular, and two for common plural. Note that the distinction between the masculine and feminine oblique forms here is one of stress. In addition, the masculine singular oblique and the common gender plural direct forms are the same in Kafteji. Kelāsi, which on the other hand has lost the distinction in gender, has only two forms in the singular, direct and oblique, and since it has also lost the distinctions of case in the plural, it only has one common form for plurals in all functions. Hence Kelāsi has a total of three unique case markers for nouns as compared to the five of Kafteji.

A special feature of gender marking in Kafteji is that a feminine noun in citation form may drop the overt feminine direct case marker, the unstressed -*æ*, leaving the noun as its bare stem. Thus *kelékæ* 'girl' appears as *kelek* in isolation. The same holds true for feminine nouns used in non-specific senses (see generic:

**Table 1:** Noun morphology, Type I.

|  |  | Kafteji | | Kelāsi |
|---|---|---|---|---|
|  |  | Masculine | Feminine | (no gender distinction) |
|  |  | 'man' | 'girl' | 'man, girl' |
| SG | DIR | *mærdǽk-Ø* | *kelék-æ* | *mærdǽk-Ø, kelék-Ø* |
|  | OBL | *mærdǽk-ə* | *kelek-ə́* | *mærdǽk-e, kelék-e* |
|  |  | Plural | | Plural |
| PL | DIR | *mærdǽk-ə, kelék-ə* | | *mærdæk-ón, kelek-ón* |
|  | OBL | *mærdæk-ón, kelek-ón* | | |

[9] 'drink water!'; indefinite: [14] 'I bought an apple' below). This *-æ* morpheme is not to be interpreted as a definiteness marker as we can see in (16b), where an indefinite noun is modified by an adjective, but retains the feminine direct case marker. My notes show so few examples, however, that it is difficult to tell whether the retention of this marker in this pattern is optional or obligatory. Conversely, demonstrative adjectives, while they render the noun definite, do not necessarily trigger a feminine direct case marker on the head noun:

(1)  Without feminine direct case marker
    *æm  bəz  čiki=a?*        (alternate: *bə́z-æ*)
    this  goat  who.OBL=3SG.F  goat-F.DIR
    'Whose nanny-goat is this?'

    With feminine direct case marker
    *æm  pærvejín-æ  čəmən  nᵉ-áya.*
    this  sieve-F.DIR  mine   NEG.BE-3SG.F
    'This sieve is not mine'.

In addition, both Kafteji and Kelāsi have a second noun type, restricted to some eight to ten kinship terms, with a different oblique singular formation in *-(æ)r* for both genders in Kafteji and the equivalent nouns in genderless Kelāsi, as shown in Table 2. Type I Kafteji gender distinctions are still retained in the direct case forms of Type II nouns.

**Table 2:** Noun forms, Type II (both dialects) (Closed set of some eight to ten kinship terms).

|  | Kafteji (MASC) | | Kafteji (FEM) | | | Kelāsi (no gender distinction) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| DIR | *bera-Ø* | *pe-Ø* | *xóv-æ* | *máy-æ* | *dét-æ* | *bera-Ø* | *pe-Ø* | *xo-Ø* | *ma-Ø* | *det-Ø* |
| OBL | *bera-r* | *pe-r* | *xo-r* | *ma-r* | *det-ǽr* | *bera-r* | *pe-r* | *xo-r* | *ma-r* | *det-ǽr* |
|  | 'brother' | 'father' | 'sister' | 'mother' | 'daughter' | 'brother' | 'father' | 'sister' | 'mother' | 'daughter' |

Note the following examples of the oblique case of Kafteji Type I masculine and feminine nouns compared with similar, but gender-undifferentiated, forms in Kelāsi:

(2) a.  Kafteji
Possessive use of the oblique case:

| *æ* | *mærdǽk-ə* | *nom* | *čiči=æ?* | (alt: *či=æ*) |
|---|---|---|---|---|
| that | man-M.OBL | name | what=3SG.M | |

'What is that man's name?'

Kelāsi

| *æ* | *mærdǽk-e* | *nom* | *či=æ?* |
|---|---|---|---|
| that | man-S.OBL | name | what=3SG |

'What is that man's name?'

b.  Kafteji

| *æ* | *zænæk-ə́* | *nom* | *čiči=æ?* | (alt: *či=æ*) |
|---|---|---|---|---|
| that | woman-F.OBL | name | what=3SG.M | |

'What is that woman's name?'

Kelāsi

| *æ* | *zænǽk-e* | *nom* | *či=æ?* |
|---|---|---|---|
| that | woman-S.OBL | name | what=3SG |

'What is that woman's name?'

Oblique case governed by a postposition:

(3) a.  Kafteji          Kelāsi

| *patíl-ə* | *delæ* | *livón-e* | *delæ* |
|---|---|---|---|
| pot-M.OBL | in | glass-S.OBL | in |

'in the pot (M)'    'in the glass'

b.  Kafteji          Kelāsi

| *av-ə́* | *delæ* | *áv-e* | *delæ* |
|---|---|---|---|
| water-F.OBL | in | water-S.OBL | in |

'in the water (F)'    'in the water'

## 2.2  Tense and the oblique case marking of verbal arguments

In Kafteji-Kelāsi, as well as most Western Iranian languages, the morphology of verbal paradigms is split into two sets of tenses,[4] the present and the past systems,

---

**4** "Tenses" is simply shorthand here for the "tense-aspect-mood" forms of the verbal paradigms.

depending on which tenses are formed from the present stem or the past stem of the verb, respectively (see also note 6). These two stems affect the morphological formation of all the tense-aspect-mood paradigms of transitive verbs and – by extension from the morphological definition of intransitive and transitive (not the semantic definition) – also trigger a tense-based split of the case marking of core arguments of the clause. In Kafteji-Kelāsi, the present system includes the simple present, present progressive, subjunctive, and imperative. The tenses of the past system include the simple past (or preterit), imperfect (or durative past/conditional), past progressive, present perfect, past perfect (or pluperfect), and subjunctive perfect. See "Notes on the tense names and usages" below for a discussion of the usages of the imperfect, the progressives, as well as the subjunctive perfect.

In the system of present tenses where we generally find accusative alignment for core arguments, the direct case is used for agents (or subjects of intransitive verbs in all tenses) and the oblique is only used for definite objects. Since both dialects have differential object marking (DOM), non-specific indefinite or generic nouns in all tenses remain unmarked, i.e., in the direct case, in the two alignment systems. In the system of past tenses, both dialects (but especially Kafteji) use the oblique case to mark both agent and definite object of transitive verbs, resulting in double oblique alignment. Further along in the process of losing features of an original ergative alignment, both dialects also sometimes use the direct case for noun agents in the past system. In Kelāsi, and to a lesser extent in Kafteji, the direct and the oblique cases of pronouns are often interchangeable for the subjects of intransitive verbs, as in the Kelāsi example (44b), and copulas, as in Kelāsi examples (39a), (39b), (43), and (46b).

The following are examples of Kafteji masculine and feminine forms of various oblique-marked core arguments, accompanied by the genderless Kelāsi equivalents (where available).

Agents of past transitives marked by the oblique case:

(4)  a.  Kafteji

    *æ*    *kot-é*[5]    *či(=š)*    *vat?*

    that  boy-M.OBL  what(=3SG₂)  said

    'What did that boy say?'

   b.  *æ*    *kelek-ə́*    *či(=š)*    *vat?*

    that  girl-F.OBL  what(=3SG₂)  said

    'What did that girl say?'

---

**5** For a discussion of a masculine stressed *-é* oblique here, see note 8.

Definite objects marked by the oblique case, present system:

(5) a.  Kafteji                                Kelāsi
        *árd-ə*        *á-xšuj!*           *bǽr-e*        *á-kæ!*
        flour-M.OBL    PVB-knead           door-S.OBL     PVB-do
        'Knead the flour (M) (into dough)!'  'Open the door!'

    b.  *av-ə́*        *á-xa!*              *áv-e*         *dé-kæ!*
        water-F.OBL    PVB-eat             water-S.OBL    PVB-do
        'Drink the water (F)!'              'Pour the water!'

Definite objects marked by the oblique case, past system:

(6) a.  *mən*    *æ*     *kot-é*        *vénd=əm.*
        I.OBL    that    man-M.OBL      saw=1SG$_2$
        'I saw that boy.' (1SG$_2$: Set$_2$ person agreement markers [PAMs] are
        defined in note 6)

    b.  *mən*    *æ*     *kelek-ə́*      *vénd=əm.*
        I.OBL    that    girl-F.OBL     saw=1SG$_2$
        'I saw that girl'.

As mentioned, however, both dialects have differential object marking and thus
for non-specific objects, the direct case of the noun is used. Contrast the follow-
ing with (5a) and (5b):

(7)  Kafteji                       Kelāsi
     *gel*    *á-xšuj!*            *av*    *ú-kæš*    *ča=der!*
     clay     PVB-knead           water   PVB-pull   well-from
     'Knead (some) clay (M)!'      'Draw (some) water from the well!'

When an indefinite or generic noun in Kafteji is feminine, the feminine direct
marker is generally dropped as discussed in Section 2.1. Note the differential
object marking in the following two contrasting examples with a feminine noun
used first in a specific sense and then in a generic sense. Note that (8)–(14) are
only Kafteji examples.

(8)  *av-ə́*        *á-xa!*
     water-F.OBL    PVB-eat
     'Drink the water (F)!'

(9)  *av      á-xa!        < áv(-æ)*  'water' (F)
     water   PVB-eat       water-F.DIR
     'Drink water!'

Additional Kafteji feminine nouns in non-specific forms (minus the direct case marker):

(10) *av      a-m-xar-əm.*
     water   PVB-DU-eat-1SG₁
     'I drink water'.

(11) *mən    æsif    me-xar-əm.     < æsíf(-æ)* 'apple (F)'
     I.OBL   apple   DU-eat-1SG₁        apple-F.DIR
     'I eat apples/I'll eat an apple'.

The next three Kafteji sentences demonstrate equivalent alternate forms with the mobile Set₂ person clitics (see 2.5 and note 6) either (a) attaching to the object, (b) optionally deleted, or (c) with the optionally retained full pronoun agent as well. (Sentences [12] and [13] will also have all three alternatives, but the actual forms don't appear in my field notes.) In any case, since the objects are feminine and non-specific, the feminine direct marker is deleted:

    Set₂ PAMs only              Full pronoun/No PAM     Full pronoun + PAM
(12) *av      ča=ræ=m        u-kæšə́st*
     water   well=from=1SG₂   PVB-drew
     'I drew water from the well'.

(13) *av=ešon    a-nə́-xa       ~ jon      av      a-nə́-xa.*
     water=3PL₂  PVB-NEG-ate     they.OBL  water   PVB-NEG-ate
     'They didn't drink water'.

(14) *æsif=əm    a-gæ      ~ mən  æsif  a-gæ.     ~ mən  æsíf=əm    a-gæ*
     apple=1SG₂  PVB-took   I.OBL apple PVB-took   I.OBL apple=1SG₂  PVB-took
     'I got/bought an apple'.

In the case of the experiencer verb *go/gost* 'want', the stimulus argument ('the apple') is always in the direct case, but when this stimulus noun is non-specific and feminine in this pattern, the feminine direct marker is also omitted (Set₂ person agreement markers are used for experiencer verbs in all tenses):

(15)  Kafteji

*mən  æsíf=əm  me-go.*  < *æsíf(æ)* (F)
I.OBL  apple=1SG₂  DU-want
'I want an apple'.

Kelāsi

*mən  æsíf=əm  me-goy.*
I.OBL  apple=1s₂  DU-want
'I want an apple'.

## 2.3 Gender and adjectival agreement in Kafteji

Within the Noun Phrase (NP) the gender of the head noun triggers gender agreement in both attributive and predicative adjectives in Kafteji. Attributive adjectives precede their head nouns and are joined to them in both dialects by an NP-dependent-marked linking morpheme *-æ*, which is also found in noun–noun compounds, e.g., Kelāsi (adjective + noun) *gæ̀rm-æ áv* 'hot water' (hot-LINK water); (noun + noun) *yùz-æ púst* 'walnut shell' (walnut-LINK skin), *æsìf-æ dár* 'apple tree' (apple-LINK tree). If an adjective in both dialects ends in a stressed *-æ* as part of its root, e.g., *pillǽ* 'large, big', the *-æ* connector morpheme then merges with the final vowel of the adjective, which is the masculine form of the adjective (or the genderless form in Kelāsi), as in (16), i.e., *pillǽ + -æ* (root + connector) yields *pillǽ*. Feminine forms of adjectives in Kafteji are marked by an unstressed feminine *-a*, after which the connector morpheme *-æ* is suppressed via sandhi rules. In the case of adjectives already ending in a stressed *-æ* as part of its root, e.g., *pillǽ*, the *-a* feminine adjectival marker also suppresses this vowel of the root, i.e., both *pillǽ + -a + -æ* (root + feminine + linker, as in (16b), attributive adjective) as well as *pillǽ + -a* (root + feminine, as in (17b), predicative adjective) yield *pill-á*. Note also that the stress now shifts from the suppressed final vowel of the root to the usually unstressed feminine *-a* marker.

Adjectival agreement

(16)  a.  Kafteji

*i  pillǽ  sævzæ-Ø  dir-əm.*  (< *pillǽ-æ*)
one  big  broom-M.DIR  have-1s₁
'I have a big broom (M)'.

Kelāsi

*i  pillǽ  bera-Ø  dar-əm.*  (< *pillǽ-æ*)
one  big  brother-S.DIR  have-1SG₁
'I have an older brother (lit: big)'.

b. Kafteji

$i$     *pill-á*    *čél-æ*          *dir-əm.*    (< *pillǽ-à-æ*)
one   big-FA   spindle-F.DIR   have-1s$_1$
'I have a big spindle (F)'.

Kelāsi

$i$     *pillǽ*   *xo-Ø*         *dar-əm.*
one   big    sister-S.DIR   have-1SG$_1$
'I have an older sister (lit: big)'.

The gender of the head noun also controls gender agreement in the predicative adjective in Kafteji:

(17) a. Kafteji

*æm*      *sævzæ-Ø*     *pillǽ*   *ní-æ.*
this     broom-M.DIR   big    NEG.BE-3SG.M
'This broom (M) is not big'.

Kelāsi

*æ*    *mærdæk-Ø*   *pir=æ.*
that   man-SG.DIR   old=3SG
'That man is old'.

b. Kafteji

*æm* *čél-æ*          *pill-á*   *n$^{e}$-á(ya).*   (< *pillǽ-à*)
this   spindle-F.DIR   big-FA   NEG.BE-3SG.F
'This spindle (F) is not big'.

Kelāsi

*æ*    *zænæk-Ø*     *pir=æ.*
that   woman-S.DIR   old=3s
'That woman is old'.

## 2.4 Demonstrative adjectives and pronouns

As we see in (2a), (2b), (4a), (4b), (6a), (6b), (17a), (17b), (24a), (24b), (42a), and (42b), demonstrative adjectives do not distinguish gender in Kafteji. Demonstratives, however, are also used independently in both dialects as the only type of third person pronouns, and as pronouns they do distinguish gender in Kafteji. Both proximal and distal demonstratives are commonly used in this function:

(proximal) Kafteji: *æm* (masculine), *ǽm-æ* (feminine), 'this one' or 'he' and 'she' respectively; Kelāsi *æm* 'this one, he, she'; (distal) Kafteji: *æ* (masculine), *a* (feminine) 'that one' or 'he' and 'she' respectively, Kelāsi *æ* 'that one, he, she'.

Demonstrative pronouns, direct case:

(18)  a.  Kafteji

| *æm-Ø* | *æmǽd-ə* | *z ᵉæ-Ø* | *ní-æ.* |
|---|---|---|---|
| this-M.DIR | PN-M.OBL | son-M.DIR | NEG.BE-3SG.M |

Kelāsi

| *æm* | *æmǽd-e* | *ziæ-Ø* | *ní-æ.* |
|---|---|---|---|
| this | PN-S.OBL | son-S.DIR | NEG.BE-3SG$_1$ |

'This (or 'he') is not Ahmad's son'.

b.  Kafteji

| *ǽm-æ* | *æmǽd-ə* | *dét-æ* | *n ᵉ-áya.* |
|---|---|---|---|
| this-F.DIR | PN-M.OBL | daughter-F.DIR | NEG.BE-3SG.F |

Kelāsi

| *æm* | *æmǽd-e* | *dét-Ø* | *ní-æ.* |
|---|---|---|---|
| this | PN-S.OBL | daughter-S.DIR | NEG.BE-3SG$_1$ |

'This (or 'she') is not Ahmad's daughter'.

In (50a) and (50b), we also see the distal forms functioning as pronouns.

Kafteji demonstrative pronouns only distinguish gender in the Direct case; the Oblique forms are of common gender for both proximal and distal 3rd SG; see also (23a and 23b):

Oblique case of demonstrative pronouns (common gender) as objects:

(19)  Kafteji

| *mən* | *ja* | *bə-zzə=m* |
|---|---|---|
| I.OBL | that.OBL | PU-hit.PAST=1S$_2$ |

'I hit him/her'.

Kelāsi

| *tæ* | *ja* | *Ø-væǽnd=i?* |
|---|---|---|
| you | that.OBL | PU-saw=2SG$_2$ |

'Did you see him/her?'
(*tæ* 'you' [SG], DIR/OBL)

## 2.5 Gender agreement in the verb phrase

The most numerous loci for encoding gender agreement with nouns in Kafteji grammar are found within the verb. We have already seen this process with the negative forms of the copula in examples (17)–(18b). In addition to the copula,

**Table 3:** Present system paradigms (both dialects)

|  | Kafteji | | | Kelāsi | | |
|---|---|---|---|---|---|---|
|  | 'to go' | 'to fall' | 'to say' | 'to go' | 'to fall' | 'to say' |
| Present/Future | | | | | | |
| 1s | meším | megænám | amvajám | meším | megeném | amvajém |
| 2s | mešíš | megæní | amvají | meší | megení | amvají |
| 3sm | mešíæ | megæná | amvajǽ } | mešé | megené | amvajé |
| 3sf | meš$^e$ǽ | megæn$^e$ǽ | amvaj$^e$ǽ | | | |
| Subjunctive | | | | | | |
| 1s | béšim | bégænem | vajám | béšim | bégenem | vajém |
| 2s | béšiš | bégæni | vají | béši | bégeni | vají |
| 3sm | béšiæ | bégæne | vajǽ } | béše | bégene | vajé |
| 3sf | béš$^e$æ | bégæn$^e$æ | vaj$^e$ǽ | | | |

all verbs in the third person singular in Kafteji show the distinction of gender in all tenses except in the past system of transitive verbs. In the latter cases, Set$_1$, or direct person agreement markers (PAMs), are not used and in place of them we find the genderless Set$_2$ person markers,[6] i.e., the optional, leftwardly mobile (or "floating") oblique enclitics used to cross-reference both agents in the past system and experiencers in all tenses. As gender is not encoded in any form in the past system of transitive verbs, these verb forms have been omitted from further discussion in this chapter – but see (4a), (4b), (6a), (6'b), (12) to (14), (19), (31a), and (31b) as incidental examples. The Set$_1$ agreement markers of past intransitive verbs are bound suffixes that are neither optional nor mobile.

I will begin the discussion of agreement marking in the verb by examining the present system. Table 3 lists the present paradigms and contrasts the 3rd SG

---

**6** Tatic languages, as with most Northwest Iranian languages, have a distinction between intransitive and transitive conjugations in the past tenses. The distinction centers around two different sets of PAMs, which I will call Set$_1$ or direct and Set$_2$ or oblique. They generally parallel the tense-based split in agent/subject marking in the alignment structures: nominative (direct cases of overt subjects/agents and Set$_1$ PAMs) in the present system and ergative (oblique cases of overt agents and Set$_2$ PAMs) in the past system. The nominative alignment and Set$_1$ PAMs are used for the present system of tenses of all verbs and for past system of intransitive verbs. The Set$_2$ PAMs, as mentioned, are leftwardly mobile. Kafteji and Kelāsi allow neither the indexing of objects (which in other conservative Tati dialects are indexed by Set$_2$ in the present and Set$_1$ in the past) in the verb nor the use of the oblique markers (Set$_2$) as pronominal possessive enclitics as we find in many Tatic languages and, indeed, various other Northwest Iranian languages. Please note that, as seen in various examples here – (6a), (6b), (10)–(12), (14)–(16b), (19), etc. – the first singular of both Set$_1$ and Set$_2$ are coincidentally identical. The third person shows a clearer distinction, e.g., Set$_1$: (20)–(26b) vs. Set$_2$: (4), (25a) and (25b) for Kafteji. Example (25b) has both types.

in Kafteji with the equivalent genderless Kelāsi paradigms. (Plural forms have been omitted from the various tables and the rest of the chapter since they do not distinguish gender.)

Gender agreement in third singular, present system; intransitive:

Present tense, intransitive

(20) a. Kafteji
*aynæ    me-šgy-é.*
mirror   DU-break.INCH-3SG.M
'The mirror will break'.

Kelāsi
*æm   mærdæk   pir=a    né-m-vi-Ø.*
this   man       old=CS   NEG-DU-become-3SG$_1$
'This man doesn't grow old'.

b. Kafteji
*čél-æ          me-šgy-ǽ.*
spindle-F.DIR   DU-break.INCH-3SG.F
'The spindle will break'.

Kelāsi
*æm   zænæk   pir=a    né-m-vi-Ø.*
this   woman   old=CS   NEG-DU-become-3SG$_1$
'This woman doesn't grow old'.

(21) a. Kafteji
*čəm   bera     kærdæ   me-ší-æ       kelas.*
my     brother   PROG    DU-go-3SG.M   PN
'My brother is going to Kelas'.

b. *čəm   xóv-æ          kærdæ   me-š$^{\varepsilon}$-ǽ       kelas.*
my     sister-F.DIR   PROG    DU-go-3SG.F   PN
'My sister is going to Kelas'.

Subjunctive, intransitive:

(22) a. Kafteji
*čəm   z$^{\varepsilon}$æ   mo-go    bó-xos-e.*
my     son   DU-must   PU-sleep-3SG.M
'My son has to go to sleep'.

Kelāsi

*ægæ    ara       b-á-yæ,* ...
if      tomorrow   PU-come-3SG₁

'If he/she comes (here) tomorrow, ...'

b.  Kafteji

*čəm  dét-æ          mo-go   bó-xos-ᵉæ.*
my    daughter-F.DIR  DU-must  PU-sleep-3SG.F

'My daughter has to go to sleep'.

Kelāsi

*monken-e    narahæt   bí-vi-æ.*
possible-is  upset     PU-become-3SG₁

'He/she may get upset'.

(23)  a.  Kafteji

*ja          vaj-en   b-íšt-e.*
that.OBL    tell-CM   PU-stand-3SG.M

'Tell (SG) him to wait!'

Kelāsi

*æ      mo-go    kiæ    b-íšt-e.*
that   DU-must   house  PU-stand-3SG₁

'He/she must stay at home'.

b.  Kafteji

*ja          vaj-en   b-íšt-ᵉæ.*
that.OBL    tell-CM   PU-stand-3SG.F

'Tell (SG) her to wait!'

Present tense; transitive:

(24)  a.  Kafteji

*æ     mærdæk-Ø  ayn-é[7]       me-šgen-é.*
that  man-M.DIR  mirror-M.OBL  DU-break.TR-3SG.M

'That man will break the mirror'.

---

**7** Type I masculine nouns whose roots end in a stressed *-ǽ*, as *aynǽ* 'mirror', *kotǽ* 'boy', and *tævilǽ* 'stable' – see also (4), (6), (49a), and (49b) – take the usual masculine oblique in *-ə*, but the two vowels merge as *-é*. Since the final suppressed vowel of the root was stressed, the oblique marker now assumes this stress.

Kelasi

*æ      hizem         me-xruš-ə́.*
that   firewood      DU-sell-3SG₁

'He/she sells firewood'.

b.   Kafteji

*æ    zænǽk-æ       ayn-é          me-šgen-ᵉǽ.*
that  woman-F.DIR   mirror-M.OBL   DU-break.TR-3SG.F

'That woman will break the mirror'.

Subjunctive, transitive:

(25)   a.   Kafteji

*či    me-go=š          Ø-vaj-ə?*
what  DU-want-3SG₂   PU-say-3SG.M

'What does he want to say?'

Kelāsi

*m-oste          zud-tær      Ø-vaj-ə́.*
DU-want.PST   soon.CMP   PU-say-3SG₁

'He/she should have said (it) sooner'.

b.   Kafteji

*či       me-go=š           Ø-vaj-ᵉǽ?*
what   DU-want-3SG₂   PU-say-3SG.F

'What does she want to say?'

(26)   a.   Kafteji

*ægæ   Ø-vin-ə́,        čə      bə́-kær-əm?*
If       PU-see-3SG.M,   what   PU-do-1SG₁

'If he sees (it), what should I do?'

b.   *ægæ   Ø-vin-ᵉǽ,       čə      bə́-kær-əm?*
If       PU-see-3SG.F,   what   PU-do-1SG₁

'If she sees (it), what should I do?'

# 3   Systemic expansion of gender: The verbs 'to be'; past intransitives

Kafteji has increased the loci of gender cross-referencing in the verbal system by extending to areas in which, although not unknown elsewhere, gender is

not generally encoded in Western Iranian languages: it has spread beyond the usual third person to the first and second person singulars as well. We see this systemic expansion into the full singular paradigm in two areas of the Kafteji verb: (1) all tenses of the past system of intransitive verbs (but not the present tenses), and (2) the verbs "to be" in all tenses,[8] including the forms of the present system.

An important additional issue regarding the verbs "to be" and the encoding of gender is that the types of be-verbs themselves have expanded from the usual two of Western Iranian – usually a simple copula and a verb of existence – to a typologically unusual *five* verbs of being. Gender is expressed throughout the entire singular paradigm of these verbs, but Be$_5$, which is only used with inanimate nouns, lacks first and second singular forms altogether.

## 3.1  Gender in first and second persons in all tenses of intransitive past system

As mentioned, Kafteji cross-references the gender of the subject in the verb in all tenses of the intransitive past system. All of these tenses – although not enough data are available on the perfect tenses to make a secure statement – have masculine and feminine forms for all three persons of the singular. The singular paradigms of the past system tenses attested in my field notes are listed in Table 4 for Kafteji and are contrasted with genderless Kelāsi.

*Notes on the tense names and usages.* The imperfect has various usages in both dialects: (a) past habitual; (b) progressive aspect when preceded by the invariable particle – *kærdæ* for Kafteji, *kǽræ/kærǽ* for Kelāsi (also used with the simple present to form the present progressive in both dialects, as in (21a) and (21b); (c) the verbs in both the protasis and the apodosis of irrealis conditional clauses.

The subjunctive perfect has several usages where we find the subjunctive equivalent of the present perfect after: (a) a modal "(I ought) *to have gone*"; (b) a main verb or other expressions that require the subjunctive "(let it be that, I hope that, God forbid that) *he has gone*"; (c) an implied, but usually unverbalized, expression whose sense is loosely translated "[so what if] *I have gone?* [what's it to you that] *I have gone?*"

Masculine and feminine forms in the past system:

---

**8**  I assume that gender is distinguished in all singular forms of all tenses of "be" since the present and past forms are attested, but my field notes don't show any examples of masculine/feminine contrasts other than third person singular forms for the subjunctive or the perfect paradigms.

**Table 4:** Past system paradigms (both dialects).

| Kafteji | | | | Kelāsi | |
|---|---|---|---|---|---|
| | 'to go' | 'to fall' | | 'to go' | 'to fall' |
| **Simple past (Preterit)** | | | | | |
| 1SM | bešém | bekǽtim | } | bešém | bekǽtim |
| F | bešᵉám | bekǽtam | | | |
| 2SM | bešéš | bekǽtiš | } | bešé | bekǽti |
| F | bešᵉáš | bekǽtaš | | | |
| 3SM | bešé | bekǽt | } | bešǽ | bekǽt-Ø |
| F | bešᵉǽ | bekǽtæ | | | |
| **Imperfect** | | 'to run' | | | 'to run' |
| 1SM | mešém | medævástim | } | mešém | medævéstem |
| F | mešᵉám | medævástam | | | |
| 2SM | mešéš | medævástiš | } | mešéy | medævésti |
| F | mešᵉáš | medævástaš | | | |
| 3SM | mešé | medævást | } | mešǽ | medævést-Ø |
| F | mešᵉǽ | medævást-æ | | | |
| **Present perfects ('to go')** | | | | | 'to fall' |
| 1SM | bešém | | | | bəkætém |
| F | ?? | | | | |
| 2SM | bešéš | | | | bəkæté |
| F | ?? | | | | |
| 3SM | bešé(yæ) | | | | bəkæté |
| F | bešiá | | | | |
| **Past perfects ('to go')** | | | | 'to go' | |
| 1SM | bešiǽ=bem | | | blšiǽ=vem | |
| F | ?? | | | | |
| 2SM | bešiǽ=beš | | | blšiǽ=ve | |
| F | ?? | | | | |
| 3SM | bešiǽ=bə | | | blšiǽ=və | |
| F | bešiǽ=bᵉæ | | | | |
| **Subjunctive perfect** | | | | | |
| | 'to go' | | | | |
| 1SM | bešᵉǽ-bi-m | | | | |
| F | ?? | | | | |
| 2SM | bešᵉǽ-bi-š | | | | |
| F | ?? | | | | |
| 3SM | bešᵉǽ-bi-(æ) | | | | |
| F | bešᵉǽ-bᵉǽ | | | | |

Simple past

(27)  a.  Kafteji
    *æmǽd-ə    de-škæsést-im.\**
    PN-M.OBL  PVB-looked-1SG.M
    'I (M) looked at Ahmad'.

    Kelāsi
    *de-kǽt-im    čal-e        delæ.*
    PVB-fell-1SG  pit-S.OBL  inside
    'I (M, F) fell into the pit'.

   b.  *æmǽd-ə    de-škæsést-am.\**
    PN-M.OBL  PVB-looked-1SG.F
    'I (F) looked at Ahmad'. \*The verb *de-škæsést-* is treated
    morphologically as intransitive.

(28)  a.  Kafteji
    *tæ    kovra    be-šé-š?*
    you  where   PU-went-2SG.M
    'Where did you (M) go?'

   b.  *tæ    kovra    be-š$^{\varepsilon}$-áš?*
    you  where   PU-went-2SG.F
    'Where did you (F) go?'

(29)  a.  Kafteji
    *æ    mærdæk-Ø  b-omé-Ø         ǽya.*
    that  man-M.DIR  PU-came-3SG.M  here
    'That man came here'.

    Kelāsi
    *i       kotæ  b-omǽ-Ø        æya.*
    one  boy   PU -came-3SG$_1$  here
    'A boy came here'.

   b.  Kafteji
    *æ    zænǽk-æ    b-om$^{\varepsilon}$-ǽ         ǽya.*
    that  woman-F.DIR  PU-came-3SG.F  here
    'That woman came here'.

    Kelāsi
    *i      kelek  b-omǽ-Ø        æya.*
    one  girl   PU -came-3SGG$_1$  here
    'A girl came here'.

(30)  a.  Kafteji
*čəm   zebon-Ø        be-vrin-ᵋ-ǽst-Ø.*
my    tongue-M.DIR   PU-cut-INCH-PT-3SG.M
'My tongue got cut'.

Kelāsi
*štæ   pe    ko    be-šé-Ø?*
your  father  where  PU-went-3SG₁
'Where did your father go?'

b.  Kafteji
*čəm   ængíšt-æ       be-vrin-ᵋ-ǽst-æ.*
my    finger-F.DIR   PU-cut-INCH-PT-3SG.F
'My finger got cut'.

Kelāsi
*štæ   ma    ko    be-šé-Ø?*
your  mother  where  PU-went-3SG₁
'Where did your mother go?'

Imperfect (progressive):
Kafteji:

(31)  a.  *æzíræ       ge    kærdæ  kᵉæ   me-šé-m,      čəm*
yesterday  SUB   PROG   house  DU-went-1SG.M  my
*bera-r          vénd=əm.*
brother-M.OBL₂  saw=1SG₂
'Yesterday when I (M) was going home, I saw my brother'.

b.  *æzíræ       ge    kærdæ  kᵉæ   me-šᵋ-ám,     čəm*
yesterday  SUB   PROG   house  DU-went-1SG.F  my
*bera-r          vénd=əm.*
brother-M.OBL₂  saw=1SG₂
'Yesterday when I (F) was going home, I saw my brother'.

Kelāsi:

(32)  *kǽræ   me-šé-m        baq.*
PROG   DU-went-1SG₁  garden
'I (male/female) was going to the garden'.

(33) *æzíræ     ko      kǽræ   me-šé-y?*
    yesterday   where   PROG   DU-went-2SG₁
    'Where were you (male/female) going?'

(34) *kǽræ   narahæt   me-v-ə.*
    PROG   upset      DU-became-3SG₁
    'He/she was getting upset'.

While not enough information is known about the perfect tenses in Kafteji to state whether gender is encoded in the first or second persons, there are contrasting pairs of masculine and feminine third person singular forms for the present perfect – (35a) and (35b) – and the subjunctive perfect – (36a), and (36b):

(35) a. Kafteji
    *šta    bera     kovra   be-š-é(yæ)?*
    your   brother   where   PU-went-3SG.M
    'Where has your brother gone?'

    Kelāsi
    *šta    bera     ko      be-š-é?*
    your   brother   where   PU-went-3SG.M
    'Where has your brother gone?'

   b. Kafteji
    *šta    xóv-æ       kovra   be-ši-á?*
    your   sister-F.DIR   where   PU-went-3SG.F
    'Where has your sister gone?'

(36) a. Kafteji
    *šayæd     be-š^e-ǽ=bi-æ.*
    perhaps   PU-went-PPL=AUX-3SG.M
    'Perhaps he has gone'. /
    'He may have gone'.

    Kelāsi
    *šayæd     naxoš=a   via=vi-Ø.*
    perhaps   sick=CS    became=AUX-3SG₁
    'Perhaps he/she has gotten sick'. /
    'He/she may have gotten sick'.

   b. Kafteji
    *šayæd     be-š^e-ǽ=b-^ɛæ.*
    perhaps   PU-went-PPL=AUX-3SG.F
    'Perhaps she has gone'.

My Kafteji consultant claimed that a feminine form is sometimes distinguished for the preterit third plural as well: *beš^e ándæ*. He offered this form without my prompting, but elicitation in other areas of grammar that coincidentally also had past tense third person plural verbs did not indicate any use of a feminine plural – (37) and (38) are only Kafteji:

(37) æ       kelék-ə    kovra   be-š-énde?     kelék-ə      b-énde?
     that   girl-P.DIR   where   PU-went-3PL₁    girl-PL.DIR   was-3PL₁
     'Where did those girls go?'                 'Was it the girls?/Were they
                                                  the girls?'

(38) kot-é      b-á-nde,       kelék-ə      hurǽ   b-íšt-ende!
     boy-PL.DIR   PU-come-3PL₁   girl-PL.DIR   there   PU-stand-3PL₁
     'The boys come (in) (and) the girls stay there!' (command forms)

Since gender is not distinguished in the noun plurals, there has been no motivation so far to extend gender agreement to the plural forms of verbs, but note that the speaker himself offered examples for plural feminine forms. If Kafteji were to continue along its natural course, it might in fact start normalizing such third plural forms and possibly even later extend them into the first and second plural as well. Such a development would probably also entail the innovation of a nominal feminine plural – not an impossibility, considering that Delijāni, a Central Plateau dialect not far from Vafs, does in fact distinguish between masculine and feminine nominal plurals (direct case). My guess is that this is a local innovation rather than an archaism.

## 3.2 Gender in first and second singular persons in all forms of "to be"

The other major domain where gender has been extended in Kafteji is in the singular forms of all tenses of the verbs "to be". As mentioned above, there are five different verbs of being in Kafteji-Kelāsi, each used in very specific functions. This proliferation of be-verb types is typical of other languages of the immediate area as well: Eastern and Western Gilaki (Stilo 2001), Southern Talyshi, and Rudbari, which also have the same types of distinctions with almost identical details of how and when each of the verbs is used. The five verbs have the base forms and general usages listed below. It should be noted that the morphology of types 3–5 is based on Be₁ and thus reflects the gender morphology of Be₁. Be₃ and Be₅ are both participial formations (of "stand" and "put" respectively). Explicit examples of each of the five verbs

are given in the following sections. Table 5 lists only the known Kafteji forms; the Kelāsi forms are generally identical to, or very close to, the Kafteji masculine forms.

1. *=e ~ =æ* (NEG: *ní-*), a general, neutral copula, enclitic in the affirmative: "Hasan is a doctor", "the house is white". This verb uses a different root for the negative.

2. *hǽst-* (NEG = Be₁), used only in the present affirmative, has two primary functions: (1) it is generally used in those places where the enclitic copula (Be₁) cannot occur: (a) independently, i.e., not enclitically, in the context of conversation: "Is he?" or (b) when stressed: "Hasan **is** a doctor"; and (c) used synonymously in place of the copula, but with no special stress for emphasis or focus; (2) it is used for the expression of existence ("there is, there are"). Since this verb has no special forms for other tenses or the negatives, the respective forms of Be₁ are used instead.

3. *išt-* (NEG: *n-išt-*) 'to stand, wait (see [23]), stay (see [38])', from which a participial formation is used to indicate "be" of location exclusively with animate subjects: "Hasan is at home; the cow is here".

4. *dér-e* (NEG: *de-ní-*) indicates containment/location-within or existence-within, mostly with inanimate subjects ("the tea is in the teapot", "there is tea in the teapot") or animate, nonhumans ("the horse is in the stable"). Be₄ may also occur with human subjects, but the rules and full distribution for this usage are beyond the scope of the present study.

5. *nehé* (NEG: *néne*), a participial formation of "to put, place", functions as a be-verb. Since *nehé* is used exclusively with inanimate subjects, it occurs only in the third person. It encodes the sense of existence ("Is there any rice to eat?") and sometimes location.

**Table 5:** The Kafteji verbs of being.

| | Be₁ usual copula | Be₂ emphatic copula | Be₃ human location | Be₄ nonhuman location inside | Be₅ inanimate existence and location |
|---|---|---|---|---|---|
| **Present** | | | | | |
| 1SM | *=im* | *hǽstim* | *ištém* | *der-im* | |
| F | *=am* | *hǽstam* | *išt ͤam* | ? | |
| 2SM | *=iš* | *hǽstíš* | *ištéš* | *der-iš* | |
| F | *=aš* | *hǽstaš* | *išt ͤaš* | ? | |
| 3SM | *=e ~ =æ* | *hǽst-e ~ hǽst-æ* | *išté ~ ištǽ* | *dér-e ~ -æ* | *nehé ~ nehéæ* |
| F | *=a* | *hǽsta* | *išt ͤá(ya)* | *dér-aya* | *neh ͤ-áya* |
| **Past** | Be₁ | Be₂ | Be₃ | Be₄ | Be₅ |
| 1SM | *bem* | (see also note 8) | *išt ͤǽ=bim* | | |

**Table 5:** (continued)

| F | $b^e am$ | ? | | |
|---|---|---|---|---|
| 2SM | $beš$ | $išt^e ǽ=biš$ | | |
| F | $b^e aš$ | ? | | |
| 3SM | $be$ | $išt^e ǽ=be$ | $dé=be$ | $nehæ=be$ |
| F | $b^e æ$ | $išt^e ǽ=b^e æ$ | $dé=b^e æ$ | ? |

Be$_1$: The general copula:

(39)  a.  Kafteji
    *æz=im.*
    I.DIR=1SG.M
    'It's me (M).'

    Kelāsi
    *mən  štæ=jen  pildæ-tær=im*
    I.OBL  your=from  big-CMP=1SG$_1$
    'I (M, F) am older than you'.

    b.  Kafteji
    *æz=am.*
    I.DIR=1SG.F
    'It's me (F).'

    Kelāsi
    *mən ~ æz  kelasi=me.*
    I.OBL ~ I.DIR  Kelasi=1SG$_1$
    'I (M, F) am a Kelasi'.
    (i.e., 'I am from Kelas')

(40)  a.  Kafteji
    *tæ=iš?*
    you.DIR=2SG.M
    'Is it you (M)?' ~ 'Is that you (M)?'

    Kelāsi
    *ara  kiē? (< kiæ=i)*
    tomorrow  house=2SG$_1$
    'Are you (M, F) home tomorrow?'

    b.  *tæ=yaš?*
    you.DIR=2SG.F
    'Is it you (F)?' ~ 'Is that you (F)?'

(41)  a.  Kafteji
          *æ          hoštén=e.*
          that.DIR    self=3SG.M
          'It's him himself'.

          Kelāsi
          *čem  bera     tehran=æ.*
          my    brother  PN=3SG₁
          'My brother is in Tehran'.

      b.  Kafteji
          *a          hoštén=a.*
          that.DIR    self=3SG.F
          'It's her herself'.

(42)  a.  Kafteji
          *æm   kæl        čiki=æ?*
          this  male.goat  who.OBL=3SG.M
          'Whose billy-goat is this?'

          Kelāsi
          *æm   gav   ški=æ.*
          this  cow   who.OBL=3SG₁
          'Whose cow is this?' (cf. Kafteji:
          *gávæ*, F)

      b.  Kafteji
          *æm   báz-æ                čiki=a?*
          this  female.goat-F.DIR    who.OBL=3SG.F
          'Whose nanny-goat is this?'

See also (2), (17a), (17b), (18a), and (18b) for other examples of the copula (Be₁).
Be₂: Emphatic/independent form of copula; existence:
Be₂ may be used as an emphatic equivalent of the copula (Be₁), in which case the first syllable of the verb takes the primary stress for the whole sentence.

(43)  a.  Kafteji                        Kelāsi
          *æz    hoštén **hǽst-im.***    *mən/æz       kelasi **hǽst-ime.***
          I.DIR  self   be₂-1SG.M        I.OBL/I.DIR   kelasi be₂-1SG₁
          'It ***is*** me (M) myself'.    'I (M, F) ***am*** a Kelasi'.

b. Kafteji               Kelāsi

    *æz*     *hoštén*   **hǽst-am.**    **hǽst-ime!**

    I.DIR   self    be$_2$-1SG.F   be$_2$-1SG$_1$

    'It **is** me (F) myself'.   'I (M, F) **am**!'

Be$_2$ is also used for the expression of existence in both dialects. When the first or second persons are used in the sense of existence, the meaning is generally not "do you exist?" (although it can rarely have that sense), but rather implies that there is a cooperative effort, such as a game, a hunt, etc., and is used with the meaning of "I am up for it" (cf. Turkish *ben varım*), "Are you in?", that is, "Are you going to join us in this project?" In this usage, the existence verb occurs most commonly either alone or at most with the pronoun subject:

(44)  a.  Kafteji (Masculine)     Kafteji (Feminine)

         *tæ*      *hǽst-iš?*    *tæ*      *hǽst-aš?*

         you.DIR  be$_2$-2SG.M   you.DIR  be$_2$-2SG.F

         'Are you (M) in?'      'Are you (F) in?'

         (i.e., Are you going to play? Are you up for this task? Are you coming along?)

    b.  Kelāsi (Gender-neutral)

         *ara*       *mən*  *kæræ*  *i*   *yaga*  *me-ši-m.*   *te-y*    *hǽst-i?*

         tomorrow  I.OBL  PROG  one  place  DU-go-1SG$_1$  you-also  be$_2$-2SG$_1$

         'I am going to go somewhere tomorrow. Are you (M, F) coming along?'

Unfortunately, my only examples of this usage of Be$_2$ in the third person, either animate or inanimate, are in Kelāsi and thus don't distinguish gender:

(45)  Kelāsi (Gender-neutral)

    *pəla*  *hǽst-e*     *bé-xar-om?*    *æhæn.*  *hǽst-e.*

    rice   be$_2$-3SG.M  PU-eat-1SG$_1$   yes    be$_2$-3SG$_1$

    'Is there rice (for me) to eat?'   'Yes, there is'.

Be$_3$: Location, human subjects:

Be$_3$ is used to indicate being or existing in a location and is used exclusively with animate (±human) subjects:

(46)  a.  Kafteji

         *mǽ-tærs!*  *æz*     *išté=m.*

         NEG-fear  I.DIR  be$_3$=1SG.M

         'Don't be afraid! I (M) am (here)'.

Kelāsi
*mǽ-tærs! æz æyæ išté=m.*
NEG-fear I.DIR here be$_3$=1SG$_1$
'Don't be afraid! I (M, F) am here'.

b. *mǽ-tærs! æz išt$^\varepsilon$=ám.*
NEG-fear I.DIR be$_3$=1SG.F
'Don't be afraid! I (F) am (here)'.

Kelāsi
*mən ara kiæ išté=m.*
I.OBL tomorrow house be$_3$=1SG$_1$
'I (M, F) will be home tomorrow'.

(47) a. Kafteji
*tæ išté=š?*
I.DIR be$_3$=2SG.M
'Are you (M) (around)?'; 'Will you (M) be (at home)?' (according to context)

Kelāsi
*ara koræ išté:? (< išté=i)*
tomorrow where be$_3$=2SG$_1$
'Where will you (M, F) be tomorrow?'

b. Kafteji
*tæ išt$^\varepsilon$=áš? (< išté=aš)*
you.DIR be$_3$=2SG.F
'Are you (F) (around)?'; 'Will you (F)
be (at home)?' (according to context)

Kelāsi
*ara kiæ išté:? (< išté=i)*
tomorrow house be$_3$:2SG$_1$ be$_3$-2SG$_1$
'Will you (M, F) be home tomorrow?'

(48) a. Kafteji
*štæ bera-Ø kovra išté(æ)-Ø.*
your brother-M.DIR where be$_3$-3SG.M
'Where is your brother?'

Kelāsi
*ja xord-æ bera-Ø ištÍ-Ø.*
his small- CONN brother-S.DIR be$_3$-3SG$_1$
'His/her younger brother is (here)'.

b. Kafteji

| *štæ* | *xóv-æ* | *kovra* | *išt^{ɛ}=á(ya)* |
|---|---|---|---|
| your | sister-F.DIR | where | be₃=3SG.F |

'Where is your sister?'

Kelāsi

| *ja* | *xord-æ* | *xo-Ø* | *ištÍ-Ø.* |
|---|---|---|---|
| his | small- CONN | sister-S.DIR | be₃-3SG₁ |

'His/her younger sister is (here)'.

(49) a. Kafteji

| *æsp-Ø* | *tævil-é* | *delæ* | *išté(æ)-Ø.* |
|---|---|---|---|
| horse-M.DIR | stable-M.OBL | inside | be₃-3SG.M |

'The horse (M) is in the stable'.

Kelāsi

| *æsp-Ø* | *tævilæ-Ø* | *ištÍ-Ø.* |
|---|---|---|
| horse-S.DIR | stable-S.DIR | be₃-3SG₁ |

'The horse is in the stable'.

b. Kafteji

| *madyón-æ* | *tævil-é* | *delæ* | *išt^{ɛ}=áya* |
|---|---|---|---|
| mare-F.DIR | stable-M.OBL | inside | be₃=3S.F |

'The mare (F) is in the stable'.

As the Kelāsi examples above show, e.g., (49), since Be₃ is used as a verb of location, an overt adposition of location is not always necessary. The noun either appears in its base (direct case) form, as in (47b) and (49), or in the oblique case accompanied by the locative postposition. The same is also true of Kafteji, as in (49a), (49b), and (53)–(54b) with postpositions and for (50) and (51b) without postpositions.

Past tense:

(50) a. Kafteji

| *æ* | *æzíræ* | *k^{ɛ}æ* | *ištæ=be-Ø.* |
|---|---|---|---|
| that.M.DIR | yesterday | house | be₃=AUX.PST-3SG.M |

'He was home yesterday'.

Kelāsi

| *æ* | *æzíræ* | *kiæ* | *ištiǽ=f-Ø.* |
|---|---|---|---|
| that.S.DIR | yesterday | house | be₃=AUX.PST-3SG₁ |

'He/she was home yesterday'.

b. Kafteji

> *a*      *æzíræ*      *kᶜæ*      *ištæ=b-ᶜæ.*
> that.F.DIR yesterday house be₃=AUX.PST-3SG.F
> 'She was home yesterday'.

(51) Kafteji

> *æzíræ*      *gáv-æ*      *tævilæ*    *n-íštæ=b-ᶜæ*
> yesterday cow-F.DIR stable NEG-be₃=AUX.PST-3SG.F
> 'The cow wasn't in the stable yesterday'.

Verb Be₄ may also be used synonymously for this function if the subject is a non-human animate, but my only examples of the latter are in Kelāsi (see below).

Be₄: Location-within/containment~existence-within, nonhuman subjects:

Be₄ is used for location-within or containment, mostly with non-human (±animate) subjects. As mentioned, Be₄ with human subjects is possible but highly restricted; hence, no examples of human subjects with Be₄, including first and second person pronouns, were elicited in my fieldwork.

(52) a. Kafteji

> *čai-Ø*      *ɣuri*      *delæ*      *dér=e.*
> tea-M.DIR teapot inside be₄=3SG.M
> 'The tea is in the teapot' ~
> 'There is tea in the teapot'.
> (Partial synonym in existence sense: be₅)

> Kelāsi

> *čai-Ø*      *ɣuri*      *delæ*      *dér=e.*
> tea-S.DIR teapot inside be₄=3SG₁
> 'The tea is in the teapot' ~
> 'There is tea in the teapot'.
> (Partial synonym in existence sense: be₅)

b. Kafteji

> *čai-Ø*      *ɣuri*      *delæ*      *de-ní-æ.*
> tea-M.DIR teapot inside be₄-NEG.BE-3SG.M
> 'The tea isn't in the teapot' ~
> 'There isn't any tea in the teapot'.
> (Partial synonym in existence sense: be₅)

> Kelāsi

> *čai-Ø*      *ɣuri*      *delæ*      *di-ní-Ø.*
> tea-S.DIR teapot inside be₄-NEG.BE-3SG₁

'The tea isn't in the teapot' ~
'There isn't any tea in the teapot'.
(Partial synonym in existence sense: be$_5$)

(53)  a.  Kafteji

> *sævzæ-Ø    k-é           delæ    dér=e.*
> broom-M.DIR  house-M.OBL  inside  be$_4$=3SG.M
> 'The broom (M) is in the house'. ~
> 'There is a broom (M) in the house'.
> (Partial synonym in existence sense: be$_5$)

> Kelāsi

> *čai-Ø     yuri   delæ   dér=e     á-xar-om?*
> tea-S.DIR  teapot  Inside  be$_4$=3SG$_1$  PVB-drink-1SG$_1$
> 'Is there tea in the teapot (for me) to
> drink?'
> (Partial synonym in existence sense: be$_5$)

    b.  Kafteji

> *sævzæ-Ø    k-é           delæ    de-ní-æ.*
> broom-M.DIR  house-M.OBL  inside  be$_4$-NEG.BE-3SG.M
> 'The broom (M) isn't in the house'. ~
> 'There isn't any broom (M) in the house'.
> (Partial synonym in existence sense: be$_5$)

> Kelāsi

> *næ.   di-ní=æ*
> no   be$_4$-NEG.BE=3SG$_1$
> 'No, there isn't'. (in answer to above)
> (Partial synonym in existence sense: be$_5$,
> cf. [49])

(54)  a.  Kafteji

> *čél-æ        k-é           delæ    der=áya.*
> spindle-F.DIR  house-M.OBL  inside  be$_4$=3SG.F
> 'The spindle (F) is in the house'.
> 'There is a spindle (F) in the house'.
> (Partial synonym in existence sense: be$_5$)

> Kelāsi

> *æsp-Ø        tævilæ-Ø     dér=e.*
> horse-S.DIR  stable-S.DIR  be$_4$=3SG$_1$

'The horse is in the stable'.
(Partial synonym in existence sense: be$_3$)

    b.  Kafteji
       *čél-æ*         *k-é*           *delæ*   *de-né-ya.*
       spindle-F.DIR  house-M.OBL  inside  be$_4$-NEG.BE-3SG.F
       'The spindle (F) isn't in the house'. ~
       'There isn't any spindle (F) in the house'.
       (Partial synonym in existence sense: be$_5$)

Be$_5$: Existence, inanimate subjects, third person only (no Kelāsi examples in my field notes):

(55)  Kafteji (Masculine)
     *sævzæ-Ø*     *av-ə́*       *delæ*    *nehé(æ)-Ø.*
     broom-M.DIR  water-F.OBL  inside  be$_5$-3SG.M
     'There is a broom (M) in the water'.
     (Partial synonym in existence sense: be$_4$)

     Kafteji (Feminine)
     *čél-æ*         *av-ə́*       *delæ*    *neh$^ɛ$-áya.*
     spindle-F.DIR   water-F.OBL  inside  be$_5$-3SG.F
     'There is a spindle (F) in the water'.
     (Partial synonym in existence sense: be$_4$)

(56)  Kafteji (Masculine)
     *sævzæ-Ø*     *av-ə́*       *delæ*    *né-ne(æ)-Ø.*
     broom-M.DIR  water-F.OBL  inside  be$_5$-NEG.BE-3SG.M
     'There isn't a broom (M) in the water'.
     (Partial synonym in existence sense: be$_4$)

     Kafteji (Feminine)
     *čél-æ*         *av-ə́*       *delæ*    *né-n$^ɛ$-aya.*
     spindle-F.DIR  water-F.OBL  inside  be$_5$-NEG.BE-3SG.F
     'There isn't a spindle (F) in the water'.
     (Partial synonym in existence sense: be$_4$)

(57)  Kafteji (Masculine)
     *pəla*       *nehé(æ)-Ø*  *bé-xar-əm?*
     rice-M.DIR  be$_5$-3SG.M    PU-eat-1SG$_1$
     'Is there rice (for me) to eat?'

Be$_5$: Location, inanimate subjects:

(58)  Kafteji (Masculine)

*štæ    æbru-Ø          kovra    nehé(æ)-Ø?*
your   eyebrow-M.DIR   where    be₅-3SG.M
'Where is your eyebrow (M)?'

Kafteji (Feminine)

*štæ    ængíšt-æ        kovra    neh^ᵉ-áya?*
your   finger-F.DIR    where    be₅-3SG.F
'Where is your finger (F)?'

# 4 Tatic, Northwest Iranian, and the fate of grammatical gender in Kafteji-Kelāsi

We have seen that on the issue of grammatical gender, Kafteji has gone in the opposite direction from Kelāsi, its closest neighbor and closest genealogical relative. While the latter has lost all traces of gender morphology in the noun and its cross-referencing with the noun elsewhere in the Noun Phrase or clause, the former is not only conservative in *retaining* gender, but has in fact *expanded* it into new grammatical domains in which most other Northwest Iranian (henceforth NWI) languages that have retained gender do not show such distinctions. Since so few NWI languages, including Tati, have the types of extensions discussed here, we can assume them to be later innovations.

Why have such opposite trends taken place in such closely related dialects from very small neighboring villages? As mentioned above, there have been different internal and external processes affecting them.

## 4.1 Factors for Loss of Gender: Kelāsi

Gender was clearly not a crucial distinction for communicative purposes, given that at least half of the Tatic family (and indeed many other Iranian languages) have lost gender distinctions throughout their grammatical systems. Many Indo-European languages – cf. English, Armenian, Persian, and others – have also lost the category of gender in nouns and the encoding of the nominal gender in other loci in the clause. This may serve as a type of *internal* motivation. One important *external* motivation, however, was probably a much larger factor acting upon the grammatical system of Kelāsi: areal phenomena. Kelāsi is within a zone with other languages that have completely eliminated gender from their grammatical

systems. Map 1 shows the area covered by the isogloss for gender retention within the Tatic family. In addition, but not included on the map, there are certain areas of western Iran where we find additional groups of Northwest Iranian languages with gender retention. On a more complete map, the gender isogloss would continue:

(i)     southeastward to the Central Plateau Group not far from the Vafs area without much of a geographic break from the isogloss shown on Map 1. Even within Central Plateau, however, only the dialects of the northwestern zone of this group, i.e., those geographically nearest to Vafsi and Southern Tati, have retained gender (see Amorei to Abuzeydabadi in Table 6 ); the vast majority of the Central Plateau dialects to the south, east, and southeast have also lost the distinction of gender;

(ii)    eastward to Semnani (but not to the other languages of the Semnan area).

Notes to Map 1:
1.    Small squares represent only points where Tatic languages are spoken. Not all Tatic-speaking locales are shown, especially outside the isogloss line, e.g., Northern Tati, most of the Talyshi-speaking areas, and most of densely populated Tāleqān.
2.    Squares inside circles represent the five points within Tati (Kafteji and four Southern Tati villages) where gender loci are extended to first and second persons in the verb.
3.    Talyshi-speaking points on the map are underlined.
4.    The village of Dikin has two distinct languages of different groups: (1) a main Tati type distinguishing gender, and (2) a Tāleqāni Tatoid type lacking gender. Oskulak also has two different ethnic groups speaking different genderless languages: (1) the southernmost (known) point of Talyshi and (2) the northernmost point of Rudbari.

(iii)   to the southwest to Gorani (e.g., Hawrāmi), but with some geographic discontinuity (possibly due mostly to the intrusion of Turkic languages into the area);

(iv)    after a larger geographic break, we also find gender distinctions in Northern Kurdish;

(v)     within Northern Kurdish-speaking territory, farther still from this Tatic gender isogloss, we also have gender in Zazaki/Dimli of eastern Turkey.

In sum, six of the eight groups of NWI languages (seven major groups and one minor group, the Semnan Sprachbund area) – that is, the majority of NWI groups – have retained gender distinctions. However, these figures regarding the conservative retention of gender in NWI are deceptive since, of the six groups
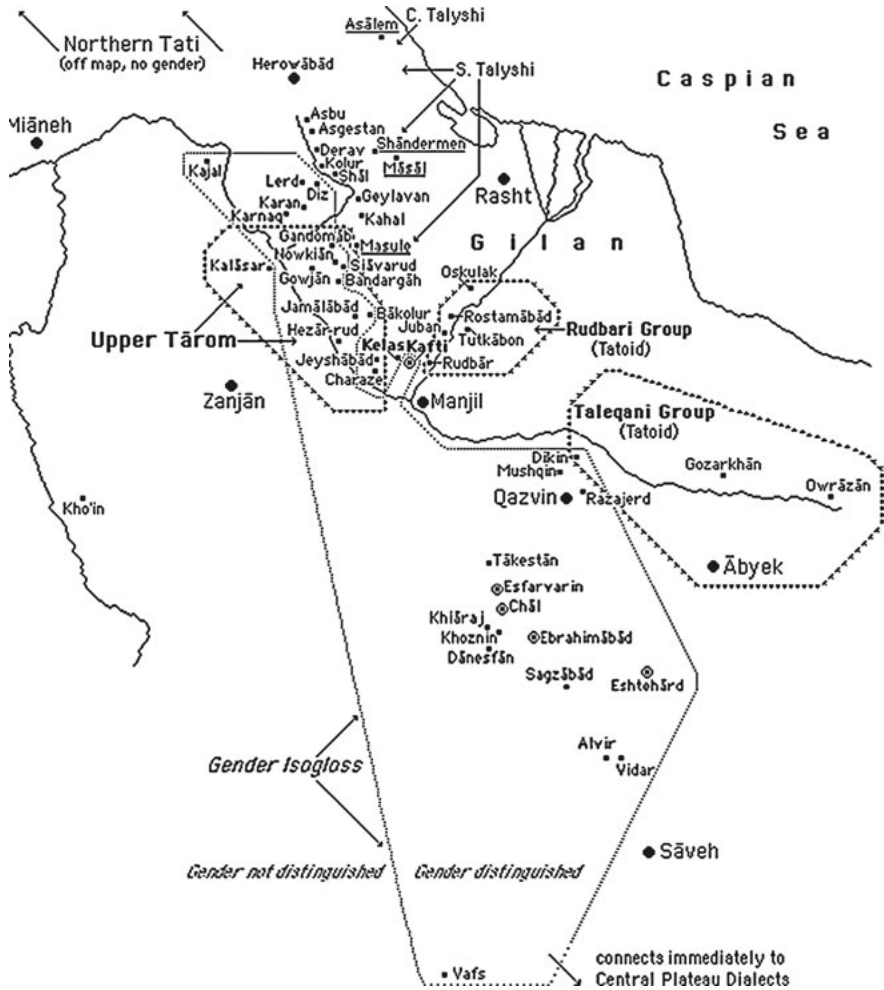
**Map 1:** Gender in the Tatic language zone.

still distinguishing gender in the noun, several of these groups really only make the distinction in about half, or fewer, of the languages and dialects within each group. Within the NWI groups that still retain gender, it has been lost at least in certain members of the groups: (1) at least half of the Tatic family (including all Talyshi and all Tatoid dialects), (2) most of the Central Plateau Dialect group, (3) Central and Southern Kurdish, (4) in most of the languages of the Semnan Sprachbund: (retained) Semnāni; (lost) Sangesari, Sorxei, Lāsgerdi, Biyābunaki, Aftari, and Shamirzādi (Caspian family). The whole Caspian family and all forms of Balochi are the only NWI groups that have completely lost all traces of gender.

In addition to a sporadic geographic distribution of gender retention, gender does not always have the same prominence in the grammar or the same distribution within the grammatical domains in each of these different languages and dialects. In some cases, even within the same group, the gender of the noun is cross-referenced in different loci within the Noun Phrase and/or the Verb Phrase. Table 6 shows the various domains where the gender of the noun is cross-referenced and their distribution in various NWI languages.

## 4.2 The crucial factor?

A key factor that may help account for the diachronic loss of gender in Kelāsi is in fact hidden in one point of Kafteji grammar. As mentioned above, overt gender marking in Kafteji feminine nouns is dropped in isolation, e.g., in citation forms and when used generically (but unmodified). Thus, bare feminine nouns at first hearing sound like masculine (or Kelāsi genderless) nouns. In my initial collection of lexical items gathered simply as a list to help me identify the dialect affiliation – and thus all provided by the consultant in citation forms – most Kafteji feminine nouns were given only with the bare root, devoid of the feminine direct case marker, e.g., *kelek* 'girl', *zænæk* 'woman', *xælæk* 'aunt (mother's sister)', *veyvæk* 'bride', *bəz* 'goat', *gav* 'cow', *av* 'water', and others. Other feminine words were offered in both forms: *æŋgišt ~ æŋgíštæ* 'finger', *kærg ~ kǽrgæ* 'chicken, hen', *šilanæk ~ šilanǽkæ* 'apricot', whereas only two, *máyæ* 'mother', *xóvæ* 'sister', both close kinship terms, were given in their full forms. Deleting the direct feminine ending is thus not obligatory in citation forms, but is usual. Several of those words later appeared in full sentences and only then in their full forms.

As shown in examples (9)–(15), Kafteji feminine nouns used in a generic ("Drink water!") or indefinite ("I bought an apple") sense (and occasionally as a definite – see example [1]) also use the bare noun root without the direct feminine marker. Hezār-rudi, a Tāromi dialect close to Kafteji both genealogically and geographically (see Map 1), also drops the direct feminine suffix when used in a generic sense (both examples from Yarshater 1970: 456):

(59) Hezar-rudi: *bez    a-gi!*      but: *æ   béz-æ      mi-ši-æ*
               goat  PVB=take      that goat-F.DIR  DU-go-3SG.F
               'Buy goats!'      'That goat (F) is going'.

But why should the direct feminine morpheme have become optional in the first place? Two plausible reasons immediately leap to mind, both having to do with stress patterns: (1) vowels (often short) following the major locus of stress

in a word, especially word-final vowels, are often reduced and then lost cross-linguistically; (2) an even more important issue regarding stress for the history of Kelāsi is the exceptional stress pattern of feminine nouns in most NWI languages that have retained gender. That is, all masculine nouns (including proper names), all adjectives, postpositions, numerals, pronouns, most adverbs, all non-finite verbal forms (infinitives, participles) have one overall stress pattern: always on the final syllable, e.g., Kafteji *berá* 'brother', *værzá* 'ox', *berénj* 'rice (uncooked)', *tævilǽ* 'stable', *zæminlærzǽ* 'earthquake', *hæsǽn* 'Hasan', *tehrán* 'Tehran', *šemǽ* 'you (plural)', *hoštén* 'self', *delǽ* 'inside, within', *kænǽ* 'old', *xorém* 'good', *dozzǽ* 'twelve', *ará* 'tomorrow', *pærirúz* 'the day before yesterday', *vatǽn* 'to say', *bešnævessǽn* 'to hear', *dæškestǽn* 'to look at', etc. Only feminine nouns in those NWI that have an overt feminine (direct) marker – the full forms of Kafteji – have a different pattern: they are stressed on the penultimate syllable.

While stress in Western Iranian languages may indeed shift away from word-final position when inflectional morphemes are at issue (e.g., unstressed possessive enclitics and indefinite suffixes and, for verbs, stress-attracting TAM prefixes, etc.), these stress patterns always appear in the context of morphology added to base forms via grammatical rules. Therefore, an exceptional stress pattern for base-form stems such as feminine nouns could be one plausible explanation for the loss of the unstressed feminine morpheme.[9] This feminine direct case marker was most likely *not* optional in NWI but later became so in some dialects. That is, there is a certain pressure within the language to normalize all stress patterns to the word-final stress pattern of the majority, hence: *zænǽkæ > zænǽk* 'woman', *æŋgíštæ > æŋgíšt* 'finger', and so forth.

## 4.3 Other treatments of the feminine marker

In other NWI languages and dialects we find various tendencies for handling the category of grammatical gender and some of these include other possible interesting clues to the fate of the feminine marker in pre-Kelāsi. The various treatments it can undergo elsewhere may also give us hints as to why this marking was dropped, eventually causing the elimination of gender throughout the

---

**9** Note also that Kelāsi has undergone another shift away from a penultimate stress pattern involving a suffix: the direct plural forms. The original forms in Kelāsi must have been the same as the current Kafteji set-up: DIR *mærdǽk-ə*, OBL *mærdæk-ón* 'men'. Rather than simply deleting the unstressed direct plural morpheme, as it did with the direct feminine singular, leaving a bare stem, the oblique plural expanded its domain to include the direct usage, neutralizing the case distinctions in the plural.

grammatical system of the dialect. The discussion here is not intended to be an exhaustive discussion of gender in Iranian or even in Tatic. The immediate goal is to look at those processes that may ultimately lead – or, in the case of Kelāsi, may have led – to the loss of the feminine marker and hence the loss of gender differentiation.

The unstressed -æ ~ -ă feminine marker in the contemporary Iranian languages that have it (there are other types of feminine markers) derives from an original final proto-Iranian *-ā. In the history of Iranian, final short and long vowels by late Old Iranian fell together to short vowels only, and subsequently most final short vowels were eliminated for phonological reasons, which was then the primary impetus for the loss of gender and cases at least in some languages by the time of Middle Iranian (Prods Oktor Skjærvø, personal communication). Some languages, however, retained some final unstressed short vowels in order not to lose the oblique singular masculine case, the direct plural (both are generally unstressed; the singular feminine oblique may have been stressed; see Stilo 2009), or the distinction of feminine gender in the direct singular.

While feminine marking clearly has a history as a suffix bound directly to the root in Iranian (as well as in various other branches of Indo-European) and still is such in most NWI languages that have it, in some of the languages of the Tatic group it has acquired a certain detachable nature as discussed above, but also in the following sense: in Kajali (a Central Tati language, spoken to the northwest of Upper Tarom, see Map 1), for example, when an enclitic possessive pronoun – the oblique Set₂ endings introduced above – is added to a feminine noun ending in unstressed -æ, the former is inserted between the root and the -æ as a type of mesoclitic in instances of what Haig calls "debonding" (Haig 2016): *howlíg-æ* 'sister (direct case)' > *howlíg=**em**-æ* 'my sister (direct case)' vs. *howlig-ón=em* 'my sisters' (*-on:* oblique plural case, stressed, not separable). In the latter case, the oblique Set₂ endings follow the case suffix. While Set₂ enclitics are not used as possessives in Kafteji and the Tāromi dialects, the same type of debonding occurs in Kajali and certain Tāromi dialects, and farther to the east in Dikin-Marāqei, when the same Set₂, as a leftwardly mobile clitic encoding the agent of past transitive verbs, moves leftward from the verb and attaches to a preceding host, most often an object. This object in an ergative construction in the past is often in the direct case and the direct feminine -æ morpheme may then become separated from the root in these dialects by the insertion of this Set₂ agent mesoclitic:

(60) Kajali: *gow=š-æ*            *ædvænda gonšár=eš.*   < *gów-æ* 'cow (F)'
            cow=3SG₂.AG -F.DIR   threw        front=3SG₂
            'He put (lit: threw) the cow in front of him' (Yarshater 1960).

(61)  Hezar-rudi:  *xeylǽk=**eš**-æ*
                    spade=3SG$_2$=F.DIR
                    *u-gært-æ.*
                    PVB=took-3SG$_1$ (3SG$_2$: Agent; 3SG$_1$: Agreement with
                    feminine object)
                    'He picked up the spade' (Yarshater 1970).

(62)  Dikini:  *æsíf=m-an*          *o-do-yan*          *öštö.*
               apple=1SG$_2$.AG-DEF.F  PVB-gave-1SG$_1$.F  you:GEN
               'I gave you the apple' (Stilo 2018: 55).

Both of these mesocliticizing-type patterns occur occasionally elsewhere in gendered Tatic languages and dialects. The point of interest here is that this phenomenon of separating the feminine marker from the root by a mesoclitic sets a precedent for allowing this marker to become detachable from the root. While there is no direct evidence for its occurrence in contemporary Kafteji[10] as a parallel to what could have happened in pre-Kelāsi, it does occur in Hezar-rudi, a Tāromi dialect not far to the west of the village of Kelas, and provides an important clue for understanding the deletion of the feminine markers from noun stems in Kelāsi.

## 4.4  A possible scenario for gender loss in Kelāsi

This pattern of long and short forms of feminine nouns in Kafteji presents us with a possible scenario for the total loss of gender in Kelāsi:

(i)   The morpheme of the feminine direct case was probably optional in pre-Kelāsi for citation forms and the generic nouns in sentences, as it is in Kafteji today.

(ii)  Feminine forms in Kafteji and other NWI languages never occur for adjectives or verbs as citation forms in isolation and are only triggered by an accompanying noun in the context of syntax or a deleted referent in the context of discourse.

(iii) While there may have been a gender distinction in the demonstrative pronouns, there probably was no such distinction in identical demonstrative

---

**10** The fronting of Set$_2$ as agent to a definite feminine object is not represented in my Kafteji field notes and thus it is possible that such a process of detachability from the root also occurs in Kafteji. For examples of the fronting of Set$_2$ as agent to a non-specific feminine object where the direct feminine marker is dropped, see (13)–(15).

adjectives as in Kafteji and thus a precedent was also set for genderless demonstratives.

(iv) Since (a) the morpheme that marks the overt feminine form of the noun – the crucial locus of gender in the grammar – may be dropped and (b) feminine forms are never used in any other grammatical category (e.g., adjectives) as isolated words, all categories that allow feminine morphology have gender-neutral instances in speech when it is not overtly expressed.

(v) The next logical steps were that gender marking: (a) became optional probably first in certain domains and then spread throughout the system; (b) went through gradual stages of erosion by becoming more and more rarely used in speech; and (c) was then dropped as a category from the language.

Thus, if we take a Kafteji sentence with a feminine noun triggering gender marking in two other loci – (17b) *æm čél-æ pill-á n$^e$-á(ya)* 'That spindle$_F$ isn't$_F$ big$_F$' – and if we extract all feminine gender marking from the sentence and pronounce each word separately as citation forms – *æm, čel, pillæ, níæ* – we get the exact equivalent of what the sentence would be (but was not actually elicited) in contemporary Kelāsi. Compare again the forms of (18b) in both dialects:

(18b)  Kafteji (Feminine)
    *ǽm-æ*     *æmǽd-ə*  *dét-æ*         *n$^e$-áya*
    this-F.DIR  PN-M.OBL  daughter-F.DIR  NEG.BE-3SG.F

    Kelāsi
    *ǽm*  *æmǽd-e*  *dét-Ø*        *ní-æ.*
    this  PN-S.OBL  daughter-S.DIR  NEG.BE-3SG
    'This (or 'she') is not Ahmad's daughter'.

The above scenario for gender loss in Kelāsi, however, does not address its retention and expansion in Kafteji. There had to be completely different processes in effect in Kafteji.

# 5 The archaic retention of gender: Kafteji

The following domains in the Kafteji grammatical system are most likely inherited as archaic features that were not eliminated from the grammar as we see elsewhere in NWI (listed with the number of languages/dialects out of thirty-two

on Table 6 that share the feature): formal expression in the nouns (twenty-nine), attributive adjectives (nine),[11] predicative adjectives (nine), demonstrative pronouns (twenty-four), the third singular of the intransitive past (twenty-six), the encoding of the gender of the object in the past system of transitive verbs (seventeen), and the third singular of the copula (twenty-two). These features can be considered archaic since Kafteji shares them with several other disparate languages at a certain geographic and genealogical distance from it. Most of them would have existed in the oldest stages of Western Iranian as well. In only two minor areas of the grammatical system has Kafteji lost the cross-referencing of gender: the demonstrative adjective and the indefinite article (i.e., the number one as an article). These features appear in seven and eleven other languages, respectively, in Table 6.

Table 6 shows another important fact: while Kelāsi has lost all traces of gender and is thus not included in the table, we see that Kafteji, on the other hand, has the most loci for gender marking (eleven of thirteen) of all the languages in the table. While Kafteji is certainly conservative in the retention of gender marking and cross-referencing, it has also added some features to the gender marking system that are probably not original to NWI. I should mention briefly again that Kafteji is also conservative in other areas, e.g., in retaining the stress contrast to distinguish the masculine and feminine oblique forms.

## 5.1 The origins and paths of the Kafteji gender expansions

It is clear that the third person distinction of masculine and feminine in the verbal domain is more common and more stable than the other loci of gender in Tatic. This gender distinction probably originally began in the past intransitive system, where we see that it is almost universal in Table 6 and therefore probably older. The origin of these past tense third person singular forms is easily explained diachronically since the tenses of the past system are all participial in origin and in most West Iranian have merged with the older copula.[12] Since this formation was originally PARTICIPLE + COPULA, it would be natural

---

**11** There are probably more occurrences of feminine marking of at least attributive adjectives in other Tati dialects that are left blank in Table 6 simply because no data exists in the published literature for a proper assessment of its occurrence.

**12** PAST STEM + COPULA was an analytic construction that became a grammaticalized form of the preterit, similar to the fusion of the original Latin infinitive with forms of "have" as a suffixal future tense marker in Romance. In some Romance dialects these elements are still separate, e.g., Neapolitan Italian *ai cantá* vs. standard Italian *canter-ai* 'you will sing'.

**Table 6:** Loci of gender in Northwest Iranian languages in the vicinity of Kafteji and the Tāromi group of Tati languages

| | Dialect | 1 Formal Marking in Nouns | 2 Attributive Adjectives | 3 Predicate Adjectives | 4 Demonstrative Adjectives | 5 Indefinite Articles | 6 Demonstrative Pronouns | 7 3rd sg. Pres. Copula | 8 3rd sg. Past Intransitive | 9 3rd sg. Present ±TR | 10 1st, 2nd sg. Past Intransitive | 11 1st, 2nd sg. Copula | 12 Gender of Object, Past Transitive | 13 3rd sg. Subjunct. ±TR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Kajali | + | | | | + | ± | + | + | + | − | − | + | + |
| | Nowkiani | + | | | | | | | | + | | | + | + |
| | Lerdi | + | | | | | + | + | + | | | | + | + |
| | Karnaqi | + | | | − | | | | + | + | | | + | |
| T | Karani | + | − | − | − | − | + | + | + | + | − | − | + | + |
| A | Dizi | + | | | − | | | | + | | | | + | |
| T | Hezar-rudi | + | − | − | − | − | + | + | + | + | − | − | + | + |
| I | Siavarudi | + | | | | | + | | | | | | | + |
| | Bakoluri | + | | | − | − | − | + | | | | | | |
| | Jamalabadi | + | | | − | − | − | + | | | | | | |
| | **Kafteji** | + | + | + | − | − | + | + | + | + | + | + | + | + |
| | Dikin-Maraqei | + | | | − | | + | + | + | + | − | | + | + |
| | Chali | + | + | + | (−) | + | + | + | + | − | + | + | − | − |
| | Ebrahimabadi | + | − | + | − | | +/(−) | + | + | + | + | + | − | − |
| | Sagzabadi | + | − | + | − | + | + | + | + | + | | + | − | − |
| | Takestani | + | − | | − | | + | + | + | + | − | − | − | + |
| | Esfarvarini | + | − | − | − | − | + | + | + | − | + | − | − | − |
| | Xiaraji | + | − | | ± | + | ± | − | + | + | − | − | − | |
| | Danesfani | + | − | | − | + | ± | − | + | − | | − | − | |
| | Xoznini | + | − | | − | + | + | − | + | + | − | − | − | |
| | Eshtehardi | + | − | + | + | + | + | + | + | + | − | + | ± | − |
| | Alviri | + | +[1] | | + | | + | + | + | + | − | − | | + |
| | Vafsi | + | + | − | − | − | − | + | − | | − | − | ± | − |
| P | Amorei | + | + | | | | + | | + | | | | | |
| L | Delijani | + | + | + | + | + | + | + | + | | − | − | ± | − |
| A | Jowsheqani | + | − | − | − | − | + | + | + | + | − | − | + | + |
| T | Abyanei | + | + | + | + | + | + | + | + | + | | | + | |
| E | Nashalji | + | − | | − | + | + | + | + | | | | (+) | |
| A | Qohrudi | − | − | | − | − | − | + | + | | | | − | |
| U | Arani | − | − | | + | − | | − | − | | | | − | |
| | Abuzeydabadi | − | + | + | + | − | + | − | + | | | − | (+) | |
| | Hawrami | + | + | + | − | + | + | + | + | − | −[2] | − | + | − |

**1** In all dialects below this point in column 2, the attributive adjective generally follows the head noun.
**2** Hawrami does not distinguish gender in first and second singular in the past, but does so in the perfect tenses.

to distinguish masculine and feminine forms of the participle on the model of ADJECTIVE_{M/F} + COPULA. This analytic preterit formation later became fused and grammaticalized as a synthetic PAST PARTICIPLE + SUFFIX, giving rise to distinct gender endings.

## 5.2 Kelāsi-Kafteji and areal factors

Given that both Kafti and Kelas villages consist of small populations in regular socioeconomic contact with other villages on the plateau, the Rudbar Valley, the Southern Talysh, the urban centers of Gilan, and, through the media and for administrative purposes, also with Persian, most (or probably all) Kaftejis and Kelāsi are quadrilingual in at least their own native dialect, plus Rudbari, Gilaki, and Persian. Most men, due to their mobility for trade purposes, also speak Azerbaijani, the lingua franca of the wider Tārom-Zanjan area, and probably also Talyshi.

While the Kaftejis have no less interaction with surrounding areas than the Kelāsis do, it is intriguing to see that the dialect has still expanded the loci of gender in its grammar, even though all three to five additional languages they speak for their multilingual communication are in fact all genderless. If contact were the issue, we would expect Kafteji to have lost gender as Kelāsi did. While the Kaftejis may have wanted to retain something emblematic of their dialect to set themselves apart, it is doubtful whether such an abstract category as grammatical gender would have been a conscious choice.

Areal phenomena may be a possible explanation for the loss of gender in the case of Kelāsi, Charazei, Jeyshabadi, and Kalāsari, which are all Tāromi dialects, but then why not for Kafteji and the other eight gendered dialects of the Tāromi Tati subgroup in the same area? True, Kafteji is in a geographic cul-de-sac (note how the gender isogloss darts northward on Map 1 just to include Kafteji), but only slightly more so than Kelāsi. On the other hand, we still see that its difficult geographic accessibility has in fact not made it immune to areal features. To take only three areal features that have been encountered above that are very particular to the Gilan-Southern Talyshi and immediately adjacent areas, both Kafteji and Kelāsi share with Eastern and Western Gilaki and Southern Talyshi (1) the five verbs of being (even including all the same intricacies and highly idiosyncratic uses of each of these verbs that appear in all these languages that have this five-way distinction); (2) they do not allow the use of the Set$_2$ oblique person markers for indexing possessives, which is a somewhat restricted phenomenon in Western Iranian; and (3) they all share the lexeme *pillæ* 'large'.

### 5.3 Factors for systemic expansion of gender: Kafteji

The internal motivations for this expansion seem to produce a more coherent system by normalizing the encoding of gender throughout the singular of the verbal system. As Table 6 and Map 1 show, this type of systemic diffusion of gender has also happened elsewhere in Tatic, cf. Southern Tati: Chali, Ebrahimabadi, and partially Esfarvarini and Eshtehardi. These cases of the spread of gender to first and second person are probably independent innovations since the morphology of these forms in Kafteji and the Southern Tati forms are somewhat different. The existence of other examples of spontaneous systemic diffusion of gender elsewhere in Tatic just further reinforces the suggestion that these innovations are language-internally motivated tendencies to even out the system.

It is even conceivable that the diffusion of gender marking in verb paradigms proceeded initially from third person to second person before spreading to the first person since this is the situation we currently find in Zazaki/Dimli. The distinction of gender in verbs in this language is found only in the second and third singular for most tense-aspect-mood paradigms (see Paul 1998: 83). The first person has one common form for both genders.

## 6 Conclusion

I firmly believe that the languages of Greater Iran participate in a very close typological and areal bond, and that subgroups within this area clearly share many linguistic and cultural commonalities with their immediate neighbors, both Iranian and non-Iranian. While stating this with great conviction and having examined tremendous amounts of data from the field to support this assertion, I do not wish to ignore the fact that not all structures respond to areal tendencies. Languages, even those spoken in small villages whose inhabitants are all multilingual, can develop in their own directions, picking up on tendencies that were nascent in the language and developing them or by following certain "natural" tendencies more consonant with the internal trends of the language now more clearly understood through the study of language universals and typology. The trends of language convergence rarely hold true for all features in convergent languages. This chapter presents a fascinating case of an extreme process of divergence on one grammatical issue in two very closely related and immediately neighboring dialects.

# 7 Interlinear glosses and other abbreviations

| | | | |
|---|---|---|---|
| 1SGF | first person singular, feminine Set 1 (direct marker) | LINK | Noun Phrase linker (genitive-N or adjective-N |
| 2SGF | second person singular, feminine Set 1 (direct marker) | M | masculine |
| 3SGF | third person singular, feminine Set 1 (direct marker) | MO | oblique case marker, masculine |
| 1SGM | first person singular, masculine Set 1 (direct marker) | NEG | negative |
| 2SGM | second person singular, masculine Set 1 (direct marker) | NEG.BE | negative of "to be" |
| 3SGM | third person singular, masculine Set 1 (direct marker) | OBL | oblique |
| 1SG | first person singular, Set 1 (common or no gender) | OBL$_2$ | oblique case marker, type II |
| 2SG | second person singular, Set 1 (common or no gender) | PAM | person agreement marker |
| 3SG | third person singular, Set 1 (common or no gender) | PL | plural |
| 3PL | third person plural, Set 1 (common or no gender) | PN | proper name |
| 1SG$_2$ | first person singular, Set 2 (oblique marker) | PPL | past participle formant |
| 2SG$_2$ | second person singular, Set 2 (oblique marker) | PROG | progressive aspect formant |
| 3SG$_2$ | third person singular, Set 2 (oblique marker) | PST | past stem of verb |
| AG | Agent | PT | past formant of past stem |
| AUX | auxiliary (be) verb formant of past perfect | PU | punctual marker |
| CM | command; imperative formant | PVB | preverb (original directional particles, lexicalized with specific verb roots) |
| CMP | comparative formant of adjective | | |
| CS | change of state morpheme (in causative and | SG | singular |
| | inchoative light verb constructions) | SO | singular oblique case |
| DEF | definite marker (feminine only) | SUB | subordinator |
| DIR | direct | TAM | tense-aspect-mood marker |
| DU | durative marker | TR | transitive form of verb |
| F | feminine | | |
| INCH | inchoative (valence-reduced) form of a transitive verb | | |

NB: Final stress should be assumed for those words not marked for stress in the text.

# Bibliography

Haig, Geoffrey. 2016. Case debonding in Kurdish and beyond. Paper presented at the Third International Conference on Kurdish Linguistics (ICKL3), University of Amsterdam, 25–26 August.

Paul, Ludwig. 1998. *Zazaki*. Wiesbaden: Reichert.

Stilo, Donald L. 1981. The Tati language group in the sociolinguistic context of Northwestern Iran and Transcaucasia. *Iranian Studies* XIV (3–4, Summer–Autumn 1981). 137–187.

Stilo, Donald L. 2001. Gilaki. In E. Yarshater (ed.), *Encyclopaedia Iranica*, vol. X, *Fascicle 6 (Germany VI – Gindaros)*, 660–668. New York: Bibliotheca Persica Press.

Stilo, Donald L. 2009. Case in Iranian: From reduction and loss to innovation and renewal. In Andrej Malchukov and Andrew Spencer (eds.), *The Oxford Handbook on Case*, 700–715. Oxford: Oxford University Press.

Stilo, Donald L. 2015. The polygenetic origins of the Talyshi language. In Uwe Bläsing, Victoria Arakelova, and Matthias Weinreich (eds.), *Studies on Iran and the Caucasus: In honour of Garnik Asatrian*, 411–453. Leiden: Brill.

Stilo, Donald L. 2018. Dikin Marāqei Tati of Alamut: An undocumented conservative Tati language. In Saloumeh Gholami (ed.), *Endangered Iranian Languages*, Reichert Verlag. 659–824.

Yarshater, Ehsan. 1960. The Tāti dialect of Kajal. *Bulletin of the School of Oriental and African Studies* 23, Part 2. 275–286.

Yarshater, Ehsan. 1970. The Tati dialect of Ṭārom. In M. Boyce and I. Gershevitch (eds.), *W. B. Henning memorial volume*, 451–467. London: Lund Humphries.

Habib Borjian
# Mazandarani: A typological survey

Mazandarani, also known as Tabari, is spoken in south of the Caspian Sea in an area of about 15,000 km$^2$. Its speakers, roughly four million in number, mostly dwell in the towns and the surrounding villages spread over the lowlands of Mazandaran, as well as in the foothills and higher valleys of the Alborz chain. The geographical extent of Mazandarani does not match with the current administrative borders of the Mazandaran Province. The Mazandarani language domain extends eastward toward the plains of Gorgān in the east and stops at the Čālus River in the west; thus, the districts of Kalār and Tonekābon (the western parts of historical Ruyān), while still within the current borders of the Mazandaran Province, have dialects with only low mutual intelligibility with either Gilaki or Mazandarani, and hence are considered a third separate language group of the Caspian language family (Stilo 2001). See Figure 1.

Within the Mazandarani language itself, it is very apt to delimitate the varieties into Mazandarani proper and the dialects that surround it. What I designate here as Mazandaran proper, and its language varieties as Mazandarani proper, is an area extending roughly from the town of Nur (former Suledeh) in the west across the littoral plains to Behšahr in the east. It embraces the four major towns of the province, namely, Āmol, Bābol, Šāhi, and the provincial capital Sāri. The peripheral varieties are spoken in the far east toward Gorgān, in the far west in the districts of Kojur and Nur, and in the southern highlands.

The varieties of Mazandarani are mutually intelligible to various degrees, more so in Mazandaran proper. None of the dialects, though, is considered standard or formal. Dialectal continuum varies in two geographical directions: along the latitude (river valley variation) and along the longitude (lowland vs. highland). Almost every locale has its own subdialect, and even nearby settlements may exhibit substantial phonological difference. Lexicon, however, is fairly uniform throughout the province (see more in Borjian 2004a). The data in this chapter, unless otherwise specified, are from my field notes collected in the district of Espivard, located between the cities of Sāri and Šāhi, thus occupying a central position within Mazandarani proper.

The intention of this chapter is to show Mazandarani's salient phonological and morphological traits, with frequent references to Persian, which has lent Mazandarani some of its current morpho-syntactic features, and to Gilaki, the sister language to Mazandarani within the Caspian linguistic zone, as well as to

---

**Note:** To the pioneer of West Iranian typology Donald Stilo.

other related languages. In the concluding section, additional typological insight will be made by featuring ten selected grammatical topics in an attempt to establish the isoglosses that bind Mazandarani to or demarcate it from other major West Iranian languages.[1]

# 1 Phonology

The consonants are essentially similar to those of Persian, but not without idiosyncrasy. A conspicuous one is the voiced uvular fricative /ɣ/ ([ʁ]), which is articulated in all positions where Persian has /q/, e.g., ɣ*ərâri* 'contract farm worker', derived from Persian *qarār* 'contract'. Mazandaranis may also disclose their linguistic identity even when speaking Persian in articulating /k/ and /g/ as velar stops in all positions, contrasting with the palatalized forms before front vowels in Persian.

The vowel system, as in the case of other Iranian languages, exhibits certain variation among the dialects. The vowel sounds, /a e i â o u/, are not fundamentally different from Persian.[2] See Table 1 for the phonemes and their allophonic ranges. This inventory is supplemented by what is shown in this study as /ə/, whose phonemic status is not quite well understood. It has relatively broad allophonic distribution, ranging among [ə], [ɐ] [ɜ], and [ɛ].[3] While the phoneme is perceived in Persian orthographic terms as a *kasra* (ranging between [e] and [ɛ]) in Sāri, those from Āmol and Bābol find it more relevant to render it a *fatḥa* ([æ]) when writing down their vernaculars. In the urban variety of Sāri one may say that /ə/ and /e/ have merged into a single phoneme.

Historical distinction of length is lost in Mazandarani, yet remnants of the front *majhul* vowel are noticed in a number of words with loss of length: *vešə* 'woods', *bedâr* 'awake', *mevə* 'fruit', *dešo* 'last night', and in the prefix *be* 'without'. The vowel /o/ has various etymological sources. It appears in words such as *xo* 'sleep' (< *xwāb*), *šo* 'night' (< *šab*), *sio* 'black' (cf. Parth. *siāw*, Pers. *siāh*), *dəryo* 'sea' (cf. Pers. *daryā*), *səron* 'singing' (cf. Pers. *sarāyān*), *jo* 'barley' (< *jaw*), *hoz* 'pool' (< *howz* < *hawz*), among others. Contrary to most Iranian languages, omission of glottal consonants is not compensated by vowel lengthening, e.g., ɣ*ar* (< *qahr*) 'sulk, grump', *jam* 'assembly' (< *jamʿ*).

---

**1** This study is not updated for the works that were published after 2008.
**2** Zav'jalova (1956) proposes a vowel inventory for Mazandarani consisting of *i*, *e*, *ɛ*, *a*, *u*, and *o* without making known the dialect or dialects she studied. See Borjian (2004a).
**3** Ṣādeqi (1990) proposed [ɐ].
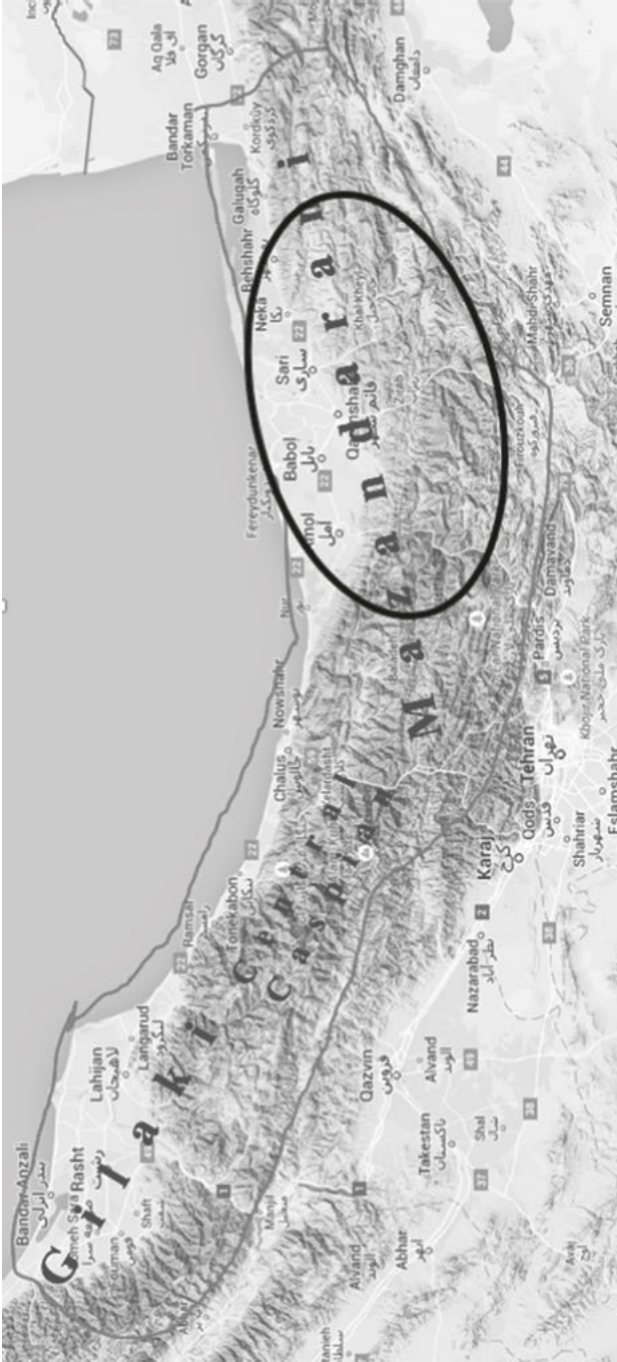
**Figure 1:** Delineation of the Caspian language family, with Gilaki in the west, Central Caspian in the middle, and Mazandarani in the east. The area bounded by the ellipse approximates Mazandarani proper.

**Table 1:** Vowel phonemes and allophones

|        | Front      | Mid                              | Back               |
|--------|------------|----------------------------------|--------------------|
| High   | /i/ [i]    |                                  | /u/ [u] ~ [ʊ]      |
| Mid    | /e/ [e] ~ [ɪ] |                               | /o/ [o] ~ [ʊ]      |
|        |            | /ə/ [ə] ~ [ɛ] ~ [ɜ] ~ [ɐ]       |                    |
| Low    | /a/ [æ]    |                                  | /â/ [ɒ]            |

When /i/ is realized as a phoneme, it is retained in phonemic transcription even when the vowel turns to /y/ due to stress shift, e.g., *baíta* [bæ'ⁱitɜ, bæ'ʰitɜ] 'he took', *bái(r)* ['bæ'ⁱjⁱ] 'take!' This rule generally holds for verbal stems, including the verbs 'to be', 'to be in', and 'to become': *bía*[4] ['biˈjɜ] 'it was', *báiə* ['bæ⁽ʰ⁾ⁱjɜ, 'bæijɜ, 'bæjjɜ] 'it became', *dáiə* ['dæˈjɜ, 'dæijɜ, 'dæjjɜ] 'it was in, existed'.

The vowel sounds /u/ and /o/ have a range between [u] and [o]; thus we have two forms (inter-dialectal and intra-dialectal) for certain words: *o ~ u* 'water', *asio ~ asiu* 'water mill', *bó-e ~ bú-e* 'that it be'; *mion ~ miun* 'middle'.

*Stress:* The stress is predictable and phonemic, for example:

(1)

| | |
|---|---|
| *davəndí* 'shoe' | *dávəndi* 'that you (SG) tie/close' |
| *dár šunə* 'he ran away' | *dar šúnə* 'he was going' |
| *dar-émbə* 'I am coming' | *dárəmbə* 'I am (in)' |
| *bətúndi* 'you can' | *bátundi* 'you cannot' (irregular negative) |

In nominals, stress patterns are similar to Persian, i.e., the stress falls on the final non-inflectional syllable (except for plural markers). Exception is a subset of words, including *ə́ttə* 'one', *xále* 'very', *vále* 'but', which happen to correspond to Persian *yéki*, *xéyli*, *váli*, also stress-initial. The vocative, as in other Iranian languages, is stressed on the initial syllable.

Verb forms require stress shift, e.g., *bávəre* 'that he takes', *várdə* 'he would take', *návərdə* 'he did not take', *bávərdə* or *bavárdə* 'he took', *bavərdá* (adj.) 'taken'. Note the stress contrast between the last two words. The verbal stress patterns can be summarized as follows.

---

**4** In *ie*, *ia*, *io*, *iu*, and *iâ* the euphonic glide *y* between the two vowels is implied.

(i) The negative morpheme always takes the stress: *náinə* (neg. of *gəinə* 'it takes'), while the substantive verbs allows stress shift: *dániə = daníə* 'it is not (there)'.

(ii) In the imperative and the subjunctive the stress falls on the first syllable, that is, the verbal prefix: *báir* 'take!', *dávoše* 'that it be'.

(iii) In the present-future the stress is normally taken by the penultimate syllable, be it the stem or the ending: *ešémbə* 'I look', *vešénə* 'he opens', *iā́nnə / yā́nnə* 'he brings', *var-ə́mbi / vár-mi* 'we carry'. An exception is the verb 'be': *hássemə* 'I am', *dárəmbə* 'I am in'.

(iv) In the forms built on the past stem, the stress falls normally on the last syllable of the stem: *hâ-kárd-i* 'you (SG) did', *vəséssə* 'it was blazing', *sar-bazúə* 'it sprang'. Nevertheless, the stress may optionally shift to the preverb. This shift is heard mostly when two vowels come together, e.g., *baímə*, *bahímə*, *báimə/báymə* 'I became'; *baíə*, *báiə/báyyə* 'it became'; *baítənə*, *báytənə*, 'they took', *bâútə*, *bā́utə* (optionally contracted to *bā́tə*) 'he said'. The stress may travel from the final vowel of the stem across another vowel to the preverb: *bakəšínə*, *bákšinə* 'they pulled', *baxəríə*, *báxriə* 'he ate'.

(v) The verbal nouns follow the nominal pattern, i.e., the stress falls on the last syllable: *basutə́* 'burned' (participial adjective), but *basút* 'id' in periphrastic perfect tenses.

## 2 Noun phrase

*Number and definiteness:* The plural markers are *-un* (< *-ān*) and *-â*; the latter is most likely from Persian *-hā*. Collective use generally does not mark plural. The indefiniteness markers are *əttâ / əttə* (lit. 'one unit') that precedes the noun, or the unstressed suffix *-e/-i*. Examples: *və əttâ zənâ badiə* 'He saw a (certain) woman'; *və dətər-i dâštə* 'He had a daughter'; *šo-e haf sâât xətə* 'He used to sleep seven hours each night'.

*Modifiers:* Possessives and adjectives precede the head noun, with a 'reverse-*eżâfe*-like' connector.[5] There is however a marked difference in Mazandarani between noun–noun and adjective–noun constructions: the connector *-e* is *unstressed* with noun modifiers and *stressed* with attributive adjectives. Examples: genitive: *piár-e səre* 'father's home'; adjective: *gat-é bâmši* 'big cat'.[6]

---

**5** The term "reverse-*eżâfe*-like connector" has been coined by Donald Stilo.

**6** See Typological Feature 1.

*Adpositions:* Preponderance of postpositions over prepositions is a characteristic of Mazandarani.[7] The object of a postposition is, generally speaking, suffixed with *-e*, except for the postpositions *-(r)ə* and *-je*, which are directly suffixed to the object. The latter two should be treated therefore as bound morphemes, and hence preceded by a hyphen in transcription. Other postpositions are transcribed separately unless the morpheme *-e* is dropped; thus *in-vari* 'this way', *kəm-vari* 'which way', *səre-vari* 'toward the house', but occasionally *səre-(y)e vari* 'id'. I treat the postposition *vesse* as a bound morpheme when it is reduced to *-esse* by dropping the initial consonant.

The object marker *-ə*, postvocalic *-rə*, may have been borrowed from New Persian, when the latter had grammaticalized Middle West Iranian *rāy* (from Old Persian *rādiy* 'because of, on account of') into a dative-accusative marker. Examples in Mazandarani are: *rikâ jəmə-re kijâ-rə hədâə* 'The boy gave the shirt (direct object) to the girl (indirect object)'; *və-rə* (ind. obj.) *un-ə* (dir. obj.) *hədâmə* 'I gave it to him'.[8] The same morpheme signifies the object of modal verbs: *mərə ənâr vesse* 'I had a desire for pomegranates'.

The polysemous postposition *-je* 'from/to, with' is mainly used as (1) ablative, but it also marks (2) the instrumental and (3) abstraction:

(i)   *sang bumsar-je dakətə* 'the stone fell off the top of the roof' *nâmə-rə me xâxər-je barəsən* 'take the letter to my sister!'
(ii)  *as-je burdəmə* 'I went by horse'
(iii) *vəne dim še mâr-e dim-je mundənə* 'her face resembles her mother's face'

Major postpositions attested in Mazandarani are: *bən* 'under', *dələ* 'in', *dəm(b)âl* 'behind', *həmrâ* 'with', *jer(i)* 'under', *kənâr* 'by', *miun* 'in', *pali* 'near, at', *pali-je* 'on one's side', *pe* 'at the foot', *pəšt* 'behind', *piš* 'by, with', *sar* 'on, over, at', *sari* 'at', *vâri* 'like, as', *vari* 'toward, in the direction of', *(v)esse* 'for, in order to'.[9]

*Personal pronouns:* Six sets of personal pronouns can be identified in Mazandarani (Table 3). One may, however, discern two basic sets: (1) The nominative or subject set descends from Middle West Iranian oblique forms, e.g., *mən* 'I' (cf. Middle Pers. *man* < Old Pers. gen.-dat. *manā*). (2) The possessive determiners,[10] all ending in *-e*, are apparently of secondary construction formed by the fusion of the nominative pronouns and the morpheme *-e*. Consequently, the stress shifts

---

**7**  See Typological Feature 3.

**8**  Cf. Eastern Gilaki *únæ únæ hádam* 'I gave it (dir. obj.) to him (ind. obj.)' (Stilo 2001: 662).

**9**  Postpositions draw a noticeable differentiation between Mazandarani proper and its peripheral dialects.

**10**  See Typological Feature 2.

away from the unstressed suffix: *əmā́* 'we' + *-e* → *áme* 'our'; *šəmā́* 'you' + *-e* → *šā́me* 'your', etc. For the singular possessive, normally two sets are employed: *me*, *te*, *ve* ... and *məne*, *təne*, *vəne* ... , sometimes within the same dialect; the latter set appears to have developed to differentiate between the increasingly merging *tə* and *te* (second singular) and *və* and *ve* (third singular). The nasalization of the possessive in the second and third singular may have occurred on the first singular model: *mən* + *-e* → *məne*, followed by *təne* and *vəne*.[11]

Other pronominal cases shown in Table 2 are formed by the agglutination of either of the aforementioned two basic sets and suffixes *-je*, *-(r)ə*, and *-(ve)sse*, as well as *-še*.[12] Enclitic pronouns are nonexistent in Mazandarani.[13]

**Table 2:** Personal pronouns.

|  |  | Nominative | Possessive (determiner) | Acc.-Dat. | Ablative | Benefactive | Possessive (pronoun) |
|---|---|---|---|---|---|---|---|
| SG | 1 | *mən* | *me, máne* | *márə* | *méje* | *mésse*[1] | *méše* |
|  | 2 | *tə* | *te, táne* | *tárə* | *téje* | *tésse* | *téše* |
|  | 3 | *və* | *ve, váne* | *várə* | *vánje*[2] | *vánesse* | *váneše* |
| PL | 1 | *əmā́* | *áme* | *əmā́rə* | *ámje* | *ámesse* | *ámeše* |
|  | 2 | *šəmā́* | *šáme* | *šəmā́rə* | *šámje* | *šámesse* | *šámeše* |
|  | 3 | *vəšún* | *vəšúne* | *vəšún(n)ə* | *vəšúnje* | *vəšúnesse* | *vəšúneše* |

[1]Also *mévəsse*, *tévəsse*, etc.; *mévə*, *tévə*/*tánevə*, etc.
[2]Also *vəneje*, etc.

*Reflexives:* The invariable pronoun *še* 'self' (any person)[14] occurs either in a reflexive sense (*və še pali bâutə* 'he said to himself') or in an emphatic sense: *(və) še bâutə* 'he said [it] himself'; *(və) še burdə* 'he went himself'. The same morpheme is used also as a possessive adjective: *še go-rə rušəmbə* 'I'll sell my cow'.

---

**11** The nasalized forms are also found in the Lori dialects of Southwestern Iran.
**12** See Borjian 2004b. A tripartite case system of direct, oblique, and possessive is proposed for Mazandarani (cf. Rastorgueva and Èdel'man 1982: 529). Cf. similar system in Gilaki laid out by Donald Stilo (2001): Western Gilaki first plural pronouns *aman(n)* (subject), *ama-ra* (object), *ame* (possessive).
**13** See Typological Feature 4.
**14** In Eastern Mazandaran, from Behšahr to Gorgān, the forms *xod* and *xaštən* are used. Cf. Pers. *xwēš*, Kurdish *že*, Aftari *jun*, Tākestāni *jâ* and *janâ*, Talyshi *čay* and *čavōn*, Semnāni masc. *žo* and fem. *žin*, etc.

Adding the reflexive morpheme *-še* or *-šene* to possessive determiners or to an interrogative pronoun results in the corresponding possessive pronoun (Table 2), e.g.,

(2)  *in jəmə meše / me-šene*       'this shirt is mine'
     *intâ kéne-še? intâ vəneše*     'whose is this? it's his/hers'

*Deixis:* Demonstratives include *in* 'this', *un* 'that', with the plurals *inâ(n)*, *unâ(n)*; *intâ*/*untâ* 'this/that one', as in *intâ xi* 'that other boar'; *hamintâ*/*hamuntâ* 'this/that very (same) one', as in *hamintâ-rə me piər-əm dâštə* 'My father too had the very same one'. Other deixis are: *injə* 'here', *unjə* 'there'; *haminjə*/*hamunjə* 'right here/there'; *ungədər* 'then, at this time'.

# 3  Verb phrase

The verbal system of Mazandarani is based on a binary set of stems (past and present); the indicative and subjunctive moods; the imperfective and perfective aspects; and four sets of personal endings. There are five simple TAM formations: the present-future, subjunctive present, imperative, preterite, and imperfect (but no present perfect). There are three periphrastic perfect tenses based on the past participle and conjugated forms of the verb "be". The progressive tenses are formed on the present and past of the locative verb. See Table 3 for person markers and Table 6 for paradigms.

*Stems:* Mazandarani generally retains a clear distinction between present (or nonpast) and past stems of verbs, with similar types of association between the two stems as in most other Iranian languages, e.g., the past stem = present stem + *-(e)ss-*, as in *mâl-*: *mâless-* 'rub'; + *-i-*, as in *kand-*: *kandi-* 'dig'[15]; + *-d-*, as in *bur-*: *burd-* 'go'; − *-n-*, as in *čin-*: *či-* 'pick'. There are many stems that do not show an overt synchronic agreement, such as *šor-*: *šoss-* 'wash', *xəs-*: *xət-* 'sleep', *tej-*: *tet-* 'run', *ruš-*: *rut-* 'sell'. Some verbs have alternate past forms with different processes, as in *kâr-*: *kârəss- ~ kâšt-* 'sow', *vand-*: *vess- ~ vandess-* 'close'. Some original verb stems have been converted to the nominal element of compound verbs, e.g., *bərmə kərd-* 'to weep', *xandə kərd-* 'to laugh' (but also the simple verb stem *xandəss-* 'to laugh').

---

**15**  The past formant *-i-* is used in causative past stems (below). See also Typological Feature 9.

*Prefixes:* The perfective and conjunctive are marked with the prefix *bV-*. Main lex-icalized preverbs *hV-*[16] and *dV-*[17] replace *bV-* and may add to the meaning of the stem. Examples:

(3)  *da-čiən*            'to arrange'
   *ba-čiən*            'to pick' (e.g., fruits from trees)
   *hâ-čiən*            'to pick from the ground'
   *da-pissən*          'to be soaked'
   *ba-pissən*          'to decay'
   *da-kətən*           'to fall'
   *hâ-kətən*           'to become useless'
   *ba-xətən*           'to sleep'
   *da-xətən*           'to bend' (e.g., a rod)'
   *da-gərdəssən*       'to rotate'
   *ba-gərdəssən*       'to search'
   *bar-da-gərdəssən*   'to return'

*Negation:* Negation is expressed by the stressed prefix *nV-* which replaces the modal-aspectual prefix: *ná-vərin* 'do not cut!', *ní-yəl* 'do not put!', *ní-igumə* 'I did/would not put', *nə-dâštəni* 'you did/would not have'. Note that the preterite (*bə-dâštəni*) and the imperfect (*dâštəni*) fall together in their negative form.

*Compound verbs:* As in most Iranian languages, Mazandarani is replete with verbs consisting of an invariant nonverbal element and a verbal element that conjugates. Common light verbs are *bakərdən*, *baiən*, *hədâən*, *baitən*, *bazuən*, *baxərdən*, *burdən*, among others. Examples:

(4)  *ba-kərdən*         'to do'
   *da-kərdən*         'to put into, put on, make'
   *hâ-kərdən*         'to do, perform'
   *yâd hâkərdən*      'to remember'
   *yâd bakərdən*      'to forget'
   *koš bakərdən*      'to scratch'
   *ši bakərdən*       'to marry'
   *las hâkərdən*      'to wait'

---

**16**  The Mazandarani preverb *hâ-* or *ha-* (cf. Eastern Gilaki *ha-*) corresponds to Western Gilaki *fa-*, Pers. *farā*, etc., thus < Old Iranian *\*frā-*. It is also proposed that it is ultimately derived from *\*hama-aiva-da* 'same time', or parts thereof (Windfuhr 1989: 256).
**17**  Apparently from the adverb *dar* 'in'.

*Personal endings:* Four sets of person markers can be identified in Mazandarani, as shown in Table 3. Set I is employed in the indicative present (present-future), Set II in the preterite and the imperfect and in the present tense of the substantive verbs, Set III in the subjunctive present, and Set IV in the imperative. The epenthetic *-ə-* appears when the stem ends in a consonant; it normally shifts to *-e-* in the vicinity of sibilants: *ešembə* 'I look', *vešenə* 'he opens', *hassemə* 'I am'.

There is a certain degree of variance in person markers among Mazandarani varieties. For instance, the first singular and plural markers are reduced to *-me* and *-mi* in some dialects and third person plural is *-(ə)nənə* in the varieties spoken in Āmol and Bābol.

Table 3 would be very crowded if we incorporated all the morpho-phonological variants. Both the ending and the present stem may alter depending on the final sound of the stem (more in Borjian 2005):

**Table 3:** Personal endings.

|  |  | I | II | III | IV |
|---|---|---|---|---|---|
| SG. | 1 | *-(ə)mbə, -(ə)mmə* | *-(ə)mə* | *-əm* | |
|  | 2 | *-(ə)ni* | *-i* | *-i* | zero |
|  | 3 | *-(ə)nə* | *-ə* | *-e* | |
| PL | 1 | *-(ə)mbi, -(ə)mmi* | *-(ə)mi* | *-im* | |
|  | 2 | *-(ə)nni* | *-(ə)ni* | *-in* | *-in* |
|  | 3 | *-(ə)nnə* | *-(ə)nə* | *-ən* | |

a.  When the stem ends in either *r* or *l*, *-mə* and *-mi* are used optionally for the first person singular and plural, respectively:

(5)  *var-* 'carry'    *var-mə* or *var-əmbə* 'I carry'
     *yəl-* 'put'       *yəl-mi* or *yel-əmbi* 'we put'

b.  When the stem ends in either *r* or *l*, these liquid consonants assimilate to the nasal consonant of the second and third singular endings:

(6)  *var-* 'carry'      *van-nə* 'he carries'
     *dâr-* 'have'       *dân-ni* 'you have'
     *yəl-* 'put'        *yən-nə* 'he puts'
     *mi(ə)r-* 'die'     *miən-nə* 'they die'

c. When the stem ends in *n*, the latter is absorbed by the nasal consonant of the first person singular and plural:

(7) *zan-* 'beat'  *za-mbə* 'I beat'
  *vin-* 'see'  *vi-mbi* 'we see'
  *kən-* 'do'  *kə-mbi* 'we do'

d. When the stem ends in *n*, the latter is absorbed by the nasal consonant of the second and third plural endings*:*

(8) *vin-* 'see'  *vi-nnə* 'they see'
  *tun-* 'can'  *tu-nni* 'you can'

e. Subsequently, a dissimilation process occurs to distinguish between the plural and singular endings. (Otherwise the second person singular *\*vin-ni* would be confused with the second person plural *vinni*, with the underlying form *vin-nni*):

(9) *vin-* 'see'  *vin-di* 'you (SG) see'
  *zan-* 'hit'  *zan-də* 'it hits'

The four-set system of person markers given above is on a purely synchronic ground. Resorting to the underlying, historical formation would greatly simplify our analysis. The extra nasal consonants in Set I endings, when compared to Set II, in Table 3 can be traced to the Old Iranian participle formant *\*-ant-* infixed between the stem and the original endings. Subsequently, Stilo (2001) considers a single set of endings for Gilaki and treats the nasal element in Set I as a present indicative marker.[18] By adopting this strategy, the personal markers will be reduced to two underlying sets in all Mazandarani varieties:

  The indicative set:  *-mə, -i, -ə, -mi, -ni, -nə*
  The subjunctive set:  *-əm, -i, -e, -im, -in, -ən*

*Simple tenses.* There are five basic verb structures in Mazandarani:

| | | | | | | |
|---|---|---|---|---|---|---|
| Present indicative | = | | | PRS stem | + | Ending I |
| Present subjunctive | = | Prefix | + | PRS stem | + | Ending III |
| Imperative | = | Prefix | + | PRS stem | + | Ending IV |

---

**18** See Typological Feature 7.

| Preterite | = | Prefix | + | PST stem | + | Ending II |
|---|---|---|---|---|---|---|
| Imperfect | = | | | PST stem | + | Ending II |

What is designated as present indicative in Mazandarani normally conveys habituality or future, e.g., *fərdâ gəmbə* 'I will say tomorrow'; similar functions exist in Persian and Gilaki.[19] An ongoing action in progress is expressed by the progressive forms (see below).

The present perfect is generally held to be absent in Mazandarani proper.[20] Its absence manifests itself in some native speakers of Mazandarani not using the present perfect correctly when speaking Persian. Nevertheless, Yoshie (1996: 36–37) proposes a present perfect for stative forms,[21] which might truly be a morpho-phonological feature. Moreover, Dargāhi (1993) suggests a differentiation between the present perfect and the preterite through stress shift (*báčime* 'I picked' vs. *bačíme* 'I have picked'). This is indeed possible, but only in a variety influenced by Persian, perhaps in the town of Šāhi, which was populated by immigrants seeking work in the newly established industries under Reza Shah Pahlavi.

*Periphrastic tenses based on the past participle:* The past participle (PST PTCP) is formed by the past stem prefixed by the modal marker *bV-*. The suffix *-ə*, added when the past participle is used as an adjective (e.g., *baxurdə́* 'eaten'), seldom appears in verb forms. The past perfect, perfect subjunctive, and pluperfect subjunctive are formed by the past participle succeeded by conjugation of the auxiliary "to be":

> Pluperfect = PST PTCP + PST of 'be' (*bi-mə, -i, -ə, -mi, -ni, -nə*)
> Perfect subjunctive = PST PTCP + SBJV past of 'be' (*bu-/bo-əm, -i, -e, -im, -in, -ən*)
> Pluperfect subjunctive = PST PTCP + *bi* (PST PTCP of 'be') + PST SBJV of 'be'

Examples with references to the parallel structures in Persian when one exists:

(10)  *burd bimə*  ~ Pers. *rafte budam* 'I had gone'
  *burd buəm*  ~ Pers. *rafte bâšam* 'that I should have gone'
  *burd bi buəm*  'I might have gone' or 'maybe I have gone'

---

**19** Cf. Western Gilaki *xâyəm guften* 'I shall say', *xayí guftən* 'you will say', etc., and Eastern Gilaki *xan búgom, xan bégi*, etc. (Stilo 2001: 665).

**20** See Typological Feature 6.

**21** As in *baxâtúe*, for Pers. *xwāb ast* or *xwābide ast*, in my documentation from Sāri. I have not come across this form in my notes from other Mazandarani varieties.

*Progressive tenses:* Mazandarani employs the existential/locative verb *daiən* 'to be in' (see below) as auxiliary to form progressive forms.[22] The main verb always conjugates (in the present or preterite tenses), while the conjugation of the auxiliary is optional:

(11)  *dar(nə) šunnə*    'they are going'
       *dai(nə) šinə*     'they were going'

*The verb "Be":* Mazandarani distinguishes two substantive verbs (Table 4 ):

(i)  The general, neutral copula, which normally appears as clitics, e.g., *və gâləš-ə* 'he is a herder', *və gâləš ni-ə* 'he is not a herder'. The clitics may take the present stem *hass-*, which is employed only in present affirmative and generally emphatic, as in *mən gâləš hassəmə* 'I am a herder', but may also be used in the same place as copula, with no special emphasis, as in *in mardi čeči márdi hassə?* 'what kind of *man* is this man?'

(ii)  The locative or existential verb, meaning 'be in, exist', as in *lavə-e dəle o darə / danía* 'there is/isn't water in the pot', *Vəndâd səre dare* 'Vendad is home'.

**Table 4:** The verbs 'be'.

|        |         | Copulative |        | Locative   |        |
|--------|---------|------------|--------|------------|--------|
|        |         | **Present** | **Past** | **Present** | **Past** |
| sg. 1  |         | *-əmə*     | *bimə* | *dar(ə)mə* | *daimə* |
|        | 2       | *-i*       | *bi(i)* | *dari*     | *dai*   |
|        | 3       | *-ə*       | *biə*  | *darə*     | *daiə*  |
| pl. 1  |         | *-imi*     | *bimi* | *darəmi*   | *daimi* |
|        | 2       | *-ini*     | *bini* | *darəni*   | *daini* |
|        | 3       | *-ənə*     | *binə* | *darənə*   | *dainə* |

*Modals:* Mazandarani modals follow similar patterns to those in Persian, i.e., they are followed by the present subjunctive of the main verb. They include: *xâ-: xâss-* 'want': *xâmmə bâurəm* 'I want to say'; *tun-: tunəss-* 'can': *natundə bure* 'he cannot go'. Impersonal forms include: *vénə* and *vésse*, followed by the subjunctive or the infinitive: *vénə burəm* 'I must go', *və navesse bâure* 'he shouldn't have

---

**22**  See Typological Feature 5.

said [it]', *vessə/venə burden* 'one must go'; *šâyed burem* 'maybe I go'; *našene bure* 'he out not to go', *našene burden* 'one shouldn't go'.[23]

*The Causative:* Suffixing the causative morpheme *-ən(d)-* to the present stem of an intransitive verb makes it transitive; the past stem is formed by agglutinating the past-stem formant *-i-*.[24] The derivation of the causative is illustrated in Table 5.

**Table 5:** Causative formation.

| intransitive ↓ | | **'dress'** | **'sleep'** |
|---|---|---|---|
| | Infinitive | *da-puši-ən* | *ba-xət-ən* |
| | past stem | *puši-* | *xət-* |
| | present stem | *puš-* | *xəs-* |
| ↓ transitive | present stem | *puš-ənd-* | *xəs-ənd-* |
| | past stem | *puš-əndi-* | *xəs-əndi-* |
| | Infinitive | *da-puš-əndi-ən* | *ba-xəs-əndi-ən* |
| | | 'dress' | 'put to bed' |

*Passive forms:* The passive is formed analytically as in Persian. Windfuhr's (1989: 257) statement that Mazandarani has preserved the inflectional passive marked with inherited *-i-* (as in Sangesari *eštende* 'he is standing up' vs. *eštinde* 'he will, is going to stand up') does not hold true for Mazandarani proper.

**Table 6:** Verb forms.

| | **Eat** | | **Sit** |
|---|---|---|---|
| | **Affirmative** | **Negative** | **Affirmative** |
| Present-future | *xənnə* | *naxənnə* | *nišennə* |
| Subjunctive | *baxəre* | *naxəre* | *həniše* |
| Imperative | *baxər!* | *naxər!* | *həniš!* |
| Present progressive | *dar(ə) xənnə* | – | *dar(ə) nišennə* |

---

**23** Šokri 1995, p. 199. Cf. Gilaki, in Stilo (2001: 666–667).

**24** It is comparable to the formation of the causative in Persian: the causative formants *-ân-* and *-ân(i)d-* are added to the plain present stem to make the causative present and past stems, respectively. See also Typological Feature 9.

|  | **Eat** |  | **Sit** |
|---|---|---|---|
|  | **Affirmative** | **Negative** | **Affirmative** |
| Preterite | *baxərdə* | *naxərdə* | *həništə* |
| Imperfect | *xərdə* | *naxərdə* | *ništə* |
| Past progressive | *daiə xərdə* | – | *daiə ništə* |
| Pluperfect | *baxərd biə* | *naxərd biə* | *həništ biə* |
| Past subjunctive | *baxərd bu* | *naxərd bu* | *həništ bu* |
| Pluperfect subjunctive | *baxərd bi bu* | *naxərd bi bu* | *həništ bi bu* |
| Infinitive | *baxərdən* | *naxərdən* | *həništən* |
| Past participle | *baxərd(ə)* | *naxərd(ə)* | *həništ(ə)* |

All conjugated forms except the imperative are in the third person singular.

# 4  Typological features

This survey of Mazandarani typology incorporates ten grammatical isoglosses, or features, deemed relevant in setting Mazandarani among West Iranian languages. The features are listed in Table 7 and further discussed below. Features 1–4 correspond to noun phrase and 5–10 to verb phrase.

**Table 7:** Grammatical features.

|  | Feature | CP | Kurd. | Tati | Gil. | Maz. | Aft. | Sem. | Bal. | Pers. |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Modifier + HN | − | − | + | + | + | + | + | + | − |
| 2 | Freestanding possess. det. + HN | − | − | + | + | + | + | + | + | − |
| 3 | Mainly postpositional | − | − | + | + | + | + | + | +/− | − |
| 4 | No pronominal clitics in NP | − | −/+ | +/− | + | + | + | + | − | − |
| 5 | Progressives with locative verb | − | − | − | + | + | + | − | − | − |
| 6 | No present perfect | − | − | − | + | + | + | − | − | − |
| 7 | Present tense formant -*n*- | − | − | − | −/+ | + | + | − | − | − |
| 8 | Imperfect without tense marker | − | − | − | −/+ | + | − | − | − | − |
| 9 | Past stem formant < -*īd*- | − | + | − | + ? | + | − | ? | + | + |
| 10 | No trace of ergativity | − | − | − | + | + | − | − | − | + |
|  | Sum of isogloss agreements with Maz. | 0 | 1.5 | 3.5 | 9 | **10** | 7 | 4 | 3.5 | 2 |

All living West Iranian languages are considered in this study, even if only eight appear in the table. First and foremost is Gilaki, another Caspian language besides Mazandarani. Gilaki's internal split into Eastern and Western Gilaki (Stilo 2001) plays a role in the outcomes of this comparative endeavor (Features 7 and 8).

In the south of Mazandaran, across the Alborz chain, rests the district of Semnan, hosting a language Sprachbund areally related to Mazandarani in one way or another. The languages in Semnan district are Semnani, which is a language isolate, Sangesari, another isolate, and Sorxe'i, Lāsgerdi, and Aftari, whose linguistic affiliations remain to be studied in detail. Among these languages Semnani (Sem.) and Aftari (Aft.) are selected for the comparative Table 7.

Another typologically related language is Balochi (Bal.), which is in reality a language group surmised to have been, on historical as well as linguistic grounds, originated in somewhere near the Caspian Sea. Tati, another southern neighbor of Caspian, has also a place in the table, owing to its grammatical similarities with Mazandarani and Gilaki. Moreover, the Central Plateau (CP) and Kurdish (Kurd.) language groups, although not geographically contiguous with Caspian, are included to stretch the spectrum of this typological investigation. Nevertheless, aiming at minimalism, the smaller languages of the northern half the West Iranian linguistic zone, namely, Talysh, Gorani, and Zazaki, are excluded from the table but included in the discussion that follows. Likewise with the languages of southern Iran, that is, the dialects of Fārs, Lārestani, and the dialects of southern Kermān, which have little bearing on our investigation. Persian, however, could not be left out since it serves the backdrop to the linguistic map of Iran.

Feature 1. *Modifiers:* In Mazandarani adjectives and possessives normally precede the head noun (HN), that is, in the reverse order of their arrangement in Persian. Examples:

(12)   possessives:   Maz. *dā́r-e valg*   vs. Pers. *bárg-e deraxt*   'leaf of a tree'
      adjectives:    Maz. *sabz-e valg*   vs. Pers. *bárg-e sabz*   'green leaf'

This feature places Mazandarani in accord with its neighbors Gilaki and Semnani, as well as with Tati and Balochi. See Stilo 2005: Isoglosses Six and Seven.

Feature 2. *Freestanding Possessive Determiners:* Mazandarani is areally bound to the languages that have freestanding possessive determiners preceding the head noun. The areal band includes the Semnan area and extends westward to Gilaki, Tatic (including Talysh), Caucasian Tat, as well as symbiotic Azeri Turkish, Armenian, Georgian, and Ossetic (see Stilo 1981: Isomap 3, p. 163). Examples are Maz.

*me bərâr*, Gil. *mi berâr*, Tati of Šāl *čemen berā* (Stilo 1981: Table Four), contrasting with Pers. *barādar-am* or *barādar-e man* 'my brother'. As demonstrated in Table 7, this feature follows a similar areal patterning as the previous feature.

Feature 3. *Adpositions:* Mazandarani lies within the geographic continuum where postpositions prevail over prepositions (Stilo 2006). This continuum entails Tati and the Semnani Sprachbund. On the other hand, in a panoramic study, Stilo (2005: Isogloss Eight) places Mazandarani and Gilaki in a buffer zone whose languages have both postpositions and prepositions. The latter classification, nevertheless, is only relative to a frame of reference that includes purely postpositional Turkmen and the languages of the Caucasus, and as such Stilo's classification does not violate the supposition in this study that Mazandarani and adjoining languages are chiefly postpositional.

Feature 4. *Pronominal clitics:* The enclitic pronouns typical to Persian (*-am*, *-at*, *-aš*, etc.) have no place in Mazandarani. The latter, like English, has a set of possessive determiners (Feature 2), abiding by the syntactic rule for genitival modifiers shown above under Features 1 and 2. Typologically, this feature unites Mazandarani with Gilaki, Talysh, Semnani and its neighbors, Zazaki, and North Kurdish but not with Central or South Kurdish, which make extensive use of the enclitic pronouns — thus a −/+ for Kurdish in the table. Tati too meets this isogloss halfway, but for another reason: while Tati has its own set of freestanding possessives, it allows pronominal clitics as well, e.g., Tati of Tākestān *čeme piar ~ piar-em ~ čeme piar-em* 'my father' (Stilo 1981: Isomap 2B). It is worth mentioning that Aftari, Sorxeʾi, Lāsgerdi, and Biābānaki in the Semnan region typologically pattern with Caspian in that none allows pronominal clitics in the noun phrase, even if they employ clitics to function as agents in the past tenses of transitive verbs, i.e., the last relic of ergativity (Feature 9) that remains in these languages.

How contiguous is the geography of this feature? Caspian, Semnan, Talysh, and Tati certainly do form a continuum. Zazaki is clearly a distant outlier in its present habitat but is hypothesized to have Caspian origins. It is also conjectured that Zazaki has heavily influenced Kurmanji syntax. This said, the distribution of this isogloss may somewhat be justified.

Feature 5. *Progressive tenses:* In a number of Iranian languages the progressive is distinguished from the habitual by employing various periphrastic forms, the most common of which employ either "have" or the locative verb as auxiliary.

The auxiliary "have" is used in Persian (*dâr-*: *dâšt-*), Semnani (*dɛr-*: *dɛrd-*), Central Plateau languages (e.g., Kupāʾi *dâr-*: *dârt-*), and the Biābānak language group (e.g., Xuri *der-*: *dâšt-*). Examples:

(13)  Pers.  *dārad mixorad / dāšt mixord*  'he is/was eating'
  Semnani  *dɛrd=eš mexā / mexord=eš*  'he was going/eating'
  Kupā'i (CP)  *dârt-em komze=m=o-vont*  'I was cutting the melon'
  Xuri  *deri deši / dâšta dešehi*  'you are/were going'

The locative verb is used as auxiliary in Mazandarani, Gilaki, and the vernaculars spoken around Semnan, including Sorxe'i, Lāsgerdi, Aftari, and Sangesari. Among them Gilaki is different in that it employs the infinitive succeeded by the auxiliary. The present stem of the locative verb is *dar-* in all of the languages, but the past stem is dissimilar, as illustrated in the following examples.

(14)  Maz.  *dar(mə) šumbə*  'I am going'
  Maz.  *dai(mə) šimə*  'I was going'
  West Gil.  *giftán-dərəm / giftán-dubum*  'I am/was taking'
  East Gil.  *gité-dərəm / gité-dəbum*  'I am/was taking'
  Aftari  *dara vāreš deyne*  'It is about to rain'
  Aftari  *daboa mon mon harf misseta*  'he was talking with me'
  Sangesari  *dab-i mi-šu-y*  'I was going'

Feature 6. *Present perfect:* In Mazandarani as well as Gilaki, the present perfect merges with the preterite. This feature makes the Caspian language family unique among the Iranian family. As is well known, the present perfect in Persian is built on the past participle, e.g., *âmadé-am* 'I have come'. A similar formation is seen in Tati, e.g., Čāli *be-ttaté-mō* 'we have run'. North Kurdish differentiates the present perfect in personal endings: *hatˤ-ime* 'I have come' vs. *hatˤ-im* 'I came'.

In terms of the present perfect tense, the Semnan area seems to offer a transition zone. We find no present perfect in Šahmirzādi, a Caspian outlier lying upstream from Semnan. Nor do we find the tense in Aftari, which is spoken on the trade route from Semnan to Mazandaran via Firuzkuh. The dialect of Sangesar, also upstream of but closer to Semnan, has distinct forms for the present perfect only in the third person singular and plural, e.g., *bəšu-ya* 'he has gone' vs. *bəšö* 'he went'. Sorxe'i does differentiate the present perfect (*be-xot-ešt-im* 'we have slept' vs. *be-xot-im* 'we slept'), but not in the third person singular. Finally, Semnani proper differentiates the tense for all persons, e.g., *be-ši-č-on* 'I have gone' vs. *bi-š-on* 'I went'.

Feature 7. *Present tense formation:* As noted above, the Mazandarani personal endings specific to the present indicative can be reanalyzed diachronically as ordinary endings accompanied by a nasal element (derived from the Old Iranian participial marker *-ant-) inserted between the stem and the ending. As shown

by Gernot Windfuhr, this nasal formant is a major isogloss that unites Mazandarani with Sangesari, Parachi, Zazaki, and Harzandi (Azami and Windfuhr 1972 : 198). To this list we should add Eastern Gilaki (e.g., *gÛ-n-ə* 'he says')[25] and Aftari, Sorxeʾi, and Lāsgerdi in the Semnan area, e.g., Aftari, Sorxeʾi *xos-enn-im*, Lāsgerdi *xos-end-im* 'we sleep'. Note that Western Gilaki, Semnani proper, and the Tati dialects other than Harzandi stay outside of this isogloss zone.

**Feature 8.** *Imperfect:* The imperfect in Mazandarani and Eastern Gilaki is built on the past stem without any further tense marker, e.g., Maz. *xurd-ə* 'he used to eat', Eastern Gilaki *xurd-∅* 'id'. This feature sets Mazandarani and Eastern Gilaki apart from the other languages employed in this comparative study, as the latter normally mark the imperfective with a prefix: North Kurdish *di-*, etc.; Central Plateau *e(t)-*, etc.; Aftari, Semnani, Tati *mi-*. The latter prefix is evidently borrowed from Persian, which have grammaticalized it from the adverb *hamē*. It is worth mentioning that the highly contagious *mi-* has even permeated Šahmirzādi, the Caspian variety spoken in the Semnan area, even if just in the past tense.

**Feature 9.** *Past stem formant:* Most New Iranian languages build their "regular" past stem on the present stem suffixed by one of these formants: (1) *-i* or *-id*, corresponding to Middle Persian *-īd*; (2) *-â* or *-âd*, corresponding to Parthian *-āδ*; and (3) *-(e)ss*, corresponding to Middle Persian *-ist*. The output of this formant can best be tested in the causative past stem (see above). A morphologically productive formation in most Iranian languages, the causative past stem requires a "regular" pattern. Examples of causative past stems:

(15) Mazandarani     *xəs-ənd-i-*            'put to sleep'
       Persian           *par-ān-(i)d-*       'make fly'
       Gilaki             *pər-an-e-*         'make fly'
       Aftari             *tāj-en-ā-*          'make run'
       Xuri (Biābānak)   *kɔn-en-â-*          'make laugh'
       Ārāni (CP)        *tej-ən-ād-*         'make run'
       Kešeʾi (CP)       *vaz-n-â-*           'make run'
       Jarquʾi (CP)      *xous-n-ā-*         'put to sleep'
       Čāli (S. Tati)     *u-xos-en-ass-*     'put to sleep'
       Koroshi (Balochi)   *časp-ed-*         'glue'

---

**25** Stilo (2001).

Feature 10. *Ergativity:* The split ergative alignment that emerged in Middle Iranian transitive past has been inherited in various capacities by most New West Iranian languages. While we find a near full rendition of split ergativity in North Kurdish and Semnani, this outstanding syntactic feature has more often survived in reduced forms. All that has remained from the original, full form of ergativity in, for instance, Sorxeʾi and Lāsgerdi, is merely a distinct set of verb endings, based on pronominal clitics, in transitive past. On the other hand, transitivity is not differentiated in Mazandarani, likewise with Gilaki and Persian. Lack of ergativity is thus a highly distinctive syntactic feature of Caspian.

Having included ergativity as an isogloss, one may reasonably ask why I did not include grammatical gender, which is equally lacking in Mazandarani, as it is in Gilaki, Aftari, Balochi, and Persian, while it exists in other languages considered in this study. The reason why gender is inapt as an isogloss is that among many varieties of Kurdish, Tati, and Central Plateau, only certain dialects of each language group possess grammatical gender. The same argument holds for the two-term case system that exists in various degrees in many Iranian languages but not in Caspian.

# 5 Conclusions

The last row of Table 7 sums up the instances of isogloss agreement with Mazandarani for each listed language. It comes as no surprise that Gilaki ranks highest by scoring nine out of ten isogloss agreements. Gilaki is, after all, sister to Mazandarani in the Caspian language family.

All other languages, as the scores reveal, fail to form as tight a typological bond with Mazandarani. Among them, Aftari shows the highest degrees of relatedness (70%). Aftari, together with Semnani, which wins the next rank though with substantial distance (40%), belong to a Sprachbund that is designated as Kōmeši or Komisenian (Borjian 2008), after the old name of the province lying immediately south of Mazandaran. They were treated by earlier philologists, first and foremost by Wilhelm Geiger (1898–1901), as Caspian,[26] even if Semnani proper is genetically distant from the Caspian group.

Next in rank are Balochi and Tati, each showing agreement in about one-third of the isoglosses. Low concordance notwithstanding, some interesting

---

**26** Geiger's notion of Caspian languages is strictly geographical; it includes Mazandarani, Gilaki, Talysh, the Caucasian Tat, and Semnani. At the turn of the twentieth century, Geiger had no knowledge of the Tati group.

inferences can be drawn if we narrow our focus down to the rewarded isoglosses, that is, the top three features of Table 7. These features alone bear significant weight in Iranian typology, owing to their correspondence with the three of the eight isoglosses included in Stilo's (2005)[27] seminal study of West Asia that cut the deepest crevices within the Iranian language zone. The central top box inside Table 7 corresponds to what Stilo sorts out as "Left Branching" or "Modified–Head" languages. Note that all the Modified–Head languages of our table are geographically contiguous except Balochi, whose geography gravitates toward the southeastern corner of the Iranian Plateau. An explanation of this unexpected association can be sought in the historical position of Balochi somewhere in the vicinity of the Caspian region. Does this mean that Balochi, despite its decisive areal shift, has successfully retained its Caspian features for a millennium?[28] Probably not. Balochi may have as well attuned its syntax with its current neighbors in South Asia. The latter scheme becomes even more likely if we cross-examine the same premise on Zazaki and Gorani, which are also believed to have originated from the Caspian region. Yet neither Zazaki nor Gorani shows agreement with Caspian on the top three features of Table 7, leading to the assumption that, whatever their provenances may have been, they must have been syntactically immersed into the surrounding Kurdish varieties.

The power of areal adjustment manifests itself in the split between the Tati and Central Plateau groups as well. The latter language group shows no agreement with Mazandarani in any of the ten isoglosses, hence carrying none of the three Modified–Head features of Table 7. This attribute may be explained by the longstanding historical contact between Tati and Caspian (and Tati and Turkic), on the one hand, and, on the other, between Central Plateau and Persian, notwithstanding the Median pedigree shared by Tatic and Central Plateau.

# Abbreviations

| : | Separates present stem from past stem |
|---|---|
| Aft. | Aftari |
| Bal. | Balochi |
| CP | Central Plateau languages |
| Gil. | Gilaki |

---

**27** Stilo's corresponding isoglosses are: Isogloss Six: adjective–noun, Isogloss Seven: genitive–noun, and Isogloss Eight: object–postposition.

**28** Historical records show that Caspian was a Modified–Head language even a millennium ago. See Borjian, forthcoming.

| HN | head noun |
|---|---|
| Kurd. | Kurdish |
| Maz. | Mazandarani |
| NP | noun phrase |
| Pers. | Persian |
| PL | plural |
| PST | past |
| PTCP | participle |
| PRS | present |
| PROG | progressive |
| Sem. | Semnani |
| SBJV | subjunctive |
| SG | singular |
| Tat. | Tati |

# References

Azami Sangesari, Cheragh Ali, and Gernot Windfuhr. 1972. *Sangesari dictionary*. Tehran: Šerkat-e sahāmi-e ketābhā-ye jibi.

Borjian, Habib. 2004a. Mazandaran: Language and people – The state of research. *Iran and the Caucasus* 8 (1). 289–328.

Borjian, Habib. 2004b. Personal and reflexive pronouns in Mazandarani. *Orientalia* (Yerevan) 2, 7–13.

Borjian, Habib. 2005. Šenāse-ye feʿl dar māzandarāni-e šarqi [Personal endings in eastern Mazandarani verbs]. *Guyeš-šenāsi* (3). 13–19.

Borjian, Habib. 2008. The Komisenian dialect of Aftar. *Archiv Orientální* 76 (3). 379–416.

Borjian, Habib. 2009. *Motun-e Ṭabari / Tabari Texts*, Tehran: Mirāṣ-e maktub.

Dargāhi, Zeyn-al-ʿĀbedin. 1993. Barrasi-e didgāhhā dar mowred-e feʿl-e goẕašte-ye sāde o naqli dar dastur-e māzandarāni [Review of the studies on the simple past and the present perfect in Mazandarani]. I. H. Ṣamadi (ed.), *Dar qalamrav-e Māzandarān* 2, Qāʾemšahr: DSL, pp. 45–53.

Geiger, Wilhelm. 1898–1901. Die kaspischen Dialekte. In W. Geiger and E. Kuhn (eds.), *Grundriss der Iranischen Philologie* I/2, 344–380. Strassburg: Trübner.

Rastorgueva, V. S., and D. I. Èdel'man. 1982. Prikaspijskie yaziki: gilanskij, mazanderanskij (s dialektati shamerzadi i velatru) [Caspian languages: Gilaki, Mazandarani, (with the dialects of Šāhmirzād and Velātru)] In *Osnovy iranskogo jazykoznanija* [III.] *Novoiranskie jazyki: zapadnaja gruppa, prikaspijskie jazyki* [The fundamentals of Iranian linguistics. New Iranian languages: Western group, Caspian languages], 447–554. Moscow: Academy of Sciences.

Ṣādeqi, ʿAli-Ašraf. [1369] 1990. Se asar darbāre-ye lahje-ye māzandarāni [A review of three monographs on Mazandarani]. *Majalle-ye zabānšenāsi* 7 (2, Autumn and Winter). 63–71.

Šokri, Giti. 1995. *Guyeš-e Sāri* [The dialect of Sari]. Tehran: Pažuhešgāh-e ʿolum-e ensāni.

Stilo, Donald. 1981. The Tati language group in the sociolinguistic context of Northwestern Iran and Transcaucasia. *Iranian Studies* 14 (3–4, Summer–Autumn). 137–187.

Stilo, Donald. 2001. Gilān x. languages. *Encyclopaedia Iranica* 10, fasc. 6: 660–668.

Stilo, Donald. 2005. Iranian as buffer zone between the universal typologies of Turkic and Semitic. In É. Á. Csató, B. I. Isaksson, and C. Jahani (eds.), *Linguistic convergence and areal diffusion*, 35–64. London: Routledge.

Stilo, Donald. 2006. Circumpositions as an areal response: The case study of the Iranian zone. In L. Johanson and C. Bulut (eds.), *Turkic-Iranian contact areas*, 134–152. Wiesbaden: Harrassowitz Verlag.

Windfuhr, Gernot L. 1989. New West Iranian. In Rüdiger Schmitt (ed.), *Compendium linguarum Iranicarum*, 251–262. Wiesbaden: Reichert.

Yoshie, Satoko. 1996. *Sāri dialect*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Zav'jalova, V. I. 1956. Novye svedenija po fonetike iranskix jazykov. Gilanskij i mazanderanskij jazyki [New information on the phonetics of Iranic languages. Gilaki and Mazandarani languages]. *Trudy Instituta jazykoznanija AN SSSR*, Moscow 6: 92–112.

Geoffrey Haig and Shirin Adibifar

# Referential Null Subjects (RNS) in colloquial spoken Persian: Does speaker familiarity have an impact?

## 1 Referential Null Subjects (RNS): Background

It is well known that languages differ considerably in the extent to which a clause requires an overt subject NP. Some languages, like Persian, tolerate clauses without overt subject constituents in a very wide range of contexts, while others (e.g., English) only permit referential null subjects (RNS) under highly constrained conditions. In the literature, two approaches to these cross-linguistic differences can be discerned: a parametric approach, and a discourse, or usage-based approach. The parametric approach goes back to Perlmutter (1971), who introduced a "pro-drop parameter", according to which a language requires, or does not require, overt expression of referential subjects. The pro-drop parameter was exclusively concerned with zero subjects, but the original either/or pro-drop parameter has since given way to more refined typologies, involving four distinct types (e.g., Holmberg 2009), and has been extended to include Referential Null Objects under the label "radical pro-drop" or "discourse pro-drop" (e.g., Neeleman and Szendröi 2008). Within parametric approaches, the presence of RNS is often linked to the presence of rich agreement morphology, which apparently licenses RNS, though what exactly constitutes rich agreement continues to be a matter of controversy (see Camacho 2013 for recent discussion of null subjects and rich agreement).

While research in the parametric tradition continues to perceive RNS as a parameter of individual grammars, a second line of research is usage, or discourse based. On this view, RNS is a locus of gradual variation, thus not entirely determined by "the grammar" of a language, but also dependent on contextual and interactional factors. Methodologically, this approach adopts empirical, quantitative methodologies, drawing on the analysis of language usage ("performance") rather than on intuitions regarding grammaticality. Within language

typology, this line of research is associated with Bickel and associates' work on Referential Density (RD) (Bickel 2003, Stoll and Bickel 2009). RD is an empirical measure of the overall density of overt argument expressions in actual discourse, and is not restricted to subjects and objects. Thus RD is not conceived as a categorical feature of the grammar of "a language", but as a graded value, characterizing a specific stretch of discourse. Unlike pro-drop and its descendants, RD does not yield an either/or parameter setting for a particular language, but exhibits a certain degree of language-internal variation, depending on text type and other factors to be taken up below, and is a tool of corpus-based, rather than grammar-based typology (Haig et al. 2011).

Probably the best-known empirical approach to RNS is within variationist sociolinguistics. Rates of overt vs. zero expression of subjects have been extensively investigated as a linguistic variable, most notably across different varieties of Spanish (see Pešková (2013: 120–121) and Carvalho et al. (2015) for discussion of the relevant literature). Among the main findings of this research is the recognition that dialects of "the same" language can vary quite considerably, a finding which is of considerable relevance in connection with a large and dialectally diverse language such as Persian. This chapter adopts a usage-based perspective on RNS in Persian, drawing on corpus of colloquial spoken Persian and applying quantitative methods to address the issue of which factors are relevant in affecting the rate of RNS in natural discourse. Thus the assumption is that RNS is a variable, rather than categorical feature, and it is the analyst's task to determine the factors which drive the variation.

With the exception of Saeli and Miller (2018)[1], there has been no comparable quantitative research on colloquial spoken Persian. In our study, we focus on the factor of "familiarity" between the interlocutors, which has been suggested as relevant in this regard, but we also consider gender of the speaker. Although the current sample is small (see Section 3), our provisional finding is that rates of RNS in Persian are not sensitive to either speaker familiarity, or gender, but in fact emerge as a relatively stable variable across a range of different speakers.

---

**1** Saeli and Miller (2018) are concerned with the impact of extra-linguistic factors on colloquial spoken Persian, including the issue of pronoun omission, based on elicited responses to a "favor-asking" task. However, the pronouns concerned in their research are second person forms, rather than the third person forms that dominate in our data. They find an effect of same vs. different gender in speaker diads, but the absolute number of second person subject forms in their data is just twenty-five (including tokens of the polite pronoun *šomā*, familiar pronoun *to*, and zero, cf. Saeli and Miller 2018: Table 2, p. 180). Nevertheless, this is a promising avenue for future research that complements the current study, both in methodology and the domain of investigation.

These findings echo to some extent the picture from research on better-studied languages, in particular Spanish, which show that with regard to RNS, it is primarily language-internal factors that determine most of the attested variation (see next section).

# 2 Factors determining rates of RNS: Previous research

The most detailed research on the factors impacting on RNS stems from the variationist sociolinguistic tradition within Hispanic linguistics (see the contributions in Carvalho et al. 2015).[2] All investigations to date confirm that the primary determinants of subject expression are language internal, with a surprisingly high degree of overlap across different studies with regard to the nature of the relevant factors. The impact of speaker-related factors (e.g., age or gender), on the other hand, has not been consistently demonstrated. Among the linguistic factors, the following are worth mentioning.

**Person and number value of the pronoun**
This appears to be the highest-ranking factor in determining rates of subject omission. Pešková (2013) notes significantly higher rates of pronoun expression in the first and second person as opposed to the third person, while Carvalho et al. (2015) state that the "broadest generalization" is that singular pronouns are more frequently overt than plural pronouns.

**Distance and role of antecedent**
As a general finding, subject omission is favored when the antecedent is subject of the immediately preceding clause, with rates of pronoun retention increasing with increasing distance of the antecedent.

---

**2** Most of these studies use some measure of pronoun omission or retention as the unit for investigating what I have termed RNS. Thus the unit of comparison is not zero subjects as a percentage of all subject NPs (as it is here), but rates of pronominal versus zero subjects. Nevertheless, both measures are ultimately concerned with the same phenomenon, though the resultant figures are not directly comparable.

### TMA morphology of the verbs

In Spanish, different tense/aspect values are associated with different types of agreement patterns. Carvalho et al. (2015) suggest that pronouns are used more frequently with verb forms with the least ambiguous agreement paradigms.

### Lexical semantics of the verb

Peškova (2013), working with elicited data, finds epistemic verbs ("know", "believe") have higher rates of pronoun retention than perceptive verbs.

The ranking of these factors varies from study to study; nevertheless, there seems to be a broad consensus that person and number of the pronoun is the most predictive factor. Turning to the speaker-related factors, the findings here are less consensual. Several studies find an effect of gender. Alvaraz (2015), based on the Spanish of Santo Domingo, finds a weak preference for pronoun retention among women, as does Orozco (2015) for Colombian Costeño Spanish, though the latter case also shows interaction with age. Peškova (2013), on the other hand, does not mention an effect of speaker gender in her investigation, which, unlike the others discussed in this paragraph, is based on a controlled production experiment. Age is also reported as relevant, but the direction of the correlation is not consistent. Orozco (2015) reports that in Mexico City, younger speakers use fewer overt pronouns, while in Puerto Rican Spanish the opposite trend is found. Genre (arguably an internal factor) is also reported as relevant: argumentation favors overt pronouns, while narration favors pronoun omission (Carvalho et al. 2015).

A further factor that has been discussed in this connection is the degree of familiarity between the interlocutors. Bickel (2011), investigating overall rates of zero argument expression (Referential Density, RD), claims an effect of degree of personal familiarity: where speaker and addressee are personally acquainted, fewer arguments receive overt expression, while lack of personal familiarity leads to higher rates of overt arguments. A related claim is made by Meyerhoff (2011), who discusses rates of subject pronoun deletion in Bislama, the English-based creole of Vanuatu. She compared two versions of the same story, one recounted by a native speaker to his extended family, and one version recounted by the same speaker to the investigator (i.e., an out-group person). Rates of subject pronoun omission were nine percentage points higher with familiar addressees than with the out-group addressee. Meyerhoff (2011: 45) suggests that the higher frequency of subject pronouns used with the out-group addressee may be motivated by the speaker's desire to provide "a non-native speaker with more overt information about who he is referring to in any given sentence".

These findings point to an intuitively plausible impact of speaker familiarity on rates of argument realization: when speakers are addressing persons with whom they are familiar, they can afford to reduce overt informational density, relying on the shared body of cultural knowledge and the addressee's assumed familiarity with the speaker's speech habits to fill in the gaps. When addressing a stranger, however, the speaker cannot assume shared cultural knowledge and familiarity with routinized speech habits, and will accordingly switch to a more explicit style, leading to an overall higher level of overt argument expression. If speaker familiarity is indeed a factor in affecting rates of overt vs. zero subject expression, this would be in line with approaches to linguistic variation which focus on accommodation to the addressee, such as Bell's "Audience Design" (Bell 2006).

# 3  Research question and data

Persian is a southwest Iranian language of the Indo-European family, and the official language of the Islamic Republic of Iran.[3] It exhibits a mixed word-order typology, with OV order in the clause, but with head-initial ordering elsewhere. With regard to RNS, it has been claimed that "Persian is also a radical pro drop language with frequent use of null arguments in both subject and object positions" (Sato and Karimi 2016: 3). However, we are unaware of any empirically based approaches to RNS in Persian to date. In this chapter, we investigate RNS in a corpus of spontaneous spoken Persian and investigate the role of a number of linguistic and speaker-related factors. The main focus is on the factor of speaker familiarity, as discussed in the preceding section: Do speakers tend to use more RNS when they are personally familiar with their interlocutors?

Although there is no previous research on this specific issue in Persian, we nevertheless considered that Persian could be a potentially interesting laboratory for investigating factors such as interlocutor familiarity, because Persian is characterized by an elaborated range of registers and styles. Speakers are highly sensitive to degrees of formality and to politeness norms, adapting phonology, lexical choices, address forms, and grammar accordingly (Jahangiri 1980, Saeli and Miller 2018). Thus it seemed a reasonable hypothesis that in a language community where speech habits are intimately tied to social status and familiarity, the likelihood of an effect of speech setting on RNS would be high. In order to test this, we compiled a corpus of spontaneous spoken Persian (see next section),

---

**3** We continue to use the term traditionally used in the western academic tradition "Persian", although the speakers refer to their language as Farsi.

under conditions that were controlled for speaker familiarity, and analyzed the resulting data quantitatively.

Finally, we note that in Persian, finite verbs obligatorily agree with their subjects via a set of six distinct person and number suffixes on the verb. There is a set of free pronouns, which may be omitted under conditions of pragmatic recoverability, and which are flagged for syntactic function in the same manner as nouns (i.e., with the accusative clitic =*rā*, or via various prepositions). With respect to "pro-drop", then, these are the relevant pronouns. Verbs do not agree obligatorily with objects, though objects may be indexed on the verb through a set of clitic pronouns. The clitic pronouns are briefly mentioned in connection with certain predicate types in examples (4)–(6), but are otherwise not relevant here (see Rasekh 2014, Mahootian and Gebhardt 2018, and Haig, under review, for discussion of clitic pronouns and agreement).

# 4 Experiment design and setting

The aim of the study is to test whether speaker familiarity has a significant impact on RNS. In order to test this, we gathered data from twenty-nine native speakers of Persian, with the speakers divided into two groups on the basis of their degree of familiarity with the interviewer (who remained the same throughout).[4] One group included only persons who were either connected to the interviewer through a kinship relationship, or a close personal friendship of at least two years. Interviews with this group took place in a relaxed domestic setting in the region of the interviewer's home town in the Mazanderan region of northern Iran, and in three cases in the speakers' apartments in south Germany. Respondents from the second group had no prior contact to the interviewer. They were recruited among students via their lecturers from the Islamic Azad University in Tehran and Behšahr University in Mazanderan Province. Interviews with these speakers were conducted in seminar rooms of the respective universities, thus heightening the contrast in settings between the two groups.

All interviews took place entirely in Persian. The methodology largely replicates that of Bickel (2003), though with minor modifications: respondents were shown the Pear Story, a six-minute video clip widely used in cross-linguistic investigations of discourse (Chafe 1980), on a laptop computer, and then asked to

---

**4** Originally thirty interviews were conducted, but one speaker did not produce a coherent narrative that would have been comparable to the other texts, and that text was excluded. This left two groups with fifteen and fourteen speakers, respectively (see Appendix A for details).

recount the story to a native-speaker interviewer.[5] The film contains no speech, but the storyline is simple and can be readily grasped by those watching the film. Pear Story retellings have been widely used in cross-linguistic studies of discourse, so that the resulting corpus of Persian is also of considerable utility for future researchers. The entire corpus with annotations is available under a Creative Commons License Agreement,[6] and is thus available for re-analysis or re-interpretation by other scholars.

The sample of respondents was intended to be representative of educated, young adult, native speakers of standard Persian, socialized in an urban environment. Prior to the recordings, all speakers provided basic information regarding age, gender, education, places of socialization, languages of communication (in and outside of the domestic setting), and language of their parents. Prior to the recordings, all speakers received the same set of instructions in Persian, provided by the interviewer, a female educated native speaker of Persian from the same age cohort.

Recordings were transcribed, translated, and syntactically annotated using the GRAID system, which provides a set of decision procedures for identifying zero arguments (Haig and Schnell 2014: 7–8; Haig and Schnell 2016). Transcriptions, translations, and annotations were entered into the software ELAN, which time-aligns annotations with the sound file.[7]

## 4.1 Issues in coding and analysis

The concept of "subject" has been variously defined at different times, and in different approaches to syntax. Whether or not all clauses, in all natural languages, should be analyzed in such a way that they "have" (at some level of analysis) a subject is an open question. But on the assumption that a very significant number of clauses in a very significant number of languages can be analyzed in this manner is sufficient justification for maintaining it as a concept of syntactic theory.[8] We thus follow mainstream practice and assume that subjects can be

---

**5** In this respect, our methodology departs from that of Bickel (2003) and Chafe (1980) in that the respondents recounted the story to the same interviewer who showed them the film, rather than to another person. Given the aims of the experiment, it was crucial to keep the identity of the interviewer constant across all groups in order to reduce the impact of factors outside of the main dependent variable, that of speaker familiarity.

**6** See https://lac2.uni-koeln.de/en/multicast/.

**7** Developed by Han Sloetjes at the MPI Nijmegen, see https://tla.mpi.nl/tools/tla-tools/elan/.

**8** Within various versions of Generative Grammar, the subject role is generally derived from a particular structural configuration, for example as the Specifier of an IP in a GB approach (Farrell 2005: 176), or in Minimalism as, e.g., an NP that is c-commanded by a finite complementizer

relatively uncontroversially identified for Persian, though we note some problematic cases below.

The basic unit of analysis is the clause unit, consisting of a predicate plus associated arguments. For each clause unit, the subject constituent is identified and coded as either full (or lexical) NP, pronoun, or zero. Example (1)[9] shows a clause unit with an overt lexical subject NP. Example (2) contains a sequence of three clause units, the first with an overt subject NP and the second and third clauses with zero subjects. Example (3) contains a sequence of clauses with zero subjects (clause (3c) also contains a zero object).

(1)  *bad   yek   pesar-i      mi-yā-yad*
    then  one   boy-INDEF   INDIC-come.PRS-3SG
    'then a boy comes by' (g1_f_08/06)

(2)  a.  *in     pesar-e   bā     dočarxe  āmad*
        this   boy-DEF   with   bike     come.PST.3SG

    b.  *Ø   rad        šod*
        Ø   passing    become.PST.3SG

    c.  *Ø   raft*
        Ø   go.PST.3SG

    a.  'This boy with the bike came along
    b.  passed by
    c.  went.' (g2_f_06/09)

---

(Radford 2004: 136), or via checking of nominal features (Farrell 2005: 181). Within LFG and related theories, the subject role is a non-derived category within the layer of structure known as F-Structure. In less formalized, but typologically inspired approaches to syntax, various "cluster-concept" notions of subjecthood have been put forward involving structural, semantic, and information-structure-related properties. These were pioneered in Keenan (1976); see Comrie (1989: 104–123) and Falk (2006: 1–21) *inter alia* for discussion. Philippine-type and syntactically ergative languages continue to pose certain challenges for a universal definition of subject, but these lie outside the scope of this chapter.

**9** All examples are sourced according to the group (g1 = familiar speakers, g2 = unfamiliar speakers), gender (m/f), and number of the recording. Abbreviations used in the examples are: ACC = accusative; ADD = additive particle; AUX = auxiliary; DEF=definite; INDEF = indefinite; INDIC = indicative; PL = plural; POSS = possessive; PROG = progressive; PRS = present; PST = past; SG = singular.

(3)  a.  *bad   Ø   mive-hā=rā    čid*
         then  Ø   fruit-PL=ACC   pick.PST.3SG

     b.  *va   Ø   āmad         pāyin  va*
         and  Ø   come.PST.3SG  down   and

     c.  *Ø   Ø   rixt          tuye  sabad*
         Ø   Ø   pour.PST.3SG  into  basket

     a.  'Then (he) picked the fruit
     b.  and came down and
     c.  (he) poured (them) into the basket.' (g1_m_04/2)

Persian has one type of clause which poses certain difficulties for identifying subjects. Semantically, these involve predicates of perception and cognition. Syntactically, they are typically lexicalized combinations of a light verb and some non-verbal element. The NP expressing the Experiencer, if present in the clause, is in the nominative case, but is obligatorily indexed through a possessive clitic attached to the non-verbal element of the complex predicate. The light verb takes the default 3SG person agreement marker. The commonest expression of this type in our corpus is *havās=aš part* šodan 'attention=3SG separated become', i.e., 'to be distracted'. Examples of experiencer predicates are found in (4), (5b), and (6b):

(4)  *Ø   češm =aš       in    sabad-hā=rā    gereft*
     Ø   eye=POSS.3SG   this   basket-PL=ACC  take.PST.3SG
     '(He) caught sight of these baskets (lit. his eye took the baskets)' (g1_f_05/5)

(5)  a.  *yek  doxtarxānum-i  dāšt         bā    dočarxe  miy-ām-ad*
         one  girl-INDEF     AUX.PST.3SG  with  bicycle  PROG-come.PST-3SG

     b.  *ke   Ø   havās =aš         be   u    part        šod*
         so   Ø   attention=POSS.3SG  to   3SG  separated   become.PST.3SG

     a.  'a girl was coming by on a bike
     b.  so his attention was distracted to her ...' (g2_m_08/07)

(6)  a.  *kolāh =aš       mi-oft-ad*
         hat =POSS.3SG   INDIC-fall.PRS.3SG

     b.  *bad  in   ham  havās =aš          part      mi-šav-ad*
         then  3SG  ADD  attention=POSS.3SG  separated  INDIC- become.PRS.3SG

a. 'His hat falls off,
b. then he gets distracted' (lit. he his.attention becomes separated)'

(g1_f_14/13)

The correct analysis of such constructions is a matter of some debate (see, e.g., Ghomeshi, forthcoming). We follow Sedighi (2010) and assume that the experiencer constituents of these predicates are subjects, because they exhibit most of the syntactic characteristics of canonical subjects in Persian, and we therefore include them in the overall counts for subjects. However, they are Non-Canonical in the sense that the nature of the agreement morphology they are associated with differs from the agreement morphology associated with canonical subjects in Persian (see Haig (2008: 19–22) for discussion of Non-Canonical Subjects with reference to Iranian languages). Rather than a verbal affix, the agreement morphology is an obligatory clitic, e.g., =*aš* in (5b), which we thus analyze as non-pronominal in this context. What this means is that in (4) and (5b) we count a zero subject, while in (6), we count the pronoun (actually a proximal demonstrative) *in* as a pronominal subject.[10]

Subordinate clauses, including relative clauses, generally involve finite syntax in Persian and are thus not significantly different from independent clauses. We have therefore included them in the data, but followed the procedure of Bickel (2003) in considering only those subject constituents that could be overtly realized, without impairing grammaticality. Where unequivocal decisions could not be reached, the string was marked as "nc" (not classifiable), and excluded from the counts.

Rate of RNS (or simply "RNS") was calculated by dividing the number of zero subject constituents in a given text by the overall number of subjects in that text, yielding a figure between zero and one. For example, the speaker g1_m_1 has an RNS value of 0.558, indicating that somewhat more than half of all clauses in his text contained an RNS. The mean value for RNS across all speakers was 0.589; see Appendix A for details.

## 4.2 Variables and hypotheses

The main dependent variable is zero versus overt expression, or more generally, rates of RNS, calculated as the rates of zero subjects against the total

---

**10** We interpret the 3sg pronoun/demonstrative *in* in this example as an overt pronominal expression of the Non-Canonical Subject in the second clause, triggered by the subject change between the first clause (*kolāh=aš* 'his hat' and the implied subject of the second clause (the boy).

number of subjects produced, yielding values between 0.0 (no subjects are zero) and 1.0 (all subjects are zero). Our main aim was to test the effects of speaker familiarity on rates of RNS, but we also considered a number of other predictor variables. These include two speaker-related factors, and two linguistic variables.

### 4.2.1 Speaker-related variables

#### Age and gender
Although the available literature yields no obvious hypothesis regarding the effects of these two variables (cf. Section 2), we include them as standard variables in variationist research.

### 4.2.2 Linguistic variables

#### Number of clause-units (CUs) in each text
Each text is an individual re-telling of the Pear Story film, produced by one speaker. The different speakers actually produced texts of very varied length, measured as the number of CUs (mean 49, SD 21). Some speakers produced an exceedingly brief, almost telegraphic, re-telling, while others were quite elaborated. We assumed a possible effect of length on rates of RNS, based on the following assumption: Given that these narratives contain approximately the same content, all other things being equal, a longer text would offer greater opportunities for zero expression, because zero expression is connected to discourse persistence; a participant to which repeated reference is made over consecutive clauses is more likely to be coded with zero, hence yielding an overall higher rate of RNS. The initial hypothesis with regard to length, then, is that length correlates with higher rates of RNS.

#### New referents per clause unit (NewRef/CU)
This variable relates to the notion of "Information Pressure" (Du Bois 1987): texts differ in the extent to which they accommodate new information (the introduction of new referents). Some texts recount the continued actions of a small number of protagonists, while others involve repeated introductions of new referents. The latter are characterized by what Du Bois (1987) refers to as "high information pressure", measured in terms of new referents per clause unit. The general assumption is that higher information pressure would correlate with

lower rates of RNS, because new referents involve overt expressions, as opposed to zero (see Stoll and Bickel 2009 for counter-examples, and Haig and Schnell (2016) for critical discussion of Information Pressure). We therefore counted for each text the number of new referents introduced, restricting ourselves to individualized entities introduced in the form of a NP, and potentially pronominalizable, yielding an absolute figure of new referents per text (mean 15, SD 4). We then divided that figure by the number of clause units (cf. preceding variable), yielding the rate of new referent introduction per clause unit for each text. The hypothesis is that high information pressure will correlate negatively with rates of RNS.

# 5 Results

The absolute figures from the twenty-nine transcribed and coded texts are provided in Appendix A. Figures 1 and 2 show the results of the linguistic variables 'Length of text in CUs' (Figure 1), and 'New Referents per CU' (Figure 2), while Table 1 provides the Pearson Correlation Tests.
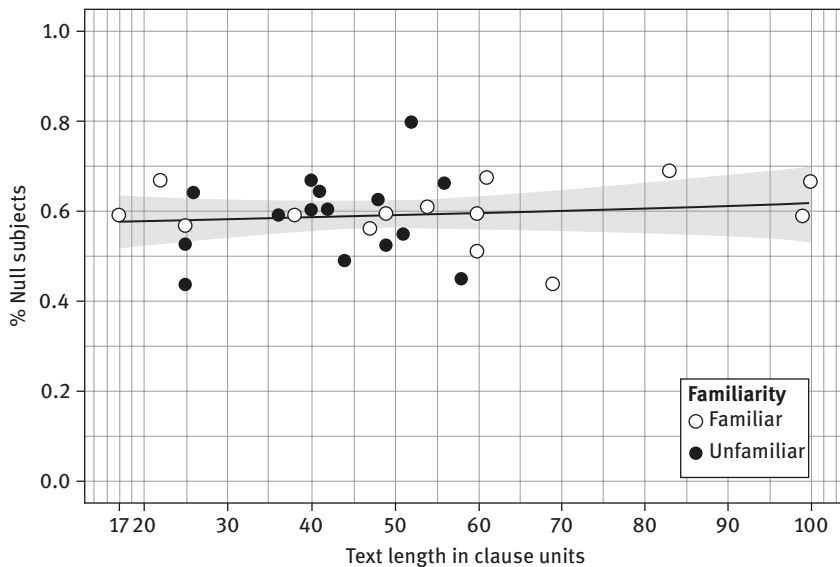


**Figure 1:** The effect of text length on RNS.

**Table 1:** Pearson correlation tests for Figures 1 and 2.

| Factor | r | p |
|---|---|---|
| RNS | text length: | 0.1212 | 0.531 |
| RNS | newRefs/CU: | −0.2020 | 0.294 |
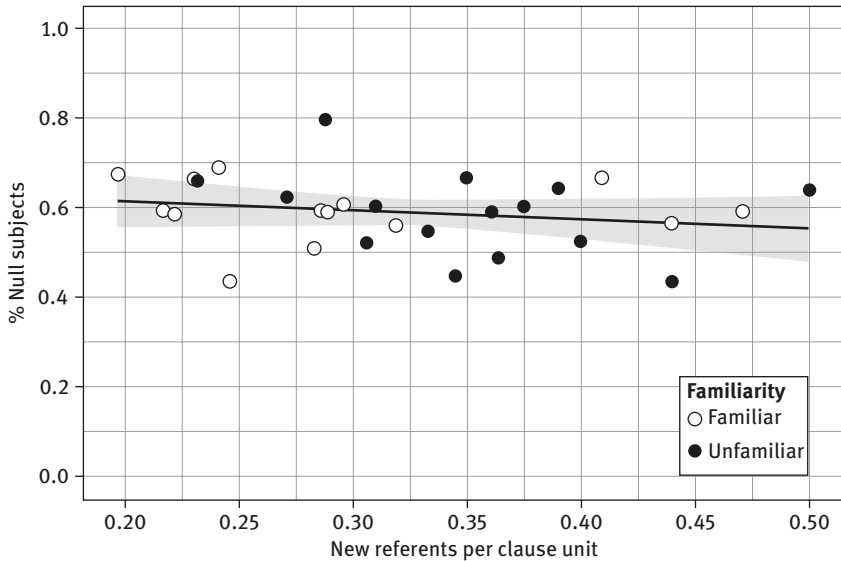


**Figure 2:** The effect of new referent per CU on RNS.

The tests suggest neither length of text, nor density of new introductions per clause unit, correlates significantly with rates of RNS. Although the weak negative correlation of New Referents with RNS indicated in Figure 2 points in the expected direction of the hypothesis, it does not reach significance.

Turning to the non-linguistic factors of familiarity, age, and gender, it likewise turns out that none of them appear to impact significantly on rates of RNS. Figure 3 provides the results based on the division into two groups, familiar and non-familiar, and Figure 4 the results according to speaker gender.

The box-plot in Figure 3 suggests that unfamiliarity leads to greater range of values than familiarity, but the overall mean of both the familiar and unfamiliar groups is similar, and the ANOVA test (Table 2) reveals no significant effect of speaker familiarity. Likewise gender does not appear to make an obvious difference. Age was also tested, but given the generally homogenous age grouping in the sample (the speakers' ages ranged from 20 to 39), age was not expected to be
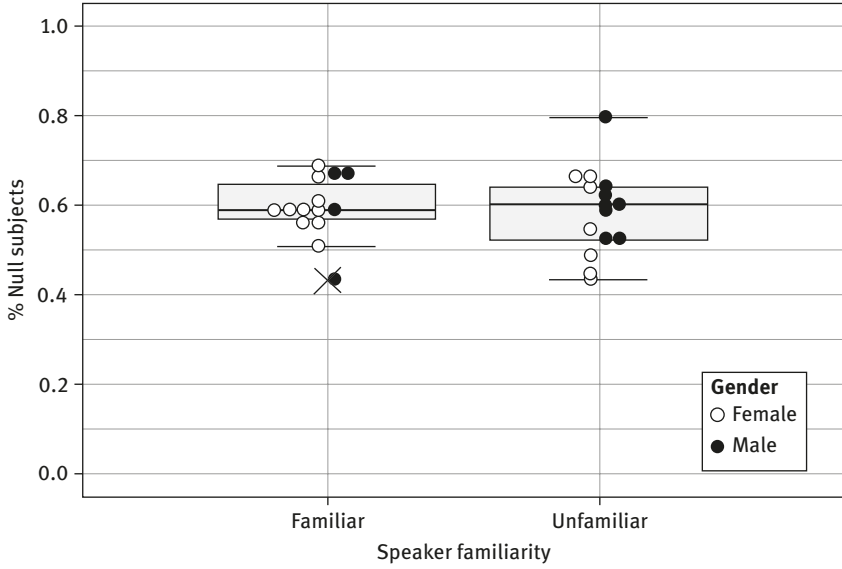
**Figure 3:** RNS according to speaker familiarity.
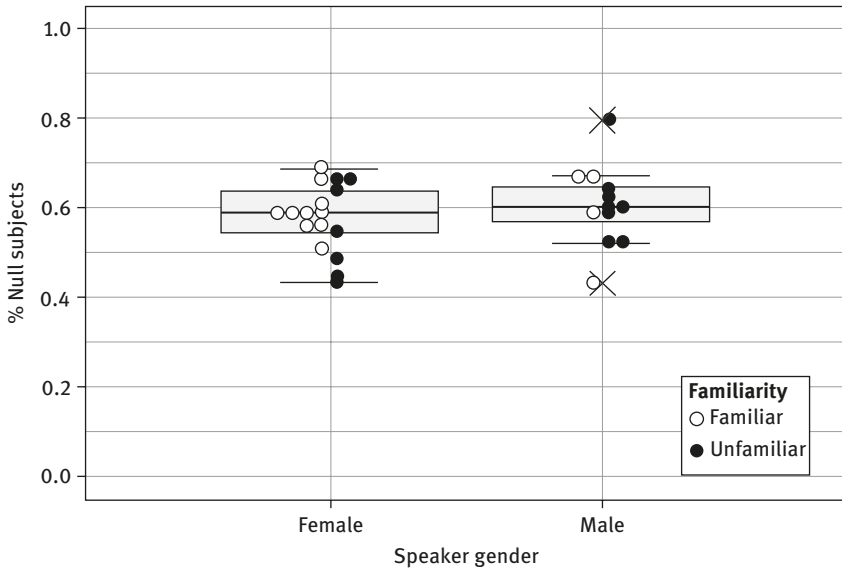


**Figure 4:** RNS according to speaker gender.

significant, and did not turn out to be (Pearson test for age: r = 0.0644, p = 0.740). The results of an ANOVA on all five variables is provided in Table 2.

**Table 2:** ANOVA of linguistic and non-linguistic factors.

| Factor | F | p |
|---|---|---|
| RNS \| length | $F(1, 27) = 0.403$ | $p = 0.531$ |
| RNS \| newR/CU | $F(1, 27) = 1.148$ | $p = 0.293$ |
| RNS \| familiarity | $F(1, 27) = 0.081$ | $p = 0.779$ |
| RNS \| gender | $F(1, 27) = 0.726$ | $p = 0.402$ |
| RNS \| age | $F(1, 27) = 0.113$ | $p = 0.740$ |

Based on these data, our conclusion is that rates of null subjects is remarkably stable across all speakers, regardless of age, gender, or degree of familiarity among the interlocutors. Thus we find no support for the hypothesis that speaker familiarity has an effect on RNS.

# 6 Discussion

Perhaps the most striking feature of the results is, disregarding for a moment two outlier values,[11] the stability of the RNS value across the sample as a whole (see, e.g., the Standard Deviation (SD) value for RNS in Appendix 1). Whether our results generalize to other experimental settings remains an open question; they may be specific to Persian, or specific to the task, or simply reflect small sample size. However, our results actually appear well in line with the findings from research on spoken language registers summarized in Biber and Conrad (2009: 261). Commenting on the results of several decades of research on variation across spoken registers, the authors note that "speech is highly constrained in its typical linguistic characteristics". Although written language displays considerable cross-register variation, "all spoken texts are surprisingly similar linguistically, regardless of communicative purpose (excluding scripted or memorized texts)". These conclusions may seem at odds with decades of research in the Labovian tradition of variationist sociolinguistics, which has sought to emphasize socially determined variation in speech, but there is an important

---

**11** The two outliers are the speakers g1_m_13, with the unusually low RNS of 0.433, and g2_m_13 (RNS = 0.796). We are unable to identify any biographic factors (e.g., bilingualism in another language) that might explain these extreme values.

difference: most research in the variationist sociolinguistics paradigm continues to focus on phonology, rather than syntax (e.g., the phonological realization of the -ING suffix of English verbs has remained a "staple of sociolinguists" since the 1950s (Hazen 2006)). Thus although we find it highly plausible that a social variable such as speaker familiarity would be reflected in phonological variation, or lexical choices, or perhaps intonation contours, it seems equally probable that syntactic features of discourse would be relatively stable, reflecting general cognitive constraints on short-term memory and instantiated through deeply entrenched and routinized patterns of delivery, mediated by language-specific morphosyntactic configurations. Biber and Conrad (2009) repeatedly point to the primacy of content and genre in determining variation in syntax. If this is indeed correct, then we would expect to find little variation across a sample of spoken texts of comparable content, regardless of setting. This prediction is borne out by our Persian data, where content was held fairly constant across all speakers.

# 7 Conclusions

Our investigation took up the challenge of investigating the factors that may impact on rates of null subjects in colloquial spoken Persian. We focused on a possible impact of speaker familiarity, hypothesizing that greater familiarity among the interlocutors may lead to higher rates of null subjects in their speech, because familiar speakers can rely on a broader expanse of "common ground" (Matić et al. 2014), and hence afford to be less explicit. Our investigation found little support for this idea, however. Rates of RNS in spoken Persian instead appear to be relatively stable, and did not significantly correlate with speaker familiarity, or with the factors of gender and age. These findings are consistent with research on morphosyntactic variation in spoken language (Biber and Conrad 2009), which points to a high degree of homogeneity in spoken language, with the main determinants of variation being content and genre. The latter were held constant in our experimental design, which may help explain the overall lack of variation. However, we note that our data is almost entirely in the third person; dialogical data, involving first and second person forms, may pattern differently; this deserves further research.

Finally, we consider our research as an initial step toward an empirical and usage-based approach to syntactic variation in spoken Persian. Recently, corpus-based studies have opened up promising avenues for issues such as word-order variation in Persian (e.g., Faghiri et al. 2014), and we expect that these developments will gather momentum in coming years. However, there is a considerable

gap between written and spoken Persian, and as yet, most researchers interested in usage-based, as opposed to formalist, analyses of Persian have concentrated on the written language as their object of study (e.g., Roberts 2014), or on "scripted spoken language", as in the film dialogues investigated in Vafaeian (2018). But with the exception of Frommer (1981) and Saeli and Miller (2018), there is very little empirical research on spontaneous colloquial spoken Persian. Our research is thus a modest attempt to develop corpus-building standards and methodologies for the future study of spoken Persian.

**Appendix A:** Raw figures from the experimental data, all speakers

| Speaker | Familiarity | Gender | Age | CUs | RNS | NewRef | NewRef/CU |
|---|---|---|---|---|---|---|---|
| g1-f-01 | Familiar | Female | 39 | 47 | 0.558 | 15 | 0.319 |
| g1-f-02 | Familiar | Female | 29 | 54 | 0.608 | 16 | 0.296 |
| g1-m-03 | Familiar | Male | 22 | 17 | 0.588 | 8 | 0.471 |
| g1-m-04 | Familiar | Male | 25 | 61 | 0.673 | 12 | 0.197 |
| g1-f-05 | Familiar | Female | 26 | 60 | 0.510 | 17 | 0.283 |
| g1-m-06 | Familiar | Male | 32 | 22 | 0.667 | 9 | 0.409 |
| g1-f-07 | Familiar | Female | 25 | 38 | 0.588 | 11 | 0.289 |
| g1-f-08 | Familiar | Female | 25 | 25 | 0.565 | 11 | 0.440 |
| g1-f-09 | Familiar | Female | 25 | 100 | 0.663 | 23 | 0.230 |
| g1-f-10 | Familiar | Female | 31 | 83 | 0.688 | 20 | 0.241 |
| g1-f-11 | Familiar | Female | 33 | 60 | 0.593 | 13 | 0.217 |
| g1-f-12 | Familiar | Female | 33 | 49 | 0.591 | 14 | 0.286 |
| g1-m-13 | Familiar | Male | 35 | 69 | 0.433 | 17 | 0.246 |
| g1-f-14 | Familiar | Female | 29 | 99 | 0.585 | 22 | 0.222 |
| g2-f-01 | Unfamiliar | Female | 20 | 58 | 0.446 | 20 | 0.345 |
| g2-f-02 | Unfamiliar | Female | 20 | 44 | 0.486 | 16 | 0.364 |
| g2-f-03 | Unfamiliar | Female | 20 | 40 | 0.667 | 14 | 0.350 |
| g2-f-04 | Unfamiliar | Female | 20 | 25 | 0.435 | 11 | 0.440 |
| g2-f-05 | Unfamiliar | Female | 21 | 26 | 0.640 | 13 | 0.500 |
| g2-f-06 | Unfamiliar | Female | 38 | 56 | 0.660 | 13 | 0.232 |
| g2-f-07 | Unfamiliar | Female | 33 | 51 | 0.545 | 17 | 0.333 |
| g2-m-08 | Unfamiliar | Male | 20 | 49 | 0.522 | 15 | 0.306 |
| g2-m-09 | Unfamiliar | Male | 22 | 42 | 0.600 | 13 | 0.310 |
| g2-m-10 | Unfamiliar | Male | 20 | 41 | 0.641 | 16 | 0.390 |
| g2-m-11 | Unfamiliar | Male | 25 | 25 | 0.524 | 10 | 0.400 |
| g2-m-12 | Unfamiliar | Male | 20 | 40 | 0.600 | 15 | 0.375 |
| g2-m-13 | Unfamiliar | Male | 20 | 52 | 0.796 | 15 | 0.288 |
| g2-m-14 | Unfamiliar | Male | 20 | 36 | 0.588 | 13 | 0.361 |

| Speaker | Familiarity | Gender | Age | CUs | RNS | NewRef | NewRef/CU |
|---|---|---|---|---|---|---|---|
| g2-m-15 | Unfamiliar | Male | 27 | 48 | 0.622 | 13 | 0.271 |
| **Mean** | | | 26.03 | 48.86 | 0.589 | 14.55 | 0.325 |
| **SD** | | | 6.00 | 20.54 | 0.082 | 3.61 | 0.081 |

# References

Alvaraz, Gabriela. 2015. Variation of overt and null subject pronouns in the Spanish of Santo Domingo. In Ana M. Carvalho, Rafael Orozco, and Naomi Lapidus Shin (eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective*, 5–18. Georgetown: Georgetown University Press.

Bell, Alan. 2006. Speech accommodation theory and audience design. In K. Brown (ed.), *Encyclopedia of language and linguistics*, 648–651. Boston: Elsevier.

Biber, Douglas and Susan Conrad. 2009. *Register, genre, and style*. Cambridge: Cambridge University Press.

Bickel, Balthasar. 2003. Referential density in discourse and syntactic typology. *Language* 79. 708–736.

Bickel, Balthasar. 2011. Putting variation center-stage: Beyond "language" (or "dialect") as the basic data unit in language typology (and elsewhere). Keynote lecture at the conference: *Variation and typology: New trends in syntactic research*. The Linguistic Association of Finland, Helsinki, Aug. 25–27, 2011.

Camacho, José. 2013. *Null subjects*. Cambridge: Cambridge University Press.

Carvalho, Ana M., Rafael Orozco, and Naomi Lapidus Shin (eds.). 2015. *Subject pronoun expression in Spanish: A cross-dialectal perspective*. Georgetown: Georgetown University Press.

Chafe, Wallace. 1980. *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*. Norwood, NJ Ablex.

Comrie, Bernard. 1989. *Language universals and linguistic typology. Syntax and morphology*. Chicago: University of Chicago Press.

Du Bois, John. 1987. The discourse basis of ergativity. *Language* 63(4). 805–855.

Faghiri, Pegah, Pollet Samvelian, and Barbara Hemforth. 2014. Accessibility and word order: The case of ditransitive constructions in Persian. In Stefan Müller (ed.), *Proceedings of the 21st International Conference on Head-Driven Phrase Structure Grammar*, University at Buffalo, 217–237. Stanford, CA: CSLI Publications.

Falk, Yehuda. 2001. *Lexical-functional grammar. An introduction to parallel constraint-based syntax*. Stanford, CA: CSLI Publications.

Farrell, Patrick. 2005. *Grammatical relations*. Oxford: Oxford University Press.

Frommer, Paul. 1981. *Post-verbal phenomena in colloquial Persian syntax*. PhD thesis, University of Southern California.

Ghomeshi, Jila. 2018. Other approaches to syntax. In Anousha Sedighi and Pouneh Shabani-Jadidi (eds.), *The Oxford handbook of Persian linguistics*. Oxford: Oxford University Press.

Haig, Geoffrey. 2008. *Alignment change in Iranian languages: A construction grammar approach*. Berlin: Mouton.

Haig, Geoffrey. Under review. The pronoun-to-agreement cycle in Iranian: Subjects do, objects don't.

Haig, Geoffrey and Stefan Schnell. Multi-CAST (Multilingual Corpus of Annotated Spoken Texts).

Haig, Geoffrey and Stefan Schnell. 2014. Annotations using GRAID (Grammatical Relations and Animacy in Discourse), version 7.0. Available at: https://lac.uni-koeln.de/corpora/Multi-CAST/guidelines/Annotations/GRAID-Manual-7.0.pdf.

Haig, Geoffrey, Stefan Schnell, and Claudia Wegener. 2011. Comparing corpora from endangered languages: Explorations in language typology based on original texts. In Geoffrey Haig et al. (eds.), *Documenting endangered languages: Achievements and perspective*s, 55–86. Berlin: Mouton.

Haig, Geoffrey and Stefan Schnell. 2016. The discourse basis of ergativity revisited. *Language* 92(3). 591–618.

Hazen K 2006. IN/ING Variable. In Keith Brown (Editor-in-Chief), *Encyclopedia of language and linguistics*, *Second edition*, volume 5, pp. 581–584. Oxford: Elsevier.

Holmberg, Anders. 2009. Null subject parameters. In Theresa Biberauer Anders Holmberg, and Michelle Sheehan (eds.), *Parametric variation: Null subjects in Minimalist Theory*, 88–124. Cambridge: Cambridge University Press.

Jahangiri, Nader. 1980. *A sociolinguistic study of Tehrani Persian*. PhD thesis, London: University College.

Keenan, Edward. 1976. Towards a universal definition of subject. In Charles Li (ed.), *Subject and topic*, 303–333. New York: Academic.

Mahootian, Shahrzad and Lewis Gebhardt. 2018. Revisiting the status of -eš in Persian. In Alireza Korangy and Corey Miller (eds.), *Trends in Iranian and Persian linguistics*, 263–276. Berlin: De Gruyter.

Matić, Dejan, Rik van Gijn and Robert D. Van Valin. 2014. Information structure and reference tracking in complex sentences: An overview. In Rik van Gijn Jeremy Hammond, Dejan Matić, Saskia van Putten, and Ana Vilacy Galucio (eds.), *Information structure and reference tracking in complex sentences*, 1–41. Amsterdam: Benjamins.

Meyerhoff, Miriam. 2011. *Introducing sociolinguistics*. London: Routledge

Neeleman, Ad and Kriszta Szendröi. 2008. Case morphology and radical pro-drop. In Theresa Biberauer (ed.), *The limits of syntactic variation*, 331–348. Amsterdam: Benjamins.

Orozco, Rafael. 2015. Pronominal variation in Colombian Costeño Spanish. In Ana M. Carvalho Rafael Orozco, and Naomi Lapidus Shin (eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective*, 19–40. Georgetown: Georgetown University Press.

Perlmutter, David. 1971. *Deep and surface constraints in syntax*. New York: Holt, Rinehart and Winston.

Peškova, Andrea. 2013. Experimenting with *pro-drop* in Spanish. *SKY Journal of Linguistics* 26. 117–149.

Radford, Andrew. 2004. *Minimalist syntax. Exploring the structure of English*. Cambridge: Cambridge University Press.

Rasekh, Mohammad. 2014. Persian clitics: Doubling and agreement. *Journal of Modern Languages* 24(1). 16–33.

Roberts, John. 2014. *A study of Persian discourse structure*. Uppsala: Acta Universitatis Upsaliensis.

Saeli, Hooman and Corey Miller. 2018. Some linguistic indicators of sociocultural formality in Persian. In Alireza Korangy and Corey Miller (eds.), *Trends in Iranian and Persian linguistics*, 163–182. Berlin: De Gruyter.

Sato, Yosuke and Simin Karimi. 2016. Subject-object asymmetries in Persian argument ellipsis and the anti-agreement theory. *Glossa: A Journal of General Linguistics* 1(1): 8. 1–31.

Schiborr, Nils N. 2016a. English. In Geoffrey Haig and Stefan Schnell (eds.), Multi-CAST (Multilingual Corpus of Annotated Spoken Texts).

Sedighi, Anousha. 2010. *Agreement restrictions in Persian*. Leiden: Leiden University Press.

Stoll, Sabine and Balthasar Bickel. 2009. How deep are differences in referential density? In E. Lieven J. Guo, N. Budwig, S. Ervin-Tripp, K. Nakamura, and Ş. Özçalişkan (eds.), *Crosslinguistic approaches to the psychology of language: Research in the traditions of Dan Slobin*, 543–555. London: Psychology Press.

Vafaeian, Ghazaleh. 2018. *Progressives in use and contact: A descriptive, areal and typological study with special focus on selected Iranian languages*. PhD thesis, Stockholm University, Faculty of Humanities, Department of Linguistics.

Ketevani Gadilia
# A typological study of (in)definiteness in the Iranian languages

## 1 Introduction

In this chapter[1] I use the technical term "(in)definiteness", which is a formal combination of "definiteness" and "indefiniteness".

The category of definiteness and indefiniteness is one of the functional (semantic and syntactic) categories of languages with a general function of noun determination and actualization, which may be conveyed not only by a special grammatical unit article, but also by various language means (like demonstratives, definite or indefinite pronouns, and the numeral one).

A fundamental monograph by Christopher Lyons (1999) is based on modern achievements of linguistics and abundant cross-linguistic data. The initial point of Lyon's work is the concept of definiteness itself, which is investigated in much depth in the book. Lyons differentiates two major groups of definiteness and indefiniteness, which are the simple and complex types. Simple definiteness and indefiniteness consist of lexical items like an English article ("a, the"), or the affix like Arabic (prefix *al-* and suffix *-n*), which indicates the definiteness or indefiniteness of the noun phrase. Much of Lyons's book is devoted to the "noun phrases whose definiteness or indefiniteness is due to something other than presence or absence of an article" (Lyons 1999: 107), a group of complex definites that includes proper nouns, personal pronouns, and noun phrases containing a demonstrative or possessive modifier.

*Definiteness* by Christopher Lyons is especially noteworthy for specialists of the Iranian languages because of the work's Persian language data, particularly the article of indefiniteness *-i* and postposition *-ra*. Lyons's most important concept is the areal character of (in)definiteness: "Marking of definiteness is often an 'areal feature'". "The languages which are geographically contiguous … may develop common characteristics" (Lyons 1999: 48). On the one hand, "The greatest concentration of languages of marking definiteness today is in Western Europe" (Lyons1999: 48) and, on the other hand, the Middle East and

---

**1** This chapter was written more than a decade ago. Since that time much new research on this subject has appeared in the scholarly literature. The author is aware of these developments but has decided to keep the chapter in its original form because she has already dealt with many of them in other publications.

Central Asia is the area, where "involvement of definiteness in areal facts concerns the combined representation of definiteness and the direct object relation" (Lyons1999: 49). Thus, according to Lyons, there are at least two major Iranian language areas: the Middle East and Central Asia.

The typological study of languages analyzes the reasons for their differences and similarities. In general, "the primary task for linguistic typologists is to identify and explain the properties that make human language what it is" (Song 2014: 3). The goal of this chapter is less ambitious and closer to another definition of language typology by Moravcsik (2013: 9) – that is, "the study of two phenomena: typologically and universally shared features of languages". This chapter investigates one category in one language group to identify typological and universal features of (in)definiteness in Iranian languages. This is the first step toward the areal investigation of category of definiteness and indefiniteness of Iranian languages.

## 2 Problem statement

The exponents of the (in)definiteness in Modern Iranian languages are complicated and diverse. Consequently, the terms and definitions that are generally used for their notification are also diverse and in some cases even contradictory (see below). Needless to say, there is much terminological diversity of the following definitions: article, numeral article, demonstrative article, indicator of definiteness, indicator of indefiniteness, particle, element expressing definiteness and indefiniteness, and suffix functioning as the article are widely used in Iranian studies.

For instance, even in Persian, the most studied Iranian language, various definitions for the article *-i* clearly demonstrate the alternative understanding of it. The terminological diversity is presented in Lazard's well-known article, "L'enclitique nominal *-i* en Persian: un ou deux morphèmes?" He describes the terms used by Iranian scholars such as *yā-ye ešārat* ('*i* of indication'), *yā-ye vahdat* ('*i* of singularity'), *yā-ye nakare/tankir* ('*i* of indefiniteness'), *yā-ye ma'refe/ta'rif* ('*i* of definiteness'), as well as the definitions presented in the research of European Iranists like "î d'unité" and "î démonstratif" (Darmsteter 1883), "indefinite and distributive -i" and "relative suffix" (Nye), etc. (Lazard 1966: 249).

Afšār (1991: 28–30) asserts that there are two articles in Persian: indefinite and definite. According to him, the indefinite article is performed by perfect and imperfect forms. The definite article is used only in constructions like *kas-i*

*ke* ('the one [who]') in compound sentences with the definite clause: *mard-i ke miayad pedare man ast*

On the other hand, according to Shafâ'i (1989: 14), *yā-ye nakare* and *yā-ye wvahdat* are the isolated morphemes indicating two exponents and two meanings, respectively, although *yā-ye nakare* is an inflectional morpheme that creates the grammatical category of selectivity (*tafkik*). Ahmad Šafā'i comes to this conclusion when comparing the following language situations:

(1)  *bâzargân-i be šahri vâred šod*
     'a merchant entered the city'

(2)  *amâ bâzargân az šahr xošaš nayâmad*
     'but the merchant did not like the city'

If (1) is followed by additional information, such as (2), the noun should not be indefinite because the participants of communication possess some knowledge about the merchant, as well as the town.

In Russian tradition, the article *-i* is regarded as the solitary, unique morpheme and is called a "selective article". This term, suggested by Rubinchik (1959: 182), is now acknowledged as the most suitable determination not only for Persian data but for related languages. Although Lazard agrees with Rubinchik's opinion, he (1966: 249–64) also devotes a special article arguing the problem of two morphemes *-i* in modern Persian. He comes to the conclusion that "il est possible de trouver à l'enclitique nominal *-i* dans tous ses emplois une valeur sémantique commune: celle d'un instrument de 'sélection'" (1966: 263).

Even less extensive studies of the nature of the article in Persian clearly illustrate the high level of ambiguity. The situation is almost the same for the other Iranian languages. Why is there such a diversity and dispersion of assessment? What does it allude to?

From our point of view, there are several reasons for the terminological lack of clarity. Firstly, there is the diversity of Iranian languages themselves and the syncretic trait of semantic and grammatical category of definiteness in this group of languages. Another possible reason is the complexity of the category of definiteness itself and the diversity of theoretical methods of determination of the category of definiteness and indefiniteness as well as its grammatical markers.

In some cases the terminological opacity is caused by the uncertain status of a morpheme – that is, whether it is a case marker or a inflectional grammatical unit article (for instance, the same Kurdish morpheme may be simultaneously

considered as the article as well as the case marker). We can assume that the reason of current dispersion is a combination of language factors, i.e., (1) the peculiarities of Iranian languages, (2) the linguistic methodology applied by scholars, and (3) the purpose of the research.

# 3 Typological attitude

In the mid-1970s, Efimov, Kerimova, Molchanova, Pireyko, Rastorgueva, and Edelman, the leading specialists in the Department of Iranian Languages at the Institute of Linguistics in Moscow, published the two-volume *Istoriko-tipologicheskoe izuchenie iranskikh yazikov* (1975). The first volume is dedicated to the morphological analysis of the processes that caused the typological transformation of Iranian languages. The Old, Middle, and Modern Iranian phonological type is described in detail, but the main goal is focused on analysis of internal transformation of the Old Iranian inflexional system, the development of analytical means, and the appearance of the secondary synthesis. The second volume is a detailed description and analysis of the grammatical categories. The authors investigate the ways of transforming the grammatical means (volume 1) as well as the morphological and syntactic categories (volume 2). Even though the category of definiteness is only considered in connection with the category of case, the investigation made an invaluable contribution in tracing the development of the category of definiteness. Even though the research was published more than forty years ago, the applied typological method of investigation and its results still preserves its significance.

This chapter outlines upcoming research on the typological attitude to the *category of definiteness and indefiniteness* in the related Iranian languages. I presume that the typological approach can answer some of these questions or at least allows us to understand this linguistic phenomenon. The combination of synchronic and diachronic data may provide an opportunity to understand the complicated problems concerning definiteness in the Iranian languages. In particular, these are four points below:

1. The typological attitude can establish the historical process and patterns of one category in the related language group (diachronic typology).
2. The typological method makes it possible to compare language data and dialect data (synchronic and diachronic typology).
3. The typological attitude can determine the dominant and marginal tendencies of the category of definiteness and ascertain the basic rules of their functioning in the modern Iranian languages (synchronic typology).

4. The advantage of a typological analysis is the ability to consider chronologically distant language stratum data, like the prepositional article in the northeastern branch of Iranian languages such as Horezmian (Middle Iranian), early Ossetic (New Iranian), and Ossetic-Digor (Modern Iranian).

# 4 (In)definiteness in the Middle Iranian languages

The category of definiteness vs. indefiniteness and its grammatical markers was developed in the Middle Iranian period, though Benvenist (1978: 233) asserted that the Avestan relative pronoun *ya* possessed all features of the article. The formation of the category of (in)definiteness is connected to the reduction of the Old Iranian multi-case inflexion system to compensate the eliminated synthetic case system. Therefore, the analytical and agglutinative system became dominant.

The markers of (in)definiteness in Middle Iranian languages are the indefinite, postpositional article in Middle Persian (East-West Group): *-ēv/-e(v)>*OP aiva*; the indefinite, postpositional article in Parthian (North-East group) *-ēv* (*'yw*); the demonstratives functioning as definite articles, preposition: hō (pl. *hīvin*), *im* (pl. *imīn).* The prepositional article in Horezmian (East [North] Iranian): *'y* [ī] masc., sing., masc., and fem. pl.; *y'* [yā] – fem. sing; the definite, prepositional article in Sogdian (East Iranian): *x-/w-, y-/m-*; articles expressed the case relations reflecting the category of gender, number, and case.

# 5 Definiteness in Modern Iranian languages

The morphological, semantic, and syntactic fields of the category of definiteness and indefiniteness in the group of Iranian languages consist of various language units such as the articles proper, deictic elements (especially demonstratives), the numeral one, definite and indefinite pronouns, case markers, postposition *rā*, particle *na(h)*, and stress.

This chapter considers more prevalent patterns and models in Modern Iranian languages and conveys the (in)definiteness based on the following language units or their combinations like articles (1), Middle Persian *rāy/rād<OT rādiy* and its reflections in Modern Iranian languages such as the postposition *-rā* and some case markers (2), and, finally, auxiliary words (3).

## 5.1 The articles and grammaticalized formal elements

The enclitic, unstressed, postpositional extraction/indefinite article -*i* in Persian is also found in Dari -*e*, Tajik -*e*, Balochi -*i*, Lori and Bakhtiari -*i*, dialects of Fârs -*i*, Talishi -*i*, Gilaki -*i*/*ə*, Tati -*ī*, and Mazandarani -*ī*. The above-mentioned element was generated in East-West Iranian languages, but presumably it is borrowed in several South-West Iranian languages like Talishi, Gilaki, Balochi, and even in East Iranian (Yagnobi), which demonstrates the strength of this morpheme.

Talishi and Gilaki worked out the prepositional elements, which behave as indefinite articles: *gʊlay* (*igla*/*glay*/*illa*) and *i*/*ta*.

The grammars based on traditional analysis, and not considering dialectology data, usually declare that the article -*i* in Persian, and its variants -*e* and -*ī* in Tajik and Dari are not binary opposition exponents. However, the presence of the definite article -*e* in the Tehran dialect indicates the opposite tendency.

The binary opposition can be found in North-West Iranian languages: Kumzari -*ō* (def.) vs. -*ē* (indef.), Gurani -*aka* (masc.)/-*aka* (fem.) (def.) vs. -*ew* (masc.)/-*ewa* (fem.) (indef.), Zazaki -*äkä* (def.) vs. -*ek*/-*äk* (indef.), Suleymani (Kurdish dialect) -*eke*/-*e* (def.) vs. -*ek*/-*e* (ind.).

Some Iranian, especially West Iranian languages, use both prepositional and postpositional elements simultaneously. For instance, in dialects of Fârs, the prepositional demonstrative -*i* combines with the case markers -*a*, -*o*, -*u* to express definiteness. A similar construction is found in Kurdish (Suleymani), where case markers -*e*/-*ye* (postposition) and demonstratives (preposition) *ew* or *em* must be combined.

In the Pamir languages definiteness is conveyed by the demonstratives in the function of the article (Wakhi, Ishkashimi). In Wakhi, adjective pronouns *ya*, *yaet*, *yaem* (unlike the proper pronoun usage) are used in the direct case, singular, regardless of the form of noun group "invariable article with variable noun" (Edelman 1990: 89). In Ishkashami, demonstratives in the function of article are used in "frozen", reduced invariable form either in direct case *ma*, *da*, *(w)a*, or in indirect case singular *am*, *ad*, *(a)w* regardless of the noun group form.

## 5.2 Middle Persian Particle rād/rāy "for the sake of" lost the lexical meaning and became a grammeme

(a)  In languages where the old category of case was completely eliminated, the prepositional element *rādiy* transformed into postposition -*rā*. It became a part of definiteness in Persian, Dari, and Tajik, where it mainly marks the definite direct object: *āb-rā biyār* (Dari).

(b)   In the languages of the North-West branch where the secondary innovation case paradigm was developed, the element *rādiy* transformed into case markers (Balochi, Tati, Gilaki, and Mazandarani).

In Balochi, the marker of objective case (Casus Dativus): *-a, -ara/-r, -ra <rād<rādiy* designates the direct object:

(3)   *sing-a zurti*
      'he took the stone'

In Gilaki (4) and Mazandarani (5), the unstressed accusative-dative case marker *−a/-ra*, *-ō/-rō* also designates the definite direct object, which is connected with the preliminary context or the situation.

(4)   *semavər-a åtəs bukun* (Rastorgueva and Edelman 1982)
      'Heat the samovar'

(5)   *mōn še pul-rō/pul-ō gum bōkōrdōm* (Rastorgueva and Edelman 1982)
      'I lost my money'

In Tati, the marker of the direct object *-rä/-ä* (Eastern dialect), *-re/e* (Northern dialect) descends from Middle Persian *rād* and designates the things that are definite or single by their origin.

(6)   *šäbi tkun-ä bäst, reft be kutan* (Grunberg and Davidova 1982)
      'Shabi locked up the shop and went away'

In cases where the thing is indefinite and single, it is marked by the numeral "one" (functioning as article), which is sometimes combined with the unstressed suffix *-i* as in (7):

(7)   *ye märd-i bire* (Grunberg and Davidova 1982)
      'there was a (one) man'

In the Fars dialects (Mâsarmi, Burunguni, Pâpuni, etc.), the suffix *-a < rād<râdiy* conveys the meaning of definiteness of a subject as well as of the objects.

Designation of definiteness is not the essential function of case markers. The agglutinative case paradigm was developed in Gilaki, Balochi, and some other Modern Iranian languages after reducing the Old Iranian multifunctional inflexion case system. Modern Iranian innovation agglutinative case category is not

directly connected and correlated with the category of definiteness, but is very interrelated and interdependent. First, the direct case denotes and distinguishes the direct and indirect objects. Then the opposition of definite direct object vs. indefinite direct object became relevant. Thus, the definite nouns are marked by direct case markers like Balochi *-ra*, *-a/-râ*, *-â* and Gilaki *-a*, while the indefinite nouns are represented by nominative unmarked case. For instance,

(8)  *ĵinik gul-â âurt* vs. *ĵinik âp âurt* (8) (Moshkalo 1991: 43)
  'a girl brought the flower' vs. 'a girl brought (some) water'

## 5.3 Numeral one and demonstratives as auxiliary words can convey the meaning of definiteness and indefiniteness

In these cases they express indefiniteness and definiteness like the similar language units in other world languages. However, some Iranian features (shortened forms, combined models, etc.) make them specific.

### 5.3.1 Numeral one

Pers. *yek*, Tajik, Dari, Baloch. *yak*, Lori and Bakht. dialects *yäk*, *yä*, dialects of Fârs *yä*, Lâri *yak*, Mazand. *ye*, *yä* (reduced form) and *yek*, *yäk* (full form), Tati *yä/y*, Gil. *i/ita* (with numerative *-ta*), Talish. *igla*, *gla(i)*, *illa*, *ila* (with particle *-la*), and Semnâni *i*, *iä*, (West Iranian group).

Sarikoli: *i(w)*, Ishkashimi *uk/ůk*, *ug/üg*; Shugni-Rushani: *y* (reduced form), *yie*, *i*, Wakhi *(y)i/yī(w)*, Ossetic *yu/iw* (Iron) and *yeu/yew* (Digor), Yagnobi *ī* (East Iranian group).

### 5.3.2 Demonstratives

In the Wakhi and Ishkashimi languages, the demonstratives of all three series indicating distance between a speaker and hearer are available in the function of the article. In Wakhi, the adjective demonstratives *ya*, *yaet*, *yaem* (in addition to the demonstrative functions) as usual are used in nominative case singular and are independent of the form of the noun. Thus, we can find the correspondence between the convertibility of nouns and the immutability of articles.

(9)  *ya nan yet kâš-ər [1] and ki* . . . (Pakhalina 1975)
  'The mother tells the little boy, that ...'

A similar rule with the Wakhi is in Ishkashim(i), where the demonstratives that function as the articles regardless of the determined noun are used in the reduced set of the singular forms, either in the nominative case *ma*, *da*, *(w)a*, or in the oblique case *am*, *ad*, *(a)w*.

(10)   *a non, ma ív duk pъ ma xafob demo?* (Pakhalina 1959)
        'Mummy, may I wash these filthy clothes in this soap-water?'


# 6  Conclusion

Based on the presented information, some typological similarities have become evident:

When expressed, the positional distribution of markers category of (in)definiteness shows that in West Iranian languages, the central element of the category is the so-called postpositive selective article.

In the East Iranian group the dominant means of expressing definiteness are the prepositional demonstratives/numeral one.

Grammaticalization of various types indicates that the historical process impacted the modern patterns of definiteness of Iranian languages.

Designating the actants, direct object marking (DOM), by case markers, is to distinguish the subject and the direct object of a transitive verb and avoid the possible polysemy. Thus, case markers and particle *rā* - in Modern Iranian languages reflect the topic-focus/subject-object relations.

Diachronic predefinition and historical continuity are valuable for understanding the Modern Iranian category of definiteness. However, history of definiteness is split in some particular processes; therefore, we are not able to talk about global diachronic constant that is equally relevant for all Iranian languages. The grammaticalization and formation of the category of definiteness is common in all the Iranian languages.

The table below classifies the elements of definiteness and indefiniteness in Modern Iranian languages:

**Table 1.** Definiteness and indefiniteness in Modern Iranian languages

| Language | Articles and similar elements | *Old Persian rādiy* and its reflections | Auxiliaries (demonstratives and numeral one) |
|---|---|---|---|
| Persian | *-i* | – | *yek* ('one') |
| Dari | *-ī* | – | *yak* ('one') |

**Table 1.** (continued)

| Language | Articles and similar elements | *Old Persian rādiy* and its reflections | Auxiliaries (demonstratives and numeral one) |
|---|---|---|---|
| Tajik | *-e* | – | *yak* ('one') |
| Balochi | *-i* | *-a, -ara/-r, -ra* | *yak* ('one') |
| Tati | | *-rä/-ä, -re/-e* | *yä/y* ('one') |
| Lori/ Bakhtiari | *-i* | – | *yäk, yä* ("one") |
| Dialects of Fârs | *-i* | *-a* | *yä* ('one') |
| Talishi | *-i, gьlay/illa* | – | – |
| Gilaki | *-i, i/ita* | *-a/-ra, -ō/- rō* | – |
| Mazandarani | *-i* | *-a/-ra, -ō/- rō* | *ye, yä/yek, yäk* ('one') |
| Kumzari | *ō* (def.), *ē* (indef.) | – | – |
| Gurani | *-aka* (def.)/*-ew, -ewa* (indef.) | – | – |
| Zâzâ | *-äkä* (def.)/*-ek, -äk* (indef.) | – | – |
| Suleymani | *-eke, -e* (def.)/*-ek/-e* (indef.) | – | – |
| Wakhi | – | – | *ya, yaet, yaem* (dem.) *(y)i/yī(w)* ('one') |
| Ishkashimi | – | – | *ma, da, (w)a/am, ad, (a)w* (Dem.) *uk/ŭk, ug/ūg* ('one') |
| Lâri | – | – | *yak* ('one') |
| Sarikoli | – | – | *i(w)* ('one') |
| Shugni-Rushani | – | – | *y, yie, i* ('one') |
| Ossetic (Iron) | – | – | *yu/iw* ('one') |
| Ossetic (Digor) | *i (def.)* | – | *yeu/yew* |
| Semnâni | – | – | *ī, iä* |
| Yagnobi | – | – | *ī* |

# References

Afšâr, Iraj. 1991. *Yā-ye nakare, yā-ye nakare* [The *-i* of indefiniteness and *-i* of definiteness]. *Âyande* 1–4. 28–30.

Benvenist, E. 1978. *Otnositelnoe predlozhenie kak problema obshchego sintaksisa. Obshchaya lingvistika* [General linguistics]. Moscow: Progress.

Collinson, W. E. 1937. *A study of demonstratives, articles and "indicators"*. Language monograph no. 17. Baltimore: Waverly Press.

Darmsteter, James. 1883. *Études iraniennes*, vol. I. Paris: F. Vieweg

Edelman, J. 1990. *Sravnitelnyaya grammatika vostochno iranskikh yazikov* [The comparative grammar of East Iranian languages]. Moscow: Nauka.

Edelman D.I., Rastorgueva V.S. 1975. *Istoriko-tipologicheskoe izuchenie iranskikh yazikov* [The historical and typological study of Iranian languages], 2 vols. Moscow: Nauka.

Grunberg, A.L., and K. Kh. Davidova. 1982. *Osnovi iranskogo yazikoznaniya. Novoiranskie yaziki: zapadnaya gruppa, prikaspiyskie yaziki* [Foundations on Iranian linguistics. New Iranian languages: West group, languages of Kaspian area]. Moscow: Nauka.

Gukhman. M. 1981. *Istoricheskaya tipologiya i problemi diakhronnikh konstant* [Historical typology and the problem of diachronic constants]. Moscow: Nauka.

Lazard, Gilbert. 1966. L'enclitique nominal -*i* en Persan: un ou deux morphemes? *Bulletin de la société de linguistique de Paris*, 61 (1). 249–264.

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

Moravcsik, Edith A. 2013. *Introducing language typology*. Cambridge: Cambridge University Press.

Moshkalo, V. V. 1991. *Osnovi iranskogo yazikoznaniya. Novoiranskie yaziki: severo-zapadnaya gruppa* [Foundations on Iranian linguistics. New Iranian languages: North-West group]. Moscow: Nauka.

Pakhalina, T. N. 1959. *Ishkashimsiy yazik* [Ishkashimi language]. Moscow: Izdatelstvo akademii nauk SSSR.

Pakhalina, T. N. 1975. *Vakhanskiy yazik* [Wakhi language]. Moscow: Nauka.

Rastorgueva, V. S., and D. I. Edelman. 1982. *Osnovi iranskogo yazikoznaniya. Novoiranskie yaziki: zapadnaya gruppa, prikaspiyskie yaziki* [Foundations on Iranian linguistics. New Iranian languages: West group, languages of Kaspian area]. Moscow: Nauka.

Rubinchik, Y. A. 1959. *Slozhnopodchinyennie predlozheniya s opredelitelnim pridatochnim v sovremennom persidskom* [Compound sentences with the definite clause in Modern Persian]. Moscow: Izdatelstvo vostochnoy literature.

Shafâ'i, Ahmad. 1989. *Mabāni-ye 'Elmi-ye dastur-e zabān-e Fārsi* [Scientific essentials of Persian grammar]. Tehran: Nowin Publishing Company.

Song, Jae Jung. 2014. *Linguistic typology: Morphology and syntax*. New York: Routledge.

Sascha Völlmin

# The quotative marker in Gilaki

Gilaki belongs to the Caspian subgroup of Northwest Iranian and can be divided into two major dialect groups, Western Gilaki with Rasht as its center and Eastern Gilaki with Lahijan as its center (Stilo 2001: 660). There are between two and three million speakers, most of whom are bilingual in Gilaki and Persian.[1]

The following is a first presentation of reported speech in Gilaki or, more exactly, the quotative marker in Rashti, i.e., the Western Gilaki dialect of the city of Rasht. The language data used in this chapter derive from a corpus of about thirty minutes of audio recordings of natural speech and dialogues or spontaneously narrated stories of Gilaki speakers raised in (or around) Rasht. Thus, the results presented here hold true only for Rashti for certain. The situation in other Western Gilaki variants or Eastern Gilaki must be left open. However, concluding from a remark of a speaker of Eastern Gilaki saying that "people in Rasht speak like this" (i.e., with the quotative marker), it is possible that no (or no similar) quotative marker exists in Eastern Gilaki.

There are not many sources available that describe Gilaki, and no descriptions of reported speech in this language, a topic often neglected in grammars. The most comprehensive grammar (Rastorgeuva et al. 1971) and the most recent (but short) description by Stilo (2001) do not treat reported speech; likewise Sartippur (1990). This chapter intends to be a contribution to fill this gap. On the other hand, however, I also hope to show that Gilaki exhibits typologically interesting features with regard to reported speech in general as well as in Iranian languages in particular.

# 1 Reported speech in general and in the region

There are two ways to report speech: direct and indirect. Generally speaking, direct speech is the unchanged quotation of an utterance, whereas indirect speech is the repetition of an utterance with shift of the deictic center from the reported speech situation to the actual speech situation. This involves in

---

**1** Different degrees of mastery of both Gilaki and Persian can be found. Nowadays (at least in Rasht and presumably also in other towns) children are often raised only in Persian. As a result, many are semi-speakers or have only passive command of Gilaki. Cf., e.g., Pakpour (2015: 19).

particular the pronouns. Depending on the language, additional changes as, for example, the TAM-forms or the choice of the complementizer may be required for indirect speech.[2]

When dealing with reported speech in European languages, a main issue often is which of these changes are necessary in indirect speech. For the languages in and around Iran, however, this question is rather secondary. In this area, the more usual way to report an utterance is the use of direct speech. Indirect speech may be possible, but in general it is not often employed or may, for example, be limited to higher registers such as the literary language. This situation is reported for Kurdish by Akin (2002: 79). In his corpus indirect speech is very rare and appears, if at all, only in written language. The same holds true for Caucasian languages such as Georgian and Abkhaz: according to Hewitt and Crisp (1986), there is a great preference for direct speech to indirect speech. In Persian, the situation seems to be a little less clear-cut. Nevertheless, according to Alavi and Lorenz (1994: 239), as well as to my own intuition, Persian "exhibits a reluctance" toward indirect speech. The direct method appears to be the more natural way of reporting speech.

Gilaki does not differ in this respect from these languages. In the corpus used for this chapter, virtually no instances of indirect speech are found. Direct speech, i.e., the quotation of an utterance without shifting the deictic center, is definitely the preferred way to report speech. There is, however, a small but essential difference that distinguishes Gilaki (or at least Rashti) from Persian and other Iranian languages and makes it typologically outstanding: the marker -ə. Due to its primary use to mark the verbs in quoted utterances, it is best named the "quotative marker". It will therefore be glossed as QUOT in the examples.

The following sections describe the form as well as the basic and secondary functions of the quotative marker in more detail. Finally, a possible origin of -ə is proposed.

---

**2** German, for example, has a shift in mood (from indicative to subjunctive) and tense (from present to past; especially to avoid ambiguous forms). In addition, the choice of the complementizer (no complementizer possible with direct speech) and word order change (from verb second to verb final as a characteristic of German dependent clauses) can distinguish indirect from direct speech (Coulmas 1986: 14–21).

# 2  The quotative marker *-ə*

In Rastorgueva (1971), instead of the proper personal ending of the first person singular *-əm*, occasionally a variant *-əmə* is found. Some paradigms (p. 147) might suggest that *-əmə* is an allomorph in the present tense and subjunctive, but at large the distribution is completely random and no explanations of its function are given there. These few examples of an extra *-ə* after *-əm* are the only traces of the quotative marker in the available sources.

   In my corpus, however, there are over two hundred instances of an *-ə* added to the personal ending of a verb. As Table 1 shows, this is not restricted to first person or present tense, but is possible in virtually all persons and TAM-forms:

**Table 1:** Examples for verb forms with *-ə*.

|       | Present    | Subjunctive  | Imperative  | Past        | Imperfect    | Past perfect   |
|-------|------------|--------------|-------------|-------------|--------------|----------------|
| 1SG   | *bər-əm-ə* | *bu-kun-əm-ə* | -           | *bu-kud-əm-ə* | *nə-nəst-i-m-ə* |              |
| 2SG   | *dan-i-ə*  | *bə-xa-i-ə*  | *bu-kun-ə*  | *bu-kud-i-ə* |              |                |
| 3SG   | *kun-e-yə* | *bə-b-ə-yə*  | -           | *b-amo-yə*  | *šo-i-ə*     | *bə-kəftə-bu-yə* |
| 1PL   | *isa-imi-ə* | *b-avər-imi-ə* | -         |             |              |                |
| 2PL   | *dar-idi-ə* | *bi-g-idi-ə* | *bu-kun-idi-ə* |          |              |                |
| 3PL   | *kun-idi-ə* |              | -           |             | *fukud-i-d-ə* |               |

The forms listed above are actual data from the corpus. The blank fields do not represent impossible forms; rather, they are simply not found in the corpus by chance. As for the transcription of the quotative marker after vowels, I have chosen the following convention: *-ə* after *i*, *-yə* after all other vowels.

   I will not discuss the verbal system of Gilaki here.[3] The important point for this chapter is that *-ə* can be added to every single (conjugated) verb form, regardless of its person, tense, or aspect, including even the imperative. There is only one crucial exception (or "special case"), which will be discussed in Section 7: the third person singular of the past.

---

**3** Consider, though, the following general remarks: no special tense-aspect-modality marker for present; imperfect has a suffix *-i*; subjunctive/imperative, past and past perfect have a prefix *bə-*. *bə-* is suppressed (a) with the negation *nə-*, (b) when the verb root already has a preverb, (c) after the change-of-state marker *-a*.

# 3 Basic pattern: Occurrence of -ə after *goftən* 'say'

The by far most frequent occurrence of verb forms with an additional -ə is in complement clauses of *goftən* 'say', normally (but not necessarily) with the subordinator *ke*. Example (1) represents a prototypical sentence for the use of -ə:

(1) *xånəm mudir   bu-goft-**ə**      ke    dånəšåmuz-i ke   kəlas-ə-mian*
lady  principal  TAM-say.PST-3SG  SUB  pupil-I      SUB  class-EZF-in
*dar-idi-**ə**         bibəzåat-idi-**ə**  az    vəz-ə           måli*
have.PRS-2PL-QUOT  destitute-be₁:3PL-QUOT  of  situation-EZF  financial
*xob    ni-idi-**ə**         muarrəfi    bu-kun-idi-**ə**.*[4]
good  NEG-be₁:3PL-QUOT  introduction  TAM-make.PRS-2PL-QUOT
'The principal said: "[Please] mention to me the pupils that you have in class [and that] are destitute [and] not well off!"'

The speaker introduces her narration with *xånəm mudir bugoftə* 'the principal said'. After the subordinator *ke* she repeats what the principal had once said to her. Every single verb following *bugoftə ke* 'said that' is marked with -ə: a present (*dar-idi-ə*), a positive (-*idi-ə*), and a negative copula (*ni-idi-ə*), as well as an imperative[5] (*bu-kun-idi-ə*). Apart from the introductory verb and the subordinator, the only difference between this reported speech and a (non-reported) direct speech is the existence of the added -ə. Therefore the (reconstructed) original sentence quoted in (1) must have been as the following:

(2) *dånəšåmuz-i  ke           kəlas-ə-mian  dar-idi        bibəzåat-idi*
pupil-I        SUB    class-EZF-in  have.PRS-2PL  destitute-be₁:3PL
*vəz-ə             måli       xob       ni-idi       muarrəfi*
situation-EZF  financial  good     NEG-be₁:3PL  introduction
*az*
of
*bu-kun-idi.*
TAM-make.PRS-2PL
'"[Please] mention to me the pupils that you have in class [and that] are destitute [and] not well off!"'

---

**4** The use of *-idi* may be confusing. This suffix is the enclitic copula and the personal ending of the verbal paradigms of both second and third person plural.
**5** This form could as well be interpreted as a subjunctive with optative meaning.

In (1) the quoted sentence is a first-hand report, i.e., the original utterance was directed toward the speaker of (1). This, however, is not a prerequisite for the use of *-ə*. The reported information can also be second-hand or third-hand:

(3)  *Mi        mår       tarif        kud-i                   ke,*
     1SG.POSS   mother    description   make.PST-IPFV:3SG   SUB
     *goft-i                ke,   i-ta*
     say.PST-IPFV:3SG       SUB   one-CLF
     *åqå-ye    i-ta      zən-i          bi-giftə-bu-**yə**.                   badən*
     sir-DEF   one-CLF   wife-INDEF     TAM-take.PP-AUX:PST:3SG-QUOT       then
     *ani                 xånəvådə*
     3SG.POSS            family
     *tarif         kud-i-**ə**                      ke    an    zəmån-i   ke    kučik*
     description   make.PST-IPFV:3SG-QUOT   SUB   3SG   time-I    SUB   little
     *bu-**yə***
     be₁.PST:3SG-QUOT
     *bə-kəftə-bu-**yə**                   ani         pišåni*
     TAM-fall.PP-AUX:PST:3SG-QUOT    3SG.POSS    forehead
     *bə-škəstə-bu-**yə**.*[6]
     TAM-break.PP-AUX:PST:3SG-QUOT
     'My mother used to tell (and she used to say) that a man had gotten married (lit. had taken a woman). (Then) his family used to tell that when he was little he had fallen down [and] his forehead had broken'.

Here, the speaker reports what her mother used to tell her. The first two verbs after the introducing verbs of saying are the mother's words (*bi-giftə-bu-yə* 'he had taken' / *tarif kud-i-ə* 'she used to tell'). The following verbs (*bu-yə* 'he was'/ *bə-kəftə-bu-yə* 'he had fallen' / *bə-škəstə-bu-yə* 'it had broken'), however, are not her mother's words but "his family's", i.e., a third person's. Nevertheless, all verbs are marked equally with *-ə* (and only one *-ə*). Thus, the source of the utterance is not important. Rather, *-ə* indicates that the utterance is not of the actual speaker, but so to speak out of somebody else's mouth. The same holds true in the following example:

---

**6** The use of past perfect is not (necessarily) a shift of tense to indicate anteriority. In Iranian languages, the past perfect is much more common than, for example, in English, so it can have been used in the original speech as well.

(4) *bu-gu          agər  qərår      bə-b-ə-yə*                    *ke   un   Məhin-a*
    TAM-say.PRS  if    agreement  TAM-be₁.PRS-3SG-QUOT  SUB  3SG  Mahin-OBJ
    *rəvån-a    kun-ə-yə,              bə-d-ə⁷              Məhin  xu-re     tənhå*
    going-CS  make.PRS-2SG-QUOT  TAM-give.PRS-3SG  Mahin  self-for  alone
    *bə-š-ə-yə.*
    TAM-go.PRS-3SG-QUOT
    'Say [to her]: "If it is agreed that he should send Mahin [on a voyage], let
    Mahin go all by herself"'.

Here again, the sentence following the introducing imperative *bugu* 'say' is not
what the speaker actually says to the hearer, but what the hearer is supposed to
say to somebody else (in the future). This example makes it clear that -*ə* indeed
is a quotative marker, as it not only occurs when reporting previously pro-
nounced utterances, but rather in all sentences that do not constitute an actual
(or original) direct speech. Thus, every time you say the words of someone else,
be it a real report or a speech that has not yet happened, the quotative marker
is used.

To recapitulate so far, Gilaki (Rashti) possesses a quotative marker -*ə*, which
is added to every single verb in quoted sentences. This is the usual method to
report speech in this language. Remember that these quotations (and therefore
reported speech) do not involve any shift of the deictic center: the personal
pronouns and endings stay the same as they are in the original sentence. One
effect of this system is that in reported speech you can say "I", "me", or "my"
without referring to yourself, or on the other hand say "you" or "your" without
meaning the person you are speaking with. Yet no confusion arises. Since the
verbs of the quotations are marked with -*ə*, it will not be understood as actual
direct speech. Who exactly you are referring to has to be determined by the
context.

In the case of the first person singular, however, there seems to be a certain
reluctance to the use of the pronouns *mən* 'I' and *mi* 'my' in quotations. Although
there is no problem to do so – cf. example (7) – this somewhat peculiar situation
can be avoided by using the logophoric pronoun *xu* 'self' instead of both *mən* 'I'
(second *xu*) and *mi* 'my' (first and third *xu*):

---

**7** For the lack of -*ə* here, see Section 6 "Absent -*ə*".

(5) *g-e* *ke* *pəs* *tu* *muvåzəb-ə* ***xu*** *zən* *bu-bu-yə!*
 say.PRS-3SG SUB SO 2SG watching-EZF LOG wife TAM-be₁.PRS-QUOT
 *an* *muddət-i* *rå*[8] *ke* ***xu*** *n-esa-m-ə* ***xu*** *zən-a*
 this while-I OBJ SUB LOG NEG-be₃-1SG-QUOT LOG wife-OBJ
 *tu* *kontorol* *bə-dar-ə!*
 2SG control TAM-have.PRS-QUOT
 'He says: "Thus you look after my wife! When I am not here, you control my wife!"'

The use of *xu* instead of *mən/mi* depends probably only on the speaker: some seem to prefer *xu*, while others rarely use it.[9] As for the personal endings of the corresponding verbs, note that they are not affected by *xu*, but remain in the first person singular (*xu nesam̱ə* 'I am not').

# 4 *-ə* without *goftən* 'say'

Quotations are most of the time introduced by *goftən* 'say' or another verb of saying. This, however, is not a condition for the use of *-ə*. The introductory verb can also be omitted. The presence of *-ə* assures that the sentence is perceived as a quotation. It is thus possible to change between direct (i.e., non-quoted) and quoted speech without announcing it:

(6) [A]*ettəfåqən* *xeyli* *ådəm-ə* *xånəvådə-i* *iss-ə.* *zən-ə* *xeyli*
 by.the.way very person-EZF family-ADJ/INDEF be₂-3SG woman-EZF very
 *xob-i-ə.* [B]*təmåm-ə* *mi* *ləbås,* *rəxt,* *hamə* *či-a*
 good-INDEF-be₁:3SG all-EZF 1SG.POSS clothes laundry all thing-OBJ
 *šor-e-yə.*
 wash.PRS-3SG-QUOT
 '[A]By the way, she is a very family-oriented person. She is a very good woman. [B]"She washes all my clothes, laundry, everything"'.

This is an extract of a conversation. First, the speaker makes a personal direct statement about a woman (marked with [A]). Then, he quotes a sentence somebody else said about the same woman (marked with [B]). There is no break

---

**8** *rå* object (specific-referential) (or maybe the whole structure *an muddəti rå ke*) is Persian.
**9** Further testing is needed to determine to which degree and in which contexts *xu* occurs in a regular pattern, or whether its use is more or less only idiosyncratic.

between the two parts and no introducing verb of saying. Nevertheless, due to the quotative marker, the hearer can easily distinguish between direct utterance and quoted utterance. It is clear from the context who uttered the quoted sentence originally, the involved person(s) having been introduced earlier in the conversation. As for the first person possessive pronoun *mi* 'my', recall that this does not point to the actual speaker of (6), but to the original speaker of the quoted sentence.

In the following example, the speaker again switches (without announcing it). This time, however, he does not report a previous utterance:

(7)  <sup>A</sup>*åfərin!*  <sup>B</sup>*mən*  *fikr*  *bu-kud-əm-ə*  *šime*  *sər-a*
    exactly  1SG  thought  TAM-make.PST-1SG-QUOT  2PL.POSS  head-OBJ
    *kulå*
    hat
    *bə-n-əm-ə.*  <sup>A</sup>*vəli*  *ištəbå*  *kud-ən-dər-ə.*
    TAM-put.PRS-1SG-QUOT  but  mistake  make-INF-be₄-3SG

'<sup>A</sup>Exactly! <sup>B</sup>He thought that he could trick us [lit. put a hat on us]. <sup>A</sup>But he is making a mistake'. (Or more literally: '<sup>A</sup>Exactly! <sup>B</sup>"I thought that I could trick you". <sup>A</sup>But he is making a mistake'.)

After the reply to somebody else's comment with *åfərin* 'exactly', there is a "quoted" sentence (<sup>B</sup>), which is completely imagined. Certainly it was never uttered this way. Rather it expresses what the speaker assumes somebody else might have thought or intended. To verbalize this assumption, he pretends that the other person had actually said it and employs the quotative marker to "quote" them. Finally, he adds his own comment, switching back to direct speech. Again, consider the different deictic centers: the first person subject in <sup>B</sup> points to the same extra-linguistic person as the third person in <sup>A</sup>. Note also that the possessive pronoun *šime* 'your' is pointing to the speaker himself (and others).

# 5  *-ə* after verbs of cognition

The marker *-ə* is labeled a quotative marker because of its predominant use in quoted sentences. However, there is an occasional extended use after verbs of cognition in general such as *fikr kudən* 'think' (8) and *mutəvəjjəh bostən* 'realize, become aware' (9).

(8) *mən    hamišə   fikr      kud-i-m              ke    Bənəfšə    bə*
1SG    always   thought   make.PST-IPFV-1SG   SUB   Banafsheh   to
*råh-ə    dur-i       š-e-yə.*
way-EZF   far-INDEF   go.PRS-3SG-QUOT
'I always thought that Banafsheh would go far away'.

(9) *mutəvəjjəh   ni-b-idi                ke    Piruz   Rəšt-i-ə-yə.*
attentive    NEG-become.PRS-3PL     SUB   Piruz   Rasht-ADJ-be₁:3SG-QUOT
'They did not realize that Piruz is a Rashti (=from the town of Rasht)'.

The difference to proper quotations is not as significant as it may seem, especially since thoughts can be regarded as communication with yourself, so when expressing them you quote yourself in a way. Nevertheless, this is in contrast to "real" quotations of oneself – cf. examples (10) and (11).

# 6  Absent *-ə*

Until now, *-ə* has been presented as an obligatory marker for all verbs in all quotations. But, as a matter of fact, it is sometimes also absent.

When quoting oneself, *-ə* is not present:

(10) *bu-goft-əm       tu    xa-i          bi-š-i,             Fəxri-aməra*
TAM-say.PST-1SG   2SG   want.PRS-2SG   TAM-go.PRS-2SG   Fakhri-with
*bu-šu.*
TAM-go.PRS
'I said: "If you want to go, go with Fakhri"'.

This is not limited to the first person singular, but also applies to the first person plural:

(11) *bu-goft-im        agə    ato-yə            tu     ti*
TAM-say.PST-1PL   if     like.this-be₁:3SG   2SG    2SG.POSS
*bår-o-kuč-a            usan*
load-and-household-OBJ   pick.up.PRS
*b-avər        ame        xånə.*
TAM-bring.PRS   1PL.POSS   house
'We said: "If it is like this, take your things [and] bring [them] to our house"'.

This is the reverse of the above characterization of the quotative marker, saying that -ə indicates that "the uttered sentence is out of somebody else's mouth"; thus, when quoting oneself, the uttered sentence is not out of somebody else's mouth. Consequently, -ə is not present, even though the sentence is not an actual utterance, but a quotation.

The invariable modals, in particular *båyəd*[10] 'must', cannot take -ə:

(12) *bu-goft-ə*      *mən  čəra  **båyəd**  təra    bə-šnas-əm-ə?*
     TAM-say.PST-3SG  1SG  why  must  2SG.OBJ  TAM-know.PRS-1SG-QUOT
     'He said: "Why do I have to know you?"'

Also, the inflected modal verbs quite often do not have the quotative marker as, for example, *xastən* 'want'; see also *bədə* 'let' in example (4):

(13) *ama  **xa-im**       kumək  ašan-a   bu-kun-imi-ə.*
     1PL  want.PRS-1PL  help  3PL-OBJ  TAM-make.PRS-1PL-QUOT
     '[She said:] "We want to help them"'.[11]

In addition to the modals, it occasionally happens that one does not affix -ə to every verb of the quotation, although it would be expected. I find it too difficult to determine any rules for omitting -ə. It seems to be rather random than systematic. For now, it suffices to state that -ə can also be missing every now and then without apparent reason.

# 7 The quotative marker and third person singular past

A somewhat anomalous situation is found in the third person singular of the past. To form the past, you need the past stem. There are past stems ending in consonants and past stems ending in vowels. After a consonant, the regular personal ending of the past is -ə. In this case, it is not possible to affix an additional -ə to mark the quotation:

(14) *bu-goft-ə*       *čəra?  åqå-jån  či    bu-kud-**ə-ø**         məgə?*
     TAM-say.PST-3SG  why  sir-soul  what  TAM-make.PST-3SG-QUOT  PART
     'She said: "Why? What has he (lit. dear sir) done?"'

---

**10** *båyəd* is a loan from Persian, but much more frequently used than the proper Gilaki *ba/va*.
**11** This is the continuation of example (1).

If the past stem ends in a vowel, however, the personal ending *-ə* is always missing. In exchange, the quotative marker can be added:

(15)  *Rəza*   *g-e*      *dai*   *fada-ø-yə*        *Amir-a.*
     Reza   say.PRS-3SG   uncle   give.PST-3SG-QUOT   Amir-OBJ
     'Reza says: "The uncle gave [it] to Amir"'.

It seems that the *-ə* of the quotation is elided after the personal ending *-ə*. This could be perfectly conceivable. On the other hand, however, the third person singular of the subjunctive (16) and the copula (17) are also *-ə*, but in these two cases the quotative marker *-ə* does not disappear:

(16)  *agər*   *qərår*      *bə-b-ə-yə*        ...
     if      agreement   TAM-be₁.PRS-3SG-QUOT
     "'If it is agreed ..."

(17)  *bu-goft-ə [...]*      *xob-ə-yə [...]*       *məsalə-i*       *ni-ə-yə.*
     TAM-say.PST-3SG   good-be₁:3SG-QUOT   problem-INDEF   NEG-be₁:3SG-QUOT
     'He said: "That's fine, that's no problem"'.

Apparently, the personal ending of the past differs (in behavior) somehow from the subjunctive and the copula, although they all are *-ə* (in Western Gilaki). This hypothesis is also supported by data from Eastern Gilaki, where the third person singular in the past is *-ə*, too, but in the subjunctive *-ı* (Stilo 2001: 661). In any case, it is not possible to add the quotative marker *-ə* to the third person singular *-ə* of the past. The probable solution for this "special case" could be found in a shared origin of both morphemes. This would explain the fact that they cannot appear together. The copula and the ending of the subjunctive, on the contrary, presumably originate somewhere else, and thus can be combined with the quotative marker *-ə*.

In this context, a glance at (colloquial) Persian might be helpful. In Persian the past does not have a personal ending in the third person singular (i.e., *-ø*). To express evidentiality (and also indirect speech as a subcategory of evidentiality) the perfect is used (cf., e.g., Jahani 2000), which is built with the past participle + auxiliary. In the third person singular the auxiliary is dropped. As a consequence, the verb form looks exactly like the participle, which ends in *-e*. (In all other persons this *-e* is elided by the following auxiliary/personal ending; only the stress remains on the last syllable.) Thus, one can reanalyze *-e* as the personal ending, or, alternatively, as the marker for evidentiality. As a matter of fact, in colloquial Persian it is in principle possible for all third person singular forms that are built on the basis of the past stem to carry this morpheme *-e* to mark evidentiality:

(18) Persian

| *Mariam* | *be* | *man* | *mi-g-e* | | *ke* | *vaqtike* |
|---|---|---|---|---|---|---|
| Mariam | to | 1SG | IPFV-say.PRS-3SG | | SUB | when |

| *dåšt-**e*** | *harf* | *mi-zad-**e*** | *yedaf'e* | *sedå* | *qat* |
|---|---|---|---|---|---|
| have.PST-**E** | letter | IPFV-hit.PST-**E** | suddenly | sound | cut |

| *shod-**e**.* |
|---|
| become.PST-**E** |

'Mariam tells me that when she was speaking on the phone suddenly the sound was cut' (Jahani 2000: 203).

The situation in Gilaki is comparable to Persian, but not identical. In Gilaki, past and perfect are said to have coalesced (Stilo 2001: 665), the one form being built analogous to the colloquial Persian perfect: the third person singular always looks like the past participle.[12] Thus, as already exemplified above, the alleged personal ending -*ə* only occurs when the past participle of the corresponding verb also ends in -*ə*. When the past participle ends in a vowel, no personal ending is present. In exchange, the latter can take the quotative marker (as every other verb form), but the former cannot – recall examples (14) and (15). It seems almost too obvious that in order to disentangle this irregularity, -*ə* simply should not be analyzed as third person singular of the past, but as quotative marker only. Then, there would not be this "special case" any more, i.e., a personal ending that prevents the quotative marker from being added. If this really was the case, both Rastorgueva et al. (1971) and Stilo (2001) would be wrong in presenting -*ə* as personal ending. This assumption is, surprisingly enough, further sustained by some instances in the corpus where the third person singular of the past actually does not have an ending. Consider the following two representative examples:

(19)

| *mən* | *yeho* | *mi* | *xun* | ***vasoxt**.* |
|---|---|---|---|---|
| 1SG | suddenly | 1SG.POSS | blood | steam.PST:3SG |

'Suddenly I boiled with rage (lit. my blood steamed)'.

(20)

| *Minå* | *harf-a* | *guš* | ***bu-kud**.* |
|---|---|---|---|
| Mina | word-OBJ | ear | TAM-make.PST:3SG |

'He listened to Mina's words'.

---

**12** Admittedly, they only "look" the same. The proper participle is stressed on the last syllable (-*ə*), whereas the past on the penultimate, i.e., the syllable before the personal ending.

Does this mean that Gilaki still possesses an opposition between perfect (with *-ə*) and past (without *-ə*)? And is one of the functions of the perfect to mark quotatives? The answer has to be no, the main reason being the overwhelming majority of the forms with *-ə*, also in contexts where no meaning of a perfect fits:

(21)    *dər   jå     bu-froxt-**ə**,       åre.*
       in     place   TAM-sell.PST-3SG   yes
       'He sold [it] right away, yes'.

Still, the ø-endings occur, even though they are not very numerous. Their existence could be explained as a structural borrowing from Persian, when a speaker sometimes "erroneously" switches to a somewhat more Persian style. This is fairly possible, as Gilaki underwent and is still undergoing (nowadays more than ever) heavy Persian influence. Nevertheless, further elicitation also showed that informants can accept both forms with and without *-ə* in almost every case (whatever reason this may have), except for the following example:

(22)    *šåh     **bə-kəft**!*
       king    TAM-fall.PST:3SG
       'The king has fallen!'

When somebody directly observes that the king falls and then comments this with surprise, the ending *-ə* is apparently not possible and consequently rejected by informants. This use could be regarded as the extreme opposite pole to the quotative marker *-ə*, which in a broad sense expresses indirect evidence; when *-ə* is missing, you signal that you have in contrast some very direct evidence, and consequently, you can confirm the information with certainty. Examples (19) and (20) are not direct comments to an incident, but they have in common with (22) that they also express surprise or suddenness and the speaker wants to confirm the surprising news.

     Whatever the correct explanation for all these examples is, it seems quite sure that there is a connection between the *-ə* of the past participle, the personal ending, and the quotative marker. Based on the description above, they are most probably cognates in both Gilaki and Persian. The situation in the two languages is comparable as far as the assumed origin of the morpheme is concerned, but the actual use (or so to speak, the grammaticalization) is not the same. Recall that in colloquial Persian the original ending *-e* of the past participle (used in the perfect) has spread to some other verb forms to express evidentiality. This, however, is limited to verbs built with a past stem and the third person singular. If the development in Gilaki had been the same, one would expect *-ə* to appear in

the same places. But -ə became completely detached from the past participle and is suffixed to the right of the personal ending so that the quotative marker occurs in all existing tenses and persons. Thus, the connection between the participial -ə and the quotative marker -ə is not as tight (anymore) as between the corresponding elements in Persian. The personal ending, too, developed away from the past participle, as it carries no stress like all the other personal endings of the past, the participle being stressed on -ə.

However, my proposition is only one possibility. I find it rather difficult to determine the exact path of grammaticalization of the quotative marker. Did it develop internally in Gilaki more or less analogous to Persian but much further grammaticalized? Or, alternatively, did language contact play a role and the quotative marker is a borrowing or calque? And if so, where from? Maybe from a language of the Caucasus, where quotations are usually also marked? Unless more data are available, the possible answers remain highly speculative.

# 8 Conclusion

Gilaki, or at least the Western Gilaki dialect of Rasht, possesses a quotative marker -ə, which is suffixed to (principally) all verbs of a quotation. The quotations are mostly, but not necessarily, introduced by *goftən* 'say' and the subordinator *ke*; it is also possible to quote someone without announcing this, as the quotative marker always assures that the utterance is conceived as a quotation.

Quotations correspond most of the time to reported speech, or from another point of view, to report speech quotations are used. Note that reported speech is not the same as indirect speech. As quotations are always direct, reported speech in Gilaki is direct (following a general tendency of the languages in and around Iran to report directly). Indirect speech might be possible, but it is not necessary: due to the fact that there is a quotative marker, Gilaki has a convenient means to report speech.

The use of quotations is somewhat wider than just reporting an utterance of somebody else. They also apply in utterances that have not been said (yet). Furthermore, you can even express the (assumed) thoughts or intentions of somebody else using quotations. Finally, -ə is occasionally also found after verbs of cognition in general.

Quotations do not entail an adaptation of the deictic elements (especially pronouns) to the actual speech situation. Thus, when saying *mən* 'I' the speaker does not point to himself and *tu* 'you' does not mean the person spoken to. In order to avoid this peculiar situation, the first person singular pronouns *mən* 'I' and *mi* 'my' are sometimes replaced with the logophoric pronoun *xu* 'self'.

In the third person singular of the past, it is not possible to add the quotative marker when the personal ending -*ə* is already present. The solution for this anomaly could be found in a shared origin of both -*ə*: the ending of the past participle.

# Abbreviations

| | |
|---|---|
| 1, 2, 3 | first, second, third person |
| ADJ | adjectivizer -*i* |
| AUX | auxiliary |
| be$_{1,2,3,4}$ | different verbs of being |
| CLF | classifier -*ta* |
| CS | change of state -*a* |
| DEF | definiteness (suffix borrowed from Persian) -*e* |
| E | evidentiality (in Persian) |
| EZF[13] | ezafe -*ə* |
| I | -*i* introducing restrictive relative clauses ('yåye ešårat') |
| INDEF | indefiniteness -*i* |
| INF | infinitive -*(ə)n* |
| IPFV | imperfective -*i* / (*mi-* in Persian) |
| LOG | logophoric pronoun *xu* |
| NEG | negation *nə-* |
| OBJ | object (specific-referential) -*a* |
| PL | plural (with personal endings) |
| PART | particle |
| POSS | possessive pronoun |
| PP | past participle |
| PRS | present stem |
| PST | past stem |
| QUOT | quotation marker -*ə* |
| SG | singular (with personal endings) |
| SUB | subordinator *ke* |
| TAM | tense-aspect-modality marker *bə-* |

# References

Akin, S. 2002. Discours rapporté et hétérogénéité discoursive en kurde. In L. Rosier (ed.), *Le discours rapporté*. Faits des Langues 19. Paris: Ophrys. 71–84.

---

**13** Both the dependent-marking ("Gilaki-type") and the head-marking ("Persian-type") ezafe are glossed EZF. In Gilaki both types occur.

Alavi, B. and M. Lorenz. 1994. *Lehrbuch der persischen Sprache*. Leipzig: Langenscheidt.

Coulmas, F. 1986. Reported speech: Some general issues. In F. Coulmas (ed.), *Direct and indirect speech*. Berlin: Mouton de Gruyter. 1–28.

Hewitt, B.G. and S. R. Crisp. 1986. Speech reporting in the Caucasus. In F. Coulmas (ed.), *Direct and indirect speech*. Berlin: Mouton de Gruyter. 121–143.

Jahani, C. 2000. Expressions of Indirectivity in spoken Modern Persian. In L. Johanson and B. Utas (eds.), *Evidentials*. Berlin and New York: Mouton de Gruyter. 185–207.

Pakpour, P. 2015. *Identity construction. The case of young women in Rasht*. Studia Iranica Upsaliensia 27. Uppsala: Acta Universitatis Upsaliensis.

Rastorgueva, V. S., A. A. Kerimova, A. K. Mamedzade, L. A. Pireiko, and D. I. Edel'man. 1971. *Giljanski jazyk*. Moskow: Izdatel'stvo "Nauka".

Sartippur, J. 1990. *Vižegihâ-ye dasturi va farhang-e våžehå-ye Gilaki*. Rasht: Našr-e Gilakån.

Stilo, D. 2001. Gīlān, X. Languages. In E. Yarshater (ed.), *Encyclopædia Iranica*, vol. X (Fisheries-Gindaros). New York: Bibliotheca Persica Press. 660–668.

Behrooz Mahmoodi-Bakhtiari and Hassan Rezai-Baghbidi
# Plural marking in the New West Iranian languages and dialects: a historical and typological approach

The nominal system of all New West Iranian languages has two numbers. The singular is often used after numerals and quantitative adjectives, but in some languages the plural is marked, especially when specific, after numerals larger than one. The choice of plural markers depends upon case (mostly direct vs. oblique), gender distinction, and the distinction between animate and inanimate.

The purpose of this chapter is twofold:
1. to study the form, function, and distribution of plural markers in a selection of New West Iranian languages and dialects,
2. to identify the source and the original meaning of New West Iranian plural markers.

## 1 Introduction

The nominal system of Old Iranian (OIr) had three genders (i.e., masculine, feminine, and neuter), three numbers (i.e., singular, dual, and plural), and eight cases (i.e., nominative, accusative, dative, instrumental, ablative, genitive, locative, and vocative). The formal differentiation of these cases was complete only in the singular, and even there only in one declension, that of the masculine *a*-stems. In the dual and the plural, there were only three and six separate forms, respectively. The plural case endings of OIr masculine *a*-stems can be reconstructed as follows:

Nom. Voc. *-āh, *-āhah*
Acc. *-ān(s)*[1]
Ins. *-āiš, *-aibiš*
Dat. Abl. *-aibyah*
Gen. *-ānām*
Loc. *-aiš,*-aišvā*

---

**1** -*s* was retained only in some *sandhi* combinations.

---

**Note:** This is a thoroughly revised version (with new additions and references) of an article with the same title published in the journal *Studies on Persianate Societies* (2004).

There is enough evidence to show that in Western Middle Iranian (WMIr), the cases had been reduced to two: direct and oblique. Such a two-case system is well attested in Middle Persian (MP) and Parthian (Parth) of the Sassanian inscriptions and the Pahlavi Psalter (PPs). The same system is also preserved, at least to some extent, in Book Pahlavi (BPah), but it has been mostly abandoned in Manichaean Middle Persian (MMP) and Manichaean Parthian (MParth), where the two cases are only distinguished in the first person singular pronoun and the old stems in -*ar*. The case endings of nouns in Western Middle Iranian are as follows[2]:

**Table 1:** Case endings in Western Middle Iranian.

|         |          |                      | *Ar*-stems |                        |
|---------|----------|----------------------|------------|------------------------|
|         | Singular | Plural               | Singular   | Plural                 |
| Direct  | -ø       | -ø                   | -ø         | -ar                    |
| Oblique | -ē       | -ān (-īn, -ūn)       | -ar        | -arān (-arīn, -arūn)   |

The direct case (from the old nominative and accusative) was used as a grammatical subject, and the oblique case (from the old genitive) was used as an agent (in a passive construction), an indirect object, a genitive, and the complement of prepositions and postpositions. Both the direct and the oblique cases could be used as direct objects. Inanimate nouns were most often left undeclined, unless there was a wish to emphasize plurality (Skjærvø 1983: 134).

(1)  Examples: Parth. *yazad* (Dir. Pl.) *amāh … dastgird karēnd ud pad yazad-ān* (Obl. Pl.) *pušt …* 'the gods made (historic present) us [their] property, and with the help of the gods … '; MP *šāh-ān* (Obl. Pl.) *šāh* (Dir. Sg.) 'king of kings'; MPth *tō ay pid* (Dir. Sg.) *čēim-īn harw-īn karišn* 'Thou art the Father of all these creations', *nimāyān pidar-ān* (Obl. Pl.) *wxēbēh* 'I will show [you] my own fathers'.

However, even in good classical WMIr texts, there are instances of the plural oblique case used as direct. Thus, the two-case system gradually disappeared from WMIr, and nouns came to be distinguished in number only, the general plural-ending being -*ān* (or rarely -*īn* or -*ūn*, from the OIr genitive endings \*-*īnām* and \*-*ūnām* for *i*-stems and *u*-stems, respectively). This is the only plural-ending in (M)Parth., but in (M)MP two other plural-endings, both of a latter date, have been identified: -*īhā* (originally an adverb-making suffix, usually added to inanimate nouns), and

---

**2**  Cf. Asha (1998: 7–10); Nyberg (1974: 277–278); Skjærvø (1983: 132–135).

-*agān* (adding a generic sense to the noun and designating a multitude of individuals). In addition, in some MParth. and (M)MP texts the abstract nouns formed with the endings -*īft* and -*īh*, respectively, are used as collectives or plural nouns.

(2) Examples: BPah *wuzurg-ān* 'the great ones', *ātaxš-ān* 'fires', *frazand-īn* 'children', *kōf-īhā* 'mountains', *jahūd-agān* 'a multitude of Jews', *harborz kōf-īh 2244 ast* 'Harborz has 2244 mountains'; MPth *brādar-ān ud wxār-īn* 'brothers and sisters', *dušmin-ūn* 'enemies', *pad harw bay-īft* 'among all gods'.

In WMIr strict concord between a plural noun and a modifying adjective, whether attributive, demonstrative, or quantitative, was not obligatory, but it often occurs, e.g., BPah *astōmand-ān mardōm-ān* 'corporeal men'; MPth *šubān-ān rāšt-ān* 'righteous shepherds'.

A noun modified by a numeral often occurs in the singular in WMIr, but the use of a plural noun after numerals larger than one is not infrequent, especially in MPth and MMP, where in most cases the agreement can be attributed to rhythmic factors (Brunner 1977: 46), e.g., MPth *dō brādar-ān* 'two brothers'.

In WMIr a predicate substantive need not agree in number with a plural subject or antecedent, e.g., MMP *drōzan ma hēb bawēnd* 'They should not be liars.'

Sometimes in a substantive phrase, the modifying noun is plural, not the regent noun, e.g., MMP *abārīg-ān dēn ī pēšēnag-ān* 'the other religions of the ancients'.

## 2  Plural marking in New Persian

The plural markers used in Early Modern Persian (EMP) are: -*ān*, -*(i)hā*, and -*(a) gān*, e.g., *nazdīk-ān* 'relatives', *derakht-ān* 'trees', *dast-ān* 'hands', *setāra-gān* 'stars', *setam-hā* 'cruelties', *sepīd-hā* 'the white ones', *sar-ihā* 'heads', *kār-ihā* 'deeds', *farzand-agān* 'children', *dōst-agān* 'friends'. One can also identify -*hān* as a plural suffix indicating approximation, e.g., *ānjā-hān* '(lit. those places) around there' (Lazard 1963: 195–196).

In classical Persian texts there are instances in which a modifying adjective has agreed with a plural noun (e.g., *ferishtag-ān i pāk-ān* 'pure angels'). In addition, sometimes the noun, especially when specific, has appeared in plural form after numbers larger than one (e.g., *har du zan-ān* 'both women').

The plural-endings of contemporary Persian (CP) are -*ān* and -*(h)ā*. The ending -*ān* is often used for human and human-related items, especially in literary registers. In addition, plural markers of Arabic origin are seen in all stages of

New Persian, mostly in Arabic loan words, such as *-ālāt* (*zinat-ālāt* 'ornaments'), *-āt* (*taʿlim-āt* 'teachings'), *-in* (*moʿallem-in* 'teachers'), *-iy(y)āt* (*qazal-iy(y)āt* 'sonnets'), *-jāt* (*nevešte-jāt* 'writings'), *-un* (*enqelābi-y-un* 'revolutionary people'), and broken plurals of different types, e.g., *kotob* 'books' (Sg. *ketāb*), *afrād* 'individuals' (Sg. *fard*), *romuz* 'secrets' (Sg. *ramz*), etc.

The addition of a plural marker to an already plural noun is occasionally seen in both EMP and CP, e.g., *khalifazāda-gān-hā/khalīfa-zāda-gān* 'descendants of the caliph' (Sg. *khalīfa-zāda*), *ahvāl-hā/ahvāl* 'conditions' (Sg. *hāl*).

# 3 Plural marking in New West Iranian languages and dialects

In modern Iranian languages and dialects, pluralization is generally performed by adding certain affixes. In the study carried out for this chapter, a group of New West Iranian languages and dialects were investigated, and the typological relationships in them were extracted. It should be added that in some of these languages and dialects the previously mentioned plural markers of Arabic origin are, under the influence of Persian, occasionally witnessed. The data is as follows:

**Abuzaydābādi:** *e* (replacing final *-a*), *-(h)ā*, *-hō*, *-un*, *pāk-* (to emphasize the concept of plurality): *espe* 'dogs' (Sg. *espa*), *pāk-reg* 'veins', *pāk-mü* 'hair' (Pl.) (Schmitt 1989: 317; Mazraʿati et al. 1374: 44).

**Abyānai** *-a*, rarely *-(h)ā*, *-hō*, *-un*: *dādā-a* 'sisters', *döt-a* 'girls', *raz-a* 'gardens', *bone-a* 'trees' (Sg. *böna*), *mesh-a* 'sheep' (Pl.); but: *kaye-a* 'houses' (Sg. *kaya*), *küye-ā* 'dogs' (Sg. *küyā*), *voche* 'children' (Sg. *vocha*). The plural form is used after numerals larger than one: *dō mere-a* 'two men', *sapūra* 'three boys' (Āhani 1993: 54; Schmitt 1989: 317; Yarshater 1983b: 404).

**Aftari:** *-(h)on* (*kāku-hon* 'uncles', *vachkā-hon* 'children', *dār-on* 'trees'). With respect to demonstratives, only the noun is pluralized: *ven mirdon* 'those men' (.f. *ven mirde* 'that man'), *an kargon* 'these hens' (cf. *anjeki* 'this woman'; Windfuhr 1984: 592; Homāyun 1992: 53, 57).

**Alviri:** *-hā* (Schmitt 1989: 314).

**Āmora'i:** *-gal* (*asb-gal* 'horses'). The feminine ending *-a* is omitted before the plural marker: *sü-gal* 'apples' (Sg. *süa*; Schmitt 1989: 314; Lecoq 1989: 956).

**Anāraki:** *-hā: chash-hā* 'eyes' (Schmitt 1989: 322; Windfuhr 1985: 3).

**Ārāni:** *-o(n)*, rarely *-hā*: *hayvon-o(n)* 'animals' *dērakht-o(n)* 'trees', *sag-o(n)* 'stones'. The plural marker *-hā* is mainly used in some personal, demonstrative, and interrogative pronouns: *hēmē-hā* 'we, us', *shēmē-hā* 'you', *nēmi-hā* 'these', *kiyā-hā* 'where, what places?' (Alijānzāda 1993: 203).

**Ardestāni:** *-(h)ā*, *-ū(n): pore-hā* 'sons', *bolobol-ū* 'nightingales' (Schmitt 1989: 322; Lecoq 1986: 387).

**Āštiyāni:** *-gal(ān)*, in rapid speech sometimes *-gel: mā-gal* 'mothers', *pūr-galān* 'sons'; the feminine ending *-a* is omitted before the plural marker: *khu-gal* 'shovels' (Sg. *khua*) (Schmitt 1989: 314; Yarshater 1987: 848).

**Āzari:** See Tāti.

**Bādrudi:** *-(h)ā*, *-hō*, *-e*, rarely *-un*: *bāl-e* 'spades', *böz-e* 'goats'. In ergative constructions when the logical direct object is plural, sometimes the verb also takes the plural marker: *dözz-e sarbāz-e-shun dar-kos-e* 'the thieves hit the soldiers'. When the verb or a referring pronoun indicates the plural number, often the singular is used for the plural: *sarbāz ö döz da:vā-shun be-ka* 'the soldiers and the thieves quarreled' (Schmitt 1989: 317; Yarshater 1989: 383).

**Bakhtiyāri:** *-(h)ā*, *-(y)ā*, *-vā* (for animals and inanimate objects); *-ū(n)*, *-gal*, *-yal* (for humans and animals): *māl-ā* 'tents', *gandom-ā* 'wheat (Pl.)', *čo-vā* 'pieces of wood', *ded-ūn/dedū-yal* 'sisters', *kor-gal/yal* 'boys', *bače-yal* 'children', *gā-w-ūn/ gā-hā* 'cows', *guar-gal* 'calves', *khar-gal* 'asses', *āδom-yal* 'people', *shākh-ā* 'horns'.

Bakhtiyāri of Ardal: *-ā*, *-v/wā* (after *-u* and sometimes *-ow*), *-yā* (after *-i*), *-un*, *-jal*: *dodar-gal* 'girls', *mish-gal* 'ewes', *mir-gal* 'men' (Sg. *mire*), *zin-gal* 'women' (Sg. *zine*), *bi-jal* 'kids' (Sg. *big*), *sag-jal* 'dogs,' *dej-jal* 'sisters,' *du-vā* 'sour milk (Pl.)', *orosi-yā* 'shoes', *low-vā* 'lips', *show-ā* 'nights'.

In the Bakhtiyāri dialect of Ardal, adjectives may agree in number with the nouns they modify: *mir-gal gap-ā* (even *mire gap-ā*) 'old men' (Sg. *mire gap*) (Schmitt 1989: 344; Windfuhr 1988: 560; Eydi 1996: 59; Sālehi 1990: 27).

**Balochi:** Obl. *-ān*, *-gal* (most frequently in Eastern Hill Balochi). Kechi: Nom. *-ø*, Obl. *-ān* (*-ānrā*, emphasizing on the direct object), Gen.*-ānī*: *mardom-ānī but* 'men's boots', *sepāhī-ānī jamag* 'soldiers' shirts' (Schmitt: 354–355; Collet 1983: 5; Elfenbein 1988: 635).

**Bashkardi:** Northern: *-ā*, *-on*, *-ōn*, *-un*, *-ūn* (Southern: *-an*): *lahar-on* 'huts', *pā-'-on* 'feet', *ā-'-ūn* 'they, those' (Schmitt: 366; Skjærvø 1988: 848).

**Behbahāni:** *-ā*, *-(h)u*: *sarbāz-ā* 'soldiers', *das-ā* 'hands', *mard-u* 'men', *zen-u* 'women', *gā-hu* 'cows'. With respect to the numerals, the nouns are pluralized: *i pay sikhe kebāb-ā* 'these five rows of *Kabāb*', *u de doluv-ā* 'those two old women' (Mortazavi 1995: 36).

**Behdini:** *-un*, *-(h)ā*, *-(h)o*: *porog-un* '(small) sons', *dotog-un* 'daughters', *vachog-un* 'children', *ziy-un* 'women', *noker-o* 'servants', *khda-(h)ā* 'houses', *chmosh-(h)ā* 'shoes'; Kermāni: *-ā*: *khuk-ā* 'pigs'; Yazdi: *-hu*: *wacha-hu* 'children' (Schmitt 1989: 322; Windfuhr 1989a: 105; Firuzbakhsh 1997: 49).

**Bīdgoli:** *-o(n)*: *yür-o* 'children'; final *-∂* drops before the plural marker: *karg-o* 'hens' (Sg. *karg∂*) (Yarshater 1990: 247).

**Biyābānaki:** *-ūn*: *vashko-ūn* 'infants' (Schmitt 1989: 309).

**Burenjāni:** *-al*, *-gal*, *-yal*: *dar-al* 'doors', *khar-al* 'asses' (Schmitt 1989: 342).

**Čarza'i:** *-on* (Yarshater 1992: 22).

**Dashti:** *-ā*: *kal-ā* 'boys', *dovat-ā* 'girls' (Hājiyāni 2002: 70).

**Davāni:** *-al*, *-gal*, *-yal*, *-(h)ā*: *ādam-gal* 'people', *bachchek-gal* 'children', *sāl-gal* 'years', *gap-gal* 'conversations', *shāh-ā* 'kings', *un-ā* 'those', *ke-hā* 'who, which people?' (Schmitt 1989: 342; Mahamedi 1994: 130).

**Dimili**: See **Zāzā**.

**Elāshti:** *-hān*, *-(h)un*, *-kun*: *mār-un* 'mothers', *dār-un* 'trees', *kerk-un* 'hens', *kijā-kun* 'girls', *gu-hun* 'cows', *zan-un/zan-hān* 'women'.

**Esfahāni Jewish:** *-(')ā*, *-vā* (after *-u* and *-o*), *-yā* (after *-i*), *-un*: *chesh-ā* 'eyes', *ru-vā*, 'days', *'abri-yā* 'eyebrows', *bu-vā* 'smells', *hāmelo-vā* 'pregnant' (Pl.), *veyθa-'ā* 'those standing'.

Vowel lengthening is seen after *-ā*: *dev-:ā* 'drugs' (*devā* 'drug').

Final *-e* drops out before *-ā*: *keδ-ā* 'houses' (Sg. *keδe*).

If *-ā* follows the diphthongs /*ou*/, /*āu*/, or /*eu*/, the /*u*/ sound changes into /*v*/ or /*w*/: *khov-ā* 'sleeps' (Sg. *xou*), *δāv-ā* 'delivered woman' (Sg. *δāu*), *tevā* 'fevers' (Sg. *teu*) (Schmitt 1989: 320; Kalbāsi 1994: 53).

**Farizandi:** *-(h)ā*, *-hō*, *-un* (Schmitt 1989: 317).

**Farvi:** *-gūn*: *geche-gūn* 'children' (Schmitt 1989: 324).

**Gazi:** *-ā*: *gul-ā* 'flowers' (Schmitt 1989: 320).

**Gilaki:** *-an*, *-ān*, *-en*: *vokht-an* 'moments' *jāy-ān* 'locations', *dehāt-en* 'villagers' (Schmitt 1989: 304).

**Gurāni:** Awromāni: *-e*, Obl. *−ā* (MacKenzie 1987): Indefinite: Masculine *-ê*, Obl. *-â*: *har-ê*, Obl. *har-â* 'asses'; Feminine *-ê*, Obl. *-a*: *mâhar-ê*, Obl. *mâhar-a* 'she asses'; Definite: Masculine *-ake*, Obl. *-akâ*: *har-ake*, Obl. *har-akâ* 'the asses'; Feminine *-kê*, Obl. *-kâ*: *mâhar-amâhara-kê*, Obl. *mâhara-kâ* 'the she-asses'.

Nouns ending in stressed *-a* and feminine nouns in stressed *-e* drop their final vowel before plural endings: *yāne*, Obl. *yānā* 'houses' (Sg. *yāna*), *kınāche*, Obl. *kınāchā* 'girls' (Sg. *kınāche*).

Nouns ending in stressed *-ā* replace it with *-ɛ* (in the direct plural) and *-āyā* (in the oblique plural): *pıɛ*, Obl. *pıāyā* 'men' (Sg. *pıā*).

Feminine nouns ending in *-i* preserve this in the direct plural case, but only in isolation. When they are joined with any other form (except a simple epithet), they are inflected as if they ended in *-a*: *kārdi* (*kārde*), Obl. *kārdā* 'knives' (Sg. *kārdi*).

The plural form is used after numbers larger than one: *shısh suāre* 'six horsemen'.

Adjectives agree with the nouns they modify: *kitebe siāwe* 'black books'.

In the oblique plural the succession of two words, noun and epithet, each with *-ā*, is avoided, the first *-ā* being reduced to *-a*: *yāne gawre*, Obl. *yāna gawrā* (instead of *\*yānā gawrā*) 'big houses' (Schmitt 1989: 337; MacKenzie 1987: 111; MacKenzie 1966: 13–15, 17–18, 23).

**Harzani:** See **Tāti**.

**Jowšaqāni:** *-(h)ā*, *-hō*, *-un* (Schmitt 1989: 317).

**Kafrāni:** *-ō*: *dendōn-ō* 'teeth' (Schmitt 1989: 320).

**Kahaki:** *-e*, *-ow*, *-iya*, -iye. The final vowels of words are omitted before the plural markers: *pur-e* 'sons' (Sg. *pura*), *khāk-e* 'sisters' (Sg. *khāka*), *zām-ow* 'sons-in-law' (Sg. *zāmā*), *varz-ow* 'oxen' (Sg. *varzā*), *zan-iya* 'women' (Sg. *zana*), *da'-iye* 'maternal uncles' (Sg. *da'i*), *sabz-iye* 'vegetables' (Sg. *sabzi*) (Schmitt 1989: 314; Shari'ati 1998: 58).

**Kāšāni Jewish:** *-(h)ā*, *-hō*, *-un* (Schmitt 1989: 317).

**Kelārdashti:** *-on*, *-kon*, *-shon*, *-eshon* (after consonants), rarely *-koneshon*, *-ville* (only in *vache-ville* 'children'): *berār-on* 'brothers', *zenā-kon/-shon* 'women', *rikā-shon* 'boys', *kuh-eshon* 'mountains', *merdā-koneshon* 'men'.

**Keringāni:** See **Tāti**.

**Kermāni Jewish:** *-ā*: *āsin-ā* 'sleeves' (Schmitt 1989: 322).

**Keshe'i:** *-(h)ā*, *-hō*, *-un* (Schmitt 1989: 317).

**Khānsāri:** *-ā*, *-yā* (after *-i*), *-vā* (after *-u*): *ādem-ā* 'people', *zemin-ā* 'fields', *māni-yā* 'mothers', *chu-vā* 'pieces of wood'; note also *vech-ā* 'children' (Sg. *vecha*), *kiy-ā* 'houses' (Sg. *kiya*) (Schmitt 1989: 316; Tavakkoli 1995: 71).

**Khuri:** *-un*, *-gun*, *-ugun*, *-un*, *-ō*: *shōn-ugun* 'women', *kitēb-un* 'books', *heyvun-un* 'animals', *chemchā-gun* 'spoons', *mardu-gun* 'men', *mī-v-ō* 'hair' (Pl.) (Schmitt 1989: 324; Farahvashi 1976: 66).

**Komjāni:** *-e*, *-ye* (after vowels): *rus-e* 'roosters', *raz-e* 'gardens', *vahi-ye* 'kids', *māli-ye* 'cats', but: *kay-e* 'houses' (Sg. *kayā*), *küy-e* 'dogs' (Sg. *küyā*), *mer-e* 'men' (Sg. *merā*) (Āhani 1993: 52).

**Kumzāri:** *-ā*, *-an*: *say-ā* 'dogs', *zank-an* 'women'. Nouns and adjectives agree in number: *zank-en gāp-an* 'big women' (Schmitt 1989: 366).

**Kurdish:** Northern: Ezāfa *-ê*, *-ên*, Obl. *-a(n)*; Eastern: Ezāfa *-ê(t)*, *-êd* (Mokri *-î da*); Mahābādi: *-ān*, *-wān* (after *-ū*), *-yān* (after all vowels except *-a* and *-ū*): *kich-ān* 'girls', *pyāw-ān* 'men', *khānū-wān* 'houses', *tarāzū-wān* 'scales', *birā-yān* 'brothers'.

Final *-a* is omitted before *-ān*: *zhinak-ān* 'the women' (Sg. *žinaka*), *birāk-ān* 'the brothers' (Sg. *birāka*). Sometimes the plural form is used after numerals larger than one: *shash ṛōzh-ān* 'six days' (Schmitt 1989: 330; Kalbāsi 1983: 18).

**Laki:** *-al, -yal*:*zhan-al* 'women', *āyil-al* 'kids', *kowerr-al* 'sheep' (Pl.).

Čahār-duli: *-ān:* *dōtak-ān* 'the girls', *kilwak-ān* 'the stones'.

Darra-shahri: *-aḷ, -yaḷ, -waḷ*: *kↄtow-aḷ* 'books', *koṛ-aḷ* 'boys', *khuri-yaḷ* 'nets', *dü-yaḷ* 'smokes', *pat-waḷ* 'blankets'.

Khājvandi-ye Kelārdasht: *-al, -gal, -yal* (after vowels), *-ān*: *det-al/det-gal/det-ān* 'daughters', *shi-al/shi-gal/shi-ān* 'husbands'.

Exceptions: *veryl* 'lambs' (Sg. *vark*); D. *bↄrā:ḷ* 'brothers' (Sg. *bↄrāḷ*); Khājvandi *ʿāylevil* 'children' (Sg. *ʿāyl*) (Izadpanāh 1988: 45–46; Mir-Cherāqi 1990: 17; Purmand 1995: 161; Sepehri 1999: 38).

**Lāri:** *-ea* (esp. in Evazi, Khonji, and Bastaki), *-ūn* (poetic), *-iyā*, *-niyā* (after vowels): *asp-iyā/asp-ea* 'horses', *govv-iyā/ gowv-ea* 'cows' (Sg. *gow*) (Schmitt 1989: 366; Eqtedāri-ye Lārestāni 1992: 486).

**Lāsgerdi:** *-on*: *dār-on* 'trees', *pur-on* 'sons' (Schmitt 1989: 309).

**Lori:** *-yā*, *-o* (for animates): *dār-yā* 'trees', *das-yā* 'hands', *dokhter-o* 'girls', *bach-o* 'children'; (Schmitt 1989: 344; Windfuhr 1990: 325; Izadpanāh 1984: 6; Moqimi 1994: 44–46).

**Boyer-Ahmadi:** *-al*, *-val*, *-yal*, *-gal*, *-un*: *bow-yal/-val* 'fathers', *merd-al* 'men', *kela-yal* 'hats', *zan-gal* 'women', *dwar-al/-un* 'daughters', *sāl-al/-un* 'years', *mãh-al/-un* 'months'; in this dialect sometimes the final /a/ changes to /e/: *bache-yal* 'children' (Sg. *bacha*).

**Feyli:** *-(ī)ā:* *bav-ā* 'fathers', *āyam-īā* 'people', *zan-īā* 'women'.

**Giyāni:** *-yā:* G. *asp-yā* 'horses', *mū̄-yā* 'hair' (Pl.).

**Kohgiluya'i:** *-yal*

**Mamasani'i:** *-al*, *-gal*, *-yal: -ā:* *asp-al* 'horses', *vazīr-al* 'ministers', *hūna-yal* 'houses'.

**Nahāvandi:** *ā* (Schmitt 1989: 344; Windfuhr 1990: 325; Izadpanāh 1984: 6; Moqimi 1994: 44–46).

**Mahallāti:** *-ā*: *jūnavar-ā* 'animals' (Schmitt 1989: 316).

**Māsarmi:** *-al*, *-gal*, *-yal*: *ādam-al* 'people' (Schmitt 1989: 342).

**Māzandarāni:** *-ho(n)*, *-un*, *-kun*: *zon-ho(n)* 'women', *per-un* 'fathers', *rikā-(k)un* 'boys'.

> **Sāravi:** *-un* (only for human beings), *-ā* (after consonants), *-yā* (after *-i*), *-hā* (after *-e*, *-ā*, and *-u*): *mari-y-un* 'men', *rikā-'-un* 'boys', *piyer-un* 'fathers', *zan-un/-ā* 'women', *vach-un/vache-hā* 'children', *dāmād-ā* 'sons-in-law', *piyala-hā* 'cups', *sikā-hā* 'ducks', *siyu-hā* 'black ones'.

> **Velātru'i:** *-an*, *-(h)ā* (Schmitt 1989: 305; Kalbāsi 1997: 38–39; Najafzāda Bār-Foruš 1989: 61; Shokri 1995: 70–71).

**Meyme'i:** *-(h)ā*, *-hō*, *-un* (Schmitt 1989: 317).

**Nā'ini:** *-ā*, *-jā*, *-wā*: *horchen-ā* 'stairs', *nemej-ā* 'felts', *ārt-ā* 'flour' (Pl.), *tshī-jā* 'things', *tī-jā* 'blades, razors', *ollu-wā* 'eagles', *poru-wā* 'sons', *jē-hā* 'barley' (Pl.), *guzā-hā* 'foods' (Schmitt 1989: 322; Pur-Ābedi-ye Nā'ini 1993: 85).

**Natanzi:** *-(h)ā*, *-hō*, *-un* (Schmitt 1989: 317).

**Pāpuni:** *-al*, *-gal*, *-yal*: *esfahān-yal* 'Esfahāni people' (Schmitt 1989: 342).

**Qohrudi**: *-(h)ā*, *-hō*, *-un* (Schmitt 1989: 317)

**Qomše'i**: See **Šahrezā'i**.

**Rāji (of Delijān):** *-ye*: *peur-ye* 'sons', *deji-ye* 'daughters' (Sg. *dejo*) (Safari 1994: 60).

**Šahrezā'i:** *-ā*, *-vā* (after *-u*), *-yā* (after *-i*): *lāħāf-ā* 'mattresses', *dokhter-ā* 'girls', *bālu-vā* 'airplanes', *patu-vā* 'blankets', *sandili-yā* 'chairs', *küdi-yā* 'marrow (Pl.)'; but: *khun-ā* 'houses' (Sg. *khune*), *bechā* 'children' (Sg. *beche*).

In descriptive genitive structures, the modifier is pluralized: *derakh tut-ā* 'mulberry trees', *late gandom-ā* 'wheat farms'. In possessive genitives, however, the head is pluralized: *dar-ā bāq* 'the doors of the garden', *rakhd-ā khārum* 'my sister's clothes' (Tāki 1993: 32).

**Sangesari:** *-hā*, *-un* (Schmitt 1989); *-ø*, Obl. *-uon*, Agent-case (for the past form of transitive verbs) *-uony*: *ke-hā*, *ke-v-un* 'houses', *pur-un* 'sons'.

Irregular plurals: *bochaw*, Obl. *bochuon* 'kids' (Sg. *bochae*), *kawtaw*, Obl. *kawtuon* 'boys' (Sg. *kuotae*), *mird*, Obl. *mirduon* 'men' (Sg. *maerkeyin*), *var(r)aw*, Obl. *var(r) uon* 'lambs' (Sg. *var(r)ae*), *zawkaw*, Obl. *zawkuon* 'children' (Sg. *zaw*), *zhinni*, Obl. *zhinnuon* 'women' (Sg. *shaekeyin*).

After numerals larger than one the plural form is used: *du zhinnuony baevaw* 'the two women said'.

**Sedehi:** *-hō*: *veche-hō* 'children' (Schmitt 1989: 320), also *−ā*: *žen-ā* 'women', *meli-y-ā* 'cats', *so-v-ā* 'apples' (Esmā'ili 2001: 98).

**Semnāni:** *-i*, *-y* (after stressed vowels), Obl. *-un* (Sotuda, *-ø*, *-ey*): *pir-i*, Obl. *pir-un* 'sons', *sāl-i*, Obl. *sāl-un* 'years' (Sg. *sāla*), *kiyé-y*, Obl. *kiye-y-un* 'houses' (Sg. *kiyá*), *astey* 'fruit stones' (Sg. *astā*); also note: *dot-i*, *dotar-i*, Obl. *dotar-un* 'daughters' (Sg. *dota*) (Schmitt 1989: 307; Sotude 1963: 7).

**Shahmirzādi:** *-(h)ā*, *-on*: *pier-hā* 'fathers', *tifl-ā* 'infants', *zan-on* 'women', *mard-an* 'men' *pul-(h)ā* 'money' (Pl.). ditto

**Sivandi:** *-gar*: *det-gar* 'daughters' (Schmitt 1989: 346).

**Somqāni:** *-al*, *-gal*, *-yal*: *mīsh-gal* 'sheep' (Pl.), *ādam-yal* 'people' (Schmitt 1989: 342).

**Sorkhe'i:** *-ā(n)*, *-hā*: *mirdi-ān* 'men', *zhiki-ā(n)* 'women', *dār-hā* 'trees' (Schmitt 1989: 309).

**Tāleshi:** Lankorāni: *-on*: *do-on* 'trees'; Māsāli: *-en*, Obl. *-un*: *dār-en* 'trees'; Māsula'i: *-en*, Obl. *-on*: *dār-en* 'trees'; Paresari: *-e*, Obl. *-un*: *dōr-e* 'trees'; Tulārudi: *-e*, Obl. *-an*: *dōr-e* 'trees'; Vizna'i: *-on*: *dō-on* 'trees'; Zida'i: *-an*, Obl. *-un*: *dār-an* 'trees' (Schmitt 1989: 299).

**Tāri:** *-(h)ā*, *-hō*, *-un* (Schmitt 1989: 317).

**Tāti:** (Āl-e Ahmad 1977: 171; (Schmitt 1989: 301, 303); Kārang 1954: 83, 104–105; Yarshater 1969: 75–83; Zokā' 1953: 38). The data to come are all taken from the subvarieties of Tati:

**Čāli:** *-e*, *-ē* (after *-á* and *-ā*), Obl. *-ō(n)*: *qoč-e* 'rams'. Also note that the feminine morpheme *-a* disappears before plural endings: *barr-e* 'spades' (Sg. *barra*); nouns ending in stressed *-a* or in *-ā* have their direct plural in stressed *-ē* and drop their final vowel before *-ō(n)*: *dad-ē* 'fathers' (Sg. *dada*); *p-ē* 'feet' (Sg. *pā*); *chuār-ō* 'sheep' (Sg. *chuārā*); nouns in *-i* tend to drop the plural ending in the direct case: *sāri(-e)* 'stars'; and nouns in *-u* have their direct plural in *-ö*: *höll-ö* 'peaches' (Sg. *höllu*).

**Dānesfāni:** *-on*.

**Ebrāhimābādi:** *-en*, Obl. *−ēn*:

**Esfarvarini:** *-end*, Obl. *-o(n)*:

**Eshtehārdi:** *-ehā*, Obl. *−un*: *esb-ehā* 'dogs'. The feminine morpheme drops out before the oblique plural ending: *mish-un* 'ewes' (Sg. *misha*); and final stressed *-a* coalesces with *-e-* in *-ehā* and changes into *-i* before the oblique plural ending: *esb-ehā*, Obl. *esbi-un* 'dogs' (Sg. *esba*).

**Harzani:** *-oy*, *-un*, Obl. *−un*: . *kin-oy* 'girls', *yan-oy* 'women', *küy-un* 'mountains'.

**Hazār-rudi:** *-e*, Obl. *−ān*: *asb-e* 'horses'.

**Kajali:** *-e*, Obl. *−o*: *berāleg-e* 'brothers', *zhaneg-e* 'women';

**Keringāni:** *-ende/-inde*, *-nde* (after vowels), Obl. *−un*: *yan-inde* 'women', *kinā-nde* 'girls'.

**Khiyāraji:** *-e(n)*, *-eyn* (for nouns ending in stressed *-a*), Obl. *-un*: In Khiyāraji final stressed *-a* drops out before the plural morpheme: *ki-eyn*, Obl. *ki-un* 'houses' (Sg. *kia*).

**Khoznini:** *-in*, *-en* (for nouns ending in stressed *-a*), Obl. *-un*.

**Sagzābādi:** *-e*, *-yn* (after *-ā*), Obl. *-un*: *quch-e* 'rams', *zumā-yn* 'sons-in-law'. In this didalect the feminine morpheme drops out in the plural; *-a* and *-ā* drop out in oblique plural; and nouns ending in stressed *-a* have their direct plural in *-(y)n*, with *-a* palatalized into *-e*: *bādie-yn* 'bowls' (Sg. *bādia*).

**Shāli:** *-ān*

**Tākestāni:** *-on*: qoch-on 'rams'. In Tākestāni *-a* and *-ā* drop out before the plural morpheme: *fer-on* 'boys' (Sg. *fera*).

After quantitative adjectives and numerals larger than one, the plural form is used in Chāli, Kajali, and Khiyāraji: Chāli: Obl. *shash suār-ō* 'six horsemen'; Kajali: *dö berāleg-e* 'two brothers'; Khiyāranji: *cand māhun* 'several months'.

**Tāti (of the Republic of Azerbaijan):** Northern: *-ho, -un: khune-ho* 'houses'; *kuk-un* 'sons'; Southern: *-hā, -un: khuna-hā* 'houses'; *mard-un* 'men' (Schmitt 1989: 297).

**Urāzāni:** *-ān*: *dār-ān* 'trees', *yālān* 'children', *sif-ān* 'apples' (Āl-e Ahmad 1970: 85).

**Vafsi:** *-ān, -e(n), -iya* (after vowels), *-iye* (after vowels): *bez-ān* 'goats', *dast-e* 'hands', *dūwār-e* 'walls', *dār-en* 'trees', *merd-iya* 'men' (Sg. *merda*), *zen-iye* 'women' (Sg. *zene*) (Schmitt 1989: 314; Moqaddam 1939: 119).

**Vāneshāni:** *-ō(n), -ū*: *ferzend-ō* 'children', *khūk-ōn* 'pigs', *dandūn-ū* 'teeth' (Schmitt 1989: 316). Also *-ā*: *merd-ā* 'men', *mi-y-ā* 'hairs', *xuā-h-ā* 'sisters' (Jalāli-Dehkordi 1995: 15).

**Varzena'i:** *-ā*: *ādem-ā* 'people' (Schmitt 1989: 320).

**Vidari:** *-hā* (Schmitt 1989: 314).

**Yaran(d)i:** *-(h)ā, -hō, -un* (Schmitt 1989: 317).

**Zāzā:** *-ē*, Obl. *-ân*; Ezāfa: Masculine *-ê, -dê, -yê*; Feminine *-ê, -ynē p'ōstālē min* 'these my shoes' (Schmitt 1989: 339; Asatrian 1995: 407).

**Zefre'i:** *-gelō, -hō*: *pūr-gelō* 'boys', *veche-hō* 'children' (Schmitt 1989: 322).

# 4 Conclusion

As illustrated in the above data, the most frequent plural markers in New West Iranian languages and dialects are the following endings, pronounced in slightly different ways in different languages: (A) *-hā*, from MP *-īhā*, functioning both as

an adverb-making suffix and as a plural-ending, but originally an abstract noun maker (Schmitt 1989: 258); (B) -*ān*, from WMIr general plural-ending -*ān* (< OIr *\*-ānām*), which originally designated the plural oblique case; and (C) -*gal*, from OIr *\*gr̥da*- 'troop' (Nyberg 1974: 48; cf. Pers. *galla* 'flock, herd'), which originally designated a multitude of things.

The study of the above-mentioned data also reveals the following interesting linguistic facts. First and foremost, almost all New West Iranian languages tend to pluralize their nominal items via suffixation. The only exception is Abuzaydābādi, in which a prefix (i.e., *pāk-*) is also used. However, it is also noteworthy that some plural markers are in fact made up of two independent plural-endings, e.g., -*galān* in Āshtiyāni, -*gelō* in Zefre'i, and possibly -*hān* in Elāshti. It was seen that the singular is often used after numerals larger than one and quantitative modifiers. Only in a few languages (Abyāna'i, Awromāni dialect of Gurāni, Behbahāni, Chāli, Kajali, Khiyāraji, Mahābādi Kurdish, Sangesari) the plural is used. In terms of concord between a plural noun and a modifying adjective, whether attributive, demonstrative, or quantitative, such a case does not seem to be frequent; however, it occurs in Awromāni dialect of Gurāni, Bakhtiyāri of Ardal, and Kumzāri. In terms of gender, the feminine morpheme often drops out before plural endings (e.g., in Āmora'i, Āshtiyāni, Chāli, Eshtehārdi, and Sagzābādi). Finally, it was shown that the addition of a nominal plural marker to a transitive verb in ergative constructions appears only in Bādrudi.

# References

Āhani, M. R. 1993. *Barrasi-ye guyesh-e Abyāne az tavābe'-Natanz* [*A Study of the Dialect of Abyāne: A District of Natanz*]. Tehran: University of Tehran MA thesis of linguistics.

Ahmadi-Nargese, R. 1996. *Sākht-e vāzhagāni-ye guyeš-e Lori-ye shahr-e Khorramābād va bakhsh-e Chegeni* [*The Morphology of the Lori Dialects of Khorramābād and Chegeni*]. Tehran: Allameh Tabataba'i University MA thesis of linguistics..

Āl-e Ahmad, J. 1970. *Urāzān* [*Urāzān*]. Tehran: Ketāb-Xāne-ye Dāneš.

Āl-e Ahmad, J. 1977. *Tātneshinhā-ye Boluk-e Zahrā* [*The Tāti inhabitants of Boluk-e Zahrā*]. Tehran: Amir Kabir.

Alijānzāda, H. 1993. *Zabān-e Kavir: Tahqiq dar bāra-ye zabān-e Ārān va Bidgol* [*The Language of the Desert: A Study on the Language of Ārān and Bidgol* ]. Tehran: The author.

Amini-Abyāna, F. 1994. *Barrasi-ye barkh-i vizhagihā-ye dasturi dar guyesh-e Abyāna* [*A Study of Some of the Features of Abyāna Dialect*]. Tehran: Allameh Tabataba'i University MA thesis of Persian.

Asatrian, G. 1995. Dimilī. In Ehsan Yarshater (ed.), *Encyclopedia Iranica*, vol. VII, 405–411. Costa Mesa, Calif.: Mazda Publishers.

Asha, R. 1998. *The Persic (Pahlavi): A grammatical précis*. Vincennes: Ermān.

Azami, C. A., and G. L. Windfuhr. 1972. *A dictionary of Sangesari with a grammatical outline*. Tehran: Šerkat-e Sahāmi-ye Ketābhā-ye Jibi.

Brunner, C. J. A. 1977. *Syntax of Western Middle Iranian*. New York: Caravan Books.

Christensen, A. 1921. *Les dialects d'Awromān et de Pāwä*. Copenhagen: Høst & søn.

Christensen, A. 1935. *Contributions à la dialectologie iranienne*. Copenhagen: Høst & søn.

Christensen, A. 1995. *Guyesh-e Gilaki-ye Rasht* [*The Gilaki Dialect of Rasht*]. Translated by J. Khomāmizāda. Tehran: Soruš.

Collet, N. A. 1983. *A grammar, phrase book, and vocabulary of Baluchi*. London: Burgess and Son Ltd.

Elfenbein, J. 1988. Baluchistan iii: Baluchi Language and Literature. In E. Yarshater (ed.), *Encyclopedia Iranica*, vol. III, 633–644. London; Boston: Routledge & Kegan Paul.

Eqtedāri Lārestāni, A. 1992. *Lārestān-e kohan va farhang-e Lārestāni* [*The Ancient Lārestān and the Lārestāni Dictionary*]. Tehran: Jahān-e Mo'āser.

Esmā'ili, M. M. 2001. *Guyesh-e Sedehi* [*The Dialect of Sedeh*]. Tehran: Islamic Azad University of Tehran (Science and Research branch) dissertation of Ancient Iranian languages and culture.

Eydi, R. 1996. *Barrasi-ye guyesh-e -e Bakhtiyāri: Guna-ye Ardal* [*A Study of the Bakhtiāri Dialect: The Dialect of Ardal*]. Tehran: University of Tehran MA thesis of linguistics.

Farahvashi, B. 1973. *Vāzha-nāma-ye Khuri* [*A Glossary o Khuri Dialect*]. Tehran: Markaz-e Mardom Šenāsi-ye Irān.

Farahvashi, B 1976. Vāzhe Nāme-ye Khuri [A Dictionary of Khuri]. Tehran: Vezārat-e Farhang o Honar (Markaz-e Mardomshenāsi-ye Irān).

Firuzbakhsh, F. 1997. *Barrasi-ye sākhtemān-e Dasturi-ye guyesh-e Behdinān-e shahr-e Yazd* [*A Study of the Grammatical Structure of the Behdinān Dialect in Yazd*]. Tehran: University of Tehran MA thesis of linguistics.

Gerāmi, S. 1992. *Tawsif-e sākhtemāni-ye goruh-e fe'li va esmi dar guyesh-e Lori-ye Boyer Ahmadi (lahja-ye Bāshti)* [*A Structural Description of the verbal and nominal phrases in the Lori dialect of Boir Ahmad (the Bāshti dialect)*]. Tehran: Allameh Tabataba'i University MA thesis of linguistics.

Hājiyāni, F. 2002. *Barrasiye towsifi, tatbiqi va risheshenākhtiye guyeshe Dashti* [*A Descriptive, Comparative and Etymological Study of Dashti Dialect*]. Tehran: University of Tehran dissertation of Ancient Iranian languages and culture.

Haqiri, M. 1992. *Barrasi-ye goruh-e fe'li va esmi dar guyesh-e Gilaki-ye Sawma'a-sarā* [*A Study of the Verbal and Nominal Phrases in the Gilaki Dialect of Sawma'a-Sarā*]. Tehran: Allameh Tabataba'i University MA thesis of linguistics.

Hātami-'ebād, A. A. 1996. *Barrasi-ye goruh-e esmi dar guyesh-e Koliyā'i-ye Sonqor* [*A Study of the Nominal Phrase in the Koliyā'i Dialect of Sonqor*]. Tehran: Allameh Tabataba'i University MA thesis of linguistics.

Homāyun, H. 1992. *Guyesh-e Aftari* [*The Aftari Dialect*]. Tehran: Pažuhešgāh-e Olum-e Ensāni va Motāle'āt-e Farhangi.

Humand, N. 1990. *Pazhuheš dar zabān-e Tabari* [*A Research on Tabari Language*]. Āmol: Ketāb-Sarā-ye Tāleb-e Āmoli.

Izadpanāh, H. 1984. *Farhang-e Lori* [*A Dictionary of Lori*]. Tehran: Āgāh.

Izadpanāh, H. 1988. *Farhang-e Laki* [*A Dictionary of Laki*]. Tehran: Mo'assese-ye Farhangi-ye Jahāngiri.

Jalāli-Dehkordi, M. 1995. *Barrasiye Guyesh-e Vāneshān* [*A Study of the Dialect of Vāneshān*]. Tehran: University of Tehran MA thesis of Ancient Iranian languages and culture.

Kalbāsi, I. 1983. *Guyesh-e Kordi-ye Mahābād* [*The Kurdish Dialect of Mahābād*]. Tehran: Mo'assese-ye Motāle'āt va Tahqiqāt-e Farhangi.

Kalbāsi, I. 1994. *Guyesh-e Kalimiyān-e Esfahān* [*The Dialect of the Jews of Esfahān*]. Tehran: Pažuhešgāh-e Olum-e Ensāni va Motāle'āt-e Farhangi.

Kalbāsi, I. 1997. *Guyesh-e Kelārdasht* [*The Dialect of Kelārdasht*]. Tehran: Pažuhešgah-e Olum-e Ensāni va Motāle'āt-e Farhangi.

Kārang, 'A. 1954. *Tāti va Harzani: Do lahja az zabān-e bāstān-e Āzarbāyjān* [*Tāti and Harzani: Two Dialects of the Ancient Language of Azarbāijan*]. Tabriz: Bongāh-e Matbu'āti-ye Vā'ez-Pur.

Khorramdel, M. 1968. *Sarf va dastur-e zabān-e Kordi-ye Surāni* [*Morphology and Grammar of the Sorani Kurdish*]. Mahābād: Movaffaqi.

Lazard, G. 1963. *La langue des plus anciens monuments de la prose persane*. Paris: Librairie C. Klincksieck.

Lecoq, P. 1986. Ardestāni. In Esan Yarshater (ed.), *Encyclopedia Iranica*, vol. II, 387–388. London: Routledge & Kegan Paul.

Lecoq, P. 1989. Āmora'ī. In Ehsan Yarshater (ed.), *Encyclopedia Iranica*, vol. I, 955–956. London; Boston: Routledge & Kegan Paul.

MacKenzie, D. N. 1966. *The Dialect of Awroman*, Copenhagen: Ejnar Munksgaard.

MacKenzie, D. N. 1987. Avromani. In Ehsan Yarshater (ed.), *Encyclopedia Iranica*, vol. III, 111–112. London; Boston: Routledge & Kegan Paul.

Māfi, K. 1995. *Barrasi-ye sākhtvāzha va sakhtgāh-e fe'l dar guyesh-e Kordi-ye tāyefa-ye Māfi-ye Qazvin* [*A Study of the Morphology and Verbal System of the Kurdish Dialect of the Māfi Tribe in Qazvin*]. Tehran: Allameh Tabataba'i University MA thesis of linguistics.

Mahamedi, H. 1994. Davānī: The Davānī Dialect. In Ehsan Yarshater (ed.), *Encyclopedia Iranica*, vol. V, 129–132. Costa Mesa, CA: Mazda Publishers.

Mazra'ati, A., and Mazra'ati, A. 1995. *Farhang-e Bizovoy (Abuzaydābād-e Kāšān)* [*A Bizovi Dictionary of Abu Zaydābād in Kāšān*]. Tehran: The Authors.

Mir-Cherāqi, S. F. 1990. *Barrasi-ye sākhtemān-e dasturi-ye guyesh-e Laki-ye Khājvandi-ye Kelārdasht* [*A Study of the Grammatical Structure of the Laki Khājvandi Dialect of Kelārdasht*]. Tehran: Parhām.

Mo'in, M. 1990. *Mofrad va Jam'* [*Singular and Plural*]. Tehran: Amir Kabir.

Moqaddam, M. 1939. *Guyeshhā-ye Vafs-o Āshtiyān o Tafresh* [*The Dialects of Vafs, Āshtiān and Tafresh*]. Tehran: Anjoman-e Irān-Vij.

Moqimi, A. 1994. *Barrasi-ye guyesh-e Boyer-Ahmad va …* [*A Study of Boir Ahmad Dialect and …*]. Shiraz: Navid.

Mortazavi, S. M. B. 1995. *Barrasi va tahlil-e vāj-shenākhti va dasturi-ye guyesh-e Behbahān* [*A Phonological and Grammatical Analysis of the Dialect of Behbahān*]. Tehran: University of Tehran MA thesis of linguistics.

Najafzāde Bārforush, M. B. 1989. *Vāzha-nāme-ye Māzandarāni* [*A Glossary of Māzandarāni*]. Tehran: Bonyād-e Neyšābur.

Naqib al-Qorrā', S. A. 1996. *Barrasi-ye guyesh-e Yazdi* [*A Study of the Dialect of Yazd*]. Tehran: Allameh Tabataba'i University MA thesis of linguistics.

Nyberg, H. A. 1974. *A manual of Pahlavi*, vol. II. Wiesbaden: Otto Harrassowitz.

Pur-Ābedini Nā'ini, H. 1993. *Barrasi-ye guyesh-e Nā'ini* [*A Study of the Dialect of Nā'in*]. Tehran: University of Tehran MA thesis of linguistics.

Purmand, K. 1995. *Barrasi-ye guyesh-e Laki-ye Darre-shahr* [*A Study of the Laki Dialect of Darre Shahr*]. Tehran: University of Tehran MA thesis of linguistics.

Safari, H. 1994. *Vāzha-nāma-ye Rāji (guyesh-e Delijān)* [*A Glossary of Rāji (The Dialect of Delijān*)]. Tehran: The Author.

Sālehi, S. A. 1990. *Il-e Bozorg-e Bakhtiyāri: Farhang-e vāzhagān-e Bakhtiāri* [*The Great Bakhtiāri Tribe: A Dictionary of Bakhtiāri Words*]. Tehran: Atlas.

Schmitt, R. 1989. *Compendium linguarum Iranicarum*. Wiesbaden: Ludwig Reichert.

Sepehri, J. 1999. *Barrasi-ye guyesh-e Laki ba lahja-ye Chahār-duli* [*A Study of Laki, the Chahārduli Dialect*]. Tehran: University of Tehran MA thesis of linguistics.

Shari'ati, M. 1998. *Barrasi-ye guyesh-e Kahak-e Tafresh* [*A Study of the Dialect of Kahak in Tafresh*]. Tehran: University of Tehran MA thesis of linguistics.

Shokri, G. 1995. *Guyesh-e Sāravi (Māzandarāni)* [*The Dialect of Sāri (Māzandarāni)*]. Tehran: Pažuhešgāh-e Olum-e Ensāni va Motāle'āt-e Farhangi.

Skjærvø, P. O. 1983. *The Sassanian inscription of Paikuli*, III/2: *Commentary*. Wiesbaden: Dr. Ludwig Reichert Verlag.

Skjærvø, P. O. 1988. Baškardi. In Ehsan Yarhater (ed.), *Encyclopedia Iranica*, vol. III, 846–850. London; Boston: Routledge & Kegan Paul.

Sotude, M. 1963. *Farhang-e Semnāni, Sorkha'i, Lāsgerdi, Sangesari, Shahmirzādi* [*A Dictionary of Semnāni, Sorkha'i, Lāsgerdi, Sangesari, Shahmirzādi*]. Tehran: Tehran University Press.

Tāheri, A. 1991. *Dastur-e zabān-e Tāti-ye Tākestān* [*A Grammar of the Tāti Language of Tākestān*]. Tehran: Allameh Tabataba'i University MA thesis of Persian.

Tāki, G. 1993. *Barrasi-ye guyesh-e Shahrezā (Qomsha)* [*A Study of the Dialect of Shahrezā (Qomsha)*]. Tehran: University of Tehran MA thesis of linguistics.

Tavakkoli, M. M. 1995. *Barrasi-ye sākht-e* āvā'i *va sarf dar guyesh-e Khānsār* [*A Study of the Phonology and Morphology of the Khānsār Dialect*]. Tehran: University of Tehran MA thesis of linguistics.

Valāyi, M. 1994. *Barrasi-ye ajzā'-e kalām dar guyesh-e Māzandarāni-ye Sharqi (lahja-ye Āmoli)* [*A Study of the Parts of Speech in the Eastern Māzandarani Dialect (The Dialect of Āmol)*]. Tehran: Allameh Tabataba'i University MA thesis of linguistics.

Windfuhr, G. L. 1984. Aftarī. In Ehsan Yarshater (ed.), *Encyclopedia Iranica*, vol. I, 592–593. London; Boston: Routledge & Kegan Paul.

Windfuhr, G. L. 1985. Anārakī. In Ehsan Yarshater (ed.), *Encyclopedia Iranica*, vol. II, 2–3. London: Routledge & Kegan Paul.

Windfuhr, G. L. 1988. Baktīārī Tribes iii: Baktīārī Dialect. In Ehsan Yarshater (ed.), *Encyclopedia Iranica*, vol. III, 559–561. London; Boston: Routledge & Kegan Paul.

Windfuhr, G. L. 1989. Behdīnān Dialect. In Ehsan Yarshater (eds.), *Encyclopedia Iranica*, vol. IV, 105–108. London; Boston: Routledge & Kegan Paul.

Windfuhr, G. L. 1990. Boir Ahmadī ii: The Dialect. In Ehsan Yarshater (ed.), *Encyclopedia Iranica*, vol. IV, 325–326. London; Boston: Routledge & Kegan Paul.

Windfuhr, G. L. 1991. Central Dialects. In Ehsan Yarshater (ed.), *Encyclopedia Iranica*, vol. V, 242–252. Costa Mesa, Calif.: Mazda Publishers.

Yarshater, E. 1969. *A grammar of Southern Tati dialects*. The Hague and Paris: Mouton.

Yarshater, E. 1983a. Abūzaydābādī. In Ehsan Yarshater (ed.), *Encyclopedia Iranica*, vol. I, 401–402. London; Boston: Routledge & Kegan Paul.

Yarshater, E. 1983b. Abyāna'ī. In Ehsan Yarshater (ed.), *Encyclopedia Iranica*, vol. I, 404–405. London; Boston: Routledge & Kegan Paul.

Yarshater, E. 1987. Āštīāni. In Ehsan Yarshater (ed.), *Encyclopedia Iranica*, vol. II, 848–849. London: Routledge & Kegan Paul.

Yarshater, E. 1989. Bādrūdi. In Ehsan Yarshater (ed.), *Encyclopedia Iranica*, vol. III, 383–385. London: Routledge & Kegan Paul.

Yarshater, E. 1990. Bīdgol and Bīdgoli dialect. In Ehsan Yarshater (ed.), *Encyclopedia Iranica*, vol. IV, 247–249. London; Boston: Routledge & Kegan Paul.

Yarshater, E. 1992. Čarza. In Ehsan Yarshater (ed.), *Encyclopedia Iranica*, vol. V, 22–24. Costa Mesa, CA: Mazda Publishers.

Zokā', Y. 1953. *Guyesh-e Karingān (Tāti)* [*The Dialect of Karingān (Tāti)*]. Tehran: Ketāb-Xāne-ye Dāneš.

Thamar Eilam
# A typological sketch of the Jewish Iranian dialects

## 1 Introduction

Jewish languages and dialects emerged with the exposure of the Jewish people to various cultures, going into exile or remaining in their land during occupation by other nations. In both these situations, they adopted languages other than Hebrew for everyday life. However, in order to survive as a religion and a nation, they continued to use Hebrew, mostly as a sacred tongue. Naturally, some Hebrew and Aramaic words and calques – in some languages mostly of religious nature, in others also mundane words – found their way into the Jewish variety of the local language. (See Morag, Bar-Asher, and Modena 1999, especially Morag and Alvarez-Pereyre. See also Alvarez-Pereyre 1997; Bar-Asher 1978, 1989; Bunis 1993; Morag 1992). The majority of Jewish languages uses the Hebrew alphabet for writing. In many cases, the Jewish variety, partly cut off from the language of the majority, remains somewhat more archaic (Levi 1979: 58).

These three generalizations, however, are neither essential nor sufficient for classifying a language as Jewish: the presence and extent of the Hebrew component varies greatly even within dialects of the same Jewish language. In some cases, some Hebrew words permeate into the common language, e.g., *Chutzpa*[1] (impudence), *Mazal Tov* (congratulations), understood and used also by non-Jewish Americans.[2] Other Hebrew words or calques infiltrated languages in the Christian world through the Bible, e.g., English *shibboleth* (see also Meillet 1927), and Hebrew words often appear in descriptions of Judaism made by Gentiles.[3] Some Jewish languages – including Iranian varieties – use a local script. The preservation of archaisms is shared by most minority and peripheral dialects around the world, regardless of religion. The only criterion, then, common to all Jewish languages is social: their use predominantly by Jews.[4]

---

**1** Throughout this chapter, transliteration appears in italics, and transcription in boldface-type characters.

**2** For example from Judezmo in Spanish, see Bunis (1993: 17). For Judaeo-Arabic, see Bar-Asher (1978: §12); Maman (1989: 191 and fn 66). Yiddish: see Wexler (1981: 101 fn).

**3** In Iran, see Fischel (1974: 301–304).

**4** https://www.theatlantic.com/national/archive/2014/09/yiddish-has-a-problem/379658/; also see Zand's comment in Rabin (1979: 55–56), which ascribes to this a more exclusive nature.

# 2 Judaeo-Iranian – General

The Jewish community in Iran, one of the oldest in the world, dates back to the eighth century BCE (Asmussen 1968: 1; Adler 1905/1970: 84; 2 Kings 17:6). While adhering to their religious tenets even under severe persecutions, Iranian Jews have readily adopted Iranian culture and literature, and regard them as their own (Netzer 1981: 26–27).

Iranian Jews speak local dialects invaluable to linguistic research,[5] as well as unique varieties of Persian with dialectal and Hebrew/Aramaic influences. The Hebrew scroll of Esther, as well as the Aramaic "proto-Esther scroll" found in Qumran (Milik 1992; Shaked 1995), suggests that already in the Achaemenian period Iranian Jews spoke a language with a substantial Persian component. The earliest written documentation of Jewish Persian, however, dates back only to the eighth century CE and all JP documents known to date are varieties of New Persian.[6]

The term "Judaeo-Persian", then, refers to an array of very similar – usually mutually comprehensible – dialects of Persian, spoken or written by Jews in greater Iran over a period of more than a millennium. It roughly divides to Early Judaeo-Persian (EJP) before the Mongol invasion,[7] and Classical and Contemporary JP, from the fourteenth century onward. The latter differ from the common tongue only by use of a Hebrew component. EJP, on the other hand, has many unique features. EJP further divides into two dialect groups – southwestern (probably Pārs and Khuzistān) and northeastern (resembling Tadjik) (see Gindin 2003b).

Most JP dialects use Hebrew letters, in square or Rashi script, although some nineteenth- and twentieth-century dialects employ Arabic, Roman, or Cyrillic alphabets. (See Zand's comment to Rabin 1979: 56; Shapira 2002) They contain fewer Hebrew words than most Jewish languages – in some cases none at all. Their degree of archaism varies: in fact, some Classical JP works reflect a more *advanced* stage of Persian than their Muslim contemporaries. The speakers refer to their language as Fārsi[8] (EJP perhaps *Pārsi).

---

**5** See a more detailed list below.

**6** Shapira (2002) and especially (1999) suggests that there is evidence of Jewish Middle Persian in the *škand-gumānīg wizār*. The original Jewish text, however, remains undocumented.

**7** That is, up to the twelfth century. For a thorough description of the language, see Paul (2002).

**8** Abraham Aminof refers to Judaeo-Tadjik as "pure Fārsi language" in his Hebrew introduction to Shim'on Hākhām's translation of his book *Liquttei Dinim* (1900–1904).

The term "Judaeo-Iranian" (JI) includes JP as well as other, non-Persian varieties of Iranian dialects spoken by Jews, but usually never written. The earliest documentation of such a dialect comes from the Cairo Geniza (Shaked 1998).

Non-Jews refer to JP and Jewish dialects as *židi*, *judi*, or *jidi*,[9] "Jewish" in a derogatory sense. The speakers usually refer to their dialects by the name of the city (Yazdi, Esfahani, etc.). Hamadani Jews refer to their language also as *'ebri*, 'Hebrew' and the name Juhuri means 'Jewish'. In some cities one can hear more than one Jewish dialect. For example, in Tehran almost all dialects are present. Three different dialects are spoken in Rasht – Judaeo-Esfahani, Judaeo-Kashani, and Judaeo-Siahkali[10]; Judaeo-Yazdi is spoken only in the northern Jewish neighborhood (the southerners, having come from Hamadan, speak Persian "with a Yazdi accent").

All JI dialects face extinction because of massive immigration to Tehran, Israel, and the United States. Most Jews remaining in their cities of origin regard their language as a source of shame. They neglect to preserve it and prefer speaking Persian to their children. For this reason, written literature is completely absent, and folk songs are relatively rare. Besides the Geniza fragment, most documentations were made by twentieth-century scholars aware of this danger. The degree of documentation varies between single word quotes or single sample passages, and a complete grammatical analysis. The documented dialects include those of Esfahan (JE) (e.g., Abrahamian 1936; Netzer 1987, 1991; Kalbasi 1994–1995), Yazd (JY) (Homayoun 1998; Gindin 2003a), Kerman (JKer) (Lazard 1981),[11] Shiraz (JSh), Hamadan (JH) (e.g., Abrahamian 1936; Sahim 1994), Kashan (JKash) (Yarshater 2002), Khunsar (JKhun), Nehavand (JN), Borujerd (JB) (Yarshater 1989), and Golpayegan (JG),[12] as well as Juhuri,[13] also known as Judao-Tat, and Lutera'I (Yarshater 1977; Lazard 1978), which is a secret "jargon" rather than a dialect.

The JI group of dialects is a social rather than a linguistic group. The majority of documented JI dialects belongs to the Median group and shares isoglosses

---

**9** See Ethnologue website.

**10** A distinct "Judaeo-Rashti" dialect had no time to evolve, due to the recency of the Jewish settlement in this city. See Netzer (1987: 20).

**11** JY and JKer are very similar because Jews came to Kerman from Yazd in the nineteenth century.

**12** Most of these dialects have been discussed by Zhukovskijj (1922), Yarshater (1974), and Netzer (1987) and mentioned in Gindin (2002, forthcoming [Language]).

**13** For example, Miller (1892a, 1892b, 1903). Vitaly Shalem was the Juhuri informant for this chapter. Juhuri is regarded by some as a variety of New Persian (e.g., Shapira 2002). However, although as a Southwestern dialect (see Ethnologue website) it is closer to Persian than the Median dialects, it still shows substantial differences, to the point of difficulty in mutual intelligibility.

with their non-Jewish relatives. The only feature distinguishing JI dialects from non-Jewish Iranian dialects pertains to having an additional stock language, namely Hebrew.

This chapter presents a brief typological sketch of the motley group of JI dialects, with some emphasis on features of JY.


# 3 Judaeo-Iranian typology

All JI dialects are already heavily contaminated with Persian. Almost every feature discussed here has exceptions that are in fact Persianisms.


## 3.1 Phonology

When Jews began to emerge from the *mahale* in the beginning of the twentieth century, they were readily identified as Jews because of their pronunciation and tone, which differed from that of their Muslim neighbors. Differences also exist between the different cities, and persist even in the speakers' accent when speaking Hebrew.

JI dialects basically share the same consonant/vowel inventory as Persian. Some isoglosses of consonant/vowel inventories include the following:
- Pronouncing the sibilants *s*, *z* as interdentals (*θ*, *δ*, respectively) in Esfahan, and in some cases Shiraz and Kashan.
- Communities close to the border with Iraq (west of Kashan, e.g., Hamadan and Golpayegan) and to some extent also Juhuri speakers, pronounce ح, ع, and ق as in Arabic. Occasionally, they pronounce laryngeals, especially ع, where they do not exist, e.g., *'asp* 'horse' (a word of Persian stock).
- The rounded vowels *ü*, *ö* in the speech of some JY speakers. Others, as well as JKer speakers, pronounce these phonemes as *i* and *e*, respectively. The Persian equivalent of the former is *ūy* (e.g., *mü* 'hair' = Persian *mūy*) and for the latter *ū* < Middle Persian (MP) *ō* (e.g., *göš* 'meat' or 'ear' = Persian *gūšt* < MP *gōšt* and *gūš* < MP *gōš*) or Hebrew *ō* (e.g., *ölam* 'world', Hebrew *ōlām*). JN also has *ü*, as a variety of *u* (the two interchange even in the speech of the same speaker).[14]

---

**14** In JH and JN, the prefix appears only after another prefix or nominal complement; in JKash it requires the same environments, but occurs only in the present.

–  In JY, Persian final *-e* becomes *-o*: in words of Iranian stock, e.g., *gondo* 'big' (colloquial Persian *gonde*), of Arabic stock, e.g., *madreso* 'school' (Arabic *madrasa*, Persian *madrase*), and in single morphemes, e.g., *košto* 'killed' (loan from Persian *košte*) and the definite article *-o*, Persian *-e*. The other dialects retain the *-e* as it is.

In general, with the exception of southwestern Shirazi, JI dialects conform to the Median consonant shifts, e.g., Proto-Indo-European (PIE) $d^{(h)v}$ θ *b* (Persian *d*), e.g., JY, JE, JB *bar*, JKash *ber* (Yarshater 2002: 458, 461)[15] 'door' (Persian *dar*); PIE $\hat{g}^{(h)}$ θ *z* (Persian *d*), e.g., JKash *zumād*, JE *zumāz* 'bridegroom, son-in-law' (Persian *dāmād*); Palatalized PIE $g^{(h)w}$ (or voiced PIE *k^*) θ *j* (Persian *z*), e.g., JY *rej*, JKash *rūj* 'day' (Persian *rūz*, JSh *reδ* = *\*rez*), JY, JKash *jen* 'woman, wife' (Persian *zan*, JSh *zen*); PIE *k^w → sp/sb* (Persian *s*), e.g., JKash *esbe*, JY *esbo* 'dog' (Persian and JSh *sag*); PIE *tr* (Proto-Iranian *θr*) → *(h)r* (Persian *s*), e.g., JN *pür*, JY *por*, JKash, JE, JH, JB, JKhun *pir*, JKer *pur* 'son' (Persian *pus*, JSh *poθ* = *\*pos*); MP *x^wa-* → *xa* (Persian *xo*, *xu*), e.g., JY, JE, JH *xab* 'good' (Persian *xub*), JY *xā-* past stem of 'to eat' (Persian *xordan*).[16]

Some additional phonological isoglosses, pertaining to syllable structure, include the following:
–  Persian *ab/āb* (and Hebrew *av-*) is pronounced as *ow-* in JY, JSh, JN, and in some cases in Jkash and JE, e.g., JY *owen* 'sin' (Hebrew *āvōn*), JY and JE *sowz*[17] 'green' (Persian *sabz*), JSh *šow* 'night' (Persian *šab*), *xow* 'dream' (Persian *x^(w)āb*), JN *xov* 'dream', *aftow* 'sun', JN, JKash,[18] JE, JH *ow* 'water' (Persian *āb*, cf. JKhun *aw* and JY *iv*).
–  *Imāla*, the modification of stressed *ā* (sometimes *a*) to a front vowel, exists in JY in *iv* 'water' (Persian *āb*) and *dir*, present stem of "to have" (Persian *dār*). The first singular verbal suffix *-in* probably derives from *-am* in a similar manner. There is also evidence of *imāla* in Arabic words in the different EJP dialects, e.g., *kitēb* 'book' (Arabic *kitāb*).
–  Dropping the last consonant of the word unless immediately followed by a vowel may be regarded as a general human tendency. It is enhanced in these dialects because they have no normative writing system and are only documented by linguists. For example, Jkash, JY, and JSh *di* – past stem of "to see"

---

**15** Yarshater (1974: 458) quotes the same word as *bar*.
**16** But note that the present stem is *-xor-* like Persian. Other dialects have a back vowel in the past as well.
**17** Be-m-xorte is a perfect (Netzer 1987: 27).
**18** According to Netzer (1987: 24) and Yarshater (2002: 459, 462). In Yarshater (1974: 459) the same word is quoted as *āv*.

(from *did*), JY *eš-gere* 'he/she took' vs. *eš-gereft o šu* 'he/she took and went', JE *be-š-gif* 'he/she took' vs. *be-š-gift o* ... 'he/she took and ... '.

– When speaking Hebrew, most Persian Jews break initial clusters by inserting the vowel *e*, e.g., *selixā* 'excuse me/I'm sorry'. JY speakers insert the vowel *a*: *salixā*.
– In JY intervocalic *d* weakens to *r*, e.g., *xorā* 'God'.

Another phonological peculiarity of JY concerns the stress – Persian and most JI dialects have ultimate stress unless morpho-syntax requires otherwise, while JY usually has penultimate stress.

## 3.2 Morpho-syntax

### 3.2.1 Noun

JI dialects, like Persian and most Median dialects, lack case distinction except direct and oblique case in the pronoun. Gender has also disappeared, and *ezāfe* has been adopted from Persian (Yarshater 1974: 460).

– Definite nouns are marked by final *-e* (JY *-o*) as in Colloquial Persian.
– The plural marker is usually *-ā*. JY also has a definite plural marker *-onā* that is unique among Iranian dialects (Gindin 2004).
– Some JI dialects use a special numerative for counting Jews: berāxā (Hebrew 'blessing'), e.g., di berāxā pir '2 (Jewish) sons' (Netzer 1987: 43 with examples from JE).

### 3.2.2 Verb

JP, even in its earliest forms, shows a New Persian verb morpho-syntax.

– All dialects except Juhuri use the same stem system as Persian, that is, the present stem may be traced back to the Proto Indo-Iranian present system and the past stem to the perfect passive participle. Juhuri uses the same past stems, but the present stem parallels the Persian infinitive (e.g., *xurdenum* 'I eat'), and the Persian perfect passive participle serves for infinitive (*xurde* 'to eat').
– Most JI dialects have a nominative present and an ergative past. In most cases, however, ergativity is only morphological while the syntax already conforms to the nominative construction (see Gindin 2003a).[19] Modal verb

---

**19** See Gindin (2003a). The same situation is found in other dialects as well.

is ergative also in the present (Gindin 2003a), e.g., JY m-en, JKash m-e-gu 'I want'. In some dialects this construction is only partly ergative: the personal pronoun comes as a suffix rather than a prefix, e.g., JE gu-m-e 'I want', guma-m-e 'I wanted', JH gu-m 'I want', gā-m 'I wanted' (Yarshater 1974: 459). An exception is Juhuri, which has nominative constructions in the past as well, e.g., xurd-um 'I ate' just like raft-um 'I went'.

– JKash, JH, JG, and JKhom have a passive formation with -i- that goes back to Old Iranian -ya- passive, e.g., JH JKhom nahmer-i-u 'is not being broken'. JSh, JY, JE, and JK lack such a passive formation (Yarshater 1974: 459). This construction is found in MP and, interestingly, in EJP as well: in the southwestern dialect, which is closer to MP, it serves as the regular passive form, and in the northeastern one, which is closer to Dari, there are only traces of it.

– The Median dialects create their secondary past stem in -ā, equivalent to Median -ād (Persian and JSh -id), 20 e.g., JY 'to escape': present boraxs-, past boraxsā(r).21

– The imperfect marker serves, as in Persian, for expressing the present, as well as a prolonged or habitual action in the past. In JSh this marker is mi- as in Persian, and in the Median dialects (as in some other central dialects), it is e (JH, JN, JKash), or a (JE, JY, JKer. Its place in the word forms an interesting isogloss, illustrated in Table 1:

| | Suffix | Prefix | Selective prefix |
|---|---|---|---|
| Dialects | JE | JY, JKer | JH, JN, JKash[a] |
| Examples | JE *teja-un* 'I ran', *teja-un-e* 'I was running / I used to run'; *be-m-xort* 'I ate', *xorte-m-e* 'I was eating'.[b] | *a-xor-in* 'I am eating',[c] *m-a-xā* 'I was eating' | JH *dārtem xorāk-em-e-xo* 'I was eating food'; JKash *ce kār-e-kere* 'what are you doing?' (Yarshater 1974: 459) |

[a]In JH and JN the prefix appears only after another prefix or nominal complement; in JKash it requires the same environments, but occurs only in the present.
[b]*Be-m-xorte* is a perfect (Netzer 1987: 27).
[c]Stressed differently it means 'I will eat'; see below.

---

**20** Yarshater (1974: 459) mentions this suffix as the past marker for passive stems. See also Gindin (2003a: 107–108; 2004: 47).

**21** Young Israelis born to JY-speaking families mistake this verb to be part of the Hebrew component (the Hebrew root is *b.r.h* and the third person singular past is *barah*. Most Israelis pronounce *h* as *x*) . In fact, this is an inchoative version of Iranian *vi?raik*, Persian *gorixtan*, *goriz-*.

- The present stem of the verb "to have" in the majority of JI dialects is related to NP *dār-*. The past stem in most dialects relates to *dāštan*, e.g., JKash *eš-dād*, JN *eš-dārt*; J *dārtem* (first person singular), JE *dārt-ot* (second person singular as auxiliary for prolonged action). Other dialects employ a construction literally meaning 'there was to him', e.g., JY *eš-bū*. Juhuri[22] and EJP use this "existence" construction also in the present, e.g., EJP **zwry 'zmrš'n nyst** 'they have no power'.
- The present stem of *kardan* is *ker-* in the Median dialects, and *kon* in JSh.
- Unique to JY/JKer is the second person singular verb suffix (present and intransitive past) *-eš*. All other dialects use *-e* or *-i* for the same purpose.
- Another peculiarity of JY is the morphological stress, distinguishing between present and future: in the present, the imperfect marker *-a-* discussed earlier appears stressed, e.g., *á-xor-in* 'I eat'. In the future, the stress moves to the root syllable, e.g., *a-xór-in* 'I will eat'.

### 3.2.3 Some lexicographic isoglosses

As mentioned earlier, the use of Hebrew words and expressions is the sole isogloss that may differentiate JI dialects, including JP, from non-Jewish Iranian dialects.[23] Most of these words have religious or socioreligious meanings, e.g., *tefilā* (JY *tafilā*) 'prayer', *gūyim* (JY sometimes *güm*) 'a Gentile'. In political Iran, most Jewish dialects are heavily infused with Persian, which in turn is contaminated with Arabic. In the Caucasus, Juhuri has also some Russian, Turkic, and Caucasian influence.

The use of words of different etymologies can distinguish between the different JI dialects:
- Dog: JY, JKer *esbo*, JKash *esbe*, JE *kuδe*, JH *kuye*, JKhun *kuya*, JSh *sag*, Juhuri *seg*.
- Cat: JH, JE, JKash, JKhun *meli*, JSh *gorbe*, JY *gorbo*, Juhuri *nazu, pishik*.
- Big: JH *masar*, JKhun *māsar*, JY *mas/gondo*,[24] JKash *go:di*, JSh *gunda*, JE *bele*, Juhuri *kele*.
- Brother: JH *berā*, JE *beδār*, JKash *berār*, JKhun *borāy*, Juhuri *biror*, JY *berā/kākā*, JSh *kākā*.
- House: JH *kiye*, JE *keδe*, JKhun *kiya*, JY *kero*, JKash and JSh *θerā*, Juhuri *xune*.

---

**22** In Juhuri, *doshte* exists in the meaning: 'to preserve, keep, save'. In EJP, it appears in compound verbs.

**23** Most of the examples are taken from Netzer (1987: 24).

**24** Netzer has only *mas*, but I have also come across *gondo*.

- Some Median verb stems occur in some of the Median dialects: *gav-* 'to want',[25] *?vak* for speech, e.g., JY *eš-bā* 'he said', JKash *vaj-* 'to say', JNeh *be-š-vāt* 'he said' (cf. JSh *mi-ge* 'he/she says', *eš-go* 'he/she said', Juhuri *guftire* 'to say').
- All dialects, including EJP but excluding Juhuri, use the Iranian root *?šyav-* (Persian *šodan*) in the sense 'to go'. Juhuri, and in some cases EJP, uses *raftan*. [Juhuri infinitive *rafte*]

# 4  Conclusion

Most of JI dialects belong to the Median group, with the exception of Shirazi, EJP, and Juhuri. Jewish dialects differ from village dialects of the same areas. They have simpler morphology (lack of gender and case distinction, using the *ezāfe*) and syntax (ergativity is only morphological), and have a higher degree of Persian contamination. For this reason, hardly any isogloss – phonological, morpho-syntactic, or even lexical – applies to *all* dialects. These dialects give an idea of the dialects spoken inside the cities before Persian took over. Since the only isogloss distinguishing them from non-Jewish Iranian dialects is the extra stock language, they should be studied in the context of their geographical neighbors rather than as a separate unit.

# References

Abrahamian, R. 1936. *Dialects des Israélites de Hamadan et d'Ispahan*. Paris: Adrien-Maisonneuve.

Adler, E. N. 1905/1970. Jewish literature and the diaspora. In E. N. Adler (ed.), *About Hebrew manuscripts*, 79–99. London and New York: Oxford University Press.

Alvarez-Pereyre, F. 1995. Hebrew and the identity of Jewish languages. In S. Morag, M. Bar-asher & M. Modena (eds.), *Vena Hebraica in Judaerum Linguis*, *Proceedings of the International conference on the Hebrew and Aramaic elements in Jewish languages (Milan, 23–26 October 1995)*, 15–37. Milan: Università degli studi di Milano, Dipartimento di scienze dell'antichità.

Alvarez-Pereyre, F. 1997. Hebrew and Jewish languages: May we speak of linguistic interference? *Masorot 9–11*, ed. M. Bar-Asher, 295–313. Jerusalem: Institute of Jewish Studies, Hebrew University.

---

**25** See examples in "Verb" above, but cf. Juhuri *xoste*.

Aminof, A. 1900–1904. *Liqqutei Dinim* [A collection of laws]. Translated to Judaeo-Takjik by Shimʿon Hākhām. Jerusalem: N.P.

Asmussen, J. P. 1968. *Jewish Persian texts: Introduction, selection and glossary*. Wiesbaden: Harrassowitz.

Bar-Asher, M. 1978. On the Hebrew elements in the colloquial Judaeo-Arabic of Moroccan Jews. *Leshonenu* 42. 272–294 (Hebrew).

Bar-Asher, M. 1989. Some aspects of the study of the Hebrew component in Neo-Judaeo-Arabic. *Masorot* 3–4. 147–169 (Hebrew).

Bunis, D. M. A. 1981. Comparative linguistic analysis of Judezmo and Yiddish. *International Journal of the Sociology of Language* 30. 49–70.

Bunis, D. M. 1993. *A lexicon of the Hebrew and Aramaic elements in modern Judezmo*. Jerusalem: Magnes Press.

Ethnologue website. www.ethnologue.com/show_country.asp?name=iran, http://www.ethnologue.com/show_language.asp?code=TAT (accessed 19.03.2019).

Fischel, W. J. 1974. The contribution of the Persian Jews to Iranian culture and literature. *Acta Iranica* 3. 299–315.

Gindin, T. E. 1998. *A partial grammar of the Jewish dialect of Yazd*. Jerusalem: Hebrew University of Jerusalem. MA thesis.

Gindin, T. E. 2002. Judeo-Iranian (non-Persian). *Jewish language research*. http://www.jewish-languages.org/judeo-iranian.html (accessed 19.03.2019).

Gindin, T. E. 2003a. Ergative constructions in the Jewish dialect of Yazd? In S. Shaked & A. Netzer (eds.), *Irano-Judaica*, vol. V, 105–119. Jerusalem: Ben-Zvi Institute.

Gindin, T. E. 2003b. The Tafsīr of Ezekiel: Four copyists or four authors? In Ludwig Paul (ed.), *Persian origins – early Judaeo-Persian and the emergence of New Persian*, 15–30. Wiesbaden: Harassowitz.

Gindin, T. E. 2003c. A unique plural form in Judaeo-Yazdi. In B. Hourcade (ed.), *Cultures et societes contemporaines*, vol. 3, *Iran: Questions et connaissances, Actes du 4eme congres de la societas Iranologica Europaea*, Paris, 6–10 Septembre 1999. Paris: Association pour l'avancement des études iraniennes

Gindin, T. E. 2009 . Judeo-Persian language. In Ehsan Yarshater (ed.), *Encyclopaedia Iranica* XV: 132–139, New York: Encyclopedia Iranica Foundation.

Homayoun, H. 1998. Peculiarities of verbal morphemes in the Jewish dialect of Yazd. In A.M. Haghshenas, K. Mazdapur, and M. Mirfakhrayi (Eds.) *Dr. Mehrdād-e Bahār Memorial*, 689–706 (Persian). Tehran: Agah.

Kalbasi, I. 1994. *The dialect of the Jews of Esfahan (an Iranian dialect)* (Persian). Tehran: Institute of Humanities and Cultural Studies.

Lazard, G. 1978. Note sur le jargon des juifs d'Iran. *Journal Asiatique* 266 (3–4). 251–255.

Lazard, G. 1981. Le dialecte des Juifs de Kerman. *Monumentum Georg Morgenstierne (Acta Iranica 21)*, 333–346. Leiden: Brill.

Levi, A. 1979. Jewish languages in the East and in the West. *Peamim* 1. 58–66 (Hebrew).

Maman, A. 1989. On identifying the Hebrew element in Judaeo-Maghrebian. *Masorot* 3–4. 171–201.

Meillet, A. 1927. Influence of the Hebrew Bible on European languages. In E. R. Bevan & C. Singer (eds.), *The legacy of Israel*, 473–481. Oxford: Clarendon Press.

Milik, J. T. 1992. Les modéles araméens du livre d'Esther dans la grotte 4 de Qumrân. *Revue de Qumran* 59. 321–406.

Miller, V. F. 1892a. *Materials for studying the Jewish-Tat language* (Russian). St. Petersburg: Academy of Sciences.

Miller, V. F. 1892b. *An outline of morphology of the Jewish-Tat language* (Russian). St. Petersburg: Academy of Sciences.

Miller, V. F. 1903. *On the Semitic element in the Tat language of Mountain Jews, Oriental antiquities. Studies of Oriental Committee of Moscow Archeological Society*, vol. III (Russian). Moscow: Lazarevsky Institute of Oriental Languages.

Morag, S. 1992. Hebrew words in Jewish languages. In S. Morag & Y. Avishur (eds.), *Miqqedem Umiyyam (East and West)*, vol. 5, 110–114 (Hebrew). Haifa and Jerusalem: The Hebrew University Magnes Press.

Morag, S. 1999. The integrated corpus of Hebrew elements in Jewish languages: Some aspects of analysis. In S. Morag, M. Bar-asher & M. Modena (eds.), *Vena Hebraica in Judaerum Linguis, Proceedings of the International Conference on the Hebrew and Aramaic elements in Jewish languages (Milan, 23–26 October 1995)*, 39–57. Milan: Università degli studi di Milano, Dipartimento di scienze dell'antichità.

Morag, S., M. Bar-asher & M. Modena (eds.). 1999. *Vena Hebraica in Judaerum Linguis, Proceedings of the International conference on the Hebrw and Aramaic elements in Jewish languages (Milan, 23–26 October 1995)*. Milan: Università degli studi di Milano, Dipartimento di scienze dell'antichità.

Netzer, A. 1981. *Iranian Jewry in our day*. Jerusalem: The Hebrew University of Jerusalem.

Netzer, A. 1987. Studies in the spoken language of the Jews of Iran. In Joseph Dan (ed.), *Culture and history*, 19–44 (Hebrew). Jerusalem: Misgav Yerushalayim.

Netzer, A. 1991. Judaeo-Isfahani blessings, curses and oaths. *Miqqedem Umiyyam* 4. 197–198 (Hebrew).

Paul, L. 2002. *Grammatical and philological studies on the early Judaeo-Persian texts from the Cairo Geniza*. Göttingen: Georg-August-Universität Göttingen, unpublished habilitation thesis.

Rabin, H. 1979. The general characteristics of Jewish languages. *Peamim* 1. 40–57 (Hebrew).

Sahim, H. 1994. The dialect of the Jews of Hamedan. *Irano-Judaica* 3. 171–182.

Shaked, S. 1998. An early Geniza fragment in an unknown dialect. In Werner Sundermann, Jacques Duchesne-Guillemin & Fereydun Vahman (eds.), *Barg-e Sabz. A green leaf. Papers in honour of Professor Jes P. Asmussen*, *Acta Iranica, Hommage et Opera Minora*, 219–235. Leiden: E. J. Brill.

Shaked, S. 1995. Qumran: Some Iranian connections. In Z. Zevit, S. Gitin & M. Sokoloff (eds.), *Solving riddles and untying knots, biblical, epigraphic, and Semitic studies in honor of Jonas C. Greenfield*, 277–281. Winona Lake: Eisenbrauns.

Shapira, D. 1999. Qissa-ye Daniel or "The story of Daniel" in Judeo-Persian: The text and its translation. *Sephunot* 22. 337–366 (Hebrew).

Shapira, D. 2001. Biblical quotations in Pahlavi. *Henoch* 23. 187–195.

Shapira, D. 2002. Judeo-Persian. In *Jewish language research website*. http://www.jewish-languages.org/judeo-persian.html (accessed 19.03.2019).

Wexler, P. 1981. Jewish interlinguistics: Facts and conceptual framework. *Language* 57. 99–149.

Yarshater, E. 1974. The Jewish communities in Persia and their dialects. In Philippe Gignoux & Ahmed Tafazzoli (ed)., *Memorial Jean de Menasce*, 453–466. Louvain: Imprimerie Orientaliste.

Yarshater, E. 1977. The hybrid language of the Jewish communities of Persia. *Journal of the American Oriental Society* 97. 1–7.

Yarshater, E. 1989. The dialect of Borujerd Jews. In L. de Meyer & E. Haerinck (eds.), *Miscellanea in honorem Louis Vanden Berghe* (*Archaeologia Iranica et Orientalis* II), 1029–1046. Gent: Peeters Press.

Yarshater, E. 2002. The Jewish dialect of Kāshān. *Jerusalem Studies of Arabic and Islam* 27 *(Studies in honour of Shaul Shaked)*. 439–462.

Zhukovskij, V. A. 1922. *Materials for studying Persian dialects*, vol. II, 390–394 (Russian).: Petrograd: Russian Science Academy.

Alireza Korangy
# Epilogue

Little can be said that has not already been elegantly stated by our esteemed Bernard Comrie in the introductory essay of this volume so I will keep it short. Upon completion of our publication one thing became abundantly clear: there is much work to be done in the field of Iranian typology but even more so as per comparative typology. The essays point to a need for scholarly appreciation of complexities furthered by a [proto-]forgetfulness that is an almost mystical (and certainly mystifying) feature of Iranian. Iranian linguistics, with all of its extraordinary contributions by scholars the likes of Oktor P. Skjaervo, Gilbert Lazarde, Martin Schwartz, Donald Stilo, Habib Borjian, Walter Henning, Kent and many others has endured an anonymity, that considering their brilliant work, is all but astounding if not quite sad. This work, as any work of its kind, is a tiny piece of a much larger puzzle that hopes to inspire. Bejewelled by some of the absolute best scholars in the field we feel we have accomplished a worthy task in the field and we attribute it all to our colleagues who weathered the tribulations of such a project and our editorially-driven persistence (and sometimes nuissance); what is more, they made the project better than intended. They are more than colleagues but also friends in every way: they became the editors' teachers and what can define a friend more than that? I am, both happy to share the joy of the completion of this book with them and you the reader and also sad to say goodbye to a memorable cooperation with colleagues who taught me so much. Last, but not least, I am most grateful to a wonderful and brilliant co-editor, Behrooz Mahmoodi-Bakhtiari whose enyclopedic range of referencing and cross-referencing, not to mention his astounding learnedness I find to be superb and a thing of legend. They all help coin the saying *quae potest ese vitae iucunditas, sublatis amicitiis?*: happy reading.

# Index