

ORIENTALIA SUECANA

VOL. LXI (2012)

Edited by

ÉVA Á. CSATÓ

JOAKIM ENWALL

MATS ESKHULT

CARINA JAHANI

ANETTE MÅNSSON

CHRISTIANE SCHAEFER

Guest editor:

GEOFFREY HAIG

UPPSALA
SWEDEN

© 2013 by the individual authors

Orientalia Suecana is an international peer-reviewed scholarly journal founded in 1952 and published annually by the Department of Linguistics and Philology, Uppsala University. The journal, which is devoted to Indological, Iranian, Semitic, Sinological, and Turkic Studies aims to present current research relating to philological, linguistic, and literary topics. It contains articles, reviews, and review articles.

Starting from vol. 59 (2010), Orientalia Suecana is a web-based only publication with open access. More information on <http://www.lingfil.uu.se/orientalia>

Submissions for publication and books for review are welcome. Books will be reviewed as circumstances permit. Publications received will not be returned. Manuscripts, books for review, orders, and other correspondence concerning editorial matters should be sent to:

Orientalia Suecana
Editorial Board
Department of Linguistics and Philology
Uppsala University
Box 635
SE-751 26 Uppsala
Sweden

E-mail: orientalia.suecana@lingfil.uu.se

ISSN 0078-6578

Typeset by
Textgruppen i Uppsala AB

Contents

Studies

<i>Alexander Andrason</i> , The Gnostic <i>qatal</i>	5
<i>Erik J. Anonby</i> , Stress-induced Vowel Lengthening and Harmonization in Kumzari	54
<i>Ute Bohnacker and Somayeh Mohammadi</i> , Acquiring Persian Object Marking: Balochi Learners of L2 Persian	59
<i>Forogh Hashabeiky and Mahmoud Hassanabadi</i> , An Elusive Particle in Persian: The Use of <i>Magar</i> in <i>Tārix-e Beyhaqi</i>	90
<i>Farideh Okati, Pétur Helgason, and Carina Jahani</i> , Diphthongization in Five Iranian Balochi Dialects	107

Iranian Linguistics, ed. by Geoffrey Haig and Carina Jahani

<i>Geoffrey Haig and Carina Jahani</i> , Introduction	123
<i>Ayat Hosseini</i> , The Prosodization of Function Words in Persian	126
<i>Martin Joachim Kümmel</i> , The Iranian Reflexes of Proto-Iranian <i>*ns</i>	138
<i>Stephen H. Levinsohn</i> , Introducing Reported Speeches in Balochi of Sistan with <i>ki</i>	146
<i>Corey Miller</i> , Variation in Persian Vowel Systems	156
<i>Maryam Nourzaei and Carina Jahani</i> , The Distribution and Role of the Verb Clitic = <i>a/a</i> = in Different Balochi Dialects	170
<i>Ergin Öpengin</i> , Adpositions and Argument Indexing in the Mukri Variety of Central Kurdish: Focus on Ditransitive Constructions	187
<i>Guiti Shokri</i> , Past and Non-past Structures in the Mazandarani Dialect Spoken by the Galesh of Ziarat	199

<i>Book reviews</i>	211
<i>List of Contributors</i>	225

The Gnomic *qatal*

Alexander Andrason, Stellenbosch University
andrason@sun.ac.za

Abstract

The present paper offers a cognitive and typological approximation to the problem of the gnomic *qatal*. It demonstrates that the gnomic sense of the *qatal* can be chained to the remaining semantic potential of the gram by making use of certain typological templates or universals, i.e. by so-called gnomic branches of the anterior path. Given that, from a cross-linguistic perspective, certain subtypes of a present perfect (inclusive, frequentative, and experiential perfects, as well as an anti-perfect) naturally generate gnomic extensions (following a development referred to as “gnomic branches”), that the dominant portion of the semantic potential of the *qatal* covers the domain of a perfect, and finally that all the examples of the BH gnomic *qatal* (if derived from active roots) may be viewed as generated in prototypical perfect contexts (the gnomic use of the *qatal* stems from its use as an inclusive, frequentative, experiential, and anti-perfect), the gnomic value of the gram may be cognitively (both conceptually and diachronically) chained to the remaining sphere of its semantic network by employing the “gnomic branch” linkage. In this manner, the gnomic value ceases to be aberrant and, on the contrary, becomes a fully rational component of the semantic potential of the gram. As a result, we propose a model (a map) that in a more consistent and more holistic manner represents the semantics of the entire *qatal* category; it accounts for all the senses, the gnomic values included.

Keywords: Biblical Hebrew, verbal semantics, grammaticalization, cognitive linguistics

1. Introduction

When analysing values of the *qatal* form in Biblical Hebrew (BH), one can hardly ignore cases where, more or less surprisingly, the gram – which typically functions as a perfect and also a (perfective) past – expresses atemporal or universal activities and situations. This is what has commonly been denominated as the “gnomic *qatal*” – a particular type of use of the suffix conjugation in which the *qatal* construction introduces general truths. In this usage, which is found to be principally abundant in the book of Proverbs, the formation allegedly corresponds to modern Indo-European simple presents, which, when employed in maxims and aphorisms, depict extra-temporal constant facts:

(1)

חֲכָמוֹת נָשִׁים בְּנִתָּהּ בֵּיתָהּ וְאִזְלֹת בִּיגִידָהּ תִּהְרָסֶנּוּ

The wise woman **builds** her house, but the foolish one tears it down with her own hands (Prov. 14.1)¹

¹ All the BH *qatal* forms that are employed in a gnomic sense (as well as grams that provide a gnomic value in other languages) will be indicated in bold formatting. For the sake of clarity, English constructions that render the *qatal* forms from the Hebrew quotes will likewise be distinguished by bold typeface.

1.1. *Grammatical tradition*

The fact that the *qatal* may provide a gnomic sense has been acknowledged in almost all important grammars and linguistics descriptions of the BH verbal system. Although the phenomenon has been well known, its explanation is far less straightforward and consistent. Among all the opinions related to this usage of the suffix conjugation, it is possible to distinguish three main trends.

First, a vast group of scholars limit themselves to a simple observation that the *qatal* may express atemporal general truths without, however, attempting to provide an explanation for such a supposedly aberrant function. For instance, Brockelmann (1956: 42), Grether (1967: 209), Lambdin (1971: 3), Jenni (1978: 265), and Van der Merwe, Naudé, and Kroeze (1999: 146) notice that when employed in poetry, proverbs, and maxims, the *qatal* is able to denote gnomic situations, general habitual activities, or atemporal experiences, and thus approximates the simple (general) present of English or the gnomic aorist of Classical Greek.

The second (and certainly the most numerous) faction of grammarians, similarly to the aforementioned position, maintains that the BH gram offers a gnomic sense but, in contrast with it, intends to relate this value to the prevalent meaning of the suffix conjugation. Among all the scholars who adhere to this view, it is possible to detect three different sub-groups that diverge in such a way that a semantic connection between the *qatal* and its gnomic sense is posited. In other words, researchers disagree as to how to justify the use of the gram as a vehicle of general truths.

Most commonly, scholars relate the gnomic value to what, in modern terminology, we could label as an experiential perfect. For example, Ewald (1863: 351) argues that the *qatal* of a general truth may be explained as stemming from the experience – a fact has been proven by experience and is therefore considered as definitive. A similar explanation has been proposed by Müller (1883: 2) and Driver (1892: 17) who argue that the *qatal*, in its gnomic function, expresses propositions of general character confirmed by experience. It is an experiential perfect where general truths correspond to facts that have actually occurred and are thus verified by their real incidence. Davidson (1902: 60) and Gesenius-Kautzsch-Cowley (1910: 312) maintain the experiential foundation of the gnomic *qatal*, and furthermore link this type of the suffix conjugation to the frequentative sense of – in their terminology – the “perfect”. That is to say, in the gnomic usage, the *qatal* is employed as a frequentative perfect of experience (as it introduces actions that have been proved *de facto* by their previous occurrence and that moreover have frequently been taking place; cf. Davidson 1902: 60) and as a recurrent perfect of experience (as it denotes facts that have formerly occurred, thus belonging to common experience, and that are still of constant recurrence; cf. Gesenius-Kautzsch-Cowley 1910: 312).² Recently, the perfectal link, albeit additionally grounded in a temporal nuance of the *qatal* form (viz. its past value), has been postulated by Joosten (2012). Joosten (2012: 204–205) claims that the gnomic *qatal* “originated in the observation of past occurrences”. Subsequently, these temporally remote past facts – viewed as truths

² The two definitions are highly similar. However, the former links the experiential perfect with the value of a frequentative perfect while the latter relates it to the sense of an inclusive perfect.

known from experience – were reanalysed as atemporal situations and proverbial states.³

Other grammarians understand the connection between the gnomic usage of the *qatal* and its prototypical value in aspectual terms. For instance, according to Joüon (1923: 296–297), the *qatal* of universal truths should be explained by means of the aspectual force of the gram, namely by its global, unique, and instantaneous character (cf. also Joüon and Muraoka 2009: 333). To be precise, the fact that the *qatal* expresses an action that is “unique ou instantanée” (Joüon 1923: 296) explains “[son] emploi [...] pour exprimer une vérité constante” (*ibid.*: 297). In a similar vein, Watts (1951: 24–25) alleges that when the *qatal* conveys typical characteristics, it presents a single event or situation that exemplifies a property of a person or a thing.⁴ In his view, although such perfects appear in present time, they extend beyond the actuality, referring to “a broad unrestricted present” (Watts 1951: 25). Similarly to Joüon (1923), Waltke and O’Connor (1990: 488, 506) claim that gnomic or “proverbial” *qatal* portrays a universal situation or activity as a single event. This global view, in turn, justifies the use of the suffix conjugation, which, according to them, is defined as a perfective aspect.⁵ Finally, in contrast with the two subgroups that explain the gnomic *qatal* as a manifestation of the taxis (perfect) or aspectual (perfective) nature of the gram, other scholars do not infer the gnomic function from the inherent semantics of the *qatal*, but rather justify it by making use of the concept of neutralization of the taxis-tense-aspect load of the gram. For instance, according to Cook (2002: 221–222), when the *qatal* (which is, in his opinion, a perfective construction) expresses timeless and omnitemporal statements, its aspectual (as well as temporal) potential is cancelled because the perfectivity (surfacing typically as past perfectivity) is incompatible with gnomic present tense.⁶

Beside the two factions presented thus far, one may also detect a third group of researchers who seemingly deny the gnomic interpretation of the gram. One of its most prominent members is Rogland, who in his in-depth analysis (cf. Rogland 2003) argues that the majority of proverbial *qatals* do not require an interpretation in terms of a gnomic atemporal present, but should rather be explained as past tenses, in accordance with the ordinary semantic definition of the gram, viz. a past tense (*ibid.*: 40, 46). In other words, proverbs that express universal truths employ *qatal* forms in its (i.e. the *qatal*’s) regular – in Rogland’s view – past function. It is a proverb that presents a gnomic statement, but not the verb itself, which, quite the contrary, preserves its inherent non-gnomic past temporal value. More specifically, it functions as a simple past of experience or observation (*ibid.*: 24–25), a global or aspectually neutral past (*ibid.*: 35, 37), and a relative past (*ibid.*: 43–44). By doing so, the *qatal* in maxims is a past tense with none of the habitual or imperfective nuances

³ Joosten (2012: 208) regards this gnomic value of the *qatal* as a diachronic and conceptual basis of modal (counterfactual) uses of the BH suffixed conjugation.

⁴ He labels this usage of the suffix conjugation a “characteristic perfect” (Watts 1951: 25).

⁵ In an analogical manner, Sasson (2001) understands the gnomic *qatal* as an expression of the inherent perfective aspectual value of the suffix conjugation.

⁶ However, in a later study, Cook (2005) rejected his previous view and, probably under the influence of Rogland’s ideas, understood most *qatal* forms in proverbs as “regular” (i.e. fully identical with the *qatal* in non-proverbial material) pasts or perfects (for details, see the next paragraph).

typical of general gnomic presents (*ibid.*: 37). Rogland's view has influenced Cook who, in his later paper (2005: 130), explains the *qatal* forms in the book of Proverbs as fully identical with the *qatal* in non-gnomic texts, i.e. as perfects or pasts. Since the *qatal* in the proverbial material does not comply with the tendency whereby present and imperfective forms are prototypically used in generic and gnomic statements, and given the fact that past tenses are cross-linguistically quite acceptable in maxims and anecdotes, Cook (2005: 130) argues that this use of the suffix conjugation should not be viewed as a gnomic present, but rather as an example of an "anecdotic" past tense, i.e. as a past tense used in maxims or proverbs. Thus, he infers, it should be translated with the same taxis, temporal, aspect, and modal (TTAM) load as the examples of the *qatal* in non-gnomic texts (*ibid.*: 131).

1.2. Problem, methodology, and research strategy

1.2.1. Problem

The explanation of the gnomic *qatal* is typically based on a derivational or inferential procedure whereby the gnomic sense (or any other value that a given linguist claims to be patent in the proverbial material; cf. Rogland 2003 and Cook 2005) is derived from the so-called main, inherent or invariant meaning of the form.⁷

Thus, grammarians commonly argue that in the gnomic usage – in accordance with the prototypical meaning of the gram identified as a perfect, a perfective, or a past – the *qatal* denotes previously accomplished, complete, or past events, respectively. Accordingly, the alleged invariant taxis, aspectual, or temporal value of the formation is seen as a basis for a semantic extension available in the gnomic usage. Depending on the definition of the gram adopted by a scholar, the gnomic use is hence understood as derived from and akin to a perfect of experience (taxis), a perfective (aspect), and a global past (tense).

As will be evident from the further discussion, the entire problem of the gnomic *qatal* (and thus its possible solution) may be envisaged from a completely different perspective, where the gram is viewed in dynamic terms as a manifestation of certain evolutionary processes.

1.2.2. "Dynamic" semantics⁸

Discussing the issue of semantics of a gram, two problems immediately emerge. What is the value of a form when it (i.e. this form) appears in a specific context, and what is the value of a form viewed as a whole, i.e. when it is considered as a component of the verbal system? In order to avoid such confusion in our analysis, the first phenomenon will be labelled as a "sense" and the other a "meaning".

⁷ This same procedure has been employed in the descriptions of gnomic uses of perfects or pasts in languages such as Greek and Akkadian that have sometimes constituted comparative evidence in studies devoted to the BH gram. On the gnomic type of the Greek Aorist, see Smyth (1956: 1931) or Humbert (1954: 124–125) and on the gnomic variety of the Akkadian *iprus*, see Mayer (1992), Metzler (2002), and Loesov (2004).

⁸ The present study is the fifth article dedicated to the semantic analysis of the BH *qatal* by this author. The four previous papers dealt with performative (Andrason 2012a), counterfactual (Andrason 2013a), prospective (Andrason 2013b), and precative senses (Andrason 2013c). As a result, some portions of the introductory sections devoted to methodological and theoretical issues (section 1.2.2.) are similar.

Sense

A sense may be defined as a concrete value which is displayed by a locution in a specific place and time, and which is “experimentally” measured by employing determined semantic domains or categories. It is thus a value a gram receives in a precise context, a value that is categorized by means of available conceptual structures. This implies that atomic senses depend on their contextual settings (i.e. on linguistic and extra-linguistic factors) and on *our* classification devices (i.e. on properties of humans conceptual categories; cf. Evans and Green 2006: 352–353, 368, Niki-foridou 2009: 17, 26). The definition of a sense may in fact be reduced to the following statement: a sense is a compatibility of a form with a concrete context. As supported by modern science, no two contexts describing the real world are perfectly identical. Quite the reverse: they invariably differ in some parameters. The detection of this dissimilarity depends on the precision in describing the environments under consideration. Coarse-grained (macroscopic) analyses typically group various contexts as equal while narrow (microscopic) studies treat previously indistinguishable milieus as different. But, whatever our level of precision, in an ultimate – i.e. the most atomic – description, two contexts are always dissimilar due to the infinite complexity of the universe. Since senses are contextual phenomena – they emerge in and/or are compatible with determined milieus – and since no two contexts are ideally duplicate, no two senses can be perfectly the same. In a certain approximation, all senses somehow differ because the contexts in which they appear – if analysed with the highest precision – are dissimilar (Auyang 1998: 344, Smith 1998: 51–67, 90–115, Wagensberg 2007: 56–57, 60, Schneider and Sagan 2009: 55). Moreover, one should note that in living languages a form appears in an indefinite number of uses and thus contexts. It is clear that two uses always constitute two different contexts because they must differ at least in temporal settings. As a result, a form may be found in an infinite number of contexts, delivering a likewise infinite number of senses. Certainly, such a fragmentary description of reality (in our case of a verbal gram) is unpractical. Therefore, we employ larger concepts that enable us to encompass various contexts and senses, and inversely to reduce the infinite amount of data to a finite and workable series. This precision is linked to our categorization technique (Auyang 1998: 344, Prigogine 2009: 213). Thus, the number of senses “observed” empirically by a scholar when providing a taxonomy of uses of a gram is closely related to how reality is divided into conceptual boxes. If the “measuring tools” (viz. concepts) are broad, a construction may seem to convey a few senses (or even if extremely wide taxes are used, a single sense in all uses). If, on the contrary, our conceptual devices are sufficiently sensitive, a gram may seem to provide ten, hundreds, or thousands of senses (if our concepts are extremely atomized, the number of sense will become infinite). Typically, the former description assures conceptual consistence of a phenomenon but is imprecise, while the latter provides a far too detailed and disordered view. It is important to emphasize that any such categorization is external to the universe (and hence to language) itself, being tied to our theoretical structuration. What is empirically certain or objective (if anything can be absolutely objective in science) is that a form may appear in a potentially infinite amount of uses and hence in an infinite number of contexts which at an ultimate lev-

el invariably differ in certain parameters due to the complexity of the real world. Consequently, if the world's complexity is envisaged, a form is inherently polysemous – the range of this polysemy is infinite but, depending on our categorization, it will appear as more (high fragmentarization) or less (low fragmentarization) extensive. Polysemy (or diversity of senses) is the norm in languages of the world. Hence a form almost regularly provides several senses that in some cases appear extremely unrelated and even contradictory to each other (Evans and Green 2006: 169, Bybee 2010: 183, 186–187).

Another important property of any polysemy – either extended (in a more atomic description) or minimal (in a more coarse-grained description) is that it is not a random cluster of disparate and accidental values. One of the most fundamental principles with respect to polysemy is the fact that diverse senses conveyed by the same form are necessarily related. Relatedness of senses is a constant feature in languages and constitutes one of the tenets of cognitive linguistics. First of all, it is commonly accepted that polysemy is a phenomenon that affects all the levels of a language (phonology, morphology, and syntax) and all of its components, be they lexical (lexical semantics) or functional (i.e. the *core*-grammar or “functional” semantics; cf. Lewandowska-Tomaszczyk 2007: 140, 147–148). Scholars also agree that polysemy – as a categorizing phenomenon – represents a form's total meaning as a solid conceptual category of distinct but related senses (Taylor 2002:98). In other words, the meaning of a grammatical entity consists of a number of wholly distinct yet demonstrably related senses (Janssen 2003: 96, Evans and Green 2006: 352). The fact that the senses of an item are connected signifies that they are somehow linked to the conceptual prototype or the central value (cf. Lakoff 1987: 12–13, Gibbs 1994: 157, Lewandowska-Tomaszczyk 2007: 140, 147–148, Evans and Green 2006: 36, 331). Put differently, there is, by definition, “a motivated [linear or non-linear] relationship between polysemous senses” from a central value to its extensions (Cruse 2004: 108; see also *ibid.*: 109–110).

In even stronger terms, linguists talk about “a cognitive fact”: polysemous meanings are related in reasonable and methodical ways (Tuggy 2003: 323–324). Polysemy “is not just a matter of being different meanings attached to a form”. On the contrary, in polysemy, the connection among the senses is inherently logical and systematic (*ibid.*: 348–350). A polysemous space constitutes a well-ordered conceptually solid whole (cf. Heine, Claudi, and Hünemeyer 1991: 224–225). The relation unifying the senses is reasonable and systematic because polysemous extensions reflect and arise from universal human cognitive mechanisms (such as metaphor, metonymy, abduction, image-schema transposition, etc.) that ensure a conceptual bond among numerous values, even the most distinct ones. More concretely, applying the aforementioned cognitively “natural” procedures, speakers expand one sense into another and in this manner construct superficially incongruent polysemous compositions (cf. Taylor 2002: 138–139, Tuggy 2003: 348–350, Evans and Green 2006: 332–352; see also Ibarretxe-Antuñano 1999: 29–30).

Meaning

Due to the invariably polysemous behaviour of grammatical entities, the understanding of the meaning of a gram as an invariant sense – i.e. as an identical semantic domain presented in all uses – can no longer be sustained. As explained, a single grammatical formation is able to provide a wide variety of senses that may be extended virtually *ad infinitum*. A form never offers only one sense which is duplicate in all contexts because all contexts – and hence all senses – always differ in some parameters. The identity of the senses provided in all contexts would only be possible in an extremely coarse-grained approximation wherein the infinite complexity of reality is deliberately disregarded. Such an approach clearly clashes with empirical facts concerning reality.

The immeasurable intricacy of the universe not only rules out understanding the meaning of a form as a value that is identical in all uses, but also shows that the application of the so-called “invariant” string (i.e. the semantic portion that is present in all uses) for determining the meaning of a gram is, in fact, dysfunctional. Typically, a semantic string that is shared by all the senses provided by a form (or that appears in all the contexts where the form appears) belongs to the most general or coarse-grained conceptual box, with which it is possible to embrace all senses. Of course, certain senses that belong to a given polysemy always have a common or collective string, but a string that would belong to all of them corresponds to the most coarse-grained interpretation possible. The problem is that in the case of highly developed polysemous structures (for instance, formations that constitute the core of the verbal system), this invariant string delivers such coarse-grained information that its epistemological or scientific validity is insignificant (on the incompatibility of the invariant meaning hypothesis with usage-based theory, see Bybee 2010: 183–193).

Since polysemy arises because a form spreads – through metaphor, metonymy, or other cognitive mechanisms – to new environments where it acquires new (previously unavailable) senses, two adjacent senses (i.e. a sense and its immediate metaphorical extension) typically share one or more semantic traits (the range of shared elements again depends on our categorization). But after a certain reiteration of the procedure of extending one sense to another, the semantic relation of an *n* extension and the original input may become extremely loose, being embraceable only under some highly coarse-grained and insignificant labels (cf. the Polish adjective *zielony* ‘green’ that in company of the noun *pojęcie* ‘idea’ acquired, by metaphorical extensions, the sense ‘null, no’).⁹ This means that the very idea of an invariant string – although sometimes relatively useful for “small” polysemies – cannot be treated as precise, because its significance for “large” polysemies is minimal or even null.

Furthermore, when searching for an invariant string, with which the meaning of a gram would be defined, a linguist disregards all the remaining values and thus ignores the non-shared portion of the polysemous space that may in fact be highly relevant for an appropriate comprehension of the form. Put differently, if a string that

⁹ Native speakers establish the following conceptual link relating these two senses: [green] > [unripe, immature] > [unimportant, insignificant] > [null].

is present in 90 per cent of the uses fails to appear in the remaining 10 per cent, it cannot be employed to define the gram although it does say something important about this construction. Similarly, if a formation appears 50 per cent of the time with a string – especially in certain types of contexts – this value would be disregarded in the total semantic definition of the gram; it is not accessible in all the uses. However, it may provide a significant insight into the form’s semantics. For example, it is a well-known fact that in various languages the gram which functions as a general present and a simple future (with both perfective and imperfective readings), is typically restricted to an imperfective force in a past time frame (see the Hebrew *yiqtol* in Andrason 2010a). This fact additionally weakens the relevance of the concept of an invariant string (or invariant semantic domain) in the demarcation of a gram’s meaning (cf. Bybee 2010: 186).

Since the meaning cannot correspond to the invariant sense and cannot be equated with an invariant string in all senses, how can it be defined? As maintained by cognitive linguistics, a form’s meaning is to be understood as its entire semantic potential. Such potential is equivalent to the total semantic space that includes all possible individual senses (shades of meaning) empirically “recorded” in specific realistic cases, i.e. in concrete contexts. Thus, the meaning of a formation is viewed as a set-theoretic union of all individual atomic senses that empirically exist in specific environments. This also means that since atomic values clearly depend on their contextual milieus, the entire meaning of a form – its total semantic space or network of interrelated values – is necessarily a contextual phenomenon (Evans and Green 2006: 352–353, 368 and Nikiforidou 2009: 17, 26). Furthermore, respecting the relatedness principle, cognitive linguistics represents the semantic and functional potential of a gram as a map where each sense is conceptually related to another, forming a network of interrelated components (Evans and Green 2006: 331–333).

As a result, the traditional structuralist dichotomy between the inherent-invariant meaning and contextual realizations must be replaced by a more realistic distinction between an empirical level analysis (concrete “experimental” data or senses displayed by a locution in a specific place and time) and their summation into a coherent whole (the gram’s total meaning) at a theoretical system’s level (cf. already Dahl 2000a: 14).

Dynamic description of verbal meaning

Since any extension from one sense to another unavoidably implies a chronological order (i.e. a more original sense is the basis for a subsequent extension), any conceptual input-output relation among components of a polysemy is not only abstract (viz. conceptual *sensu stricto*) but also – and, in fact, necessarily – diachronic. In this manner, a synchronic polysemous network reflects a realistic historical progression: by incorporating new senses or abandoning older ones, it expands or reduces the range of an available semantic space. The connection among the elements of a polysemous grid is invariably historical because polysemy reflects a historical change; it “explain[s] synchronic variation as resulting from diachronic change” (Lawandows-

ka-Tomaszczyk 2007: 140). In other words, conceptual connections which exist among components of a given polysemy represent diachronic processes whereby older senses gave rise to new senses due to the spread of the form to new contexts (cf. Heine, Claudi, and Hünemeyer 1991: 221–225, 227–228, 260–261, Bybee, Perkins, and Pagliuca 1994: 15–19, Tyler and Evans 2003: 344–346, Van der Auwera and Gast 2011: 186–188, Bybee 2010: 198–199).

Since a given synchronic – in principle atemporal – variation (e.g. different senses of a form) stems from a historical process as “a temporary outcome of an ongoing-change” (Sadler 2007:33) or is a typical by-product of grammaticalization, it can logically be described in dynamic terms, making use of a diachronic process (Heine, Claudi, and Hünemeyer 1991: 261). This dynamic representation of a polysemy is provided by employing models based upon grammaticalization chains or grammaticalization paths. Such chains and/or paths are mainly viewed as diachronic principles that schematize the evolutionary pattern of a given grammatical “taxonomical” class. However, since a synchronic variation is a temporary result of unceasing processes and fluctuations, a diachronic rule or (statistical) universal can be used to map – and to relate – different senses synchronically offered by a form. To be precise, the models of grammaticalization paths offer two dimensions: a diachronic dimension (they portray diachronic phenomena showing the evolution of a given form), and a synchronic dimension (i.e. they constitute dynamic relational patterns which represent the internal organization of a synchronic polysemy specifying the connection among its components; Heine, Claudi, and Hünemeyer 1991: 221–222). As a result, these evolutionary templates can be used to account for the synchronic structure of a language (*ibid.*: 252). In particular, all the senses displayed by a form can be viewed as corresponding to stages located along the chains. Some of them are less grammaticalized; these are less advanced stages on the chain which are also historically older. Others are more grammaticalized, corresponding to more advanced stages on the chain and developing later (cf. Heine, Claudi, and Hünemeyer 1991: 227–228). By linking all the sense of a form by means of a diachronic, typologically universal scenario, the path or grammaticalization chain model represents the polysemy of an entity as cognitively (both conceptually and diachronically) solid and consistent (cf. the concept of panchrony in Heine, Claudi, and Hünemeyer 1991: 260–261).

For certain modern languages and certain grams, it is possible to trace the extensions of senses from one to another and thus to establish a map of a semantic network that is built on direct diachronic data. Employing tangible evidence (e.g. texts) we can see how the form has been modifying its semantic space by adding new senses and/or losing others. In such cases, the order and logic of a synchronic semantic network can be determined by making use of diachronic processes. These diachronic processes show how conceptual extensions have actually occurred and establish the real order of a spread of the gram to new contexts. Put it simply, a synchronic variety of senses are arranged to match a concrete realistic development as documented by palpable evidence. The connection among senses, available synchronically, merely copies the historical development of the polysemy as testified by the existing data or tangible linguistic “fossils” (cf. the mapping of the semantic

potential of the Roman compound and simple perfect and past tenses which can be traced from Latin to the present; for details see Squartini and Bertinetto 2000).

However, in many cases, linguists do not have access to direct diachronic data that could establish the historical – and hence conceptual – linking among components of a polysemous map. In these cases, one may use common typological tendencies – or following a stronger view, universal laws – that codify a semantic evolution, and thus meaning extensions, typical for determined “species” of grams. In respect to the verbal system, these developmental principles (labelled “paths”) offer a model of exemplary evolution, specifying how aspects, tenses, and moods are shaped, how they develop and what their ultimate outcomes can be (Bybee, Perkins, and Pagliuca 1994, Dahl 2000b, Andrason 2011a). In other words, they predict – with a margin of error, which is inevitable in such studies – how verbal grams acquire new values and what spectrum of possible semantic extensions of a form can exist. With these, to some extent, universal moulds or templates in hand, we can propose a typologically plausible organization of semantic potentials that have been measured synchronically. Put differently, we employ paths as matrices for mapping senses offered by a gram and propose a logical (both diachronic and conceptual) ordering of the components of a semantic network. We compare the synchronic variety of senses provided by a formation with universal developmental scenarios and arrange them (i.e. these synchronic senses) in such a way that they would match a given evolutionary pattern (*vide* Andrason 2010a: 1–63; 2011b: 1–50 and 2011c: 30–34). As a result, the gram’s semantic potential (its set-theoretic union of polysemous senses) – and hence its total meaning – is understood as a fragment of a cline or a cluster of them (*vide* Van der Auwera and Gast 2011: 186–188, 281 and Andrason 2010a: 22, 2011a: 69–73, 2011b: 30–31, 2011c: 30–31, 34).

In this manner, all superficially unrelated or incompatible values – in accordance with the relatedness principle – are mapped within the same network and are chained by means of typological universals. They are comprehended as “frozen” vestiges of certain diachronic movements that, although not directly documented, are typologically plausible. Consequently, senses echo concrete diachronic stages (hypothesized by employing not concrete diachronic data but rather typological laws) during which older values were expanded to novel contexts and assumed new functions.

Since every synchronic property of a form corresponds to a precise stage of a certain diachronic phenomenon, the method has been labelled as “panchronic” (a combination of synchrony and diachrony) or “dynamic” (the present state of affairs depicted as a dynamic process; cf. Heine, Claudi, and Hünemeyer 1991: 248, 251 and 259, Nichols and Timberlake 1991, Łozowski (2000: 32), Andrason 2010a: 18–19, 2011a: 69–73, 2011b: 28–34, 2011c: 17–21, and 2012a: 15, 18–20).¹⁰ In the panchronic or dynamic approach we employ evolutionary templates (i.e. typological tendencies, dynamic laws, or developmental universals) and/or make use of concrete historical evolutions to posit a synchronically valid representation and classification of the meaning of a grammatical entity in a dynamic, process-like manner (cf. Andrason 2011b: 31–34, 2011c: 19–20, 2012b: 10–17).

¹⁰ Grygiel (2005: 98) affirms that panchrony constitutes the most objective representation of languages understood as a spatial and temporal continuum.

1.2.3. Research strategy

Complying with the relatedness principle as well as with the constraint of non-derivability of less prototypical values from an allegedly inherent meaning, in the present study we aim to offer a typological solution to the problem of the gnomic *qatal* and show how the gnomic sense of the suffix conjugation should be chained to the remaining semantic potential of the gram. That is to say, by making use of certain typological universals, we will demonstrate how the gnomic value of the *qatal* has most probably arisen and how it should be networked to the remaining semantic load of the category. As a result, the gnomic use will cease to be aberrant and instead will be incorporated as a fully rational component of the grid that has emerged following certain typologically universal evolutionary scenarios. This will, in turn, lead to a more consistent and holistic understanding of the entire *qatal* category.

Since the *qatal* has been defined as a manifestation of the resultative path (Andersen 2000: 31, Cook 2002: 209–219, Andrason 2011a, 2013b) and as a materialization of a modal contamination path that the original resultative expression has followed (Andrason 2011a, 2012b, 2013c), the gnomic value has most probably arisen as a stage of this trajectory or as a stage on a cline that branches from the resultative track and its modally contaminated varieties.

Let us explain this assumption in more detail. It has recently been demonstrated that the statistical nucleus of the semantic potential of the *qatal* may be contained in its totality and viewed as a portion of the anterior track (Andrason 2010b: 610, 2011a: 281, 305–307, 2012a: 14–15, 38–41).¹¹ In this manner, present perfect (inclusive, resultative, frequentative, and experiential), indefinite, and definite past, as well as perfective and simple past functions, can be viewed as fully compatible and congruent – all of them being matched with consecutive stages on the anterior path.¹² It has moreover been argued that resultative-stative, stative, and present temporal values can be explained by employing a network of the simultaneous path (cf. Andrason 2011a: 282–283, 305–307, 2011b: 41–43, 2012a: 39)¹³ while sporadic instances where the gram provides an evidential sense may be rationalized as having

¹¹ The anterior path is a sub-cline within the resultative track. Generally speaking, the anterior path predicts that resultatives evolve into perfects (first inclusive and resultative present perfects, later experiential and indefinite varieties) and then into past tenses (initially recent and discursive and subsequently general, remote and narrative). Moreover, during the development from a present perfect into a definite past, formations may experience another change, receiving an explicit aspectual making as a perfective (cf. Bybee, Perkins, and Pagliuca 1994, Dahl 2000b, Cook 2002; for a far more detailed treatment of the anterior path and its relation to the resultative trajectory with all its sub-tracks, see Andrason 2011a: 35–45, 2011b: 10–16 and forthcoming (a)).

¹² Here belong pluperfect uses and certain future senses as well (Andrason 2011a and 2013b).

¹³ The simultaneous path, another sub-cline in the resultative track, shows the manner in which resultative proper grams develop into present tenses (cf. Maslov 1988: 70–71, Bybee, Perkins, and Pagliuca 1994: 74–78, Drinka 1998: 120, and especially Andrason forthcoming (a)). To be precise, the cline predicts that certain resultative proper grams evolve into simultaneous resultative presents (the main emphasis is put on the resulting state while the prior action is only merely suggested), subsequently into stative presents (resultative undertones become unavailable and the only remaining sense corresponds to a static quality or situation) and finally into simple present tenses (for a more detailed discussion of the simultaneous path, see Andrason 2011a: 40–45, 2011b: 13–15 and forthcoming (a)).

arisen due to the evidential path¹⁴ (cf. Andrason 2010b: 623–624, 2011a: 282). Finally, certain modal functions of the gram – e.g. real counterfactual, unreal counterfactual, and real factual (precative) – have been classified as a manifestation of the modal contamination path of the original resultative input.¹⁵

In light of the dynamic definition of the *qatal* (i.e. as portions of the three sub-trajectories of the resultative path), it is highly plausible that the same cline constitutes the basis for the gnomic usage. Thus, one may tentatively assume that there exists a conceptual and diachronic link between post-resultative grams (i.e. grams that develop along a resultative path reaching more and more advanced sections) and the gnomic function. In order to employ such a link as a binding mechanism that coordinates the gnomic *qatal* with the rest of the senses of the suffix conjugation, we must demonstrate the universal character of such a relation between the value of gnomicity and post-resultative grams. In particular, we are compelled to provide typological evidence showing that originally resultative grams and their diachronic successors are somehow predisposed to undergoing semantic extensions into the gnomic domain. Once the evolutionary relation between the post-resultatives and the gnomic sense has been established and its universal character explained, we will be able to use it as a matrix for networking the gnomic value of the *qatal* to the remaining space of its semantic map.

We will, however, start our study by making certain basic clarifications concerning the concepts of a “gnomic sense” and a “proverb” and by establishing a clear distinction between the two phenomena. More specifically, in the subsequent section (2.1.), we will discuss a purely formal textual phenomenon of proverbs, maxims, and anecdotes. Afterwards, in section 2.2., the gnomic sense will be defined and the most typical cross-linguistic means of its expression (grams that belong to the imperfective-present continuum) will be presented. Next, we will offer a detailed typological analysis of the relation that exists between the post-resultatives and the gnomic sense (section 2.3.).¹⁶ After that, the gnomic use of the suffix conjugation will be studied. First, all

¹⁴ In accordance with this path (the third sub-cline within the resultative track), certain resultative proper forms evolve into evidential grams following the following subsequent stages: a) inferential, based upon resulting visible traces; b) inferential, based upon general assumption and hearsay; and c) broad non-first-hand evidential (cf. Lindstedt 2000, Johanson 2000, 2003, Aikhenvald 2004: 112–117, 279–281, Andrason 2010b).

¹⁵ Modal contamination codifies a process during which indicative formations (because of their consistent use in clearly modal contexts) gradually adopt the modal meaning of their environment as their own and are finally converted into genuine moods (Dahl 1985: 11, Hopper and Traugott 2003: 82, Bybee, Perkins, and Pagliuca 1994: 25–26, Andrason 2011a: 300–304; cf. also Andrason 2011d: 6–8).

¹⁶ It is important to note that an exact evolutionary connection between grams developing along the anterior cline and the gnomic sense has not yet been posited. For instance, Bybee, Perkins, and Pagliuca (1994) do not establish any particular path linking the idea of gnomicity with resultatives, perfects, perfectives, and/or past tenses. This means that we cannot merely use a typological “gnomic cline”, because such a cline has not yet been proposed. Certainly, scholars have long been aware of the relation between the gnomic value and resultatives, perfects, perfectives, and pasts, however, they have not designed an evolutionary model (a cline) where this value would be explicitly and precisely located. Therefore, the task of “discovering” a diachronic (and hence conceptual) connection between the gnomic value and verbal constructions evolving in accordance with the anterior path falls on the author of the present study. In other words, the current paper must include a special typological section dedicated to a universal relation between resultative diachronies (grams that develop along the anterior path) and the gnomic sense. We must clarify how the gnomic value is related to grams located at a given stage of the anterior path (or spanning certain por-

examples of the gnomic *qatal* in Proverbs will be introduced and classified in accordance with the typological evidence (section 3.1.), and then cases of the gnomic value of cognate forms in other Semitic languages will be discussed (section 3.2.). Finally, we will propose a typologically plausible link chaining the gnomic *qatal* to the remaining semantic potential of the gram. In this manner, a more holistic semantic map of the category will be posited, i.e. a map which includes both the characteristic senses as well as the less prototypical gnomic value (Section 4.).

2. Gnomic sense and its expressions

When discussing the issue of the gnomic sense, important distinctions must be made. First, one needs to differentiate between the formal level (proverbial texts, a literary genre that expresses timeless truths) and the semantic level (a gnomic sense or the value of universal truth).¹⁷ Furthermore, one must acknowledge that the gnomic value itself can be conveyed by various grammatical constructions. Given the inherent polysemy of grams, a semantic domain – the gnomic value, included – can be expressed by several types of forms. Put differently, since grams are regularly polysemous, they almost by definition overlap in certain areas and hence parts of their semantic maps intersect.¹⁸ In our research, we will describe two exemplary types of formations that are cross-linguistically employed to convey the gnomic value: gnomic imperfectives or broad presents (grams that are regarded as the most prototypical expressions of the concept of gnomicity; cf. section 2.2.) and post-resultative grams that frequently can transmit the idea of gnomicity (cf. section 2.3.).

2.1. Proverbs and their “tenses”

Proverbs and maxims constitute an environment that is particularly propitious for the use of forms in a gnomic sense. They are defined as a “traditional, conversational, didactic genre with [...] a potential free conversational turn [and] preferably with figurative meaning” (Norrick 1985: 78) which typically expresses general principles that may be treated as omnitemporal rules (Frykenberg 1996: 98–99). It is also important to notice that proverbs function as almost complete “small” texts. This means that they are not grammatically tied to the remaining part of the utterance or discourse, being instead only related to it at a conversational or pragmatic level (Friedman 1999: 140). *De facto*, maxims may constitute short stories or, in an ex-

tensions of it). We must specify the exact location of the gnomic sense and its stage on the anterior trajectory. Only once the gnomic sense has been incorporated into the model of the anterior path, and the exact position of the gnomic stage in the entire map has been established, will it be possible to employ this path as a template for chaining the gnomic *qatal* to the remaining portion of its semantic potential. As a result, the needed, extensive and purely linguistic typological study will lengthen the present article.

¹⁷ Rogland (2003) incorrectly claims that both terms lack precision.

¹⁸ For instance, the concept of futurity may be conveyed by properly future grams, “present” tenses (i.e. formations that are usually labelled as present tenses or that typically function as such), agentive modal expressions and subjunctives (syntactic moods), or even by post-resultative constructions (i.e. formations that derive from resultative proper grams).

treme case, elaborated anecdotes relating certain events that are not directly connected to the rest of the text. Such anecdotes may thus communicate events and situations – disconnected from the main narration or discourse line – presenting them as successive historical actions. Only a proverb, maxim, or, especially, anecdote in its entirety – viewed as a message in its totality – is relevant to the text. Its “internal” temporal organization may, on the contrary, be completely independent of the argument of the discourse or narration.

Given their grammatical “independency”, proverbs, maxims, and anecdotes may employ virtually any verbal taxis, tense, aspect, or mood with all possible taxis, aspectual, temporal, and modal values. As already explained, proverbs may constitute short descriptions or small texts on their own. The message they carry or the situation they create – truths derivable from the whole story – is omnitemporal and gnomic, but this is not necessarily the case for the verbal forms themselves that are employed in a proverbial fragment. Thus, examples where futures, perfects, pasts, or modal formations are used in maxims or small complete anecdotic texts do not *per se* trigger a gnomic interpretation of such forms. As a result, it is not necessary that grams employed in a gnomic genre express a gnomic value. Quite the opposite, they may function as genuine futures, perfects, pasts, or moods.¹⁹

2.2. *Gnomic sense and its expressions*

Among all senses that may be conveyed by grams in proverbs, maxims, and anecdotes, it is possible to distinguish one which could be labelled as properly gnomic. This value or semantic domain has been defined as “a (subjectively assumed) universal truth”, i.e. a universal truth that may be subjective.²⁰

The gnomic value is most commonly conveyed by certain imperfective grams, especially by so-called “gnomic imperfectives”²¹ (Bertinetto and Lenci 2010: 17, 29, see also Friedmann 1999, Rogland 2003, Cook 2005). According to Bertinetto and Lenci (2010: 18, 24, 28), a gnomic imperfective gram expresses law-like generalizations with a strong normative character which represent exemplary properties of an individual or of a class of individuals and which are valid for a determined period of time. However, gnomic imperfectives – besides its central gnomic cord – typically provide other “collateral” values, closely related to the idea of a universal truth. In particular, gnomic imperfectives constitute *de facto* a collection of more specific senses such as habitual (2.a), attitudinal-potential (2.b-c), generic (2.d), and

¹⁹ Rogland (2003: 22) correctly criticizes the opinion that proverbs should use normally a form that conveys general truths, viz. a present tense, because they *per definitionem* express general truths. In his opinion, this is an oversimplification, because proverbs in various languages commonly use past tenses. Thus, although the present tense is cross-linguistically frequent in proverbs, it cannot be defended that it is the proper tense for sayings and maxims. According to Rogland, languages are very lax in selecting tenses in gnomic statements and general truths therefore can be found in the three temporal spheres. Hence, they are expressed not only by present tenses but also by past and future grams (*ibid.*: 22).

²⁰ This definition may be derived from Bertinetto and Lenci (2010) and was suggested to me by Pier Marco Bertinetto in an e-mail exchange on 14.08.2012.

²¹ Broad or general presents constitute a sub-class of imperfective grams. These are imperfective constructions narrowed to the present time frame.

individual-level²² (2.e) value. Inversely, the four previously mentioned semantic nuclei are regularly extended to gnomic readings. This also means that if the four senses (habitual, attitudinal-potential, generic, and individual-level) have been grammaticalized as independent grams, such gram types typically develop a gnomic force.²³

(2) a.

John easily **gets** angry with his colleagues

b.

John **smokes** cigars

c.

John **speaks** Swahili

d.

Dogs **have** four legs

e.

Elina **is** Finnish

It should be noted that the activity or situation conveyed by the predicate in gnomic imperfective formations is characteristically bound by “quasi-universal” quantifiers, such as the adverbs *always*, *normally*, *typically*, *usually*, etc. (Bertinetto and Lenci 2010:25). This “quasi-universality” implies that general gnomic statements admit of exceptions (Krifka *et al.* 1995). They rather refer to a potential capacity or role of the individual and hence may be violated in a concrete actualization while still remaining valid for normal or prototypical circumstances (Bertinetto and Lenci 2010: 25–26 and Boneh and Doron 2010: 352 and 355). This property of gnomic imperfectives has led to the conclusion that such formations are best explained by making use of intensional models (Bertinetto and Lenci 2010: 28). Thus, gnomic sentences do not express propositions concerning the actual world, but rather “statements that need to be evaluated with respect to a contextually determined set of possible worlds or situations” (Bertinetto and Lenci 2010: 28). In that manner, the gnomic imperfectives – the most typical expression of the gnomic sense – display the following characteristics: (1) they tolerate exceptions, (2) express law-like truths, and (3) by accounting for potential functions of the subject, fail to necessitate a concrete actualization (*ibid.*: 27–28).

²² An individual-level predicate is true throughout the existence of an individual. When an individual-level predicate occurs in past tense, it gives rise to what is called a lifetime effect (Carlson 1977, Kratzer 1989 and Chierchia 1995, Carlson and Pelletier 1995).

²³ It should be observed that all of these values/grams also imply a degree of durativity and stativity.

2.3. *Anterior path grams as expressions of the gnomic sense*

As already mentioned, given the principle of polysemy, it should be possible to express the semantic domain of gnomicity – apart from conveying it by gnomic imperfectives – also by means of other types of grammatical formations. This being said, an immediate question arises: what other grams can convey gnomic nuances?²⁴

Having defined the concept of gnomicity and having described gnomic imperfectives – constructions that constitute the main device in expressing the gnomic meaning – we can propose a set of properties which should characterize a gram that is employed with a gnomic value. First, formations that are used with a gnomic force should denote potentially subjective universal truths. That is to say, they are required to introduce personal or universal normative generalizations, specifying exemplary traits of individuals that hold for a determined period of time. Moreover, their meaning is likely to be somehow related to the nuances of habituality, potentiality, and characterization (generic and individual-level domain) that typically co-occur in gnomic imperfectives. By doing so, such expressions – although still viewed as universal truths – would tolerate exceptions and fail to require actualizations. Since habituals, potentials, generics, and individual-level formations typically develop gnomic readings, if a gram is employed with a habitual, potential, and characterizing force, it is probable that it could also convey a gnomic value. For instance, taking into consideration the fact that habituals commonly create intensional situations and lend themselves to gnomic readings (Boneh and Doron 2010: 360, cf. also Palmer 2001: 179 and Hellenthal 2007: 31), one may assume that grams that convey any type of a habitual value would be suitable for gnomic extensions; in certain contexts, they may be interpreted as gnomic formations (Boneh and Doron 2008: 321; vide also see Carlson 1977 and Krifka *et al.* 1995).

Other types of grammatical constructions which – alongside the gnomic imperfectives and broad present tenses – have been reported to express gnomic nuances are perfects, perfectives, and past tenses (i.e. grams that develop along the resultative cline and, in particular, along the anterior cline). This property of perfects, perfectives, and past formations seems to be quite well-documented cross-linguistically (cf. Friedman 1999 and Rogland 2003; cf. also Norrick 1985:3, Spasov, Topolińska

²⁴ Given that the chaining is required to be based on typological “universals” (or under a less categorical assumption, on common evolutionary scenarios), in this section, we will analyse typological data from a broad spectrum of non-Semitic languages. What we are trying to do in this section is to show a common developmental and thus conceptual link between grammatical categories developing along the anterior cline and the semantic domain of gnomicity. To be precise, we will demonstrate that grammatical categories (concrete grams from distinct languages) which otherwise evolve along the anterior past (resultative proper constructions > young anteriors > old anteriors > past tenses), besides gradually acquiring values typical of these categories, also (although with a different frequency and intensity) convey gnomic values.

It is important to clearly distinguish between a semantic domain, on the one hand (i.e., a sense, a specific value or a piece of information conveyed by a form), and a grammatical category (or a form), on the other. For example, the label “future” can refer to two distinct phenomena. First, it can indicate a concrete sense conveyed by a given form which also provides other senses (observe that the sense of futurity can be conveyed by grams that are defined as present tenses, moods, or even past tenses). Second, it can refer to a grammatical category of a future tense, i.e. to a category whose most typical uses cover the domain of futurity (observe that future tenses besides conveying the idea of future also possess other common semantic properties).

and Spasov 1986:10, 47). However, no typological explanation of the phenomenon has been proposed thus far. First, scholars have failed to determine the exact location of the stage where the extension from a post-resultative sense to a gnomic value takes place. And secondly, they have likewise failed to define the nature of this “surprising” trait of post-resultatives; that is to say, they have not clarified the difference in the compatibility with the gnomic sense displayed by non-advanced (resultative proper or young perfects) and advanced grams (old perfects, and (perfective) pasts).²⁵

As previously explained, the anterior and simultaneous tracks constitute two major developmental trends within the resultative path. This inversely means that resultative proper inputs typically evolve along two different clines: the anterior one and the simultaneous one. When developing in accordance with the latter scenario (characteristic of non-dynamic or adjectival roots), they transmute into broad, general present tenses, passing through the stage of statives (cf. Andrason forthcoming (b)). Thus, they are typologically suitable to acquire functions characteristic of gnomic imperfectives. Especially, in their last stage – when they operate as general present tenses indistinguishable from presents that have arisen from properly imperfective grams (cf. the preterite-present verbs in Germanic languages) – they are entirely compatible with the gnomic domain. This indicates that the connection between these types of grams and the gnomic value is “natural” and evident. Developing towards the stage of a broad general (imperfective) present and having previously acquired the stative sense (also typical of gnomicity), they unsurprisingly lend themselves to gnomic extensions (cf. a similar observation in Rogland 2003:24). The interesting and troublesome matter is the relation between the other type of post-resultative formation (i.e. those that are situated along the anterior path) and the gnomic domain. Why can resultative proper grams, perfects, and perfective or simple past tenses express the gnomic value? Is it possible to posit a universal law that connects such grams to the gnomic sense and thus to establish a solid conceptual-diachronic linkage that would explain the gnomic sense of the BH *qatal*?

In this section of the paper, we will provide extensive typological evidence that will enable us to posit a systematic connection between post-resultatives (resultative proper 2.3.1., young anteriors 2.3.2., old anteriors 2.3.3., and pasts 2.3.4.) and the gnomic sense. In this way, we will propose a universal evolutionary scenario – a path – that links such grams and the idea of gnomicity. Consequently, an “aberrant” behaviour of perfects and pasts will be explained and cognitively justified. Having ascertained the exact nature of this commonly noticed connection, we will be able to explain the relation between the anterior path *qatal* (i.e. the dominant portion of the semantic space of the BH suffix conjugation) and its gnomic variety.

2.3.1. Resultative proper

Resultative proper grams (such as the English *is written* or the Spanish *está escrito*) constitute diachronic inputs of the anterior path; it is from them that perfects and (perfective) past tenses arise. In their most prototypical usage, such constructions

²⁵ These labels will be explained later in this section.

convey a complex sense or a two-fold piece of information; they express static non-dynamic qualities of a being or thing, viewed as resulting from previously performed activities (Maslov 1988: 64, Jaxontov 1988: 101, Sil'nickij 1988:88, 96–97). Additionally, resultative proper grams typically display an intransitive and (if derived from dynamic and transitive roots) de-transitive force (Nedjalkov 2001: 929).

However, typological data teach us that resultative proper formations also are commonly employed in a gnomic function, that is, to express general truths, habitual states, or permanent – potentially universal – situations. In this usage, resultative proper grams may appear with typical quasi-universal quantifiers such as *always*, *usually*, *typically*, *normally*, or *never* and refer to timeless everlasting properties, typical of an individual or a class. Although the idea of the prior action which has triggered the present state is still available, the temporal duration of this state is hereby expanded from the ongoing present to a general and omni-temporal present. The resultant static condition is not only current but also permanent.²⁶

As indicated by evidence provided by several Indo-European branches such Germanic (Icelandic (3.a), English (3.b), and Vilamovicean (3.c)), Slavic (Polish (3.d) and Russian (3.e)) and Romance (Spanish (3.f) and French (3.g)), as well as from the Niger Congo family (Mandinka (3.h)), it possible to argue that there is virtually no restriction in the usage of resultative proper constructions in gnomic statements. Resultative proper locutions – when situated in a present time sphere – may indicate not only actual or transitory characteristics (current results of former activities) but also, if the context requires and the enunciator wishes, stable and permanent qualities derivable from previous actions. These invariant or permanent conditions may subsequently be viewed as typical, general, and universal, thus giving rise to gnomic uses.

(3) a.

Spánverjar	eru	alltaf	komnir	seint
spaniards	are	always	come	late
Spaniards always come late				

b.

Russians **are** usually **drunk**

c.

Dy	oüta	zajn	gyrjyt	diöh	dy	benzyn
the	cars	are	propelled	by	the	gas
The cars are propelled by gas						

²⁶ It is important to note that resultative proper constructions commonly offer three main senses. As already mentioned, they are extensively employed with the sense of a resultative proper – two portions of the semantic information (i.e. the prior action and the resulting state) are equally important. However, they also can convey the value of a resultative stative and stative with non-resultative nuances (see for instance the Akkadian *parsäku*; Huehnergard 2005: 19 and Andrason 2011: 186–206). Again, one should clearly distinguish between a grammatical category and the senses it can convey. The three values offered by the *parsäku* and other resultative proper grams correspond to cross-linguistically common semantic domains with which resultative proper construction are usually compatible.

d.

Chłopcy z mojej klasy są zwykle zmęczeni
 boys from my class are usually tired
 The boys from my class are usually tired

e.

Магазины всегда закрыты во время рождественских праздников
 shops always closed in time Christmas holidays
 Shops are always closed during the Christmas holidays²⁷

f.

Los finlandeses normalmente están deprimidos
 the Finns typically are depressed
 Finns are typically depressed

g.

Les igloos sont formés de blocs de neige compacté
 the igloos are made of blocks of snow compacted
 Igloos are made of compacted snow blocks

h.

M be sabatirig Basse
 I am lived Basse
 I live in Basse

It must also be observed that, just like exemplary gnomic imperfectives, this type of formation tolerates exceptions (cf. a Spanish example in 4.a), does not require a concrete actualization (cf. Polish and Mandinka examples in 4.b-c), and expresses potential-attitudinal activities (4.d). This almost “innate” compatibility of resultative proper grams and the gnomic domain becomes evident if one considers their highly frequent usage in overt generic statements, e.g. in definitions, classifications, and exemplifying descriptions (cf. the English expressions *is defined as*, *is agreed that*, *is called*, *is labelled*, *is referred to*, *is characterized by*, etc.; 4.e). This usage of resultative constructions in definitions, permanent prescriptions, and laws is extremely common, being documented in an impressive number of languages: e.g. Polish (*Ta ulica jest nazwana imieniem słynnego pisarza* ‘This street is named after a famous writer’, *Jest wszem znane* ‘It is known to everyone’), Russian (*известно всем* ‘It is known to everybody’), Spanish (*Está prohibido fumar en las estaciones de tren* ‘smoking is forbidden at train stations’), French (*Il est convenu d’être indulgent envers les médias locaux* ‘It is expected/accepted to (i.e. one should) be indulgent towards the local media’, *Il est issu de la famille des Bourbons* ‘He comes from the family of the Bourbons’, *Il est interdit d’interdire* ‘it is forbidden to forbid’), Icelandic (*það er viðurkennt að* ‘it is accepted that’) and many others.

²⁷ This example shows that the auxiliary *be*-type verb is not obligatory in resultative formations in general and in their gnomic usage in particular.

(4) a.

Aunque este Juguete funcione,
 although this toy works
 Although this toy works,

los juguetes de China **están** siempre **rotos**
 the toys from China are always broken
 Chinese toys are always broken

b.

Banki **sa zamknięte** wieczorem ale ten jest teraz otwarty
 banks are closed in-the-evening but this is now open
 Banks are [typically] closed although this one is open now

c.

Luŋ-wo-luŋ a **be saasaariŋ**.
 every day he is been.sick
 He is sick every day.

Bari a maŋ saasaa saayiŋ
 but he is.not sick now
 But today he is not sick

d.

Humans **are given** free choice and free will

e.

A triangle **is defined as** a 3-sided polygon

2.3.2. Young anteriors

Young anteriors are grams that are typically employed in dynamic (and transitive, if possible; Nedjalkov 2001: 928–929, 932, 937–938) present perfect senses. More specifically, they function as inclusive, iterative, resultative, experiential, and indefinite perfects.²⁸ Amongst all of the archetypal perfect uses, it is possible to distinguish three that are particularly propitious for generating gnomic readings; as will be demonstrated in the following discussion, inclusive, iterative, and experiential perfects almost regularly prompt gnomic extensions.

An inclusive perfect expresses an activity or a state that persists without disruption from a determined moment in the past to the present temporal point (Jónsson 1992: 129–145). By indicating a permanent state or an activity that originated in the past but that, since then, constitutes an invariant – or rule-like – phenomenon in the subject's world, the value of an inclusive perfect is close to the concept of habituality and gnomicity. This constancy – or quasi-universality – can be overtly expressed by means of adverbs and expressions such as *always*, *usually*, or *since ever* (5.a-c). Moreover, formations that are employed with an inclusive perfectal force, in their

²⁸ These labels will be explained later in this section.

habitual and gnomic readings, do not require actualization but tolerate exceptions and indicate potential situations (5.d).

(5) a.

God **has** always **loved** men (= God loves men)

b.

Humans **have** always **had** two legs and two arms (= Humans have two legs)

c.

Humans **have** always **lived** on Earth (= Humans live on earth)

d.

I **have** always **been** a teacher (= I am a teacher although I am not teaching right now)

In cases where telic verbs are employed, the perfect does not express a continuous state or uninterrupted activity, but rather implies a habit of repeated independent actions that, however, may be viewed as an incessant custom of achieving something. This is what has been referred to as frequentative perfect (e.g. the Portuguese perfect in Squartini and Bertinotto 2000:409). Being semantically akin to the habitual domain, such perfects commonly introduce characteristic properties of individuals, presenting them as generally valid or universal (6.a–b). Again, exceptions are tolerated and a nuance of potentiality is available (6.c–d).

(6) a.

In Gambia houses **have been made** of mud and straw (= People build houses in that manner)

b.

The justice system **has killed (been killing)** innocent men for ages (= The justice system kills innocent men)

c.

In this country, parents **have** always severely **punished** wicked children; but not this couple

d.

Tom **has smoked** his entire life (= He smokes even though he is not smoking at this precise moment)

Likewise, the experiential perfect can be employed with a gnomic force by suggesting that the subject possesses a given quality because a certain activity or situation constitutes his or her experience. Experiential perfects typically indicate that the subject has had the experience of performing a certain activity (e.g. *I have read*

Principia Mathematica; Jónsson 1992: 129–145). This action may have taken place in a distant past although this past time moment cannot overtly be expressed. In the gnomic function, experiential perfects denote stable, constant, general, and (subjectively) universal qualities. However, in contrast with the universal and frequentative varieties, the universality is related not to the perfectal activity itself, but to the idea of current relevance typically conveyed by such anteriors. Namely, it corresponds to the slot of the information that is inferred from the previous action and that has been lasting from the moment the event occurred until the present time. This inferential portion of the meaning – and not the experienced event itself – may be viewed as invariant, typical, habitual, generic, and hence gnomic (7.a–d). This means that contrary to inclusive and frequentative perfects, the activity conveyed by a verbal root in this type of anterior is incompatible with the intensional domains of non-actualization or potentiality as well as with the value of habituality. It corresponds to a single event, an undeniable fact that has taken place. The inferences drawn from this experienced episode may, however, be understood as habitual and permanent possibly, thus triggering gnomic interpretations.

(7) a.

I **have seen** God (= I believe in God)

b.

They **have seen** the evil of the world (= They know it)

c.

This person **has killed** a man (= He is a murderer)

d.

My child **has graduated** from law school (= He is a lawyer)

In certain contexts, present perfects introduce independent – each time fully accomplished – events that have occurred on more than one occasion, without constituting a solid continuous sequence (cf. inclusive and frequentative perfects) and without triggering inferences that could provide such a habitual reading (experiential perfect). This is a type of anterior which could be labelled as “iterative indefinite perfect”,²⁹ e.g. *each time he has come he has asked for my sister* (compare with the frequentative perfect *he has been coming here many times today/this month...*). Obviously, due to the fact that they indicate a habit of performing individual actions, such iterative indefinite perfects, may be extended to a gnomic usage. The repetition of an event can be viewed as habitual and thus as forming a stable invariant routine in a person’s life:

²⁹ The indefinite perfect expresses events that clearly have occurred in the past. The main emphasis is placed on the past action itself without, however, situating it at a definite moment in the past. The verbal form is not accompanied by any overt past temporal specifier (Lindstedt 2000: 369 and 379).

(8)

Each time humans **have tried** to defy the gods, they **have been punished**

Additionally, negative varieties of the above mentioned types of the anterior, so-called “anti-perfects”, commonly express permanent states or generic activities, and thus are compatible with a gnomic usage:

(9) a.

Humans **have** never **lived** on Jupiter (= Humans do not live there)

b.

Houses in Gambia **have** never **been built** out of bricks (= People in Gambia do not build houses out of bricks)

c.

Humans **have** never **flown** (= Humans do not fly)

d.

Humans **have** never **been** to Jupiter (= Humans do not live on Jupiter)

Having described semantic properties of young anteriors, it is evident that these grams almost *per definitionem* incline themselves towards habitual, generic, and characteristic – and hence gnomic – uses. Since young anterior formations commonly offer inclusive, frequentative, experiential, indefinite, iterative, and anti-perfect senses – values that naturally give rise to gnomic extensions – they are automatically suitable for the gnomic usage as well. In order to further illustrate this phenomenon, in addition to the English examples introduced previously, one may provide data from Icelandic (10), Spanish (11), and Classical Greek (12).

The Icelandic perfect – a young anterior gram employed as an inclusive, frequentative, resultative, experiential, and indefinite perfect – can convey habits, customs, characteristic states, and conditions that may be potential and non-actualisable:

(10) a.

Ég	hef	starfað	sem	þjónn	alla	æfi
I	have	worked	as	waiter	all	life

I have worked (been working) as a waiter all my life

b.

Hann	hefur	reykt	síðan	alltaf
he	has	smoked	since	always

He has always smoked (been smoking forever)

c.

Hann	hefur	alltaf	verið	heimskur
he	has	always	been	stupid

He has always been stupid

d.

Íslendingar	hafa	alltaf	varið	landið	sitt
Icelanders	have	always	defended	country	their

Icelanders have always defended their country

e.

Hann	hefur	þekkt	hamingju
he	has	known	happiness

He has known happiness

f.

Hann	hefur	syngdað
he	has	sinned

He has sinned

The young anterior in Spanish – a *have*-type perfect – may also sometimes be employed with a gnomic force. It should be noted that the Spanish gram has advanced somewhat along the anterior path. It namely fails to be used in a prototypical inclusive perfect sense, having instead acquired the value of a hodiernal (today) and/or hesternal (yesterday) past. This means that a common gnomic extension of the inclusive perfect use is lacking in Spanish and thus its accessibility to the domain of gnomicity has been reduced.

(11) a.

Siempre	ha	cometido	el	mismo	error
always	he.has	made	the	same	mistake

He has always made the same mistake (= He does it even now)

b.

Este	hombre	ha	matado
this	man	he.has	killed

This man has killed (= He is a murderer)

c.

He	estado	en	China
I.have	been	in	China

I have been to China (= I know this country)

d.

Nunca	he	pecado
never	I.have	sinned

I have never sinned (= I am not a sinner)

e.

Las	mujeres	nunca	han	tenido
the	women	never	they.have	had

Women have never had

los	mismos	derechos	que	los	hombres
the	same	rights	as	the	men

the same rights as men (= They still lack some rights)

The Perfect in Classical Attic Greek – a resultative proper and present perfect gram – may also be employed with a gnomic value, and thus denote universal truths. This universality is based upon the experience of having performed an action (experiential perfect) or stems from a continuity/habituality of situations or actions that have been occurring since a given moment in the past (inclusive, frequentative, and iterative experiential perfect; cf. Goodwin 1893, Smith 1920, Rydberg-Cox 2000)

(12) (Xenophon, *Anabasis* 3.1.38)

ἡ	ἀταξία	πολλοὺς	ἤδη	ἀπολώλεκεν
the	non-discipline	many	already	has ruined

The lack of discipline has already ruined [has been ruining] many men (Rydberg-Cox 2000)³⁰

2.3.3. Old anteriors: past tenses with common perfect uses

Old anteriors are constructions that function as prototypical perfects as well as definite past tenses. In other words, they are past tenses which have not yet been reduced to solely past functions. In various languages, such formations are able to convey gnomic nuances, although this capacity is usually limited to typical perfectal contexts, i.e. when the gram is employed as an inclusive, frequentative, or experiential anterior as well as an anti-perfect. In other words, since old perfects may preserve some of their exemplary perfectal uses, such perfectal uses may, in turn – as was the case for young anteriors – give rise to gnomic extensions. In these instances, the form conveys gnomic value with all of its co-strings such as habituality, generic characterization of individuals and classes, potentiality, and non-actualization (i.e. tolerance of exceptions).

For example, in French, the *passé composé* is an archetypal present perfect and perfective past, especially as far as the spoken language and discursive texts are concerned. In environments that are typical of present perfects, the locution may provide gnomic values, particularly if it is accompanied by overt quantifiers such *toujours* ‘always’, *souvent* ‘often’, or *jamais* ‘never’. In these cases, the gram expresses permanent situations, generic properties, or truths that are subjectively valid and/or omnitemporal (13; Grevisse 1975: 727–728)

(13) a.

J'	ai	travaillé	toujours	dans	cette	usine
I	have	worked	always	in	this	factory

I have always worked in this factory (= it is my job)

b.

Un	verre	de	vin
a	glass	of	wine

A glass of wine

³⁰ The Greek quotes have been obtained via the Perseus Project at Tufts University (www.perseus.tufts.edu).

n' a jamais **fait** de mal à personne
 not has never done some³¹ harm to anyone
 has never done harm to anybody (= it is good)

The Akkadian language also possesses in its verbal repertory a gram – referred to as the *iprus* – that in most of its uses functions as a present perfect (especially in negatives) and past (either perfective or simple [preterite]; Huehnergard 2005: 19, 158 and Andrason 2010c: 333–337). Besides these common uses, the *iprus* may likewise be employed with a gnomic extratemporal value, introducing invariant or permanent situations, generic properties and typical activities (cf. Mayer 1992, Metzler 2002: 380, 743, Loesov 2004: 426–428).

(14) (YOS 11, 2:1–6)

a-ra-aḫ-ḫi ra-ma-ni a-r[a]-a-ḫi pa-ag-ri
 I.fertilize myself I.fertilize body.my
 I fertilize myself, I fertilize my body

ki-ma na-ru-um **ir-ḫu-u₂** ki-ib-ri-ša
 like river (has.)fertilized banks.her
 Like the river has fertilized [*or* has been fertilizing] its banks (Loesov 2004: 428)

Similarly, the so-called *l* past tense in Polish, both in its perfective and imperfective variety, may convey gnomic nuances. The Polish *l* past tense is a prototypical old anterior most commonly employed as a present perfect and past tense (either perfective or imperfective). In various perfectal uses, the gram expresses habitual currently relevant activities, permanent characteristics, or stable truths, thus giving rise to gnomic readings:

(15) a.

Zawsze **mieszkałem** tutaj
 always I.lived-IMPF here
 I have always lived here (= This is my residence place)

b.

Człowiek zawsze **grzeszył**
 man always sinned-IMPF
 Man has always sinned (= This is his nature)

c.

W moim życiu **poznałem** co to ból i rozpacz
 in my live I.knew-PRF what it pain and despair
 I have known pain and despair in my life (= I know them now)

d.

Widziałem zło tego świata
 I.saw-IMPF evil of.this world
 I have seen the evil of this world (= I know it now)

³¹ The lexeme *de* is in fact a preposition ‘of, from’ but in negatives it is used with the force of an indefinite article.

e.

Nigdy	nie	zgrzeszyłem
never	not	I.sinned-PRF

I have never sinned (= I am a sinless man / I am not a sinner)

f.

Zwierzęta	nigdy	nie	mówiły	ludzką	mową
animals	never	not	they.spoke-IMPF	human	language

Animals have never spoken in a human language (= They do not talk)

It should be noted that the accessibility of such old anteriors to the gnomic sense strongly depends on their tolerance of specific perfectal contexts. This means that if the gram is still commonly employed as an inclusive and frequentative perfect, the possibility of it being used with a gnomic force is high. However, more typically, old anteriors are only common in the experiential perfect use. As already explained, gnomic readings of experiential perfects are quite particular because they stem from the inferential slot of meaning conveyed by an experiential perfect (cf. section 2.3.2.). For this reason, the gnomic usage of old anteriors seems to be more restricted than that of young perfects.

2.3.4. Past tenses

The last category located on the anterior path corresponds to prototypical past tenses, i.e. grams that in most cases function as definite pasts, while perfectal uses are normally expressed by other, younger formations. It must be emphasized that, although they are defined in grammar books as exemplary past tenses, these grams quite regularly provide certain additional semantic strings, especially those that reflect their perfectal “prehistory” (i.e. usages where, at earlier developmental stages, they functioned as present perfects). In other words, various past tenses admit some – unquestionably, highly restricted and infrequent – present perfect uses. Most commonly, such perfectal functions correspond to an experiential perfect and to an anti-perfect. This compatibility (which is certainly highly limited) with such experiential and “anti-perfectal” contexts enables a formation – in most of its functions approximating a definite past – to express gnomic truths. In so doing, it indicates stable, invariant, habitual, and typical situations, activities, or properties.

For instance in Classical Greek, the Aorist Indicative is a verbal formation that, in its most prototypical use, functions as a perfective past (observe that it is marked by a past marker, i.e. by the augment). Accordingly, it indicates past perfective (complete and/or completed) actions and events (Humbert 1954:141–144, Rijksbaron 1984: 12–15, 20, Hewson and Bubenik 1997: 28–34, 43–44). In light of the preponderance of this function, it is typically classified as a perfective past gram contrasting with the Imperfect (an imperfective past gram); both are marked by the augment (the past marker) and their main distinction is aspectual (Humbert 1954: 138, Rijksbaron 1984: 12, Hewson and Bubenik 1997: 29; see also Crespo, Conti, and Maquieira 2003:275–285). However, the Greek Aorist Indicative may also (especially in maxims and proverbs) express customary truths, omnipresent habits, permanent qualities, or activities typical of a class or an individual (Gilder-

sleeve 1900: 109–110, Humbert 1954: 145–146 and Rijksbaron 1984: 3–32). It should be noted that in such cases, the Aorist does not refer to the past but – given various contextual and pragmatic factors and as an exemplary young or old anterior gram – belongs to the present temporal sphere and indicates permanent and/or universal properties or situations (Thompson 1902: 320–321, Chantraine 1953: 178, Ruijgh 1971). This means that the universal or gnomic present value – besides being conveyed by the Present (a broad present tense) – is also expressed by the Aorist in Classical Greek.³²

(16) a. (Hesiod, *Works and Days* 218)

παθὼν	δέ	τε	νήπιος	ἔγνων
he.suffered	but	and	fool	learned

But a fool has [always] learned from experience (But once he has suffered the fool realizes this)

b. (Isocrates, *To Demonicus* 1.6)

κάλλος	μὲν	γὰρ	ἢ	χρόνος	ἀνήλωσεν	ἢ	νόσος	ἐμάρῃνε
beauty	indeed	for	either	time	wasted	or	disease	withered

For beauty is either wasted by time or withered by disease (Smyth 1956: 432)

c. (Isocrates, *To Demonicus* 1.6)

ῥώμη . . .	μετὰ	μὲν	φρονήσεως	ὠφέλησεν,
strength	with	indeed	good.sense	profited

Strength with judgment does good,

ἄνευ	δὲ	ταύτης	πλείω	τοὺς	ἔχοντας	ἔβλαψε
without	but	her	more	these	having	harmed

but without, it does greater harm to those that possess it

d. (Plato, *Laws* 720D)

ὁ	δὲ	ἐλεύθερος	διδάσκει	τὸν	ἀσθενοῦντα	αὐτόν,
the	but	free.man	instructs	the	being.sick	himself

The physician who is free instructs the patient himself

καὶ	οὐ	πρότερον	ἐπέταξε
and	not	prescription	put

and does not give a prescription

πρὶν	ἂν	τῇ	ξυμπεῖσῃ
until	PART ³³	him	may.have.convinced

until he in some way succeeds in convincing him

Similarly, in French one may find a highly advanced past tense that allows gnomic uses. The *passé simple* – commonly defined as a narrative simple past tense – is sometimes employed in order to express atemporal general universal truths or cus-

³² Perhaps, in imitation of Greek conventions, Latin in certain cases uses its Perfectum (a prototypical present perfect and perfective or simple past) as a gnomic perfect.

³³ The lexeme ἂν is a particle indicating contingency.

tomary habits (Grevisse 1975: 725–726). As expected, this usage is particularly common in perfectal environments, especially with adverbs such as *toujours* ‘always’, *jamais* ‘never’, and *souvent* ‘often’. In most such cases, the *passé simple* is restricted to gnomic genres such as maxims or anecdotes. It is *de facto* only in this proverbial usage that the gram reappears in the spoken language.

(17) a. (Grevisse 1976: 726)

Qu’	un	dîner	réchauffé	ne	valut		jamais	rien !
that	a	dinner	warmed-up	not	was.worth		never	nothing

A warmed-up dinner has never been worth anything (= it is worthless)

b. (*ibid.*)

Qui	ne	sut	se	borner	ne	sut	jamais	écrire
who	not	learned	himself	limit	not	learned	never	write

The one who has not learned how to limit himself, has never learned how to write³⁴ (= he does not know)

c. (Grevisse 1968: 1483)

Jamais	gourmand	ne	mangea	bon	hareng
never	greedy.one	not	ate	good	herring

A greedy person has never eaten a good herring (= he never eats it)

d.

Jamais	avare	ne	fut	riche
never	stingy.one	not	was	rich

A stingy person has never been rich(= he is never rich)

In a similar vein, albeit very sporadically, the simple past in Peninsular Spanish – a broad simple past tense – may denote extratemporal truths, habitual facts, or constant situations:

(18)

Quien	tuvo	retuvo
Who	had	had

Who has had [something], has had [it] (= someone who has had a quality, has always had some of it)

The fact that highly advanced past tenses gradually tend to disappear from the spoken language, and thus from any discursive perfectal contexts, and that their “spoken” and discursive character is typically preserved in frozen expressions such as proverbs or maxims (which, as already mentioned, are especially likely to contain gnomic imperfectives or broad present tenses), gives the impression that a past tense functions as a present tense. Put differently, being equivalent to the gnomic use of imperfectives and presents (by expressing habitual, generic, typical, and stable situations and activities) and having lost their properly perfectal nature (by failing to be

³⁴ The English translations are as literal as possible. As already mentioned, these examples are proverbs and hence their total conversational meaning is metaphorical.

employed as present perfects in discourse anymore), the perfectal uses of highly advanced past tenses, available in proverbs, are perceived as presents. However, they represent an older stage of the gram where it was commonly used with an inclusive, frequentative, experiential, and/or anti-perfect force, thus making it compatible with the domain of gnomicity.

2.3.5. Results of the typological study

Our data demonstrate that the gnomic sense is commonly conveyed by formations located on the anterior path. Resultative proper grams are especially frequent in the gnomic function. The number of resultative proper expressions used for gnomic statements is extremely large. They almost naturally express general truths, habitual states, or permanent – potentially universal – situations because the resultant state may be viewed not only as currently present or actual but also as permanent and thus universal. In this manner, resultatives overlap with gnomic imperfectives and broad presents; all of them are typically – and probably without restriction – means of expressing the semantic domain of gnomicity. This predisposition of resultative proper grams for gnomic extension stems from the fact that they typically include a nuance of stativity in their senses. As explained above, resultative proper formations are twofold semantic complexes in which a state is portrayed as acquired. This means that the nuance of a state is present in resultative proper constructions in its two most original uses as a resultative proper (two portions of the meaning are equally important) and resultative stative (the relevance of the resultant state is emphasized while the importance of the prior action is less patent; cf. Andrason forthcoming (a)). Since the acquired (due to a prior action) state can be expanded to larger periods of time, resultative proper grams may indicate not only current resulting conditions but also permanent ones. Such permanent states (at least theoretically, acquired due to a previous action) can thus be employed to express invariant truths.

Young anteriors are other formations that virtually *per definitionem* lend themselves to habitual, generic, and characteristic, and, hence, gnomic uses. This stems from the fact that these grams commonly offer inclusive, frequentative, and experiential perfect senses (as well as their negative counterparts; cf. their use as an anti-perfect) which naturally give rise to gnomic extensions. The inclusive and frequentative perfects express universal truths, portraying them as rules that have been holding (inclusive) or repeating (frequentative) since a moment in the past to the present. The experiential perfect does not represent an activity, expressed by the verb, as habitual or constantly occurring. This type of a perfect rather provides some inferences that, since there are currently relevant, may be extended from the present actuality to a permanent view.

Old anteriors may also be employed in a gnomic function, usually in typical perfectal contexts where they function as inclusive, frequentative, experiential, and anti-perfect. This signifies that the gnomic use of the old anteriors is limited to perfectal contexts, especially to those which characteristically stimulate gnomic extensions. Consequently, the accessibility of old anteriors to the gnomic value depends on their tolerance of a specific perfectal environment. Finally, since past tenses ad-

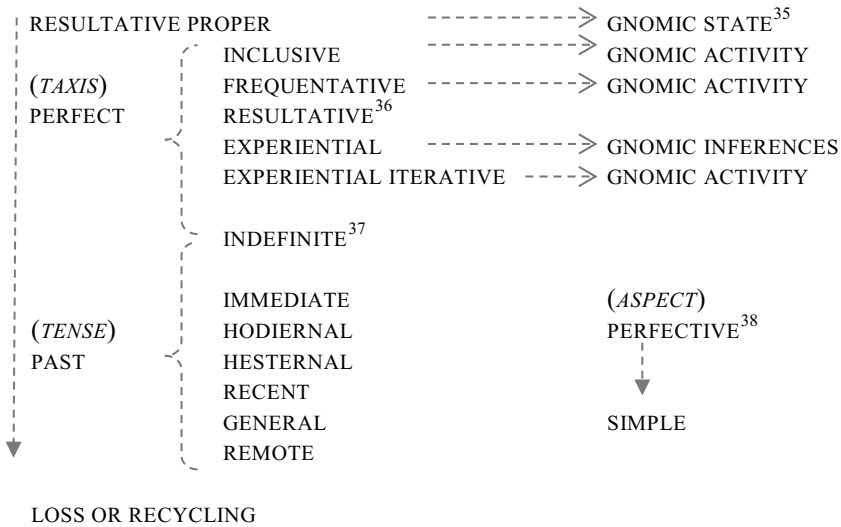


Figure 1. the anterior path and its gnomic “branch”³⁹

mit some restricted and infrequent present perfect uses (especially, the experiential and anti-perfect one), they can express gnomic truths, indicating stable, invariant, habitual, and typical situations, activities, or inferred properties. Constituting an equivalent to the gnomic use of imperfectives and presents in proverbs or maxims, but having lost properly perfectal nature in discourse, such uses of the past tenses are regarded as “presents” instead of being linked to the inclusive, frequentative, experiential, and anti-perfect nature of the formation that may be preserved in proverbial texts.

³⁵ Due to the stative component in the meaning, and because of a typically intransitive and especially de-transitive effect, resultative proper constructions give rise to gnomic states and situations, and not to dynamic activities. In this regard, they contrast with perfects that introduce dynamic actions and can trigger dynamic gnomic senses.

³⁶ The resultative anterior introduces dynamic events, portraying them as highly relevant for the present state of affairs, e.g. *I cannot come to your party – I have caught the flu* (McCawley 1971).

³⁷ The indefinite perfect (labelled also as an indefinite past) is located between the present and past time spheres. It indicates clearly past events, without, however, specifying their temporal location. As for the former property, the gram approximates a past tense. However, given the latter characteristic, the formation behaves as a typical present perfect.

³⁸ As a definite past, the gram may undergo two (to some extent independent) types of evolution. One consists of increasing the temporal distance from the speaker’s here-and-now: immediate > hodiernal (the same day or one day past) hesternal (yesterday’s past) > recent > general (a person’s life’s past) and remote (historical and ancient) past. The other includes a transformation of the anterior into a perfective past and then into a simple past. This change is facultative and occurs in determined types of verbal systems. In should be noted that there is no precise stage-to-stage equivalence between the stages which link the indefinite past and various subcategories of the definite past on the one hand, and the development of the perfective past into its aspectually neutral variant, on the other.

³⁹ The vertical arrows in this figure symbolize the diachronic progression of resultative inputs on the anterior cline. The horizontal arrows stand for gnomic extensions (both conceptual and diachronic) of certain values developed in accordance with the anterior path.

To sum up, the propensity or compatibility of a gram with the gnomic domain seems to decrease with its progression on the anterior cline; it is typical and unrestricted for resultative proper grams, highly common for young anteriors, relatively widespread for old anteriors (as long as their present perfect uses are preserved), and rather infrequent – and conceptually disconnected – for past tenses. In light of the provided evidence, we may argue that the gnomic value constitutes a typical extension of the following senses located along the anterior path: resultative proper, inclusive perfect, frequentative perfect, and experiential perfect (as far as its inferences are concerned), as well as anti-perfect, which corresponds to negative perfect uses. This type of a semantic “enlargement” could be embraced under a common label, viz. “gnomic branch” within the anterior path, an extension that characterizes the evolution of original resultative formations which, during the stages of a resultative proper and a perfect (in particular, inclusive, iterative, and experiential perfects), develop the ability to be employed with a gnomic value.

3. Evidence

3.1. Biblical Hebrew – gnomic *qatal* in proverbs

As already explained, gnomic genres (e.g. proverbs), gnomic values (a semantic domain of subjectively assumed truth) and verbal forms that either appear in gnomic genres (it is possible to use virtually any taxis, tense, aspect, and mood in such this variety of text) or express a gnomic sense (e.g. gnomic imperfectives or broad present tenses) constitute different – although connected – phenomena. Proverbs certainly convey gnomic messages but the grams that appear in them may employ virtually any taxis, tense, aspect, and/or mood. The whole message – the situation depicted by a proverb – is atemporal, but not every verbal gram used in such “a small text” is necessarily so. Therefore the fact that a formation is used in proverbial texts does not *per se* trigger its understanding as a vehicle of the gnomic semantic domain. Only in some cases, do certain formations (most typically, gnomic imperfectives and/or broad present tenses) that appear in a gnomic genre *de facto* convey gnomic meaning.

Consequently, in our analysis of a gnomic use – whether alleged or real – of the BH *qatal*, only those instances of the suffix conjugation which, when appearing in an exemplary proverbial genre (e.g. in the book of Proverbs), express subjectively assumed universal truths (with typical co-strings, explained in Section 2. above), and which thus constitute semantic equivalents of a gnomic imperfective or a broad present of Indo-European languages, are worthy of analysis. In this section, we will show that one may encounter various instances where the *qatal* – found in Proverbs – can be understood as denoting subjective universal truths. This will demonstrate that the gram does (in certain cases) convey a gnomic value in a gnomic genre.⁴⁰

⁴⁰ In our review, examples with stative verbs will be disregarded, because these predicates, when employed in a post-resultative form (such as the BH *qatal*), develop present – both actual and permanent, and hence universal – readings in accordance with the evolution along the simultaneous path. Put differently, since non-dynamic verbs in originally resultative constructions typically follow the simultaneous cline (acquiring senses corresponding to a stative present and present tense), their use for gnomic purposes is evident in contrast with the gnomic nature of certain perfects, perfectives, and past tenses that derive from resultative proper grams following the anterior cline.

In various cases, the gnomic value of the *qatal* form apparently stems from the fact that the gram is employed as an inclusive perfect (see examples 19.a–k, below). In such cases, the formation expresses an unceasing activity that has been going on for a period of time. This durative situation may be viewed as habitual, and thus give rise to generic or gnomic extensions. In this manner, an uninterrupted action that has begun in the past but which has been continuing to the present moment constitutes a permanent fact about the individual's life. It is interesting to note that such universal-perfect gnomic *qatals* usually alternate with clearly durative *yiqtol*s (19.a–e), participial forms (19.f–h) or iterative *weqatal*s (19.i). Furthermore, in a manner analogous to the *yiqtol* grams or participles, in all of the quoted examples, the *qatal* construction exemplifies a typical property of an individual, e.g. of a wise woman, the earth, God, oppressor, mocker, adulterer, parent, labourer, and scoffer.

(19) a.

חַכְמוֹת נָשִׁים בְּנִתָּה בֵּיתָה וְאִזְלֹת בְּגִדֶיהָ תִּהְרָסֶנּוּ

The wise woman **builds** [**< has incessantly/continuously/habitually⁴¹ been building**] her house, but the foolish one tears it down with her own hands (Prov. 14.1)

b.

תַּחַת שְׁלוֹשׁ רָגְזָה אֶרֶץ וְתַחַת אַרְבָּע לֹא־תוּכַל שָׁאֵת:

Under three things the earth **trembles** [**< has incessantly/continuously /habitually been trembling**], under four it cannot bear up (Prov. 30.21)

c.

עֵינֵי יְהוָה נֹצְרוּ דַעַת וְיִסְלֹף דְּבָרֵי בָגָד:

The eyes of the Lord **watch** [**< have incessantly/continuously/habitually been watching**] over knowledge; but he frustrates the words of the unfaithful (Prov. 22: 12)

⁴¹ The English translations with forms in the Present Perfect or Present Progressive Perfect (and with adverbs such as *incessantly* and *continuously* as well as with their extension towards the idea of gnomicity, the lexeme *habitually*) are principally employed in order to make a given perfect sense explicit. They are used as overt indicators of a determined perfect value from which the gnomic sense may have derived. Of course in various cases, such direct or literal “perfectal” renderings sound awkward and should hence be omitted in more refined translations. For instance, in example 19.a, the Present Progressive Perfect form, accompanied by the adverbs *incessantly*, *continuously*, and *habitually*, suggests that the gnomic value of the *qatal* may have its roots in the inclusive perfect senses. Thus, the English expressions in parentheses make an overt reference to the continuity of the action of building and taking care of one's house as opposed to someone who destroys it. This activity has supposedly started in the past and has, without interruption, been continuing until the narrator's present.

The English Present Perfects have also been employed in our translation because, just like the BH *qatal*, these grams evolve along the anterior path and are compatible with the idea of gnomicity. However, these English anterior-path constructions are not always the most adequate because, in English, the main means of conveying the gnomic value is by far the Simple Present tense. The Present Perfect and Present Progressive Perfect are certainly less acceptable. They have thus been chosen in our translations in order to preserve a possible link between certain perfectal senses/stages (inclusive, frequentative, and experiential perfect), on the one hand and the sense/stage of gnomicity, on the other. Again one should make a clear distinction between a sense (a semantic domain) and a grammatical category that conveys it. For further discussion of the issues related to translations, see section 4.2. below.

d.

גָּסוּ וְאִי־רָדָף רָשָׁע וְצַדִּיקִים כְּכַפִּיר יִבְטַח:

The wicked man **flees** [< **has** incessantly/continuously/habitually **been fleeing**] though no one pursues, but the righteous are as bold as a lion (Prov. 28.1)

e.

מִכִּין בְּקִיץ לַחֲמָה אֲגָרָה בְּקֶצֶיר מֵאֲכָלָה:

[The ant] stores⁴² its provisions in the summer. It **gathers** [< **has** [incessantly/continuously/habitually **been gathering**] its food in the harvest time (Prov. 6:8)

f.

עֹשֶׂה־דָּל תַּרְף עֲשֹׂהוּ וּמְכַבְּדוֹ חֲנֹן אֲבִיוֹ:

He who oppresses the poor **taunts** [< **has** incessantly/continuously/habitually **been taunting**] their Maker but whoever is kind to the needy honours God (Prov. 14:31)

g.

לֵיעַג לְרֹשׁ תַּרְף עֲשֹׂהוּ שְׂמִיחַ לְאִיד לֹא יִנָּקָה:

He who mocks the poor **taunts** [< **has** incessantly/continuously/habitually **been taunting**] their Maker; whoever gloats over disaster will not go unpunished (Prov. 17:5)

h.

חֹשֶׁךְ שִׁבְטוֹ שׂוֹנֵא בְנוֹ וְאֹהֵבוֹ שִׁחֲרוּ מוֹסֵר:

He who spares the rod hates his son, but he who loves him **seeks** [< **has** incessantly/continuously/habitually **been seeking**] to discipline him (Prov. 13:24)

i.

כֵּן דֶּרֶךְ אִשָּׁה מְנַאֲפֶת אֹכֵלָה וּמַחֲמָה פִּיהָ וְאָמְרָה לֹא־פָעַלְתִּי אֲוֹן:

This is the way of an adulteress: She **eats** [< **has** incessantly/continuously/ habitually **been eating**] and wiping her mouth and saying [used to say]: “I have done nothing wrong” (Prov. 30.20)

j.

גִּפְשׁ עֲמַל עֲמָלָה לֹו כִּי־אָכַף עָלָיו פִּיהוּ:

The labourer's appetite **works** [**has** incessantly/continuously/habitually **been working**] for him; his hunger **urges** [< **has** incessantly/continuously /habitually **been urging**] him on (Prov. 16.26) (frequentative and experiential interpretations are also possible here)

⁴² The *yiqtol* likewise expresses the idea of continuity and gnomicity. However, since it is a gram evolving along the imperfective path, a translation with the Simple Present (an English imperfective path gram) has been provided in this case.

k.

בַּקֶּשׁ-שֶׁלֹץ חֲכָמָה וְאִינוֹ יִדְעֶת לִנְכוֹן נֶמֶל

A scoffer **seeks** [< **has** incessantly/continuously/habitually **sought**] wisdom in vain, but knowledge is easy for one who understands (Prov. 14.6)

Sometimes the *qatal* does not express an uninterrupted situation or habitual activity, but instead introduces a repetitive custom of independent actions, approximating the category of a frequentative perfect. It namely presents a generic characteristic of a class of individuals, stating that they have separately been behaving in such and such a manner; the situation on the whole is inclusive and universal, although it consists of multiple independent acts. Again, this iterativity and “inclusivity” may give rise to habitual and hence gnomic extensions:

(20) a.

דְּבָרֵי גִּרְסָן כְּמַתְלָהִים יִהְיוּ יָרְדוּ חֲדָרֵי-בֶטֶן

The words of a gossip are like choice morsels; they **go** [< **have** repeatedly **gone**] down to a man’s inmost parts (Prov. 26.22 and Prov. 18.8)

b.

שִׁתּוֹ רָעִים לִפְנֵי טוֹבִים וְרָשָׁעִים עַל-שַׁעְרֵי צְדִיק:

Evil men **bow** [< **have** repeatedly **bowed**] in the presence of the good, and the wicked at the gates of the righteous (Prov. 14:19)

c.

נִחְלוּ כְּתֹאִים אֲגָלַת אֲעֻרּוּמִים יִכְתְּרוּ דַעַת

The simple **inherit** [< **have** repeatedly **inherited**] folly, but the prudent are crowned with knowledge (Prov. 14:18)

d.

יִרְאַת יְהוָה רֵאשִׁית דָּעַת חֲכָמָה וּמוֹסָר אֲוִילִים בָּזוּ:

The fear of the Lord is the beginning of knowledge, but fools **despise** [< **have** repeatedly **despised**] wisdom and discipline (Prov. 1:7)

e.

כִּי־רַבִּים חַלְלִים הִפְלִיחַ אֶעֱצֹמִים כָּל־הַרְגִּיקָ:

Many are the victims she **brings** [< **has** repeatedly **brought**] down; her slain are a mighty throng (Prov. 7.26)

f.

עָרוֹם רָאָה רָעָה נִסְתָּר פְּתָאִים עָבְרוּ וְעֹנָשׁוּ:

The prudent **see** [< **have** repeatedly **seen**] danger and **hide** [< **have** repeatedly **hidden**], but the simple **go on** [< **have** repeatedly **kept going**] and suffer [< **have** repeatedly **been punished**] for it (Prov. 27:12)

g.

טָמַן עֵצָל יָדוֹ בַּצִּלְחָת גַּם-אֶל-פִּיהוּ לֹא יָשִׁיבָנָה:

The sluggard **buries** [**< has repeatedly buried**] his hand in the dish; he will not even bring it back to his mouth! (Prov. 19.24 and 26.15)

Occasionally, the *qatal* gives rise to gnomic interpretation not because of the continuity, stability, or habituality of a situation or similar activities, but rather due to the experience of having performed an action. Functioning as an experiential perfect, the inference of the event's current relevance expressed by the gram is sensed as stable and invariant (cf. sections 2.3.2., 2.3.3. and the use of experiential perfects for gnomic purposes). In other words, the event expressed by the *qatal* has already occurred; what can be understood as universal are its currently relevant effects.⁴³ Again, such *qatal* forms typically characterize a class and individual (e.g. a man, a rich one, a poor one, and a wicked one), presenting their characteristic universal properties. It should be noted that this type of gnomic extension (i.e. gnomic values linked to and/or expanded from the experiential perfect) usually occurs with telic verbs that express ideas of finding, encountering, completing, or perishing.

(21) a.

כִּי מֵצֵא אֶת מֵצָא חַיִּים וַיִּפְקֵם רָצוֹן מִיְהוָה:

For whoever finds me **finds** [**< has found**] life, receives favour from the Lord (Prov. 8:35)

b.

אֲשֶׁר־יֵאָדָם מֵצָא חֲכָמָה וְאֵדָם יִפְקֵם תְּבוּנָה:

Blessed is the man who **finds** [**< has found**] wisdom, the man who gains understanding (Prov. 3:13)

c.

מֵצָא אִשָּׁה מֵצָא טוֹב וַיִּפְקֵם רָצוֹן מִיְהוָה:

He who **finds** [**< has found**] a wife **finds** [**< has found**] what is good and receives favour from the Lord (Prov. 18:22)

d.

עֲשִׂיר וָרֶשַׁת נִפְגְּשׁוּ עִשָּׂה כֻּלָּם יְהוָה:

Rich and poor **encounter** [**< have encountered**] this in common: The Lord is the Maker of them all (Prov. 22:2)

e.

עֹצֶה עֵינָיו לַחֲשׁוֹב תְּהַפְּכוֹת קִרְצִי שְׁפָתָיו כָּלֵה רָעָה:

He who winks with his eye is plotting perversity; he who presses his lips (is pinching his lips) **completes** (**brings / is bent on**) [**< has completed**] evil (Prov. 16:30)

⁴³ This means that in this usage, the experiential *qatal* may also be viewed as related to the domain of a resultative perfect if a resultative perfect is extended from the expression of qualities that are currently acquired (and, hence, transitory, or actual) to the expression of characteristics that (although acquired at some point) are now permanent.

In certain instances, the gnomic value stems from the anti-perfect use of the *qatal*. In this function, the gram indicates that a given activity or situation has not taken, or has not been taking place. This non-occurrence is viewed as a rule because it spans from a moment in the past to the present time. In such cases, one may render the universality of the statement by means of the quantifier *never*.

(22) a.

בֶּן חָכָם מוֹסֵר אָבִי לֹא־שָׁמַע גְּעָרָה:

A wise son heeds his father's instruction, but a mocker **does not listen** [**has not been listening / has never listened**] to rebuke (Prov. 13:1 and 13.8)

b.

שְׁתֵּי כְּנוֹת־הָבָה הֵב שְׁלוֹשׁ הֵנָּה לֹא תִשְׂבַּעֲנָה אֲרִפֶּה לֹא־אָמְרוּ הֵן:

There are three things that are never satisfied, four that **never say** [**have never said**]: "Enough!" (Prov. 30:15)

Finally, in a limited number of examples, a gnomic nuance may be explained as a string developed from the value of an iterative experiential perfect; a certain event has occurred on more than one occasion, without however constituting a solid continuous sequence. In this manner, the iterativity of independent actions may be viewed as a rule overtly quantified by the expressions such as *each time when....* It is interesting to note that in this function, the *qatal* is usually accompanied by a *wayyiqtol* form (23.a–c).⁴⁴

(23) a.

בְּאִי־זָדוֹן וַיָּבֹא קָלוֹן וְאֵת־צְנוּעִים חָכְמָה:

[Each time] when pride **comes** [**has come**], then comes [has come] disgrace, but with humility comes wisdom (Prov. 11:2)

b.

צָדִיק מִצָּרָה נִחְלָץ וַיָּבֹא רָשָׁע מִתְחִיָּו:

[Each time] the righteous man **is rescued** [**has been rescued**] from trouble, and it comes [has come] on the wicked instead (Prov. 11:8)

c.

עִיר גְּבֻרִים עָלָה חָכָם וַיִּרֶד עַד מִבְּטָחָהּ:

[Each time] a wise man **attacks** [**has attacked**] the city of the mighty, he pulls down [has pulled down] the stronghold in which they trust (Prov. 21:22)

d.

בְּמוֹת אָדָם רָשָׁע תֵּאבֵד תִּקְוָה וְתוֹחֶלֶת אוֹגִים אָבְדָה:

When a wicked man dies, his hope perishes; and [each time] the hope of power **perishes** [**has perished**] (Prov. 11:7)

⁴⁴ This value is of course closely related to the frequentative perfectal sense.

To conclude this review of the gnomic *qatal* forms in Proverbs, an exemplary fragment (31:11–29) – an “anecdote” that constitutes a small textual unit – can be quoted in which various inclusive, frequentative, and experiential *qatals* trigger gnomic extensions. In this manner, the following passage illustrates that the gnomic use of the *qatal* may have its roots in different perfectal functions. However, in all of such cases, regardless their exact cognitive and conceptual origin, the suffix conjugation describes constant, generic, habitual, typical properties of a good wife:

(24)

בַּטַח בָּהּ לֵב בַּעֲלָהּ וְשֵׁלֵל לֹא יִחָסֵר:

Her husband **trusts** [stative] in her and lacks nothing of value (Prov. 31:11)

וּמַלְתָּהּ טוֹב וְלֹא-רָע כָּל יְמֵי חַיֶּיהָ:

She **brings** [< **has been bringing** (inclusive)] him good, not harm, all the days of her life⁴⁵ (Prov. 31:12)

דֶּרֶשָׁה צֹמֶר וּפְשָׁתִים וְתַעֲשֶׂה בְחִפְזָן כַּפִּיָּה:

She **selects** [< **has been selecting** (frequentative)]⁴⁶ wool and flax and works with eager hands (Prov. 31:13)

זִמְמָה שָׂדֶה וּתְקַחְהוּ מִפְּרִי כַפֶּיהָ גִּטָּע כָּרֶם:

She **considers** [< **has been considering** (frequentative)] a field and buys it; out of her earnings she plants a vineyard (Prov. 31:16)

חֲגִירָה בְּעוֹז מְתַגֵּיחַ וְתַאֲמִיץ זְרַעוֹתֶיהָ:

She **sets** [< **has been setting** (inclusive)] about her work vigorously; her arms are strong for her tasks (Prov. 31:17)

טֹעֲמָה כִּי-טוֹב סַחֲרָה לֹא-יִכָּבֶה בַּלַּיִל נֵרָה:

She **perceives** [< **has perceived** (experiential)] that her trading is profitable, and her lamp does not go out at night (Prov. 31:18)

יָדֶיהָ שְׁלָחָה בְּכִישׁוֹר וְכַפֶּיהָ תִּמְכוּ פָלֶה:

In her hand she **grasps** [< **has been grasping** (frequentative)] the distaff and **holds** [< **has been holding** (inclusive/frequentative)] the spindle with her fingers (Prov. 31:19)

בָּפֶה פָּרְשָׁה לְעַנִּי וְיָדֶיהָ שְׁלָחָה לְאַבְיוֹן

She **opens** [< **has been opening** (frequentative)] her arms to the poor and **extends** [< **has been extending** (frequentative)] her hands to the needy (Prov. 31:20)

⁴⁵ Observe the use of the temporal phrase *all the days of her life*.

⁴⁶ In examples 115–17, the frequentative sense is probably better rendered by the English Present, i.e. *she selects wool, she considers*, and especially *she sets about her work*.

לֹא־תִירָא לְבֵיתָהּ מִשָּׁלֵג כִּי כָל־יִצְיָתָהּ לְגֵשׁ שָׁנִים:

When it snows, she has no fear for her household; for all of them **are clothed** [stative] in scarlet (Prov. 31:21)

מִרְבָּדִים עֲשֵׂתָהּ־לָהּ שֵׁשׁ וְאַרְגָּמָן לְבוּשָׁה:

She **makes** [**< has been making** (frequentative)] coverings for her bed; she is clothed in fine linen and purple (Prov. 31:22)

סָדְיוֹ עֲשֵׂתָהּ וְתַמְכָּר וְחִגּוֹר וְתִגָּה לְכַנְעָנִי:

She **makes** [**< has been making** (frequentative)] linen garments and sells them, and has been **supplying** [frequentative] the merchants with sashes (Prov. 31:24)

בִּיהָ פִּתְחָהּ בְּחָכְמָה וְתוֹרַת־חָסֵד עַל־לְשׁוֹנָהּ:

She **speaks** [**< has been speaking** (frequentative)] with wisdom, and faithful instruction is on her tongue (Prov. 31:26)

קָמוּ בָנֶיהָ וַיְבָרְכוּהָ וַיְשַׁבְּחוּהָ וַיְהִלְלֶהָ:

Her children **rise** [**< have been rising** (frequentative)] and call her blessed; her husband also, and he praises her: (Prov. 31:28)

רַבּוֹת בָּנוֹת עָשׂוּ חָיִל וְאַתְּ עָלִית עֲלֵיהֶן כָּל־כֹּלָה:

“Many women **do** [**< have done** (frequentative / experiential iterative)]⁴⁷ noble things, but you **surpass** [**< have surpassed** (experiential)] them all” (Prov. 31:29)

Our data suggest that certain *qatal* forms of non-static roots found in a gnomic genre (Proverbs) convey values that are compatible with the idea of gnomicity, and thus convey (subjective) universal truths. All such gnomic readings seem to derive from the uses and contexts in which the *qatal* gram is (or could be understood as being) employed with the force of a dynamic perfect such as inclusive perfect, frequentative perfect, experiential perfect, iterative experiential perfect, or anti-perfect. In all of them, just like the underlying perfect sense, the gnomic activity is dynamic and active. It is never de-transitive, in contrast with the gnomic stativity arising from resultative proper constructions. This means that the gnomic value appears as a co-string of certain perfect senses which, as explained, are typologically propitious for gnomic extensions. In that manner, the BH *qatal* adapts to the typological rule detected in Section 2..

⁴⁷ Due to the proximity of the frequentative and iterative experiential perfect, two readings are possible in this example. One stresses the inclusivity of various events (in sum, they form an uninterrupted situation: *women have been doing noble things* or *each woman has been doing noble things*), while the other emphasizes the independence of such events (they do not form an uninterrupted sequence but are viewed as an iteration of separate actions: *each woman has done a noble thing*).

3.2. Cognate formations

The phenomenon whereby the Semitic suffix conjugation is able to express a gnomic value can be detected not only in Biblical Hebrew (as demonstrated by the examples introduced previously) but also in other members of this family of languages. Put differently, cognate formations of the BH *qatal* can denote (subjective) universal truths.

For instance, in Arabic, the *qatala* – a gram that most commonly functions as a present perfect and (perfective) past – can sometimes appear in proverbs and express general or atemporal activities and situations (Wright 1964: 1 and Danecki 1994: 153). It is usually argued that in such cases the construction portrays an event as having often taken place and as still doing so (see, for example, Wright 1964: 1). This signifies that the formation approximates the category of an inclusive or frequentative perfect (cf. 25.a as well as the first verb in the *qatala* in example 25.b). In other instances, a permanent characteristic or a generic activity stems from inferences related to the nuance of current relevance, typical of an experiential perfect (cf. 24.c).⁴⁸ Apart from Hebrew and Arabic, the Semitic suffix conjugation can also be found with a gnomic value in Ugaritic (cf. Tropper 2000: 715), Aramaic (see Rogland 2003 and the references therein), and Ethiopian (Dilmann 1974 [1907]: 553) where cognate constructions of the BH *qatal* and Arabic *qatala* – although typically employed as present perfects and (perfective) past tenses – may additionally introduce habitual activities, rule-like situations, and universal truths.

(25) a.

اتفق المفسرون

Commentators have been agreeing (Wright 1964: 1) (= they agree)

b.

انجز الحر ما وعد

A noble man has been keeping what he has promised (Danecki 1994: 153) (= he always keeps it)

c.

عاش من عرف قدره

He who know his value has survived (Danecki 1994: 153) (= is alive)

The Akkadian language provides additional interesting facts. In Akkadian there is a gram – labelled *parsāku* – that is a cognate formation of the BH *qatal* or Arabic *qatala*. Just like the other varieties of the suffix conjugation, it derives from a Proto-Semitic analytical expression built on the verbal adjective (or resultative participle) and an originally independent personal pronoun (cf. Kienast 2001 and

⁴⁸ The second verb in the *qatala* form in example 24.b may be understood either as an experiential perfect “what he has (once) promised” or as an iterative experiential perfect “what he has promised (on various occasions).”

Lipiński 2001; cf. also Andrason 2011a: 199–205, 287–288).⁴⁹ The Akkadian formation is most typically used as a resultative proper and is defined as a non-advanced portion of the resultative (both anterior and simultaneous) path (cf. Andrason 2011a). Since the *parsāku* frequently functions as a resultative proper it can likewise appear with an extra-temporal gnomic force (see example 26 below as well as the sentence in 3.h which has been quoted in section 2.3.1.; cf. also Mayer 1992 and Loesov 2004: 430–431).

(26) (A 3525:7 as quoted and translated in CAD Š/1 347a (an OB letter))
 kīma ša-mu-u₂ u eršetum **dārû** bēlī lu dāri⁵⁰
 like heaven and earth last lord.my may last
 Like heaven and earth last, may my lord be lasting

It is consequently possible to argue that the gnomic use of the suffix conjugation is a quite regular phenomenon in the Semitic family and affects the formation at all stages of its typological advancement, from the original (cf. the resultative proper stage as documented by the Akkadian gram) to the highly advanced phase (cf. the stage of an old anterior that allows even narrative past uses, as illustrated by the Arabic *qatala*). All such gnomic readings appear in uses that are typologically favourable for gnomic extensions and that correspond to stages on the following stages on the anterior path: resultative proper, inclusive perfect, frequentative perfect, experiential perfect (also iterative experiential perfect), and anti-perfect.

4. Conclusion

4.1. The gnomic sense in the semantic map of the *qatal*

Having analysed the phenomenon whereby certain post-resultative grams tend to prompt gnomic extensions, and having described the nature of the gnomic *qatal* itself, we may now proceed to positing a cognitive (conceptual and diachronic) explanation of the gnomic sense offered by the suffix conjugation and, thus, its accommodation within the entire semantic network of this BH category.

As explained previously, the *qatal* is a formation which historically derives from a resultative proper input. Its most frequent uses, which are typical for dynamic predicates, have been mapped by means of the anterior cline. Values such as present

⁴⁹ The term “cognate” means “genetically” and (from a dynamic perspective) “typologically related”. As explained above, both constructions derive from the same PS input, i.e. a PS **qatVI-* formation built on a formally analogical verbal adjective (i.e. on the **qatVI* employed in a predicative function). Of course, the status of the Hebrew and Akkadian successors (either functional or semantic), and their place in the verbal systems of the two respective languages are clearly different. Nevertheless, the two forms reflect distinct stages of grammaticalization of the same (from a diachronic and typological point of view) construction which underwent a typical evolution in the two languages developing along the anterior and simultaneous path. This means that both constructions cover different portions of the anterior cline and therefore have different semantic and morphosyntactical properties (cf. a typical de-transitivity of the *parsāku*; for a panchronic comparison of the two formations, see Andrason 2011: 199–206, 281–290, 299–300, 305–306).

⁵⁰ The lexeme *lu* is an optative particle. The words *dārû* and *dāri* are two *parsāku* forms employed in the 3ms.pl and 3ms.sg, respectively.

perfect (inclusive, frequentative, resultative, experiential, and indefinite), perfective past, and simple past⁵¹ ideally correspond to this universal developmental scenario (Andrason 2011a).⁵² Given that, on the one hand, inclusive, frequentative, and experiential perfects (as far as inferences are concerned) as well as anti-perfects (which correspond to negative perfect uses) naturally generate a sub-development designated by us as a “gnomic branch”, yielding gnomic uses, and given that, on the other hand, the examples of the BH gnomic *qatal* are regularly found in contexts of inclusive, frequentative, experiential (also iterative experiential perfect), and anti-perfect values, the gnomic sense of the gram can be explained – i.e. linked, both conceptually and diachronically – by employing the “gnomic branch” linkage. In this manner, the chaining, which originates in the aforementioned perfect values, justifies the possibility of using the category with a gnomic force and presents the entire semantic potential of the formation as a consistent whole – that is to say, as a rational and coherent network in which each sense is connected to another, being either its (conceptual and diachronic) ancestor or successor (cf. Figure 2).⁵³

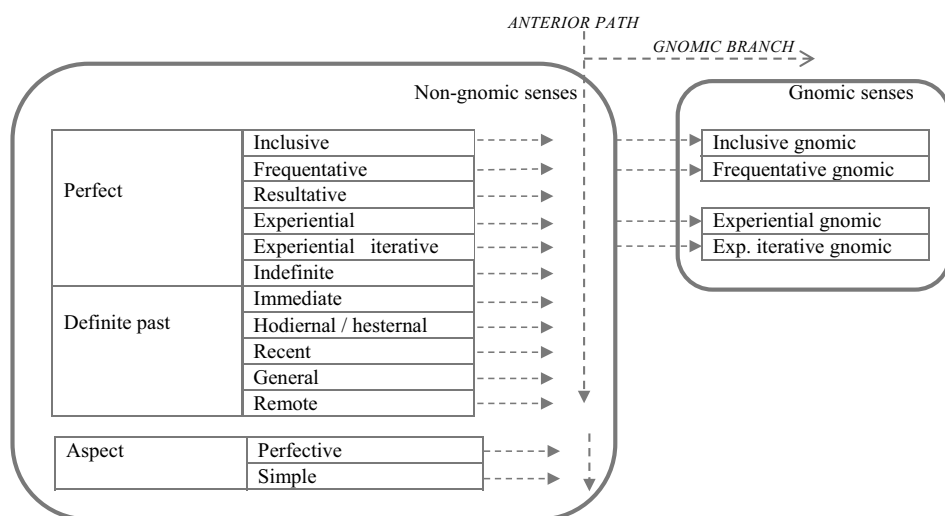


Figure 2. Map of the perfect-perfective-past senses of the BH *qatal* and their gnomic extensions (adapted from Andrason 2011a: 281)⁵⁴

⁵¹ These two latter senses are usually found in discourse, and do not appear in narration.

⁵² As already has been mentioned, various senses provided by static verbs reflect another sub-cline of the resultative path, viz. the simultaneous track.

⁵³ The resultative proper sense is unavailable at the time of Biblical Hebrew. In contrast with the original de-transitive, non-agentive, and patientive construction (still preserved in the Akkadian as the form *parsāku*), the *qatal* – when derived from active roots – is always a dynamic gram, having progressed on the anterior path.

⁵⁴ In this article, the BH *qatal* is treated as if its semantic potential has been “measured” at a single time t_0 . This means that we consider Biblical Hebrew to be a historically static language – a synchronically consistent phenomenon. This is of course an approximation given the fact that the biblical text was composed over the course of several centuries, and thus different books may represent distinct diachronic stages of the language (see, Andrason 2011d: 24, 49; furthermore, it can also include certain dialectal variations). It shall be noted that for the sake of simplicity, the sense of an anti-perfect and its gnomic extension are omitted because they may be understood as negative variants of the other perfectal uses.

The results of our study are fivefold. First, we have detected a typological evolutionary scenario that links the grams evolving along the anterior path with the gnomic value. To be exact, we have determined the precise stages where the extension towards gnomicity can take place. In this manner, a solid typological, conceptual, and diachronic rationale behind the gnomic uses of anterior path grams has been provided. Second, the findings strongly recommend a gnomic interpretation of certain uses of the *qatal* form. In contrast with Rogland (2003), typological and cognitive evidence suggests that the gnomic *qatal* does exist. In this function, the *qatal* offers a sense clearly distinct from other, more prototypical, uses. Simply put, in Biblical Hebrew, the *qatal* may be employed as a vehicle of the idea of gnomicity. Third, the study offers cognitive, typological, and diachronic arguments in favour of the “perfectal link” between the *qatal* and the gnomic sense, and thus argues against the aspectual foundation of the gnomic *qatal* (cf. Joüon (1923: 296–297, Watts 1951: 24–25, Waltke and O’Connor 1990: 488, 506; cf. a similar hypothesis in Rijksbaron 1984:32 concerning the origin of the Greek gnomic Aorist). Fourth, our explanation – although slightly coinciding with theories that relate the gnomic *qatal* to its perfectal uses – refines the connection between the perfect and the gnomic sense. We have namely demonstrated that gnomic extensions originate not only in an experiential perfect (cf. Ewald 1863: 351, Müller 1883: 2, Driver 1892: 17, Gesenius, Kautzsch, and Cowley 1910: 312) but also in frequentative perfect and especially inclusive perfect senses and uses. Fifth, in contrast with all the explanations formulated thus far, our model avoids deriving the gnomic force of the *qatal* from an invariant total meaning or semantic string. There is no need to posit any invariant meaning, sense, or value since our map incorporates all possible semantic diversity of the BH form, while at the same time presenting it as a harmonious and logical phenomenon: a network fastened by means of determined universal evolutionary clines. As a result, the profound polysemous nature of the *qatal*, including its use with a gnomic sense, is both preserved and holistically represented in a cognitively plausible map.

4.2. *Gnomic qatal and its translation*

Once the relation of the gnomic sense to the remaining – and dominant – portion of the semantic potential of the *qatal* form has been established, a final question arises: how should the gnomic *qatal* be translated? In Section 2., we pointed out that a given semantic domain (e.g. futurity, inclusive perfect, perfective past, etc.), or a given specific atomic value, may be conveyed by various grammatical constructions.⁵⁵

⁵⁵ For example, futurity may be expressed by simple present tenses, progressive presents, various modal constructions, or prototypical moods, as well as present perfects or even grams that are typically used as perfective pasts (cf. the *passé composé* in French, a broad present perfect and past tense that can also express certain values related to the idea of futurity). The sense of an inclusive perfect can be conveyed by present perfects, imperfective past tenses, and simple and progressive present tenses. The value of a perfective past can be expressed by perfective past tenses, simple past tenses (preterites), and even futures (e.g. the Polish Future Tense, which under special circumstances may introduce perfective past events in narration).

Consequently, semantic maps of such formations partially overlap; the semantic potentials of these constructions cover and/or are compatible with domains and senses that are similar to some extent.

As far as the idea of gnomicity is concerned, it is most typically expressed across languages by gnomic imperfectives or broad present tenses, although prototypical resultative proper grams and present perfects are also extensively used for gnomic purposes. Consequently, virtually any gram which is able to denote gnomic activities or situations can be employed in a translation in a target language, be it a gnomic imperfective, a broad present perfect, a resultative proper (although these constructions are typically intransitive and/or de-transitive), a present perfect (both young and old anterior), or even a highly advanced past tense that has preserved certain perfectal uses in proverbial genres. All of these grams may be used because all of them are compatible – albeit to distinct degrees – with a semantic domain of gnomicity.

However, since our data indicate that the instances where the *qatal* is employed with a gnomic force in the Hebrew Bible correspond to gnomic extensions of certain present perfect uses, we could argue as follows. If a target language possesses in its verbal inventory a post-resultative construction which has developed along the anterior path, acquiring various perfect stages that allow gnomic extensions, such a young or old anterior gram should preferably be used. This stems not only from a typological similarity with the gnomic *qatal* that has arisen during the evolution of the gram along the anterior path (cf. inclusive, frequentative, experiential, experiential iterative perfects, and anti-perfect), but also from the fact that grammatical formations regularly “colour” a given atomic semantic sense (a piece of information they convey in a specific place and time) with the remaining portion of their semantic potential. Thus, although present tense grams (i.e. grammatical formations developing along the imperfective cline) and perfects or past tenses (i.e. formations developing along the anterior cline) are both able to convey the idea of gnomicity, they do so in a distinct manner. First, their predispositions to express a gnomic value are distinct. For example, it is typical and unrestricted for gnomic imperfectives (i.e. highly advanced imperfective diachronies) while it is exceptional for past tenses (i.e. for advanced resultative diachronies). Second, grams evolving along the two clines have completely different semantic baggage associated with them. These two classes of grams possess clearly distinct scopes of polysemies; the former group reflects stages of the imperfective cline while the latter mirrors the stages of the anterior cline. Hence, they will cause two different sorts of associations when employed to express the gnomic sense in a specific time and place.

A gram can figuratively be understood as a piece of multicolour gum. At an exact historical point, and in a concrete context, only one of the possible senses is activated as being compatible with a certain semantic domain – a message to be conveyed. Metaphorically speaking, it is removed from the multicolour sphere of possible senses so that it can harmonize with the precise semantic domain shaped by the context. Nevertheless, as we select this exact sense from the gram’s polysemy in order to render a specific piece of information, and, metaphorically, tug a fragment of the gum, other values-colours that the gram potentially possesses necessarily follow;

they are conceptually tied to the selected fragment which is being pulled, because all of them, jointly, constitute the gram's total meaning. Consequently, senses that reflect consecutive diachronic stages or conceptual extensions may be understood both as atomic, individual, and sharply separated – they correspond to distinct contexts or abstract semantic domains – and, at the same time, as intrinsically connected components which cannot be separated from the remaining semantic load, i.e. from the gram viewed as a dynamic evolutionary whole and, thus, a portion of a path. This internal connection of all senses conveyed by a gram – the polysemous network of a concrete grammatical construction – will cause the gnomic sense expressed by imperfective-path grams and anterior-path grams to be slightly different. Given that the gnomic *qatal* (when used with dynamic roots) is an anterior-path-type gnomic formation, one could recommend translations based on anterior-path grams.

On the other hand, however, it must clearly be stated that, due to the infinite complexity of real-world systems such as languages, no two grams (i.e. their semantic potentials) in two distinct languages can perfectly match each other. Quite the opposite, the polysemies of grams which are typologically highly similar always somehow differ. As a result, even though a language, for instance English, possesses in its grammatical inventory a present perfect formation that is able to express continuous, habitual, repetitive, durative activities, thus giving rise to gnomic extensions and uses, it is not mandatory that all gnomic examples of the BH *qatal* should be rendered in translations by means of the Present Perfect. This means that the English Present Perfects (*has done* or *has been doing*) – despite their typological and semantic proximity, and despite our definition of the BH construction as a resultative-path gram – cannot be regarded as the sole and exclusive possibility in rendering the gnomic *qatal*. In some cases, the use of the English Present Perfect is fully correct, but in others it seems to be awkward. This stems from the already explained fact that although a gram is compatible with a given semantic (abstract) domain, it also carries additional information (the remaining portion of its semantic potential) that impregnates the value provided in a specific context. This “impregnation” causes certain gnomic examples to tolerate the English Present Perfect with a distinct intensity or in a more appropriate or less appropriate manner. The dissimilarity of the English and Hebrew verbal systems is evident. Most importantly, while Biblical Hebrew possesses five or six central grams (*qatal*, long *yiqtol*, *wayyiqtol*, short *yiqtol*, *weqatal*, and possibly *qotel*) that must cover all the semantic domains cognitively available to humans, the English verbal organization is compounded of a much greater number of constructions. As a result, the polysemy of English grams is by definition less than that of BH formations. Inversely, verbal grams in English are more specialized. In this manner, although the English Present Perfect can be employed to convey gnomic nuances, it is not a prototypical vehicle of gnomicity. The exemplary means to express gnomic ideas are the Simple Present and Resultative Proper (*be* + participle). The semantic potential of the English Present Perfect formation is much more reduced and specialized (being typically used in perfectal functions) than the polysemy of the BH *qatal*. These facts argue in favour of translations based on grams that are more typical vehicles of the concepts of gnomicity

such as the aforementioned Simple Present tense or the Resultative Proper form (in the case of intransitive, de-transitive, or stative constructions).

References

- Aikhenvald, A. 2004. *Evidentiality*. Oxford: Oxford University Press.
- Andersen, D. 2000. The Evolution of the Hebrew Verbal System. *Zeitschrift für Althebraistik* 13/1, 1–66.
- Andrason, A. 2010a. The panchronic *yiqtol*. Functionally consistent and cognitively plausible. *Journal of Hebrew Scriptures* 10/10, 1–63.
- Andrason, A. 2010b. The “guessing” *QATAL* – the BH suffix conjugation as a manifestation of the evidential trajectory. *Journal for Semitics* 19/2, 603–627.
- Andrason, A. 2010c. The Akkadian *Iprus* from the unidirectional perspective. *Journal of Semitic Studies* 55/2, 325–345.
- Andrason, A. 2011a. *Qatal, yiqtol, weqatal y wayyiqtol. Modelo pancrónico del sistema verbal de la lengua hebrea bíblica con el análisis adicional de los sistemas verbales de las lenguas acadia y árabe*. Madrid: Publicaciones de Universidad Complutense.
- Andrason, A. 2011b. The Biblical Hebrew *Wayyiqtol* – a dynamic definition. *Journal of Hebrew Scriptures* 11/7, 1–50.
- Andrason, A. 2011c. The Biblical Hebrew verbal system in light of grammaticalization – the second generation. *Hebrew Studies* 52, 19–53.
- Andrason, A. 2011d. The BH *weqatal*. A homogenous form with no haphazard functions. Part 1. *Journal of Northwest Semitic Languages* 37/2, 1–25.
- Andrason, A. 2012a. Making it sound – Performative *qatal* and its explanation. *Journal of Hebrew Scriptures* 12/8, 1–58.
- Andrason, A. 2012b. Grammaticalization paths and non-equilibrium: Thermodynamics of the Biblical Hebrew verbal system. D. Eades (ed.), *Grammaticalisation in Semitic. Journal of Semitic Studies Supplement Series*, 145–163. Oxford: Oxford University Press.
- Andrason, A. 2013a. Against floccinaucinihilipilification of the counterfactual sense of the BH suffix conjugation – or an explanation of why the “indicative” *qatal* expresses conditions, hypotheses and wishes. *Old Testament Essays* 26/1, 20–56.
- Andrason, A. 2013b. Future values of the *qatal* are “logical” – How to chain future senses of the *qatal* to the core of its semantic network? *Hebrew Studies* 54, 261–292.
- Andrason, A. 2013c. An optative indicative? A real factual past? Towards a cognitive-typological approach to the precative *qatal*. *Journal of Hebrew Scriptures* 13/4, 1–41.
- Andrason, A. Forthcoming (a). From resultatives to present tenses – simultaneous path of resultative constructions. *Italian Journal of Linguistics*.
- Auyang, S. 1998. *Foundations of Complex-system Theories*. Cambridge: Cambridge University Press.
- Bertinetto P. M., and Lenci, A. 2010. Iterativity vs. habituality (and gnomic imperfectivity). *Quaderni del Laboratorio di Linguistica* 9/1, 1–46.
- Boneh, N., and Doron, E. 2008. Habituality and the Habitual Aspect. S. Rothstein (ed.), *Theoretical and Crosslinguistic Approaches to the Semantics of Aspect*, 321–348. Amsterdam: John Benjamins.
- Boneh, N., and Doron, E. 2010. Modal and Temporal Aspects of Habituality. M. Rappaport Hovav, E. Doron and I. Sichel (eds), *Lexical Semantics, Syntax, and Event Structure*, 338–363. Oxford: Oxford University Press.
- Brockelmann, C. 1956. *Hebräische Syntax*. Neukirchen Kreis Moers: Verlag der Buchhandlung des Erziehungsvereins.
- Bybee, J. 2010. *Language, Usage and Cognition*. Cambridge: Cambridge University Press.
- Bybee, J., Perkins R., and Pagliuca, W. 1994. *The Evolution of Grammar*. Chicago: The University of Chicago Press.
- Carlson, G. 1977. *Reference to Kinds in English*. Ph.D. dissertation, University of Massachusetts, Amherst.
- Carlson, G., and Pelletier, F. J. (eds). 1995. *The Generic Book*. Chicago: Chicago University Press.

- Chantraine, P. 1953. *Grammaire Homérique. Vol. II*. Paris: Klincksieck.
- Chierchia, G. 1995. Individual-level predicates as inherent generics. G. Carlson and F. J. Pelletier (eds), *The Generic Book*, 176–223. Chicago: Chicago University Press.
- Cook, J. 2002. *The Biblical Hebrew Verbal System*. Ph.D. dissertation, University of Wisconsin-Madison.
- Cook, J. 2005. Genericity, Tense, and Verbal Patterns in the Sentence Literature of Proverbs. R. L. Troxel, K. G. Friebe and D. R. Magary (eds), *Seeking Out the Wisdom of the Ancients: Essays Offered to Honor Michael V. Fox on the Occasion of his Sixty-Fifth Birthday*, 117–133. Winona Lake: Eisenbrauns.
- Crespo, E., Conti, L., and Maquieira, H. 2003. *Sintaxis del Griego Clásico*. Madrid: Gredos.
- Cruse, A. 2004. *Meaning in Language. An Introduction to Semantics and Pragmatics*. Oxford: Oxford University Press.
- Dahl, Ö. 1985. *Tense and Aspect Systems*. Oxford: Blackwell.
- Dahl, Ö. 2000a. The Tense and Aspect Systems of European Languages in a Typological Perspective. Ö. Dahl (ed.), *Tense and Aspect in the Languages of Europe*, 3–25. Berlin/New York: Mouton de Gruyter.
- Dahl, Ö. (ed.). 2000b. *Tense and Aspect in the Languages of Europe*. Berlin/New York: Mouton de Gruyter.
- Danecki, J. 1994. *Gramatyka języka arabskiego*. Warszawa: Wydawnictwo Akademickie Dialog.
- Davidson, A. B. 1902. *Hebrew Syntax*. Edinburgh: T. & T. Clark.
- Dillmann, A. 1907 (reprint 1974). *Ethiopic Grammar (Enlarged and Improved by C. Bezold, translated by J.A. Crichton)*. London: Williams & Norgate.
- Drinka, B. 1998. The evolution of grammar. Evidence from Indo-European Perfects. M. Schmid, J. Austin and D. Stein (eds), *Historical Linguistics 1997*, 117–133. Amsterdam/ Philadelphia: John Benjamins.
- Driver, S. R. 1892. *A treatise on the use of the tenses in Hebrew and some other syntactical problems*. Oxford: Clarendon.
- Evans, V., and Green, M. 2006. *Cognitive linguistics: an introduction*. Edinburgh: Edinburgh University Press.
- Ewald, H. 1863. *Ausführliches Lehrbuch der Hebräischen Sprache*. Göttingen: Dieterich.
- Friedmann, V. 1999. Proverbial Evidentiality: On the Gnomical Uses of the Category of Status in Languages of the Balkans and the Caucasus. *Mediterranean Language Review* 2, 135–155.
- Frykenberg, R. E. 1996. *History and Belief: The Foundations of Historical Understanding*. Grand Rapids: Eerdmans with the Institute for Advanced Christian Studies.
- Gesenius, W., Kautsch, E., and Cowley, A. E. 1910. *Gesenius' Hebrew Grammar*. Oxford: Clarendon Press.
- Gibbs, R. 1994. *Poetics of Mind. Figurative Thought, Language and Understanding*. Cambridge: Cambridge University Press.
- Gildersleeve, B. L. 1900. *Syntax of Classical Greek from Homer to Demosthenes*. New York/Cincinnati: American Book Company.
- Grether, O. 1967. *Hebräische Grammatik für den akademischen Unterricht*. München: Claudius Verlag.
- Grevisse, M. 1975. *Le bon usage*. 10^e ed. Gembloux: Duculot.
- Grevisse, M. 1986. *Le bon usage*. 12^e ed. Gembloux: Duculot.
- Grygiel, M. 2008. *From Semantic Change to Conceptual Blending*. Rzeszów: Wydawnictwo Uniwersyte-u Rzeszowskiego.
- Heine, B., Claudi, U., and Hünnemeyer, F. 1991. *Grammaticalization. A Conceptual Framework*. Chicago: University of Chicago Press.
- Hellenthal, A-C. 2007. Modality properties of sentence type markers in Sheko. P. Hendriks, F. Rau, K. Soucková and J. van der Wal (eds), *Leiden Papers in Linguistics* 4.2, 17–32. Leiden: Leiden University Press.
- Hewson, J., and Bubenik, V. 1997. *The Tense and Aspect in Indo-European Languages: Theory, Typology, Diachrony*. Amsterdam: John Benjamins.
- Hopper, P., and Traugott, E. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.
- Huehnergard, J. 2005. *A Grammar of Akkadian*. Winona Lake: Eisenbrauns.
- Humbert, J. 1954. *Syntaxe Grecque*. Paris: Klincksieck.

- Ibarretxe-Antuñano, I. 1999. Metaphorical mapping in the sense of smell. W. Gibbs and G. Steen (eds), *Metaphor in Cognitive Linguistics*, 29–45. Amsterdam: John Benjamins.
- Janssen, T. 2003. Monosemy versus polysemy. H. Cuyckens, R. Dirven and J. Taylor (eds), *Cognitive Approaches to Lexical Semantics*, 93–122. Berlin: Mouton de Gruyter.
- Jaxontov, S. 1988. Resultatives: Classification and Calculus. V. Nedjalkov (ed.), *Typology of resultative constructions*, 101–106. Amsterdam/ Philadelphia: John Benjamins.
- Jenni, E. 1978. *Lehrbuch der hebräischen Sprache des Alten Testaments*. Basel, Stuttgart: Helbing & Lichtenhahn.
- Johanson, L. 2000. Viewpoint operators in European languages. Ö. Dahl (ed.), *Tense and Aspect in the Languages of Europe*, 27–187. Berlin/New York: Mouton de Gruyter.
- Johanson, L. 2003. Evidentiality in Turkic. A. Aikhenvald and R. M. W. Dixon (eds), *Studies in evidentiality*, 273–291. Amsterdam: John Benjamins.
- Jónsson, J. G. 1992. The Two Perfects of Icelandic. *Íslenskt mál* 14, 129–145.
- Joüon, P. 1923. *Grammaire de l'hébreu biblique*. Roma: Institute Biblique Pontifical.
- Joüon, P., and Muraoka, T. 2009. *A Grammar of Biblical Hebrew*. Roma: Gregorian & Biblical Press.
- Kienast, B. 2001. *Historische Semitische Sprachwissenschaft*. Wiesbaden: Harrassowitz Verlag.
- Kratzer, A. 1989. Stage-level and individual-level predicates. E. Bach, A. Kratzer and B. H. Partee (eds), *Papers on Quantification*, 1–73. Amherst: University of Massachusetts.
- Krifka, M. et al. 1995. Genericity: An Introduction. G. N. Carlson and F. J. Pelletier (eds), *The Generic Book*, 1–124. Chicago: University of Chicago Press.
- Lakoff, G. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago: University of Chicago Press.
- Lamdin, T. O. 1971. *Introduction to Biblical Hebrew*. New York: Scribner.
- Lewandowska-Tomaszczyk, B. 2007. Polysemy, Prototypes, and Radical Categories. D. Geeraerts and H. Cuyckens (eds), *Handbook of Cognitive Linguistics*, 139–169. Oxford: Oxford University Press.
- Lindstedt, J. 2000. The perfect – aspectual, temporal and evidential. Ö. Dahl (ed.), *Tense and aspect in the languages of Europe*, 365–383. Berlin/New York: Mouton de Gruyter.
- Lipiński, E. 2001. *Semitic Languages Outline of a Comparative Grammar*. Leuven/Paris/ Sterling: Uitgeverij Peeters en Departement Oosterse Studies.
- Loesov, S. 2004. Review of K. A. Metzler. *Tempora in altbabylonischen literarischen Texten*. L. Kogan (ed.), *Babel and Bibel 2: Memoriae Igor M. Diakonof*, 393–434. Winona Lake, Indiana: Eisenbrauns.
- Łozowski, P. 2000. *Vagueness in Language: From Truth-Conditional Synonymy to Unconditional Polysemy*. Lublin: UMCS.
- Maslov, J. 1988. Resultative, Perfect and Aspect. V. Nedjalkov (ed.), *Typology of resultative constructions*, 63–85. Amsterdam/Philadelphia: John Benjamins.
- Mayer, W. R. 1992. Das "gnomische Präteritum" im literarischen Akkadisch. *Orientalia* 61/4, 373–399.
- McCawley, J. 1971. Tense and Time Reference in English. C. Fillmore and D. T. Langendoen (eds), *Studies in Linguistics and Semantics*, 96–111. New York: Holt, Rinehart and Winston.
- Metzler, K. A. 2002. *Tempora in altbabylonischen literarischen Texten*. Münster: Ugarit-Verlag.
- Müller, A. 1883. *Outline of Hebrew Syntax*. Glasgow: James Maclehose & Sons.
- Nedjalkov, V. 2001. Resultative constructions. M. Haspelmath et al. (eds), *Language Typology and Language Universals: An International Handbook. Vol. 2*, 929–940. Berlin: Walter de Gruyter.
- Nichols, J., and Timberlake, A. 1991. Grammaticalization as retextualization. E. Traugott and B. Heine (eds), *Approaches to Grammaticalization. I Vol.*, 129–146. Amsterdam/ Philadelphia: John Benjamins.
- Nikiforidou, K. 2009. Constructional analysis. F. Brisard, J-O. Östman and J. Verschueren (eds), *Grammar, Meaning and Pragmatics*, 16–32. Amsterdam/ Philadelphia: John Benjamins.
- Norrick, N. 1985. *How Proverbs mean: Semantic Studies in English Proverbs*. The Hague: Mouton.
- Palmer, R. F. 2001. *Mood and Modality*. Cambridge: Cambridge University Press.
- Prigogine, I. 2009. *¿Tan solo una ilusión?* Barcelona: Tusquets.
- Rijksbaron, A. 1984. *The Syntax and Semantics of the Verb in Classical Greek*. Amsterdam: Gieben.
- Rydberg-Cox, J. A. 2000. *Overview of Greek Syntax*. www.perseus.tufts.edu.

- Rogland, M. 2003. *Alleged Non-past Uses of Qatal in Classical Hebrew*. Studia Semitica Neerlandica. Assen: Van Gorcum.
- Ruijgh, C. J. 1971. *Autour de te-épieque*. Amsterdam: A.M. Hakkert.
- Sadler, M. 2007. *Grammar in Use across Time and Space: deconstructing the Japanese "dative subject" Construction*. Amsterdam: John Benjamins.
- Sasson, V. 2001. The Waw Consecutive/Waw Contrastive and the Perfect. *Zeitschrift für die Alttestamentliche Wissenschaft* 113, 602–617.
- Schneider, E., and Sagan, D. 2009. *La termodinámica de la vida*. Barcelona: Tusquets.
- Sil'nicki, G. 1988. Structure of verbal meaning and the resultative. V. Nedjalkov (ed.), *Typology of resultative constructions*, 87–100. Amsterdam/ Philadelphia: John Benjamins.
- Goodwin W. W. 1893. *Syntax of the Moods and Tenses of the Greek Verb*. Boston: Ginn & Company.
- Smith, P. 1998. *Explaining Chaos*. Cambridge: Cambridge University Press.
- Smyth, H. W. 1956. *A Greek Grammar for Colleges* (Revised edition by Gordon Messing). Cambridge: Harvard University Press.
- Spasov, L., Z. Topolińska, and Spasov, A. 1986. *Opis na klasifikacija na pogovorkite kak avtonomen semiotičen jazičen mikrosistem*. Skopje: MANU.
- Squartini, M., and Bertinetto, P. M. 2000. The simple and compound past in Romance languages. Ö. Dahl (ed.), *Tense and Aspect in the Languages of Europe*, 403–440. Berlin/New York: Mouton de Gruyter.
- Taylor, J. 2002. *Cognitive Grammar*. Oxford: Oxford University Press.
- Thompson, J. 1902. *Greek Grammar. Accidence and Syntax*. London: John Murray.
- Tropper, J. 2000. *Ugaritisch*. Münster: Ugarit-Verlag.
- Tuggy, D. 2003. The Nawatl verb *kīsa*: A case study in polysemy. H. Cuyckens, R. Dirven and J. Taylor (eds), *Cognitive Approaches to Lexical Semantics*, 323–362. Berlin: Mouton de Gruyter.
- Van der Auwera, J., and Gast, V. 2011. Categories and prototypes. J-J. Song (ed.), *The Oxford Handbook of Typology*, 166–189. Oxford: Blackwell.
- Van der Merwe, C. H. J., Naudé, J., and Kroeze, J. 1999. *A Biblical Hebrew Reference Grammar*. Sheffield: Sheffield Academic Press.
- Wagensberg, J. 2007. *Ideas sobre la complejidad del mundo*. Barcelona: Tusquets.
- Watts, J. W. 1951. *A Survey of Syntax in Hebrew Old Testament*. Grand Rapids: Eerdmans.
- Waltke, B. K., and O'Connor, M. 1990. *An Introduction to Biblical Hebrew Syntax*. Winona Lake, Indiana: Eisenbrauns.
- Wright, W. 1964. *A Grammar of the Arabic Language. Vol.2*. Cambridge: Cambridge University Press.

Stress-induced Vowel Lengthening and Harmonization in Kumzari

Erik J. Anonby, Carleton University, Ottawa, and Uppsala University
erik_anonby@carleton.ca

Abstract

This paper presents the morphophonological effects of four suffixes on noun stem vowels in Kumzari: the simple plural suffix *-an*, the existential plural clitic *-in*, the definite suffix *-ō*, and the indefinite suffix *-ē*. Two of the suffixes (*-an* and *-ō*) have an effect on the stress placement of their host stem. This results in an array of alternations which may be generalized as lengthening of the stem's final vowel and, in the case of *-ō*, harmonization of the stem's final vowel to that of the suffix. The remaining two suffixes (the clitic *-in* and the suffix *-ē*) are included in this study to emphasize the role of stress placement in vowel alternations in Kumzari: although *-in* and *-ē* are segmentally and semantically analogous to the first two suffixes – and, in the case of *-ē*, morphosyntactically equivalent – they have no effect on stress placement and consequently do not precipitate vowel lengthening or harmonization.

Keywords: Kumzari, languages of Arabia, morphophonological processes, vowel harmony, stress-length interaction

1. Introduction¹

Kumzari is an endangered language spoken by about 5000 people in Oman, the United Arab Emirates and Iran. Speakers of the main dialect are found on the Musandam Peninsula of Oman and in small groups in cities along the Gulf coast of the United Arab Emirates. Laraki, a closely related dialect of the language, is spoken across the Strait of Hormuz by a single community on Larak Island in Iran (Anonby & Yousefian 2011, Lewis 2011).

The Kumzari language (endonym: *kumzārī*²) was identified by Jayakar (1902), and a brief grammar sketch and lexicon appeared in Thomas (1930). Although it is often referred to as a mixed language, its core vocabulary and verbal morphology support Skjærvø's (1989) classification of Kumzari within the Southwestern group

¹ An earlier version of this paper was presented at the 1st International Conference on Languages and Dialects in Iran, University of Sistan and Baluchestan, October 28–31, 2008. The present version was made possible through research-related functions of the author's position at Carleton University and the Uppsala University research project UFV-PA 2010/2580.

I would like to thank the many people who have made this research possible. I would first of all like to thank Christina van der Wal Anonby for encouragement and support during the writing of this article. In Oman, I have appreciated the oversight and input of Dr Nafla al-Kharousi and Dr Amel Salman at Sultan Qaboos University. Finally, I am grateful to all those among the Kumzari who have shown us kindness, in particular the local authorities and several people who have shared many insights about their own language. For this article, the contributions of Noufal Mohammad Ahmed al-Kumzari, Malallah Sulaiman Muhammad al-Kumzari and Ali Hassan Ali al-Kumzari have been particularly valuable:

تاتم شکره قادر کمطاریه تکم، ایشنه آر قبیلی بورن وجوانن بمناء، املو مه چدن جوان. حس چدم این ائده خانغونیم. شکره شه تکم خیکی. سادو مه چدن و وقفه بورن واما و شجه مه چدن. و حبه تکم شکره قادر آدمی تکم، خاصت ریشدن و آدمی آر تیسکن دچینه، حافظو جسین بتقالید خو اولین. و حبه تکم شکره ایشنه آر سادو مه چدن ائده بحث مه بنا، خاصتن نوفل پس حمد و مانله پس سلیمین و ایل حس ایلکو. قصرو چمی نه! بارک الله فیکم!

of Iranian languages. A large proportion of the overall vocabulary, however, as well as many linguistic structures, may be traced to influence from Arabic, including the neighbouring Shihhi dialect of Arabic (cf. Bayshak 2002, Anonby 2011). Still, Kumzari is highly differentiated from its Southwestern relatives and neighbouring Arabic varieties. In particular, Kumzari morphophonology is distinctive and complex.

One process characteristic of Kumzari is pervasive vowel alternation within stems, which is motivated by certain suffixes. This process, referred to as “umlaut” by Skjærvø (1989: 365), is observable from the data in Thomas (1930). However, it is not documented there in a systematic way, and for this reason the regularity and extent of this phenomenon, as well as its motivation, are difficult to determine.

In the present paper, I re-examine this phenomenon by introducing the morphophonological effects of four suffixes on root vowels in Kumzari:

1. the simple plural suffix *-an*
2. the existential plural clitic *-in*
3. the definite suffix *-ō*
4. the indefinite suffix *-ē*

Two of these suffixes (*-an* and *-ō*) have an effect on the stress placement of their host stem. This results in an array of alternations which may be generalized as: a) lengthening of the stem’s final vowel and, in some cases, b) harmonization of the stem’s final vowel to that of the suffix.

The remaining two suffixes (the clitic *-in* and the suffix *-ē*) are included in this description to emphasize the role of stress placement in vowel alternations in Kumzari. Although *-in* and *-ē* are segmentally and semantically analogous to the first two suffixes – and, in the case of *-ē*, morphosyntactically equivalent – they have no effect on stress placement and consequently do not precipitate vowel lengthening or harmonization.

The description given here is based on an analysis of 4500 lexical items as well as a number of longer texts collected from various Kumzari speakers (al-Kumzari (ms.), Anonby & van der Wal Anonby, in preparation). Because of the complexity of the alternations, this paper introduces the behaviour of these affixes with a common but very specific subset of noun shapes, namely two-syllable nouns with a CVC.CVC shape (e.g., *gambil* ‘hole’, etc.). While a historical and comparative analysis of affixation and stress patterns in related and neighbouring languages is important in its own right, it is beyond the scope of this paper.

² In the phonological orthography used in this article, the following symbols require explanation:

- a macron over a vowel indicates vowel length (e.g. *ā*, *ē*, *ī*, *ō*, *ū*);
- a dot below a consonant indicates emphasis (velaro-pharyngealization) (e.g. *ṣ*, *ẓ*);
- *q* represents a voiceless uvular stop;
- *ḡ* represents a voiced uvular fricative;
- *ʔ* represents the glottal stop, which is contrastive in Kumzari; and
- an apostrophe (‘) placed before the relevant syllable marks stress.

2. Patterns of stress-induced lengthening and harmonization in four suffixes

2.1. *The simple plural suffix -an*

The suffix *-an* (Modern Standard Persian = MSP reflex: *-ān*) is the basic plural marker with nouns in Kumzari.

sg.	pl.	
<i>dēw</i>	<i>dēwan</i>	‘demon(s)’
<i>gōz</i>	<i>gōzan</i>	‘walnut(s)’
<i>sūr</i>	<i>sūran</i>	‘wedding(s)’

In the data above, it is evident that pluralization can be achieved through the simple suffixation of *-an* to the noun stem. This is true of one-syllable nouns with a long vowel. However, the vast majority of nouns in the language have more than one syllable and contain a short vowel in their final syllable; for these nouns, the morphophonology is more complex. Consider pluralization of the following two-syllable nouns with a CVC.CVC shape, where the last vowel of the stem is lengthened:

sg.	pl.	
<i>birmī</i>	<i>birmītan</i>	‘hard candy(ies)’
<i>gambīl</i>	<i>gambīlan</i>	‘hole(s)’
<i>rištaḡ</i>	<i>rištāḡan</i>	‘thread(s)’
<i>ʔaqrab</i>	<i>ʔaqrāban</i>	‘scorpions(s)’
<i>barnuṣ</i>	<i>barnūṣan</i>	‘blanket(s)’
<i>xarkuk</i>	<i>xarkūkan</i>	‘parrotfish (sg./pl.)’

2.2. *The existential plural clitic -in*

The existential plural clitic *-in* (MSP reflex: *-and*) is segmentally and semantically analogous to the noun plural suffix *-an*. In fact, the two are confounded in Thomas (1930: 788–790). However, their syntactic distribution as well as their phonetic realization are contrastive, as is their phonological behaviour: in particular, *-in* does not cause the final vowel of the stem to be lengthened.

sg.	pl.	
<i>birmī</i>	<i>birmītin</i>	‘hard candy / they are hard candies’
<i>gambīl</i>	<i>gambīlin</i>	‘hole / they are holes’
<i>rištaḡ</i>	<i>rištaḡin</i>	‘thread / they are threads’
<i>ʔaqrab</i>	<i>ʔaqrabīn</i>	‘scorpion / they are scorpions’
<i>barnuṣ</i>	<i>barnuṣīn</i>	‘blanket / they are blankets’
<i>xarkuk</i>	<i>xarkukīn</i>	‘parrotfish / they are parrotfish’

The key to differences in the changes caused by these two suffixes is in their effect on the stress of the stem to which they are attached.

Recall that in MSP, stress is typically on the last syllable of a word, especially for nominals (Windfuhr and Perry 2009: 429–430). When suffixes are added in MSP, they may be classified as stressed (= tonic) or unstressed, depending on whether or not they carry word stress (among others, see Lambton 1953: 96). In Kumzari, how-

ever (similar to a common pattern in Arabic), stress is by default placed on a word's penultimate syllable: *'birmit* 'hard candy', *'gambil* 'hole'.

Word-final monosyllabic suffixes, then, are never stressed in Kumzari. However, in some cases they cause word stress to shift toward the end of a word: while the simple plural suffix *-an*, introduced above, causes stress to shift to the second-last syllable of the resulting word, the existential plural clitic *-in* does not. (The data is repeated from above, with stress also being represented.)

	sg.	simple pl.	existential pl.
'hard candy'	<i>'birmit</i>	<i>bir'mītan</i>	<i>'birmitin</i>
'hole'	<i>'gambil</i>	<i>gam'bīlan</i>	<i>'gambilin</i>
'thread'	<i>'rištaḡ</i>	<i>riš'tāḡan</i>	<i>'rištaḡin</i>
'scorpion'	<i>'ʔaqrab</i>	<i>ʔaq'rāban</i>	<i>'ʔaqrabin</i>
'blanket'	<i>'barnuṣ</i>	<i>bar'nūṣan</i>	<i>'barnuṣin</i>
'parrotfish'	<i>'xarkuk</i>	<i>xar'kūkan</i>	<i>'xarkukin</i>

Since *-an* is a typical noun suffix and *-in* is a clitic, the question remains as to whether the difference in behaviour is morphosyntactically driven. However, a comparison between the definiteness-related noun suffixes *-e* and *-o* below shows that this is not the case.

2.3. The definite suffix *-ō*

The definite suffix *-ō* (no clear MSP reflex; possibly related to the definite direct object marker *-(r)ā / -(r)o³*), like the plural suffix *-an*, causes the last vowel of the stem to be lengthened. In addition, as can be seen from the second set of words below, the low vowel *a* harmonizes with the suffix, becoming *ō*.

unmarked	definite	
<i>birmit</i>	<i>birmitō</i>	'hard candy / the hard candy'
<i>gambil</i>	<i>gambīlō</i>	'hole / the hole'
<i>rištaḡ</i>	<i>rištōḡō</i>	'thread / the thread'
<i>ʔaqrab</i>	<i>ʔaqrōbō</i>	'scorpion / the scorpion'
<i>barnuṣ</i>	<i>barnūṣō</i>	'blanket / the blanket'
<i>xarkuk</i>	<i>xarkūkō</i>	'parrotfish / the parrotfish'

2.4. The indefinite suffix *-ē*

The indefinite suffix *-ē* (MSP reflex: *-i*) is segmentally, functionally and morpho-syntactically analogous to the definite suffix *-ō*. However, it does not cause the final vowel of the stem to be lengthened.

unmarked	indefinite	
<i>birmit</i>	<i>birmitē</i>	'hard candy / a hard candy'
<i>gambil</i>	<i>gambilē</i>	'hole / a hole'
<i>rištaḡ</i>	<i>rištaḡē</i>	'thread / a thread'
<i>ʔaqrab</i>	<i>ʔaqrabē</i>	'scorpion / a scorpion'
<i>barnuṣ</i>	<i>barnuṣē</i>	'blanket / a blanket'
<i>xarkuk</i>	<i>xarkukē</i>	'parrotfish / a parrotfish'

³ Theodore Beers (pers. comm.) has pointed out this possibility.

As has been shown for the first pair of suffixes, a suffix which causes stress shift also induces lengthening. The correspondence between stress shift and lengthening is similarly evident in the data on *-ō* and *-ē*, which is presented again here and marked for stress:

	unmarked	definite	indefinite
'hard candy'	<i>'birit</i>	<i>bir'mītō</i>	<i>'biritē</i>
'hole'	<i>'gambil</i>	<i>gam'bīlō</i>	<i>'gambilē</i>
'thread'	<i>'rištağ</i>	<i>riš'tōğō</i>	<i>'rištağē</i>
'scorpion'	<i>'raqrab</i>	<i>raq'rōbō</i>	<i>'raqrabē</i>
'blanket'	<i>'barnuṣ</i>	<i>bar'nūṣō</i>	<i>'barnuṣē</i>
'newborn'	<i>'xarkuk</i>	<i>xar'kūkō</i>	<i>'xarkukē</i>

3. Conclusion

In conclusion, these data show that stress shift motivates the lengthening and, in some cases, harmonization of stem vowels in Kumzari. The fact that some suffixes (*-an*, *-ō*) cause stress shift suggests that they are phonologically more intimately associated with the stem than suffixes which do not (*-in*, *-ē*), regardless of similarities in the semantic and morphosyntactic connection to their hosts.

References

- Anonby, Erik J. (2011). "Illustrations of the IPA: Kumzari." *Journal of the International Phonetic Association* 41:3, pp. 375–380.
- Anonby, Erik J., and Yusefian, Pakzad (2011). *Adaptive Multilinguals: A Survey of Language on Larak Island* [Studia Iranica Upsaliensia, 16]. Uppsala: Acta Universitatis Upsaliensis.
- Anonby, Erik J., and van der Wal Anonby, Christina (in preparation). *Ktēb Majma Kumzārī* [Kumzari Dictionary].
- Bayshak, Maryam S. (2002). "Are There Traces of Sassanian in the Language of the Shihuh, and is Kumzari Among the Affected Varieties? The Shihhi Dialect in the Light of Linguistic Science." *Al-Khaleej* (Arabic) 8541, p. 12, 17 October 2002.
- Jayakar, Atmaram S. G. (1902). "The Shahee Dialect of Arabic." *Journal of the Bombay Branch of the Royal Asiatic Society* 21, pp. 246–277.
- al-Kumzari, Ali Hassan Ali. (Ms.). *Kumzārī Lexicon and Grammar Notes*.
- Lambton, Ann K. S. (1953). *Persian Grammar*. Cambridge: Cambridge University Press.
- Lewis, Paul (ed.) (2011). *Ethnologue: Languages of the World*, 16th edition. Dallas, Texas: SIL International. Online at: <http://www.ethnologue.com>
- Skjærvø, Prods Oktor (1989). "Languages of Southeast Iran: Lārestānī, Kumzārī, Baškardī." In: Schmitt, Rüdiger (ed.), *Compendium Linguarum Iranicarum*. Wiesbaden: Ludwig Reichart Verlag, pp. 363–369.
- Thomas, Bertram (1930). "The Kumzari Dialect of the Shihuh Tribe (Musandam), Arabia, and a Vocabulary." *Journal of the Royal Asiatic Society* 62, pp. 785–854.
- Windfuhr, Gernot, and Perry, John R. (2009). "Persian and Tajik." In: Windfuhr, Gernot (ed.), *The Iranian Languages*. London and New York: Routledge, pp. 416–544.

Acquiring Persian Object Marking: Balochi Learners of L2 Persian

Ute Bohnacker, Uppsala University

ute.bohnacker@lingfil.uu.se

Somayeh Mohammadi, Azad University of Bam

somayehmohammadi77@yahoo.com

Abstract

This paper investigates the second language acquisition of Persian object marking by 237 native Balochi-speaking children from age eight to eleven. The combination of Balochi and Persian has not been studied from an acquisitional perspective, although it widely occurs in southeastern Iran. This paper studies the second language (L2) development of direct and indirect object marking after two and three years of exposure to Persian at school, vs. a control group of 133 monolingual native Persian children of similar age and socio-economic background. Beginning learners predominantly use non-targetlike object marking constructions reminiscent of Balochi, whilst there is a clear trend towards Persian-style object marking in the more advanced learners. These findings suggest that L2 learners initially transfer core aspects of their native grammar to their interlanguage Persian. The picture is complicated by the existence of differential object marking in Balochi and Persian and the somewhat different writing patterns of native and non-native learners.

Keywords: Balochi, Persian, children, L2 acquisition, object marking, *rā*, differential object marking (DOM), specificity, written composition, L1 transfer, preposition, indirect object, direct object, discourse pragmatics

1. Introduction

This paper investigates the second language (L2) acquisition of Persian direct and indirect object marking by children with Balochi as their first language (L1). Balochi is a northwestern Iranian language closely related to Persian, but differing from it in several aspects of morphology and syntax. For example, Persian marks indirect objects with a preposition and direct objects postnominally, whilst Balochi employs one and the same suffix on indirect and direct objects. Both languages exhibit differential object marking, which means that the overtness of object markers is influenced by semantic and discourse-pragmatic factors. Moreover, the overtness of object markers in Balochi, but not in Persian, is influenced by split ergativity.

The language pair Balochi/Persian has not been studied empirically from an acquisitional perspective, although it widely occurs in the southeastern provinces of Iran where Balochi is the native language but schooling takes place in Persian only. A pilot study found Persian object marking to be a particularly problematic area for learners. The present paper investigates this issue more systematically, by studying the development of indirect and direct object marking in the L2 Persian writing of 237 Balochi L1 school children aged eight to eleven. Two groups of learners with different lengths of L2 exposure are compared with each other and also with a control group of 133 monolingual L1 Persian children of similar age and socio-economic background. Large differences between learner groups are found. For the beginning L2 learners, non-targetlike indirect and direct object marking constructions

predominate, most of which are reminiscent of Balochi, whereas the more proficient learners show a clear trend towards targetlike, Persian-style object marking. These findings suggest that the elementary L2 learners transfer core aspects of their L1 Balochi grammar to their interlanguage Persian, whilst only vestiges of such transfer remain in the productions of the more advanced learners. This picture is somewhat complicated by the existence of differential object marking in Balochi and Persian and by slightly different writing patterns in native and non-native learners.

The present paper is structured as follows. In order to familiarize readers with the morpho-syntax of object marking in the two languages, Section 2 describes the Persian constructions of indirect object marking and direct object marking, including differential object marking, in light of the existing literature. Section 3 then outlines the Balochi object marking constructions, based on scholarly publications and our own observations. Section 4 summarizes the little research there is on the acquisition of Persian object marking. Section 5 provides background and methodological information on our current study. Section 6 analyses, compares and discusses the object marking results in the beginning L2 learners, more advanced L2 learners, and native Persian controls, concerning overall frequencies (6.1), indirect objects (6.2), and direct objects (6.3). Section 7 concludes the paper.¹

2. Persian

Persian is a null-subject language with rich verbal morphology and subject-object-verb (SOV) word order. The only case marker is postnominal *rā* (discussed below) which signals the grammatical function of direct object. Subjects are not marked for case, and roles of arguments other than subject or direct object are marked by prepositions (see below). In nominal expressions, modifiers typically follow the nouns they modify. Nouns and pronouns are invariant in form in Persian. There is no grammatical gender and nominal morphology is limited to number marking (singular vs. plural) and optional indefinite marking.²

¹ We would like to thank Carina Jahani for many useful comments and suggestions and for encouraging us to write this paper in the first place. Thanks also to an anonymous *Orientalia Suecana* reviewer for suggestions.

² In Persian, (in)definiteness is not always morphologically marked, and the interpretation of noun phrases as indefinite or definite is often context-dependent. For instance, *sib xarid-i?* (apple bought.PAST-2SG) with the bare noun *sib* can mean 'Have you bought apples?' in an all-new context, but also 'Have you bought the apples?' in a context where apples or apple-buying had been discussed before (cf. Rahimian and Hajiani 2009: 400). Indefiniteness, if overtly marked, can be indicated by prenominal *yek* 'one, a' in the singular, or by indefinite enclitic *i* ('a, some'), either on its own or in combination with *yek*. (Thus *ketāb-i*, *yek ketāb* and *yek ketāb-i* all mean 'a book'.) Unlike *yek*, *i* may combine with both singular and plural nouns (e.g. Lazard 1992; Lyons 1999; Ghomeshi 1996, 2003; Ganjavi 2007: 142–143 fn. 67). The enclitic *i* is mobile and can attach either to the noun or to an adjective if there is one in the noun phrase: *ketāb-i gerān* / *ketāb-e gerān-i* 'an expensive book' (*-e* is a linking vowel here). Definiteness tends not to be morphologically marked in Persian. There is no definite article, and bare noun phrases allow specific, definite interpretations (e.g. *lebās* 'the dress'). A demonstrative (*ān/un/in* 'that/this') can be added for deixis and definiteness, and in colloquial speech, the optional enclitic *e* may appear to signal definiteness (e.g. *man un lebās-e-rā xarid-am* 'I that dress-DEF-OBJ bought.PAST-1SG' 'I bought the dress.', cf. Samian 1983; Ganjavi 2007:58; Keshavarz 2007: 259). Such definite *e* does not, however, appear in writing or in the more formal varieties of spoken Persian.

2.1. Indirect object marking in Persian

In Persian, one preposition, *be* ‘to’, is used to signal the grammatical function of indirect object, which often expresses the role of goal or beneficiary (e.g. to tell someone, to give someone something). *Be* is used with all indirect objects, whether they are noun phrases headed by a proper noun, common noun, or pronoun, as illustrated in (1)–(4). This holds for spoken and written present-day Persian alike. *Be* is used irrespective of whether the indirect object has specific reference (as in (1)–(3)), unspecific or generic reference (as in (4)).

- (1) Per. (man) {*be Sārā/be moʔalem/be mādar-am/be u*} *goft-am*
 I to Sara/to teacher/to mother-my/to (s)he told.PAST-1SG
 ‘I told Sara/the teacher/my mother/her.’
- (2) Per. *u* {*be Sārā/be dokhtar/be u/be man*} *pul mi-dah-ad*
 (s)he to Sara/to girl/to (s)he/to I money IMPF-give.PRES-3SG
 ‘She gives Sara/the girl/her/me money.’
- (3) Per. *quri-rā* {*be Sārā/be mādar-am/be u*} *dād-am*
 teapot-OBJ to Sara/to mother-my/to (s)he gave.PAST-1SG
 ‘I gave Sara/my mother/her the teapot.’
- (4) Per. *doktor-hā be mardom komak mi-kon-and*
 doctor-PL to people help IMPF-do.PRES-3PL
 ‘Doctors help people.’

2.2. Direct object marking in Persian

Persian direct objects are not marked by a preposition but by postnominal *rā*, sometimes referred to as a suffix, sometimes referred to as a postposition or enclitic particle (e.g. Bossong 1985; Windfuhr 1979: 47–57; Dabir-Moghaddam 1990; Lazard 1992 [1957]: 74–76, 183–193; Karimi 1990, 1996). *Rā* in its written form is *را*, which is pronounced [rɑ] in more formal varieties, and [ro], [rə] or [o] or [ə] in more colloquial style, depending on whether it attaches to a word that ends in a vowel or a consonant (Lazard 1970: 74; Karimi 1990: 139; Ganjavi 2007: 9, 108). In writing, *rā* appears separately or joined to the preceding word. *Rā* is used in both written and spoken varieties of Persian, and it appears on direct objects irrespective of whether they are headed by a proper noun, common noun, or pronoun, as shown in (5)–(6). When the direct object consists of a more complex noun phrase, such as one containing an adjective or a possessive enclitic, *rā* attaches to the end of the phrase, as in (7)–(8).

- (5) Per. (man) {*Sārā-rā/moʔalem-rā/mādar-am-rā/u-rā*} *did-am*
 I Sara-OBJ/teacher-OBJ/mother-my-OBJ/she-OBJ saw.PAST-1SG
 ‘I saw Sara/the teacher/my mother/her.’
- (6) Per. *ki-rā* *did-i*
 who-OBJ saw.PAST-2SG
Sārā-rā/moʔalem-rā/mādar-am-rā/u-rā
 Sara-OBJ/teacher-OBJ/mother-my-OBJ/she-OBJ
 ‘Who did you see? – I saw Sara/the teacher/my mother/her.’

- (7) Per. *quri-ye bozorg-rā* *kharid-am*
 teapot-LINK big-OBJ bought.PAST-1SG
 ‘I bought the big teapot.’
- (8) Per. *mādar-e Leilā-rā* *did-am*
 mother-LINK Leila-OBJ saw.PAST-1SG
 ‘I saw Leila’s mother.’

However, not every direct object in Persian is marked by *rā*, neither in the written nor in the spoken modality. As illustrated in (9)–(10), *rā* is obligatory on proper nouns and personal and demonstrative pronouns, which are inherently definite and specific.

- (9) Per. *Sārā*(-rā)* *did-am*
 Sara-OBJ saw.PAST-1SG
 ‘I saw Sara.’
- (10) Per. *u*(-rā)* *did-am*
 (s)he-OBJ saw.PAST-1SG
 ‘I saw him/her.’

For direct objects other than proper nouns and pronouns, i.e. noun phrases headed by a common noun, there is variability in the use of *rā*. This variability is also known as differential object marking, where a number of semantic and discourse-pragmatic factors influence the realization or non-realization of object marking (e.g. Lazard 1992: 183–194; Bossong 1985: 3, 57–67; Windfuhr 1989: 533, 1992: 31).

Linguists are not in agreement as to which factors contribute most to *rā* being realized on the direct object in Persian. Both intrinsic semantic properties of the argument (humanness, animacy) and extrinsic semantic and discourse-pragmatic factors, such as information status (e.g. identifiability) and referentiality (definiteness, specificity) may be involved (cf. Browne 1970; Lazard 1982, 1984, 1992; Windfuhr 1989: 533; Karimi 1990, 1996, 2003; Dabir-Moghaddam 1990: 32–35; Ghomeshi 1996, 1997, 2003; Shokouhi and Kipka 2003; Ganjavi 2007: 109–113, 142–150; Rahimian and Hajiani 2009). Accounts of differential object marking in Persian differ from publication to publication, and the authors’ introspective grammaticality judgments often diverge for decontextualized examples with or without *rā* (see Shokouhi and Kipka 2003: 953–957 for an overview).

Lazard (1982, 1984, 1992) voiced the view that extrinsic semantic factors (definiteness, specificity) play a larger role than humanness or animacy in Persian. According to him, definite objects are virtually always *rā*-marked, regardless of animacy; specific indefinites are, regardless of animacy, occasionally *rā*-marked; and nonspecific indefinites/generics are normally not *rā*-marked. For the latter, Lazard suggests that the use of *rā* is influenced by the animacy of the object; *rā* is preferred for humans and dispreferred for inanimates (Lazard 1982: 181–186; 1984: 278–283; 1992: 185). These claims have not yet been tested empirically via corpus studies or experimental data. Lazard’s idea that definiteness and specificity are the main fac-

tors behind *rā* marking has however led scholars to claim that *rā* exclusively appears on definite objects and functions as a definiteness marker, or that *rā* exclusively appears on specific objects and functions as a specificity marker (cf. e.g. Browne 1970: 362; Comrie 1989: 132–135; Karimi 1989, 1990). Specificity here refers to the selection of a particular individual from a set of individuals (Karimi 1990: 142–145). Other scholars have however pointed out examples where direct objects with a specific reading occur without *rā*, or alternatively where *rā*-marked objects allow indefinite and/or nonspecific interpretations (e.g. Dabir-Moghaddam 1990; Shokouhi and Kipka 2003; Rahimian and Hajiani 2009).³

We will not go further into this debate here, but simply treat *rā* as an object marker. Based on our understanding of the literature and our own observations of Persian, we believe it is fair to describe the general tendency concerning the use of Persian *rā* on direct objects as follows: The higher the argument is on the animacy scale (e.g. + human) and/or the more identifiable the argument is in context (and note that identifiability often goes together with specificity and definiteness) the more likely it is that it will carry the object marker *rā*.⁴

For instance, on direct objects denoting specific humans like Sara or the teacher or she as in (8)–(10) above and (11) below, *rā* is basically obligatory.

- (11) Per. *moʔalem*(-rā)* *did-am*
 teacher-OBJ saw.PAST-1SG
 ‘I saw the teacher.’

By contrast, the less specific, the less definite, the less given and identifiable a direct object is in context, the less likely it will be that *rā* is used. For instance, when expressing the notion of having children, as in (12), the direct object *bache* ‘children’, whilst denoting humans, does not refer to any identifiable, specific children in context. In such a case, *rā* is not used. Similarly, when expressing the notion of giving money to the poor, as in (13), the direct object *pul* ‘money’ does not refer to any identifiable, specific money in context. Again, *rā* is not used.

- (12) Per. *man* *bache(*-rā)* *dar-am*
 I child have.PRES-1SG
 ‘I have children.’

- (13) Per. *man* *be fogharā* *pul(*-rā)* *mi-dah-am*
 I to poor.PL money IMPF-give.PRES-1SG
 ‘I give money to the poor.’

³ Moreover, *rā* cannot be a generalized definiteness or specificity marker, since arguments that are not direct objects, i.e. subjects, predicative nominals, and objects of prepositions, do not carry *rā* even if they are specific (Karimi 1990).

⁴ Some authors working within the generative Minimalist framework try to derive the distributional realization of *rā* from the abstract internal syntactic structure of Persian nominal expressions. One such approach is Ganjavi (2007), who assumes that only DPs (Determiner Phrases), i.e. nominal expressions with what she regards as a complete setup of functional projections, are *rā*-marked, whilst other nominal expressions are seen as lacking some abstract functional projections and therefore lack *rā* marking as well. In the present paper, we will not discuss whether *rā* marking on direct objects should be seen as a by-product of the abstract syntactic structure of nominal expressions, as this is influenced by the author’s predilection and choice of syntactic model.

The use of *rā* on the direct object is influenced by the identifiability of the object in context. Thus, when expressing the notion of buying a house and referring to a specific, particular and identifiable house as in (14a), *rā* must be used on *khāne* ('the house'); whilst *khāne* without *rā* is the preferred way of expressing 'a house, an unspecified house' as in (14c).^{5, 6}

- (14) a. *mi-khāh-am* *khāne*(-rā)* *be-khar-am*
 IMPF-want.PRES-1SG house-OBJ SBJ-buy.PRES-1SG
 'I want to buy the house.' (a specific house which the speaker assumes to be known/identifiable by the listener)
- (14) b. *mi-khāh-am* *khāne-i-rā* *be-khar-am*
 IMPF-want-1SG house-INDEF-OBJ SBJ-buy.PRES-1SG
 'I want to buy a (certain) house.' (a specific house known to the speaker but assumed to be unknown/not identifiable by the listener)
- (14) c. *mi-khāh-am* *khāne (*-rā)* *be-khar-am*
 IMPF-want-1SG house SBJ-buy-1SG
 'I want to buy a house.' (some unspecific house as yet unknown)

Publications on *rā* are typically based on unquantified decontextualized single-sentence examples. We have not been able to find any studies that could inform us about the frequency and extent to which direct objects are marked with *rā* in Persian. This would be most illuminating, both concerning different modalities (spoken, written) and different genres. One rare corpus study is Shokouhi and Kipka (2003), who analysed six hours of colloquial spoken dialogue between young expatriate Persian-speaking academics in Australia (5000 intonation units). They found 233 instances of direct object marking with *rā*, and the authors' breakdown by information type shows that the majority of these tokens were informationally given or accessible referents (88%). 12% of the tokens with *rā* encoded new referents, but even these overwhelmingly turned out to be identifiable in context (2003: 958–962). Shokouhi and Kipka also looked at all direct objects that were informationally new in the context they occurred in and found that 87% of these objects were not marked with *rā*. This suggests that there is a correlation between identifiability and overt *rā* marking of the direct object, but not a one-to-one correspondence (2003: 962). However, as Shokouhi and Kipka (2003) did not investigate all the direct objects without *rā*, their study cannot provide any information on the frequency of *rā* marking in Persian as such.

⁵ *Rā* is not restricted to definites, but can cooccur on direct objects with the indefinite postnominal *i*, as in (14b) and/or with the indefinite prenominal *yek* 'one, a'. According to Browne (1970: 361) and Ghomeshi (2003), such overtly marked indefinite direct objects with *i* and *rā* are interpreted as identifiable, specific indefinites.

⁶ There are two types of *rā* constructions in Persian. Apart from direct object marking, *rā* can also be used to mark sentence-initial aboutness topics of the type "As for X, ...". As a topic marker, *rā* can occur on elements that are not direct objects, and as a consequence, there may be sentences that contain two instances of *rā*, one a topic marker and one a direct object marker. Persian topic marking with *rā* will not be dealt with any further here (for discussion see e.g. Karimi 1990: 143–158; Shokouhi and Kipka 2003; Ganjavi 2007: 106–109, 135–166).

3. Balochi

Balochi is a northwestern Iranian language closely related to Persian. It is the principal language of the Baloch of Balochistan in Iran, Pakistan, and southern Afghanistan. Balochi does not have a standardized language or a standardized writing system, and it mostly remains an oral language, consisting of several regional dialects. Education in Balochi-speaking areas invariably takes place in a second language such as Persian or Urdu, which means that Balochi is largely restricted to the informal language domains of the home and neighbourhood and to traditional occupations (Jahani and Korn 2009: 635). We will here concern ourselves only with the Balochi variety spoken in Iranshahr (in the Sistan and Balochistan province of southeastern Iran), because this is where our empirical study on Balochi learners of Persian was carried out. Descriptions of Balochi can be found in Jahani (2003) and Jahani and Korn (2009) and references cited therein.

Like Persian, Balochi is a null-subject language with rich verbal morphology and subject-object-verb word order. Many dialects of Balochi have split ergativity. In the present/future tense (i.e. non-past), a nominative-accusative system is used, whilst the past tense has an ergative system. Consequences of this split for object marking are discussed below.

In Balochi nominal expressions, attributive adjectives, possessives, and quantifiers precede the nouns they modify. Balochi has no grammatical gender, and nominal morphology is limited to number marking (singular/plural), optional indefinite marking,⁷ and some case-marking. In the nominative-accusative system, the suffix *-a* (discussed below) often occurs to signal the grammatical function of object, whilst subjects are not overtly marked (Jahani 2003: 118).

3.1. Indirect and direct object marking in Balochi as compared to Persian

Balochi uses the same object marker on both indirect and direct objects, and not only on direct objects as in Persian. Object marking in Balochi may thus be likened to some older versions of Persian, where *rā* marking did occur on both indirect and direct objects (e.g. Dabir-Moghaddam 1990: 32 for Early Modern Persian). In present-day Persian, object marking with *rā* is however restricted to direct objects. There are some prepositions in Balochi but, in contrast to Persian, they are not used to mark indirect objects. The Balochi object marker *-a* has a number of allomorphs, *-a*, *-ya*, *-ra*, depending on whether it is preceded by a consonant or a vowel (Windfuhr 1992: 31; Jahani 2003: 115, 118; Jahani and Korn 2009: 651–652). Indirect objects are illustrated in (15)–(16), direct objects in (17)–(18). As shown in the examples, the morphological marking on indirect and direct objects is identical.

⁷ Indefiniteness can be marked on Balochi noun phrases by enclitic *e* ('a'), which attaches to nouns in the singular (e.g. *kitab-e* 'a (certain) book'), or by the prenominal indefinite article/numeral, *ya* 'a, one'. *Ya* and *e* can also co-occur (e.g. *ya kitab-e* 'a book'). According to Jahani and Korn (2009: 667), overt marking with *e* induces an indefinite specificity reading ('a certain book').

Balochi indirect objects

- (15) Bal. {*Sārā-ya/mallem-a/wtī māt-a/āyī-a*} *gwash-on*
 Sara-OBJ/teacher-OBJ/my mother-OBJ/she-OBJ tell.PRES-1SG
 ‘I tell Sara/the teacher/my mother/her.’
- (16) Bal. *āyī* {*Sārā-ya/mallem-a/wtī māt-a/āyī-a/man-a*} *zarr-a dant*
 (s)he Sara-OBJ/teacher-OBJ/my mother-OBJ/she-OBJ/ money-IMPF
 I-OBJ give.PRES.3SG
 ‘She gives Sara/the teacher/my mother/her/me money.’

Balochi direct objects

- (17) Bal. *kay-a* *gend-e*
 who-OBJ see.PRES-2SG
 {*Sārā-ya/mallem-a/wtī māt-a/āyī-a*} *gend-on*
 Sara-OBJ/teacher-OBJ/my mother-OBJ/ see.PRES-1SG
 (s)he-OBJ
 ‘Who do you see? – I see Sara/the teacher/my mother/her.’
- (18) Bal. *Maryam* {*jāmag-a/ketāb-a*} *gī*
 Maryam dress-OBJ/book-OBJ buy.PRES.3SG
 ‘Maryam buys the dress/the book.’

Readers may have noticed that the above examples of object marking in Balochi (15)–(18) are exclusively in the present tense. In all tenses formed with the present stem, sentences in Balochi are constructed following a nominative-accusative system, and here overt object marking with *-a* is used. In the past tense, however, sentences are constructed ergatively and direct objects are not marked overtly (Korn 2009; Jahani and Korn 2009: 669). This is illustrated by the following minimal-pair examples contrasting Balochi and Persian, (19) for non-past, and (20) for past tense.

(19) Present tense: overt object marking in both Balochi and Persian⁸

- a. Bal. {*Sārā-ya/mallem-a/wtī māt-a/āyī-a*} *gend-on*
 Sara-OBJ/teacher-OBJ/my mother-OBJ/(s)he-OBJ see.PRES-1SG
- a'. Per. {*Sārā-rā/moʔalem-rā/mādar-am-rā/u-rā*} *mi-bin-am*
 Sara-OBJ/teacher-OBJ/mother-my-OBJ/(s)he-OBJ IMPF-see.PRES-1SG
 ‘I see Sara/the teacher/my mother/her.’
- b. Bal. *āyī* {*Sārā-ya/mallem-a/wtī māt-a/āyī-a*} *zarr-a dant*
 (s)he Sara-OBJ/teacher-OBJ/my mother-OBJ/(s)he-OBJ money-IMPF
 give.PRES.3SG
- b'. Per. *u* {*be Sārā/be moʔalem/be mādar-am/be u*} *pul mi-dah-ad*
 (s)he to Sara/to teacher/to mother-my/to (s)he money IMPF-
 give.PRES-3SG
 ‘(S)he gives Sara/the teacher/my mother/her money.’

⁸ Balochi person and number markings on the verb (as *-on* in (19a)) are often homophonous with pronominal clitics (CL) that attach to other elements such as noun phrases (as *-(y)on* in (20a)). Pronominal clitics often specify the person and number of the agent (Jahani and Korn 2009: 654; Korn 2009: 57–60).

(20) Past tense: no overt direct object marking in Balochi, but overt marking in Persian

- | | | |
|----------|---|---------------|
| a. Bal. | { <i>Sārā-yon/mallem-on/wtī māt-on/āyī-on</i> } | <i>di(st)</i> |
| | Sara-1SG.CL/teacher-1SG.CL/my mother-1SG.CL/
(s)he-1SG.CL | saw.PAST.3SG |
| a'. Per. | { <i>Sārā-rā/mo?alem-rā/mādar-am-rā/u-rā</i> } | <i>did-am</i> |
| | Sara-OBJ/teacher-OBJ/mother-my-OBJ/(s)he-OBJ
'I saw Sara/the teacher/my mother/her.' | saw.PAST-1SG |

There is thus an important difference concerning object marking between Balochi and Persian: Direct object marking can be overt in all tenses for Persian, but in Balochi, due to split ergativity, object marking is only overt for tenses formed from the present stem. Irrespective of tense, there are also some other ergative constructions in Balochi where a pronominal clitic (e.g. *-on*, *-ī*) attaches to the direct object whilst the object marker *-a* vanishes, as in (20) above.⁹ We need to keep this in mind when considering the acquisition of Persian object marking by Balochi pupils.

In connection with direct object *rā* in Persian, we previously discussed differential object marking (DOM). Like Persian, Balochi also exhibits DOM, and thus *-a* on the object can be realized or dropped depending on certain semantic and discourse-pragmatic factors. DOM has not been studied systematically for Balochi (Farrell 1990: 65; Korn 2009), and there are no corpus (frequency) studies of the language as yet. However, Jahani and Korn (2009: 669–670) cite examples from a number of Balochi dialects that suggest that overt *-a* marking on the direct object goes together with specificity and context-identifiability of the argument. Generic, non-specific and inanimate objects, on the other hand, tend to be unmarked (Jahani and Korn 2009: 669). Our own informal observations of Balochi (S. Mohammadi) support this; see the examples in (21)–(25). Thus, DOM of direct objects in Balochi appears to pattern similarly to what has been observed for DOM in Persian.

DOM: Overt object *-a* marking on human/animate and specific/context-identifiable objects

- | | | |
|-----------|--|-----------------------------|
| (21) Bal. | <i>man</i> { <i>wtī dōst*(-a)/āyī*(-a)</i> } | <i>har rōč-a gend-on</i> |
| | I my friend-OBJ/her-OBJ | every day-IMPF see.PRES-1SG |
| | 'I see my friend/her every day.' | |

⁹ Consider the following Balochi example (i) from Iranshahr (S. Mohammadi). The direct object carries an object marker *-a* when the agent (*-on*) is marked on the verb (a-version), but this object marker *-a* disappears when the pronominal clitic agent attaches to the object (b-version).

- | | |
|-------------|--|
| (i) a. Bal. | <i>Maryam-a dusta dar-on</i> |
| | Maryam-OBJ love have.PRES-1SG |
| b. Bal. | { <i>Maryam-(*a)-on/wtī mat-(*a)-on</i> } <i>dust-en</i> |
| | Maryam-1SG.CL/my mother-1SG.CL love.PRES-COP |
| | 'I love Maryam/my mother.' |
| (ii) Per. | { <i>Maryam-rā/mādar-am-rā</i> } <i>dust dār-am</i> |
| | Maryam-OBJ /mother-my-OBJ love have.PRES-1SG |
| | 'I love Maryam/my mother.' |

In sum, when comparing Balochi and Persian, at first glance the two languages do not appear very different from each other as far as direct object marking is concerned. Both use morpho-phonologically similar markers, *rā* (Persian) and *-a* (Balochi), and both do so in the same position, postnominally. Both languages allow differential object marking (DOM), and although the last word has not been said concerning the factors that determine the realization or non-realization of overt *rā*/*-a*, both intrinsic semantic properties of the noun (humanness, animacy) and discourse-pragmatic features (definiteness, specificity, givenness/identifiability) appear to play a role. Whilst we acknowledge the need for a more systematic study of these factors in the two languages, our impression is that Persian and Balochi do not greatly differ concerning DOM. In both languages, direct objects denoting a human, specific referent typically carry overt marking, irrespective of the type of the nominal expression (proper noun, common noun phrase, personal pronoun). On the other hand, direct objects which denote inanimate, nonspecific, and new-information referents are typically not marked overtly with *rā* (Persian) or *-a* (Balochi). Recall however that due to split ergativity, Balochi overtly marks objects only in the nominative-accusative paradigm.

Concerning the marking of indirect objects in Balochi and Persian, there is a clear morpho-syntactic difference between the two languages: Persian uses the preposition *be* and Balochi the suffix *-a*, though, as will be recalled, Balochi only uses this overt object marking in the nominative-accusative paradigm, particularly in the present tense system.

4. Previous studies of acquisition of object marking in Persian

Monolingual children appear to acquire Persian *rā* marking early on. A recent doctoral thesis (Foroodi-Nejad 2011) on four-to-seven-year-old Persian-speaking children in Iran includes an oral elicitation experiment concerning direct object *rā* marking, amongst other topics. Foroodi-Nejad found that *rā* is mastered relatively early by typically developing children, whilst omission of *rā* from obligatory contexts may be a clinical marker of language impairment in monolingual children. Typically developing monolinguals regularly use *rā* in their speech, and in the experiment they use it in a targetlike fashion in 97% of obligatory contexts already by 4 years (and probably earlier, but 4 years was the youngest age studied, cf. Foroodi-Nejad 2011: 83–103).¹⁰ Unfortunately, there is no corresponding data concerning the acquisition of object marking in bilingual or child L2 learners of Persian.

To our knowledge, object marking in Balochi-speaking second language learners of Persian has not yet been researched. In general, there are few studies of L2 Persian, and there is only one study that we are aware of that deals with the L2 acquisition of Persian object marking (Mirza 2000). In her MA thesis, Mirza looked at the case system in Armenian learners of Persian who had been exposed to Persian since early childhood. In spite of the fact that Armenian and Persian objects are marked

¹⁰ Foroodi-Nejad (2011) only elicited *rā* in obligatory contexts. Therefore, her children's 97% provision of *rā* cannot be taken to mean four-year-olds *rā*-mark 97% of all their direct objects (cf. 2.2.).

very similarly (and both languages have DOM), Mirza found that object marking was a consistent problem for a group of Armenian learners of L2 Persian. In a grammaticality judgment task, untutored L2 learners (illiterate adults) accepted Persian sentences with a direct object with a specific reading without *rā* marking. Native Persian speakers and Armenian adults who had had L2 schooling in Persian and more exposure to Persian rejected the same sentences as ungrammatical, requiring the direct object to be marked with *rā* (as in (5)–(11)).¹¹ Mirza's untutored L2 group also "corrected" grammatical Persian stimulus sentences with overt *rā* by removing *rā* from direct objects with a specific reading, which resulted in ungrammaticality. By contrast, the tutored Armenian learners of Persian kept overt *rā*. Mirza's study concerns adults only, and we have not been able to find any studies of child L2 Persian with which we could compare the results of our own study. However, Mirza's and Foroodi-Nejad's (2011) studies do suggest that Persian object marking might be a problematic area for certain language learners.¹²

5. The current L2 acquisition study: Background and method

This study investigates the acquisition of direct and indirect object marking in Persian as a second language by Balochi-speaking school children in southeastern Iran. We became interested in this issue as all literacy training and school teaching for Balochi children takes place in Persian only, and the children are expected to be able to write in Persian from second grade. In a pilot study, Mohammadi (2009) noted that one area that Balochi school children struggle with is Persian object marking. The present study investigates this issue more systematically by looking at the development of indirect and direct object marking in the L2 Persian compositions of 237 Balochi L1 school children between the ages of eight and eleven. Two groups of learners with different lengths of L2 exposure are compared with each other and also with a group of 133 age-matched monolingual L1 Persian children, all of them writing a composition on the same topic under similar conditions.

The Balochi L1 children that participated in the study had started school at the age of seven and were attending unisex primary schools in an urban area of southeastern Iran (Fatemeh Zahra Primary and Zakiye Primary, both in Mohammadan town, Iranshahr city). The children were all girls. Care was taken to keep their language background as homogeneous as possible. Therefore, only children who came from monolingual Balochi homes were included. Concerning the social background of the children, most came from low literacy homes. The majority of their parents

¹¹ Here are two of Mirza's examples:

- | | | |
|------|---|--|
| (i) | <i>Nāhid Maryam(-rā) be kelas āvard</i>
Nahid Maryam-RA to class brought
'Nahid brought Maryam to class.' | (obligatory <i>rā</i> in native Persian, but optional for L2ers) |
| (ii) | <i>Mas'ud be Ali in pul(-rā) dād</i>
Masood to Ali this money-RA gave
'Masood gave this money to Ali.' | (obligatory <i>rā</i> in native Persian, but optional for L2ers) |

¹² We are also aware that the acquisition of differential object marking in other languages, such as Spanish, is a problematic area for bilingual/L2 learners (see e.g. Bowles and Montrul 2009).

were illiterate; some had attended primary school (age 7–11) and/or middle school (up to age 13), but only a few had gone on to secondary education or received a diploma (the Iranian equivalent of having attended school for 12 years).

The children in the study had only sporadically come in contact with languages other than their native Balochi before entering school. Balochi is the language of the home and the surrounding community, in which not only family and friends, but also daycare staff, shopkeepers, and officials speak Balochi. The children may have had limited exposure to Persian via radio and television, where broadcasts are largely in Persian. Regular and extended exposure to Persian did not however occur until primary school. Here, the medium of instruction is Persian, as is the rule in Iran. Teachers are either native speakers of Persian or proficient L2 speakers of Persian (with Balochi as their L1). At primary schools in Balochistan in Iran, Persian is taught for five hours a day for approximately 28 weeks (or 7 months) per year. In first grade, the focus is on learning the basics of the language by means of listening, speaking, and literacy exercises. From second grade, there is an increased focus on the training of reading and writing. Towards the end of the second year, the children are expected to be able to write simple compositions in Persian.

All texts and literacy artifacts the children encounter are in Persian, and they are taught to read and write in Persian only. In class, the children are exclusively exposed to Persian and are encouraged to use Persian in class themselves, whilst Balochi is spoken during breaks and after school. Whilst there may be some variation in the amount of Persian that the children encounter outside school via Persian-language media and friends, we are reasonably confident that the amount of Persian input at school is relatively similar for the children of a particular age in the study.

Written data in the form of a hand-written composition on the topic “Mother” was collected from second- and third-grade pupils at the end of the school year (some in May 2009, some in June 2011). This topic was chosen because it is culturally universal and unlikely to cause problems with school authorities. By that time, the second-graders (age 8–9) had had approximately 1000 hours of exposure to Persian at school. The end of second grade was chosen as this is the earliest point at which the pupils could be expected to be able write a composition on their own. There were 107 pupils in the second-grade group. One of the authors (S. Mohammadi) gathered these data with the help of teachers and administrators at two primary schools. The pupils were given the same topic by their teachers and asked to write a free composition in class with a time limit of about one hour. Care was taken not to put pressure on the pupils and not to present the composition as an exam paper or an assignment that would be graded by the teacher.¹³

The same kind of composition was collected from third-graders (age 10–11), who by then had had ca. 1500 hours of exposure to Persian at school. There were 130 pupils in that group. The pupils were unaware that object marking would be investigated in their texts.

As a control group, 133 monolingual Persian age peers were included in the study. Hand-written compositions on the same topic were collected from third-grade

¹³ We also collected compositions on the topic ‘Prayers’ that was set as an end-of-term exam, but did not include data from these in the study.

ers (age 10–11) with Persian as their L1. The writing task was administered in the same manner as for the L2 pupils. For the L1 control group, care was taken to only include pupils from monolingual Persian homes and to keep the social background of the children as similar as possible to that of the Balochi pupils. Some of the parents of the L1 Persian children were illiterate; others had primary, and some middle school education. As it was easier to find monolingual Persian children outside Balochistan, 55 compositions were collected from pupils attending a primary school in the city centre of Teheran (Arjantin area, Zeynabiyeh Primary) in May 2009. We subsequently deemed it necessary to enlarge this L1 data set, and so 78 additional compositions were collected in 2011 from children from monolingual Persian homes who were attending schools in Iranshahr. As no differences could be detected between the L1 Persian compositions from Teheran and Iranshahr, these were combined to make up a control group of 133 L1 Persian texts.

Table 1 provides an overview of the data. As can be seen in Table 1, group sizes, and thus the number of compositions per group, are not completely even. No more than 107 L2 grade two compositions could be collected and used, since a number of children from non-exclusively Balochi-speaking homes had to be excluded. As is typical for elementary learners, the L2 texts in grade two were shorter on average than the texts in the other two groups.

Participant groups	Compositions	Individual text length	Word total for group
L1 Persian controls, grade 3	N - 133	average 150 words, range 100–200 words	19,950
L2 Persian, grade 2	N - 107	average 80 words, range 50–150 words	8,600
L2 Persian, grade 3	N - 130	average 140 words, range 70–200 words	19,500

Table 1. Overview of L1 and L2 data

Copies of the handwritten L1 and L2 texts were analyzed, word-counted, and marked up for indirect and direct objects. These objects were classified into and coded for sub-types depending on grammatical function and morpho-syntactic form, i.e. the use of *be*, *rā*, and learner variants thereof. One of the authors, a speaker of Persian and Balochi, carried out the coding and counting by hand. To increase coding reliability, some data were initially also coded by an experienced Iranian linguist and speaker of Persian and Balochi (C. Jahani). Coding disagreements were resolved through mutual discussion. The data were then entered into Excel spreadsheets, which formed the base for the subsequent analysis.

6. Analysis and results

6.1. Overall frequencies

All three groups produced more direct objects in their texts than indirect ones, as shown in Table 2. This should not come as a surprise as direct objects are generally more frequent in both speech and writing.

	Indirect objects		Direct objects		Word total
	Total	Per text	Total	Per text	
L1-3 rd years N - 133	106	0.8	605	4.5	19,950
L2-2 nd years N - 107	186	1.7	405	3.8	8,600
L2-3 rd years N - 130	266	2.0	866	6.7	19,500

Table 2. Occurrence of indirect and direct objects in the children's compositions

What did come as a surprise however was the fact that the L2 learners produced more indirect objects than the L1 controls. This tendency was pronounced both for the low proficiency learners (grade 2) and the higher proficiency learners (grade 3) who used more than twice as many indirect objects per text than did the native controls (L2ers: 1.7 and 2.0 instances, L1ers: 0.8 instances). The L2 second-graders used direct objects twice as often as indirect objects (405:186), whilst the L1 controls produced six times as many direct objects as indirect objects (605:107). The L2 third-graders produced many more direct objects per text (6.7 instances) than the L1 controls (4.5 instances). Differences in group size or text length cannot account for these ratio differences.

A qualitative analysis of the data showed that some of the L2 learners wrote down the same sentence several times in their text about "Mother", and this often was a sentence in which the mother was the indirect object, such as *be mādar-am ehterām mi-gozār-am* (to mother-my respect IMPF-put-1SG) 'I respect my mother'.¹⁴ This sentence and versions thereof were often used twice and sometimes three or four times in a text by the learners, which increased the number of indirect objects in the L2 groups to a level far beyond that of the L1 controls. We are not quite sure why such repetition of sentences occurred in the L2 texts only. Young school children sometimes have the mistaken impression that a composition is only good or "finished" once the sheet is filled, and it was the youngest learners who repeated sentences the most, even though our task instructions did not invite or urge pupils to do so. Another reason might be that processing demands were simply so high for the L2 learners that consciously or unconsciously relief was sought through the repeated use of the same phrase or sentence, a phenomenon not unknown in second language studies. Note that the L1 pupils did not repeat their own sentences in the texts.

A qualitative analysis of the texts also provided an explanation for the high number of direct objects in the third-grade L2 texts. These often contained short, minimally varied sentences, with the same subject and/or verb repeated again and again, with only the direct object having been changed (e.g. My mother buys a bag for me. My mother buys a dress for me. My mother buys shoes for me. My mother sews a dress for me.). This writing pattern drove up the number of direct objects per text (6.7) as compared to the L1ers (4.5), who did not make much use of such repetitive language.

¹⁴ The Persian for 'I respect my mother' (lit. I put respect to my mother) is:

(i) *man be mādar-e khod ehterām mi-gozār-am*
I to mother-LINK self respect IMPF-put.PRES-1SG

The L2 Persian learners used this sentence repeatedly, although the second-graders often substituted *rā* for *be*:

(ii) *man mādar-e khod-rā ehterām mi-gozār-am*
I to mother-LINK self-RA respect IMPF-put.PRES-1SG

In both the L1 and the L2 texts we find sentences with an indirect object only, sentences with a direct object only, and sentences with both indirect and direct object. In the following sections, indirect and direct objects will be investigated separately. We will start with indirect objects.

6.2. Acquisition of indirect object marking

When looking at the morphological marking of indirect objects, we find large differences between the groups, which are summarized in Table 3 and Fig. 1 and discussed below.

The morphological form of the indirect objects in the 133 L1 Persian pupils' compositions is as would be expected from reference grammars of Persian: They mark virtually all of their indirect objects with the prenominal preposition *be* 'to' (96%, 102/106). Only four indirect objects diverge from this, being non-targetlike uses of postnominal *rā* (i.e. a direct object marker) or both *rā* and *be* on the indirect object. Due to the very low frequency of these non-adultlike structures we may dismiss them as simple writing mistakes, although we cannot entirely rule out that they could also be vestiges of an earlier developmental stage, as we do not have any L1 data from grade two, when these native Persian children started learning to write. Overall, however, the native Persian pupils' indirect object marking with *be* can be considered adultlike.

	<i>be</i>	Simultaneous <i>rā</i> and <i>be</i>	Zero marking	<i>rā</i> instead of <i>be</i>
L1-3 rd yrs N - 133	96.2% (102/106)	0.9% (1/106)	0% (0/106)	2.8% (3/106)
L2-2 nd yrs N - 107	15.0% (28/186)	13.4% (25/186)	3.2% (6/186)	68.3% (127/186)
L2-3 rd yrs N - 130	85.0% (226/266)	1.9% (5/266)	0% (0/266)	13.2% (35/266)

Table 3. Indirect object marking in the compositions

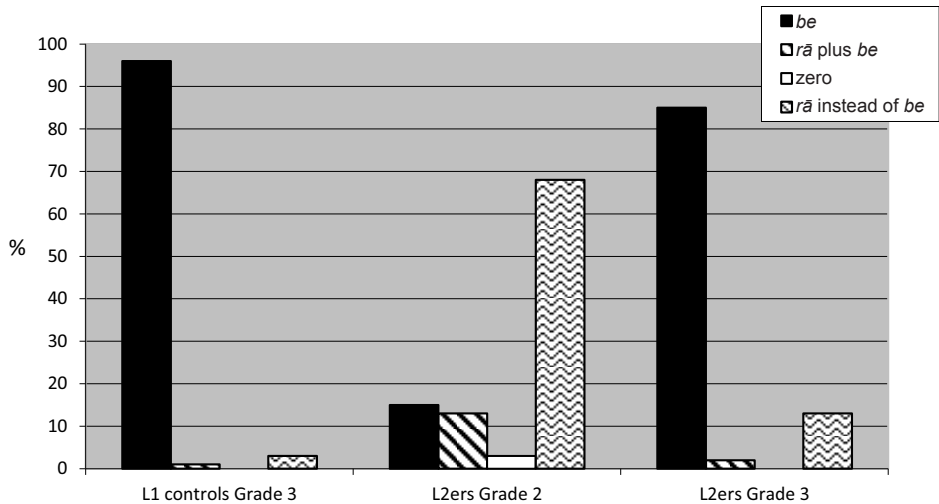


Figure 1. Indirect object marking in Persian

6.2.1. Indirect object marking in the second-grade L2 learners

In stark contrast to the native Persian pupils, the L2 learners in grade two rarely use prepositional *be* to mark indirect objects in their writing (15%, 28/186). The 107 learners only produce 28 instances of adultlike prenominal *be*-marking altogether. Instead, they mostly employ non-targetlike postnominal *rā* (68%, 127/186), as illustrated in examples (30)–(33), where the a-examples give the L2 version and the b-examples the Persian target version.

- (30) a. *mā* *bāyad* *moʔalem-rā* *salām kon-im*
 we must teacher-RA greeting do.PRES-1PL
 ‘We must greet the teacher.’ (L2 learner, grade 2, non-target)
- (30) b. *mā* *bāyad* *be moʔalem* *salām kon-im*
 we must to teacher greeting do.PRES-1PL
 (Persian target)
- (31) a. *moʔalem* *mā-rā* *dars mi-āmuz-ad*
 teacher we-RA lesson IMPF-teach.PRES-3SG
 ‘The teacher gives/teaches us a lesson.’ (L2 learner, grade 2, non-target)
- (31) b. *moʔalem* *be mā* *dars mi-āmuz-ad*
 teacher to we lesson IMPF-teach.PRES-3SG
 (Persian target)
- (32) a. *man* *moʔalem-e khod-rā* *kādo mi-dah-am*
 I teacher-LINK self-RA gift IMPF-give.PRES-1SG
 ‘I give a gift/gifts to my teacher.’ (L2 learner, grade 2, non-target)
- (32) b. *man* *be moʔalem-e khod* *hedye mi-dah-am*
 I to teacher-LINK self gift IMPF-give.PRES-1SG
 (Persian target)
- (33) a. *moʔalem* *bage-hā-rā* *mi-guy-ad*
 teacher child-PL-RA IMPF-tell.PRES-3SG
 ‘The teacher tells the children...’ (L2 learner, grade 2, non-target)
- (33) b. *moʔalem* *be bage-hā* *mi-guy-ad*
 teacher to child-PL IMPF-tell.PRES-3SG
 (Persian target)

A further 13% (25/186) of indirect objects produced by the second-grade L2 learners are simultaneously marked with prenominal *be* and postnominal *rā*, as illustrated in (34)–(36). This is not acceptable in native Persian. Simultaneous *be* + *rā* may suggest a beginning awareness of prenominal *be*, at a time when postnominal marking of indirect objects has not yet been expunged from the learners’ interlanguage system.

- (34) a. *be mā-rā* *mi-guy-ad*
to we-RA IMPF-tell.PRES-3SG
'She tells us.' (L2 learner, grade 2, non-target)
- (34) b. *be mā* *mi-guy-ad*
to we IMPF-tell.PRES-3SG
(Persian target)
- (35) a. *khodā* *be u-rā* *komak mi-kon-ad*
God to (s)he-RA help IMPF-do.PRES-3SG
'God helps her.' (L2 learner, grade 2, non-target)
- (35) b. *khodā* *be u* *komak mi-kon-ad*
God to (s)he help IMPF-do.PRES-3SG
(Persian target)
- (36) a. *be mā-rā* *dars* *mi-dah-ad*
to we-RA lesson IMPF-give.PRES-3SG
'She teaches us/gives us a lesson.' (L2 learner, grade 2, non-target)
- (36) b. *be mā* *dars* *mi-dah-ad*
to we lesson IMPF-give.PRES-3SG
(Persian target)

We thus find that the L2 learners in second grade greatly overuse *rā*: They use it to mark indirect objects, whilst it is restricted to direct objects in Persian.

From a contrastive perspective of Persian and Balochi, these results are to be expected if we take an L1 transfer-driven approach to second language learning. The use of postnominal *rā* on most indirect objects produced by the second-grade learners can be explained as structural and morphological transfer from the L1 to the L2 (e.g. Jarvis and Odlin 2000; Lardiere 2009). Recall that L1 Balochi marks indirect objects with a suffix, not with a preposition. And so the Balochi learners of Persian also latch on to a suffix to mark indirect objects in their L2 Persian. And they do not just latch onto any suffix, but one that is phonologically similar to the one used in L1 Balochi, *-a* (which in some cases allomorphs as *rā*). It is thus not surprising that the learners employ *rā* for indirect object marking in their L2 Persian in analogy with *-a* in their L1. The learners appear to have largely persianised or relexified their Balochi grammar but not yet acquired the target Persian grammar of *be*-marking indirect objects. Consider the (non-target) L2 utterances in (30') and (33') and the corresponding Balochi forms, both with suffixal object marking:

- (30') a. *mā bāyad* *moʔalem-rā* *salām kon-im*
we must teacher-RA greeting do.PRES-1PL
'We must greet the teacher.' (Persian L2 learner, grade 2, non-target)
- (30') b. *mā bāyad-ē* *mallem-a* *salām be-day-ē*
we must-COP.PRES.3SG teacher-OBJ greeting SBJ-give.PRES-1PL
(Balochi)

- (33') a. *moʔalem* *bachē-hā-rā* *mi-guy-ad*
 teacher child-PL-RA IMPF-tell.PRES-3SG
 'The teacher tells the children.' (Persian L2 learner, grade 2, non-target)
- (33') b. *mallem* *chokk-ān-a* *gwash-i*
 teacher child-PL-OBJ tell.PRES-3SG
 (Balochi)

The L2ers' *rā* on indirect objects may also be regarded as an overextended default marking. The second-graders make ample use of *rā* for direct objects (see section 6.3). Direct objects are more frequent than indirect ones, and overgeneralizing *rā* to the latter would not be an unusual acquisitional route to take. It is likely, though, that such overextension is not so much motivated by the system of Persian, but by the L1 system of Balochi where there is one morpheme (-a) for both direct and indirect object marking. Under a frequency-driven approach to second language acquisition, the learner structure may be attributed to *rā* being a frequent morpheme in Persian, which presumably is frequent in the input as well, although we currently lack frequency counts for Persian corpora in general and for teacher talk (the main source of Persian input for our learners) in particular. *Rā* is invariant in form and word-final, which may contribute to its salience, and it occurs on direct objects (which are more frequent than indirect ones) as well as aboutness topics (see fn. 6).

At the same time, we can also see that the L2ers have some modest awareness of *be*-marking in Persian. Recall Table 3, and the 28 instances (15%) of targetlike *be*-marking, produced by 19 of the 107 learners. (Six of these learners also produce non-targetlike Balochi-style *rā* on indirect objects.) Recall also the 25 instances (13%) of simultaneous *be* and *rā*, produced by 17 of the 107 learners (as exemplified in (33)–(35) above). We thus see the beginnings of a development towards the Persian target structure.

Interestingly, there is hardly any zero marking of indirect objects in the learner data (3%, 6/186 instances); that is, pronouns and nouns used as indirect objects do not occur bare but nearly always with some (albeit non-targetlike) object marker. In studies of second language production, bare and uninflected forms are usually frequent, especially in the beginning stages of acquisition, and functional morphology (such as case marking) is often found to be missing. By contrast, our beginning learners of Persian do not just omit *be* but actually use overt suffixal marking on indirect objects, which suggests that they indeed transfer their L1 Balochi morphological object marking to their L2 Persian.

6.2.2. Indirect object marking in the third-grade L2 learners

Turning to the more advanced learners, the pupils in third grade (Table 3), we can see that they mark indirect objects most of the time with prepositional *be* (85%, 226/266). This percentage is very different from the second-grade L2ers (15%), but relatively close to the provision of *be* by the age-matched native L1 controls (96%). Some of the L2 third-graders still produce Balochi-style *rā* instead of *be* (13%, 35/266). This is done by 25 out of 130 learners, and is exemplified in (37)–(38). There

are however very few children in this group that still only use Balochi-style *rā* instead of *be*; nearly all the L2 third-graders (125 out of 130 children) use targetlike indirect object marking with prepositional *be* at least once.

- (37) a. *mādar-e khod-rā* *komak mi-kon-am*
 mother-LINK self-RA help IMPF-do.PRES-1SG
 ‘I help my mother.’ (L2 learner, grade 3, non-target)
- (37) b. *be mādar-e khod* *komak mi-kon-am*
 to mother-LINK self help IMPF-do.PRES-1SG
 (Persian target)
- (38) a. *moʔalem* *mā-rā* *khāndan o neveshtan āmukht*
 teacher we-RA reading and writing teach.PAST.3SG
 ‘The teacher taught us how to read and write.’
 (L2 learner, grade 3, non-target)
- (38) b. *moʔalem* *be mā* *khāndan o neveshtan āmukht*
 teacher to we reading and writing teach.PAST.3SG
 (Persian target)

No cases of zero marking occur; there are however still a few cases of simultaneous *be* and *rā* on indirect objects (2%, 5/266) in the third-grade L2ers. Apart from these non-target forms, the third-graders have largely mastered Persian indirect object marking, as far as can be determined from their compositions.

We do not have the longitudinal data for the same individuals needed to say for sure that they indeed go from non-target *rā* marking (and optionally via interim simultaneous *be* and *rā*) to targetlike exclusive *be*-marking. Nevertheless, the comparison of the two learner groups (grades 2 and 3) suggests a clear developmental path towards the native Persian pattern.

6.3. Acquisition of direct object marking

All three groups of school children produce many instances of direct objects in their compositions (recall Table 2). When looking at the morphological marking of direct objects, we do not find quite as large differences between the groups as we did for indirect objects, but the differences between groups are still sizeable. These concern the different morphological forms, as well as extent to which *rā* marking is overt. To start with, let’s look at the different forms of morphological marking in the data, summarized in Table 4 and Figure 2.

	<i>rā</i>	zero marking	multiple <i>rā</i>	<i>be</i> instead of <i>rā</i>
L1-3 rd years	75.5% (457/605)	23.8% (144/605)	0.5% (3/605)	0.2% (1/605)
L2-2 nd years	34.6% (140/405)	49.6 % (201/405)	10.1% (41/405)	5.7% (23/405)
L2-3 rd years	65.7% (569/866)	33.3% (288/866)	1.0% (9/866)	0.0% (0/866)

Table 4. Direct object marking in the compositions

The L1 Persian pupils in third grade mark a large majority of their direct objects (76%, 457/605) with the postnominal *rā*. The group is relatively homogeneous here;

most L1 children produce such overt *rā*. The remaining 24% of their objects are not suffixed (“zero marking”), an option Persian has due to DOM. We may assume that this distribution of 76% overt *rā* and 24% zero marking on direct objects is a typical distribution for Persian and may serve as a benchmark, at least for the type of written texts that we are dealing with here. Only 0.7% of the L1ers’ direct objects are morphologically non-adultlike: There are 3 instances of multiple *rā* (i.e. more than one *rā* on the same object), and 1 prepositional *be* instead of *rā*. Considering the extremely low frequency of these cases, they are most likely to be simple writing mistakes.

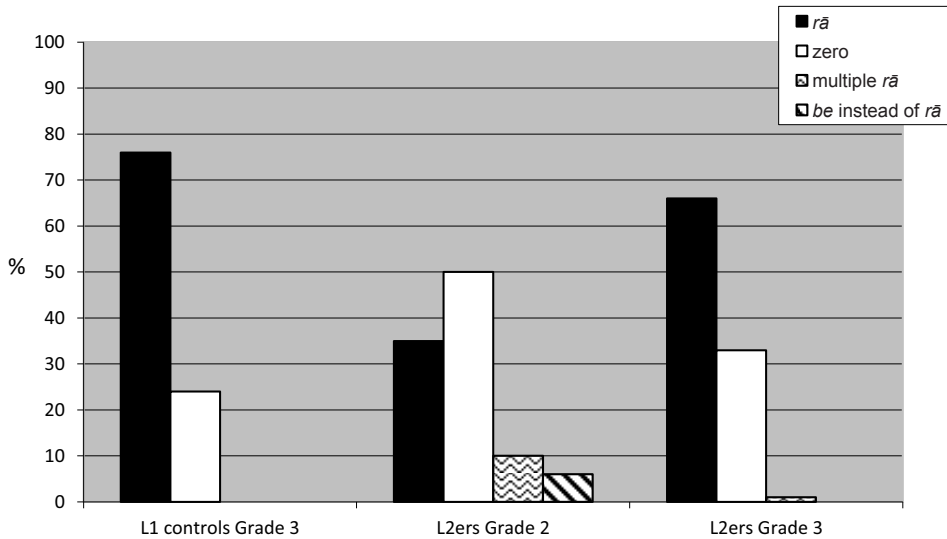


Figure 2. Direct object marking in Persian

6.3.1. Direct object marking in the second-grade L2 learners

The L2 data are quite different from the native controls (Table 4). Only a third of the second-grade L2ers’ direct objects are overtly marked with *rā* (35%, 140/405); the majority of their direct objects are zero marked (50%, 201/405). Whilst zero marking certainly is an option in Persian due to DOM, it is nevertheless striking that the learners make so much more use of zero marking than the L1 controls. We will get back to these zero markings in the discussion below (section 6.3.3).

The L2ers’ 35% *rā* on direct objects is still a sizeable percentage, and the group is relatively homogeneous here: Most of the second-graders (83%, 89 out of 107) produce *rā* on direct objects at least once. This is quite different from their indirect object marking, where only 18% (19 out of 107 children) produced an instance of targetlike indirect object marking (recall section 6.2.1).

Apart from 35% *rā* and 50% zero marking, the second-grade L2ers also produce some deviant overt direct object markings: 10% (41/405) of their direct objects are marked twice or three times with *rā*, as in (39)–(41). This is not an option in native Persian. Whilst one might want to discount the L2ers’ multiple *rā* cases as simple writing mistakes, it is nevertheless noteworthy that so many mistakes of this kind

occur in the L2 data (10%), compared to only 0.5% in the L1ers. 34 out of 107 second-grade L2 children produce multiple *rā* at least once.¹⁵

- (39) a. *moʔalem* *dars-rā-rā* *mi-dah-ad*
 teacher lesson-RA-RA IMPF-give.PRES-3SG
 ‘The teacher teaches the lesson.’ (L2 learner, non-target multiple *rā*)

- (39) b. *moʔalem* *dars-rā* *mi-dah-ad*
 teacher lesson-RA IMPF-give.PRES-3SG
 (Persian target)

- (40) a. *mā na-bāyad* *u-rā u-rā* *aziyat kon-im*
 we NEG-must (s)he-RA (s)he-RA annoy do.PRES-1PL
 ‘We must not annoy her.’ (L2 learner, non-target multiple *rā*)

- (40) b. *mā nā-bāyad* *u-rā* *aziyat kon-im*
 we NEG-must (s)he-RA annoy do.PRES-1PL
 (Persian target)

- (41) a. *man* *u-rā kheyli-rā* *dust dār-am*
 I (s)he-OBJ much-RA love have.PRES-1SG
 ‘I love her very much.’ (L2 learner, non-target multiple *rā*)

- (41) b. *man* *u-rā kheyli* *dust dār-am*
 I (s)he-OBJ much love have.PRES-1SG
 (Persian target)

Another non-target form that the second-grade L2ers use is prepositional *be* instead of suffixal *rā* (6%, 23/405), as in (42). Only a minority of children (17 out of 107) use this form.

- (42) a. *mādar-am* *be man* *ne-mi-zan-ad*
 mother-my to I NEG-IMPF-hit.PRES-3SG
 ‘My mother does not beat me.’ (L2 learner, non-target *be* instead of *rā*)

- (42) b. *mādar-am* *man-rā* *ne-mi-zan-ad*
 mother-my I-OBJ NEG-IMPF-hit.PRES-3SG
 (Persian target)

The use of *be* instead of *rā* suggests that a few second-grade children have noticed the occurrence of *be* in connection with objects in Persian in their input and are beginning to use the form themselves. They appear to be doing so, however, without

¹⁵ Sometimes there is double *rā* on the same noun, and sometimes *rā* appears twice on two adjacent words. We have at present no good explanation for this high occurrence of non-target multiple *rā* in the L2 second-graders. Could processing play a role here? These pupils are the youngest and least experienced/fluent in writing. Presumably they devote proportionally more of their processing capacity to transcription and orthographic coding than the third-grade groups need to. Temporary cognitive overload might therefore lead to more repetitions and multiple *rā*. In order to shed light on this, one should compare the second-grade L2ers with a group of age-matched L1 native Persian pupils who have as little experience in writing as they do. We leave this matter for future research.

having fully understood the Persian pattern, where *be* is restricted to indirect objects and *rā* is restricted to direct objects. As the children's L1 Balochi does not make a formal distinction between indirect and direct object marking, it is to be expected that it will take the learners some time to figure out the distributional pattern of Persian and implement it in their own L2 Persian.

The relatively low use of *rā* in general, coupled with unwarranted *rā* doubling on some direct objects, and the non-targetlike use of *be* with direct objects suggests that there is considerable confusion about Persian object marking in the second-grade L2 learners at this stage. Recall also that *be* marking on indirect objects only rarely occurs at this stage. It thus appears that a few learners have noticed the form *be* and added it to their lexicons but are not yet able to use it in a targetlike fashion, resulting in *be* overuse with direct objects and severe underuse with indirect objects.

6.3.2. Direct object marking in the third-grade L2 learners

The picture is quite different in the third-grade L2 compositions. Here, a majority of direct objects, 66% (569/866), are marked with overt *rā*, not unlike the L1ers (who, as will be recalled, used 76% overt *rā*). Nearly everyone in the group (125 out of 130 children) produces such overt *rā*. Only very few of the L2 third-graders' direct objects contain non-target multiple *rā* (1%, 9/866, produced by 7 out of 130 children).

The other non-target option that occurred in the second-grade learners, namely using prepositional *be* instead of suffix *rā*, is not attested in grade three. The L2 third-graders are thus fairly similar to their age-matched native Persian controls. Even though we do not have longitudinal data for individual learners, the comparison of the second-grade and third-grade learner groups indicates a clear development towards target Persian.

6.3.3. Zero marking on direct objects: DOM revisited

As shown in Table 4 and Fig. 2, there is a considerable difference between the groups concerning the overtness of *rā*: 76% of all direct objects produced by the L1ers show overt *rā*, but only 35% for the L2 second-graders,¹⁶ going up to 66% for the L2 third-graders. This warrants further investigation.

Recall from our discussion of DOM (section 2.2), that *rā* in adult Persian is mostly overt when the direct object is human and/or when it is specific and identifiable from context. By contrast, inanimate and unspecific or generic direct objects are only very rarely marked with *rā* in adult Persian. Our native L1 Persian third-graders fully adhere to this pattern. They use *rā* on animate direct objects with specific reference, for instance ones expressed by personal pronouns e.g. *to-rā* 'you', as in (43), as well as on inanimate direct objects with specific reference, e.g. noun phrases with a possessive determiner, e.g. 'my foodstuffs' (44), 'our clothes' (45). By contrast, the native Persian children do not use *rā* on inanimate and unspe-

¹⁶ Even if we add the second-grade L2ers' non-target multiple *rā* (10%) to their single *rā* (35%), their proportion of overt *rā* marking on direct objects (45%) falls substantially below that of the native Persian children (76%).

cific or generic direct objects, e.g. ‘good things’, ‘lunch’, or ‘delicious food’, as illustrated by the authentic examples in (46)–(48).

Overt *rā* on direct objects with specific reference:

- (43) *mādar-am* *man* *to-rā* *dust dār-am*
 mother-my I you-OBJ love have.PRES-1SG
 ‘Oh my mother, I love you.’ (L1 third grade)
- (44) *mādar-am* *khorāki-hā-yam-rā* *dar kif-am* *mi-gozār-ad*
 mother-my foodstuffs-PL-my-OBJ in bag-my IMPF-put.PRES-3SG
 ‘My mother puts my foodstuffs in my bag.’ (L1 third grade)
- (45) *mādar* *lebās-hā-ye mā-rā* *tamiz* *mi-kon-ad*
 mother clothes-PL-LINK our-OBJ clean IMPF-do.PRES-3SG
 ‘Mother cleans our clothes.’ (L1 third grade)

Zero marking (no overt *rā*) on direct objects with unspecific or generic reference:

- (46) *mādar-am* *barā-ye mā* *chiz-hā-ye khub* *mi-khar-ad*
 mother-my for-LINK we thing-PL-LINK good IMPF-buy.PRES-3SG
 ‘My mother buys good (i.e. beautiful) things for us.’ (unspecific) (L1 third grade)
- (47) *vaghti* *az madrese* *mi-āy-am*
 when from school IMPF-come.PRES-1SG
 mādar-am *nāhār* *āmāde mi-kon-ad*
 mother-my lunch prepare IMPF-do.PRES-3SG
 ‘When I come back from school, my mother prepares lunch.’ (unspecific)
 (L1 third grade)
- (48) *mādar-am* *barā-ye man* *ghazā-ye khoshmāze* *mi-paz-ad*
 mother-my for-LINK I food-LINK delicious IMPF-cook.PRES-3SG
 ‘My mother cooks delicious food for me.’ (unspecific) (L1 third grade)

For the L1ers then, the distribution of overt *rā* vs. zero marking is 76% vs. 24% (Table 4), and when looking at their direct objects in context, this distribution appears to be adultlike. We have not been able to find any ungrammatical examples with *rā* in the L1 Persian compositions, and only very few instances where an overt *rā* would have been preferable to the zero marking used by the L1 children. This is very different from the L2 learners.

The L2ers, especially in second grade, produce many more zero-marked direct objects (50%) than the L1 Persian children do (24%). In theory, it could be that the L2ers write about many more inanimate and/or unspecific objects than the L1ers, which would explain the difference in overt and zero marking between the two groups. However, an analysis of the L2ers’ productions in context indicates that this is not the case. Rather, they also write about animate objects and objects with specific reference just as the native L1 controls do.

The L2 learners zero-mark many direct objects high up on the animacy hierarchy (typically referring to humans) that have specific reference in context. Such objects would usually occur with overt *rā* in native Persian, and we will therefore refer to

them as *rā* omissions. Some examples are given in (49)–(55), where _ indicates the missing *rā*. As can be seen, the learners omit *rā* both from full noun phrase objects and from pronominal objects (e.g. *u-rā* ‘her/him’, *madare khod-rā* ‘my mother’, *man-rā* ‘me’, *bacheha-rā* ‘the children’, *moaleme khod-rā* ‘my teacher’).

- (49) a. *man* *u_* *kheyli dust dār-am*
I (s)he much love have.PRES-1SG
‘I love her/him very much.’ (L2 learner, grade 2, nontarget)
- (49) b. *man* *u-rā* *kheyli dust dār-am*
(Persian target)
- (50) a. *man* *mādar-e khod_* *dust dār-am*
I (s)he-LINK self love have.PRES-1SG
‘I love my mother.’ (L2 learner, grade 2, nontarget)
- (50) b. *man* *mādar-am-rā* *dust dār-am*
(Persian target)
- (51) a. *pedar-am* *man_* *be Chābahār bord*
father-my I to Chababar took.PAST.3SG
‘My father took me to Chababar.’ (L2 learner, grade 2, nontarget)
- (51) b. *pedar-am* *man-rā* *be Chābahār bord*
(Persian target)
- (52) a. *mā na-bāyad* *bache-hā_* *ayizat kon-im*
we NEG-must child-PL annoy do.PRES-1PL
‘We must not bother the children (i.e. the other students at school).’
(L2 learner, grade 2, nontarget)
- (52) b. *mā na-bāyad* *bache-hā-rā* *ayizat kon-im*
(Persian target)
- (53) a. *har vaght* *moʔalem-e khod_* *did-am ...*
every time teacher-LINK self saw.PAST-1SG
‘When I saw my teacher ...’ (L2 learner, grade 2, non-target)
- (53) b. *har vaght* *moʔalem-e khod-rā* *did-am ...*
(Persian target)
- (54) a. *khodā* *u_* *dust na-dār-ad*
God (s)he love NEG-have.PRES-3SG
‘God does not love her/him.’ (L2 learner, grade 3, non-target)
- (54) b. *khodā* *u-rā* *dust na-dār-ad*
(Persian target)
- (55) a. *mā bāyad ghadr-e* *moʔalem-e khod_* *be-dān-im*
we must appreciate-LINK teacher-LINK self SBJ-know.PRES-1PL
‘We must appreciate our teachers.’ (L2 learner, grade 3, non-target)
- (55) b. *mā bāyad ghadr-e* *moʔalem-e khod-rā* *be-dan-im*
(Persian target)

As shown in Table 5, non-target *rā* omissions are frequent in the beginner L2 group (grade 2), who leave out *rā* in a non-targetlike manner 30% of the time (120/405), compared to 1% in the L1 group (7/605). The learners thus underuse *rā*. Whilst not every second-grade L2 child omits *rā*, the majority of them do so some of the time (58 out of 107 children). Sentences of the type illustrated in (49), where *rā* has been omitted from a pronominal direct object, are very frequent in the L2 compositions.

	Adultlike <i>rā</i>	Adultlike zero marking	Non-adultlike zero marking (- <i>rā</i> omission)
L1-3 rd years	75.5% (457/605)	22.6% (137/605)	1.2% (7/605)
L2-2 nd years	34.5% (140/405)	19.7% (80/405)	29.8% (121/405)
L2-3 rd years	65.7% (569/866)	29.4% (255/866)	3.8% (44/866)

Table 5. Zero marking on direct objects

The third-grade L2ers also occasionally omit *rā* from direct objects (4%, 44/866), 20 out of 130 children do so. 4% is higher than the native controls (1%), but much less frequent than the second-grade L2ers (30%). Apart from these few non-target forms, the third-graders appear to have mastered Persian direct object marking. Whilst we do not have longitudinal data for the same individuals over time, a comparison of the two learner groups suggests a developmental path towards the native Persian distribution.

Readers may wonder why the 130 third-grade L2ers produce so many more direct objects in their texts (866) than the 133 native L1 controls in theirs (605), and why the L2ers produce a higher proportion of acceptable zero-marked direct objects (29%) than the native L1 controls (23%). An inspection of the compositions suggests that these differences in frequency (but not in grammaticality) are due to slightly different ways of writing (recall section 6.1): The third-grade L2ers often write short sentences with minimal variation, repeating the same subject and/or verb again and again, and only changing the direct object or the verb. The direct object here usually is an unspecific inanimate noun phrase that does not require *rā* (DOM, recall section 2.2).¹⁷ This is grammatically correct but textually repetitive. The L1ers produce such repetitive writing to a much lower degree.

But the greatest difference between the L2 learners and native controls does not concern the third-graders, but the second-graders. As mentioned above, these often omit *rā* from direct objects in a *non-targetlike* manner (Table 5). How can we explain these learner results? Some answers are suggested in the next section.

6.3.4. Discussion of direct object marking results

From a contrastive perspective, it is somewhat surprising that the beginning L2 learners (grade 2) diverge so much from the native Persian controls, since Balochi and Persian are relatively similar in their marking of direct objects. Recall from sections 2 and 3 that both languages mark direct objects with suffixes that are mor-

¹⁷ Corresponding examples in English would be: ‘My mother buys a bag for me. My mother buys a dress for me. My mother buys shoes for me. My mother sews a dress for me.’

pho-phonologically quite similar. Both languages have DOM, where objects high on the animacy hierarchy/scale (especially humans) and objects that are easily context-identifiable (specific reference) are typically marked overtly, whilst inanimate objects with unspecific reference are not.

Now, if learners simply were to transfer their Balochi grammar to Persian (relexification), we would expect to see more of the Persian pattern already in the elementary learners, since Persian direct object marking does not much differ from Balochi. By contrast, indirect object marking in the two languages is not similar, and here transfer results in 82% non-target Balochi-style and Balochi-influenced constructions in the learners (section 6.2).

What about the direct object marking in the elementary learners? Several possibilities spring to mind. Firstly, one could appeal to universal tendencies in second language acquisition to explain the learners' underuse of *rā* marking on direct objects. It is well known that learners, especially at the beginning stages, often omit functional/grammatical markers. It is also well known that learners exhibit morphological variability and optionality in their use of nominal and verbal inflection and associated lexical items. Thus, during the course of acquisition, there is a long stretch of time during which functional markers are sometimes present and sometimes absent in circumstances where they would be (more or less) obligatory for native speakers. So this might also hold for Balochi learners of Persian. Still it is noteworthy that the very same learners who optionally omit *rā* from direct objects make extensive use of all manner of functional morphology in their productions. They produce 45% *rā* on direct objects, and 82% *rā* in another context, namely on indirect objects (where Persian requires *be*, and *be* only). Thus we cannot explain the high number of zero marked direct objects as a lack of functional morphology in their interlanguage grammar or a lack of the functional morpheme *rā* in particular.

One might want to attribute the optional use of *rā* on direct objects to processing problems of some sort. Morphology must be learnt, and the L2ers have clearly learnt the form *rā*. Then why do they not use it consistently? The acquisition literature speaks of a mapping problem, a temporary breakdown between the syntax and the lexicon (e.g. Prévost and White 2000). In the context of text writing, one might further suppose that novice writers like our second-grade learners are prone to cognitive overload, under the strain of L2 sentence formulation, text planning, and handwriting, and sometimes forget to write down *rā* at the end of noun phrases. Whilst this kind of breakdown cannot be ruled out, we would expect the learners to also omit other elements (morphemes, words) in their writing. As far as we can tell, there is no such pattern.

It has been suggested to us (C. Jahani, personal communication) that the learners' omission of direct object marking might be explained via the existence of DOM. Both Balochi and Persian have obligatory overt marking of indirect objects, and transferring this requirement from the L1 to the L2 would result in constant provision of indirect object marking, which is indeed what we find in the learners, albeit largely in non-target form (*-rā* instead of *be*). By contrast, Balochi and Persian do not have obligatory but differential marking of direct objects, and transferring this 'non-obligatoriness' from the L1 to the L2 would result in optional provision of di-

rect object marking. However, if it is true that DOM in the two languages is governed by similar semantic and discourse-pragmatic constraints (sections 2.2 and 3), it remains unclear why Balochi learners of Persian would sometimes omit *rā* on Persian direct objects where *-a* is obligatory in Balochi, such as on context-identifiable animate/human objects.

Alternatively, one could entertain the possibility that Balochi and Persian are not so similar in their marking of direct objects after all, and that subtle contrastive differences between the two languages lie at the heart of the learners' *rā* omissions. Recall that in Balochi, overt object marking (*-a*) can only occur in constructions in the nominative-accusative system (present tense), but not in the ergative system (mainly past tense). On an L1-transfer approach, this split might carry over to the productions of Balochi beginning learners of L2 Persian. There might thus arise an interaction between tense and overt object marking in L2 Persian, where *rā* would be omitted more often in a clause in the past tense than in a sentence in the present tense. Why? Because the L1 would have overt direct object marking in the present tense, but not in the past tense.

We have explored this possibility post-hoc in our data, but the results are inconclusive. As the compositions were written on the topic "Mother", the children predominantly wrote in the present tense, describing their mother, her characteristics and activities, their own school-day, and their return home, etc. For this reason, the compositions only contain a few clauses in the past tense. In those that exist, we do find *rā* omissions (e.g. (51) 'My father took me to Chabahar.', (53) 'When I saw my teacher ...'). There does appear to be a higher proportion of *rā* omissions in L2 past tense clauses than in L2 present tense clauses, but the small number of past tense clauses overall precludes any meaningful statistical investigation.

The great majority of clauses in the compositions are in the present tense, for native Persian pupils and L2 learners alike. In these present tense clauses, we also find *rā* omissions for the learners (e.g. (49), (50), (52)), but in other present tense clauses, *rā* is provided, or a non-target form (multiple *rā* or *be* instead of *rā*) is used. In short, we have not been able to find a straightforward interaction of tense and overtness of *rā* marking in our learner data, but then again, the compositions are not ideal for testing this possibility.¹⁸ We aim to investigate this issue further in future work. A new learner corpus is currently being collected, with compositions on the topic "How did you spend your last summer?", which is intended to induce pupils to write predominantly in the past tense. We hope that a comparison of these compositions with the current material (predominantly in the present tense) will enable us to do more justice to the question of tense and overtness of *rā* marking.

¹⁸ Some expressions in Balochi can be constructed ergatively irrespective of tense (e.g. the verb *dust* 'love', recall footnote 8), and the direct object in these constructions is not overtly marked. If we extend the idea of L1 transfer from these constructions to L2 Persian, we would expect that the learners should not use any overt direct object marker when constructing the corresponding sentences in Persian, which would result in non-target *rā* omissions. It is indeed the case that the second-grade learners often omit *rā* from sentences with *dust* 'love' (see e.g. examples (49)–(50), and the third-grade learners sometimes do so too (e.g. (54)). Overall however, there are too few examples to draw any strong conclusions.

7. Conclusion

In this paper, we have investigated the acquisition of object marking in the L2 Persian writing of Balochi-speaking children aged eight to eleven years. 107 elementary learners in second grade (after 1000 hours of exposure to Persian), 130 more proficient learners in third grade (after 1500 hours of exposure), as well as 133 native Persian L1 children of similar socio-economic background wrote a composition on a set topic. In these compositions, an analysis of the frequencies and types of indirect and direct object marking revealed large differences between the groups. The native Persian children basically performed according to adult norms of object marking (for 96% of the indirect objects and 98% of the direct objects).

This was in stark contrast to the second-grade L2ers, who predominantly employed non-target forms of object marking in their Persian, but which closely resembled Balochi constructions: 68% of their indirect objects were not marked with prepositional *be* but with suffixal *rā* (resembling Balochi *-a*). Nearly every second-grade learner used this non-target form. Only in a minority of cases (15%) did the learners use targetlike prepositional *be* as the only object marker, whilst some created novel combinations of simultaneous *be* and *rā*. We suggested that these findings could be interpreted as initial transfer of the L1 Balochi suffixal object marking construction to the interlanguage grammar, with subsequent incipient acquisition of the target prepositional construction by a few learners. The more advanced third-grade L2ers used prepositional *be* marking on 85% of their indirect objects. This was interpreted as the learners having largely rid themselves of Balochi-style suffixal marking and having acquired the Persian target form at this stage, at least as could be determined from their L2 writing.

Concerning direct object marking, the second-grade L2ers again differed starkly from the native controls: Whilst the native Persian children mostly used overt *rā* marking (76%) on their direct objects, the L2 learners only did so 35% of the time. The finding for native Persian is interesting, because it, for the first time as far as we are aware, provides figures for differential object marking (DOM), i.e. for the frequency distribution of overt direct object marking in Persian: 76% overt vs. 24% zero, at least as can be ascertained from the ca. 20,000 words of the native written genre sampled here. The elementary L2 learners exhibited a different pattern: Whilst L1ers and L2ers used the same form they did so to very different degrees. This was surprising, as Persian and Balochi are usually said to have the same type of differential object marking (DOM) on direct objects: suffixal *rā/-a*, its overtness being constrained by semantic and discourse-pragmatic factors, with human/animate and specific/context-identifiable objects marked overtly, and inanimate, unspecific objects typically remaining bare. On an L1-transfer approach to acquisition, one would predict few learning problems in an area where the two languages are strictly parallel. However, our learners presented a different picture: Elementary second-grade L2ers often omitted direct object marking in a non-targetlike fashion (30%) or, less commonly, created novel ways of marking objects, different from their L1 Balochi and different from native Persian. The more advanced third-grade L2ers, however, used direct object marking in a fashion

much like their native Persian peers. We discussed a number of alternative explanations for the surprising omissions of object marking in the elementary learners and suggested that DOM in Balochi and Persian might not be as parallel after all. Due to its split ergativity system, Balochi only marks direct objects in certain tenses, which reduces the frequency of overt object marking compared to Persian (which does not have split ergativity). As our learner data were mainly in the present tense, we could not fully test the explanatory strength of this tense/object marking interaction, but hope to do so in future work.

Overall, we could confirm that object marking, as has been suggested in the literature, may be a problematic area for L2 learners of Persian, but we could also see a major improvement between the two learner groups with just one extra year of schooling. However, all our data concern written Persian. It remains to be seen how quickly Persian object marking is acquired and mastered by second language learners in the spoken modality, and how learners would perform on grammaticality or acceptability tests that involve object marking.

Abbreviations

1	first person	NEG	negation
2	second person	NOM	nominative
3	third person	OBL	oblique
Bal.	Balochi	OBJ	object
CL	clitic	PAST	past
COP	copula	Per.	Persian
DEF	definite	PL	plural
DOM	differential object marking	PRES	present
IMPF	imperfective	SBJ	subjunctive
INDEF	indefinite	SG	singular
LINK	linking particle (ezāfe)		

References

- Bossong, Georg (1985). *Empirische Universalienforschung: Differentielle Objektmarkierung in den neu-iranischen Sprachen*. Tübingen: Gunter Narr.
- Bowles, Melissa, and Montrul, Silvina (2009). "Instructed L2 acquisition of Differential Object Marking in Spanish." In: Leow, Ronald, Campos, Héctor, and Lardiere, Donna (eds), *Little Words: Their History, Phonology, Syntax, Semantics, Pragmatics and Acquisition*. Washington, DC: Georgetown University Press, pp. 199–210.
- Browne, Wayles (1970). "More on Definiteness Markers: Interrogatives in Persian." *Linguistic Inquiry* 1, pp. 359–363.
- Comrie, Bernard (1989). *Language Universals and Linguistic Typology*. 2nd ed. Oxford: Blackwell.
- Dabir-Moghaddam, Mohammad (1990). Pirāmun-e rā dar zabān-e fārsi [On postposition *rā* in Persian]. *Iranian Journal of Linguistics* 7, pp. 2–60.
- Farrell, Tim (1990). *Basic Balochi: An Introductory Course*. Naples: Istituto Universitario Orientale.
- Foroodi-Nejad, Farzaneh (2011). *Towards the Identification of Linguistic Characteristics of Specific Language Impairment in Persian*. PhD thesis. Edmonton: University of Alberta.
- Ganjavi, Shadi (2007). *Direct Objects in Persian*. PhD thesis. University of Southern California.
- Ghomeshi, Jila (1996). *Projection and Inflection: A Study of Persian Phrase Structure*. PhD thesis, University of Toronto.

- Ghomeshi, Jila (1997). "Topics in Persian VPs." *Lingua* 102, pp. 133–167.
- Ghomeshi, Jila (2003). "Plural Marking, Indefiniteness and the Noun Phrase." *Studia Linguistica* 52, pp. 47–74.
- Jahani, Carina (2003). "The Case System in Iranian Balochi in a Contact Linguistic Perspective." In: Jahani, Carina and Korn, Agnes (eds) *The Baloch and their Neighbours: Ethnic and Linguistic Contact in Balochistan in Historical and Modern Times*. Wiesbaden: Reichert, pp. 113–132.
- Jahani, Carina and Korn, Agnes (2009). "Balochi." In: Windfuhr, Gernot (ed.), *The Iranian Languages*. London and New York: Routledge, pp. 634–694.
- Jarvis, Scott and Odlin, Terence (2000). "Morphological Type, Spatial Reference, and Language Transfer." *Studies in Second Language Acquisition* 22(4), pp. 525–556.
- Karimi, Simin (1989). *Aspects of Persian Syntax, Specificity, and the Theory of Grammar*. PhD thesis, University of Washington.
- Karimi, Simin (1990). "Obliqueness, Specificity, and Discourse Functions: *rā* in Persian." *Linguistic Analysis* 20(3–4), pp. 139–191.
- Karimi, Simin (1996). "Case and Specificity. Persian *rā* Revisited." *Linguistic Analysis* 26, pp. 174–194.
- Karimi, Simin (2003). "Object Positions, Specificity and Scrambling." In: Karimi, Simin (ed.), *Word Order and Scrambling*. Oxford: Blackwell, pp. 91–125.
- Keshavarz, Mohammad Hossein (2007). "Morphological Development in the Speech of a Persian-English Bilingual Child." *Journal of Psycholinguistic Research* 36, pp. 255–272.
- Korn, Agnes (2009). "The Ergative System in Balochi from a Typological Perspective." *Iranian Journal of Applied Language Studies* 1(1), pp. 43–79.
- Lardiere, Donna (2009). "The Contrastive Analysis of Features in Second Language Acquisition." *Second Language Research* 25(2), pp. 173–227.
- Lazard, Gilbert (1982). "Le morpheme *rā* en Persan et les relations actanciennes." *Bulletin de la société de linguistique de Paris* 73(1), pp. 177–208.
- Lazard, Gilbert (1984). "Actance Variations and Categories of the Object." In: Plank, Frans (ed.), *Objects: Towards a Theory of Grammatical Relations*. London: Academic Press, 269–292.
- Lazard, Gilbert (1992). *A Grammar of Contemporary Persian*. Mazda: Costa Mesa, CA and New York. [English translation of *Grammaire du persan contemporain* 1957].
- Lyons, Christopher (1999). *Definiteness*. Cambridge: Cambridge University Press.
- Mirza, Zahra (2000). *Yādgiri-ye maqulhā-ye kār-kardi dar zabānhaye fārsi va armani tavasot-e faragirān-e an zabānhā* [The acquisition of functional categories in Persian and Armenian by learners of those languages] MA thesis. Esfahan: Esfahan University.
- Mohammadi, Somayeh (2009). *Tahlil-e khatāhā-ye nahvi-ye dāneshāmuzān-e Baluchzabān dar yādgiri-ye zabān-e fārsi* [The analysis of syntactic errors of Baloch pupils learning Persian]. MA thesis. Tehran: Allāme Tabātabā'i University.
- Prévost, Philippe and White, Lydia (2000). "Missing Surface Inflection or Impairment in Second Language Acquisition? Evidence from Tense and Agreement." *Second Language Research* 16, pp. 103–134.
- Rahimian, Jalal and Hajiani, Farrokh (2009). "Semantic-Pragmatic Function of *rā* in Persian: A Diachronic and Synchronic Study." *Journal of Indo-European Studies* 37(3–4), pp. 1–22.
- Samiian, Vida (1983). *Structure of Phrasal Categories in Persian: An X-bar Analysis*. PhD thesis. Los Angeles: UCLA.
- Shokouhi, Hussein and Kipka, Peter (2003). "A Discourse Study of Persian *rā*." *Lingua* 113, pp. 953–966.
- Windfuhr, Gernot (1979). *Persian Grammar: History and State of its Study*. New York: Mouton.
- Windfuhr, Gernot (1989). "Persian." In: Comrie, Bernard (ed.), *The World's Major Languages*. London: Routledge, pp. 523–546.
- Windfuhr, Gernot (1992). "Cases." In: Yarshater, Ehsan (ed.), *Encyclopaedia Iranica* 5. New York: Bibliotheca Persica Press, 25–37.

An Elusive Particle in Persian: The Use of *Magar* in *Tārikh-e Beyhaqi*

Forogh Hashabeiky, Uppsala University

forogh.hashabeiky@lingfil.uu.se

Mahmoud Hassanabadi, University of Sistan and Baluchestan, Zahedan

mahmoud.hassanabadi@yahoo.com

Abstract

Morphologically, *magar* is made up of *ma*+*agar*, which literally means “not+if”. It fulfils several syntactic and pragmatic functions which have shifted over the course of time. This paper focuses on the use of *magar* in *Tārikh-e Beyhaqi* (11th century). The study will be conducted within the framework of a functional-structural approach with a pragmatic orientation. We argue that *magar* is an elusive particle, the precise syntactic function of which is governed by the context. We also argue that some instances of *magar* fulfil a pragmatic function that fits into the theory of indirect speech acts.

Keywords: Classical Persian, *magar*, *Tārikh-e Beyhaqi*, functional-structural approach, indirect speech acts

1. Introduction

Morphologically, the particle *magar* is made up of the prohibitive prefix *ma*- and the conditional conjunction *agar*.¹ Literally *magar* means “not+if”. As a part of speech, it has variously been labelled a preposition, an interrogative word, a conjunction, and an adverb. We have chosen to describe *magar* as a particle, in accordance with Crystal’s definition by which a particle is an invariable item that has a grammatical function, but cannot readily fit into a standard classification of parts of speech Crystal (1997: 279–280).

Our preliminary observations of the use of *magar* in Contemporary Persian indicate that *magar* is an elusive particle with several syntactic and pragmatic functions. Yet this particle has not been the subject of any detailed study. It has been briefly mentioned in some grammars, textbooks, and lexicographical works written in Persian and in other languages. From some of these writings, it also appears that the functional domain of *magar* has changed over the course of time. We believe that a detailed description of the use of *magar* in Classical Persian texts is the first step towards further diachronic and synchronic descriptions of this particle, and is a prerequisite for understanding the shift in the functional domain of *magar*.

This study aims to investigate the use of *magar* in *Tārikh-e Beyhaqi* by Abu al-Fazl Beyhaqi (995–1077). *Tārikh-e Beyhaqi*, also known as *Tārikh-e Mas’udi*, mainly covers the historical events and personalities of the reign of Sultan Mas’ud of Ghazna (1030–1041). We have chosen *Tārikh-e Beyhaqi* for our study as there is a general consensus among scholars that *Tārikh-e Beyhaqi* is one of the best representatives of Early Classical Persian. It is also considered the main prose work to

¹ The prefix *ma*- has developed from the Middle Persian prohibitive particle *ma* ‘not’ and *agar* ‘if’ [NP *magar*] and means “perhaps, lest” (MacKenzie 1971: 6, 53).

mark the threshold of the stylistic change of Persian prose from *mursal* ‘bald’ to *fanni* ‘artistic’. However, what has made *Tārikh-e Beyhaqi* an especially interesting object for our study is that it is a historiographical work with literary features. It is full of detailed descriptions and lively conversations with elements from the colloquial language of its time.² These conversations are especially interesting for a linguistic study with a pragmatic orientation.

For our study, we have used the latest critical edition of *Tārikh-e Beyhaqi* by Mohammad Jafar Yahaqqi and Mehdi Sayyedi (1388). This edition is based on seventeen manuscripts and the Calcutta edition (1862) by Morley, as well as on the Adib Pishavari edition (1307), employing the eclectic method.³ The editors have also taken into consideration the Saeed Nafisi edition in three volumes (1319, 1326, 1332), the Ghani and Fayyaz edition (1324) by Qasem Ghani and Ali Akbar Fayyaz, and especially the Fayyaz edition (1350) by Ali Akbar Fayyaz.

The retrieved instances of sentences including *magar* in this edition have been cross-checked with the same sentences in the critical edition of this work by Fayyaz (1350). We have found several such sentences with two different readings in these two editions. None of these different readings have anything to do with the use of *magar*.⁴ As for the translation of the given instances, we have checked our translation against that of *The History of Beyhaqi*, by C. E. Bosworth and revised by Mohsen Ashtiany (2011), partly to find out how *magar* has been interpreted by other readers, and partly to polish our own translations.

2. Theory and method

The study employs a structural-functional approach, within which various syntactic functions of *magar* will be explored. Some instances of *magar*, are, however, interesting even from the point of view of pragmatics. We believe that a better understanding of such instances of *magar* will be possible with reference to pragmatics and the theory of indirect speech acts as presented by John R. Searle.

The term “speech act” is derived from the work of the British philosopher John L. Austin and the American philosopher John R. Searle. According to the theory of speech acts, many utterances do not convey information, but perform actions instead. This means that an act is performed as soon as the utterance has been made. Austin applies the term illocutionary acts to such acts. A promise takes place when someone says: “I promise”. Making statements, asking questions, issuing orders, promising, giving reports, greeting, and warning are some examples of speech acts. In addition to a set of necessary and sufficient non-linguistic conditions, a successful

² For more information on the style and language of *Tārikh-e Beyhaqi*, see Khatibi (1366: 130–135); Bahar (1337: 67–87); (Beyhaqi 2011 vol. I: 70–79); and (Daniel 2012: 127–129). Together with *Gulistān of Sa’di*, the prose style of *Tārikh-e Beyhaqi* is considered to have influenced generations of Persian writers even in modern times. The modern Iranian poet Ahmad Shamlu, for example, is usually described as having been inspired by *Tārikh-e Beyhaqi*, see Langarudi (1377: 151–152).

³ For detailed information on the extant manuscripts of *Tārikh-e Beyhaqi*, the quality of the manuscripts used in this edition, and the method employed, see Beyhaqi (1388: 111–134).

⁴ A number of these different readings have been mentioned in footnotes and in relation to instances given in this paper.

speech act requires a set of devices indicating linguistic function like word order, stress, intonation, punctuation, the mood of the verb, and a set of so-called performative verbs. Performing the speech act of promising, for example, demands the performative verb of “promise” in first person present indicative and usually at the beginning of the sentence (Austin 1961: 106–109; Searle 1965: 115–118). In a direct speech act:

The speaker utters a sentence and means exactly and literally what he says. In such cases the speaker intends to produce a certain illocutionary effect in the hearer, and he intends to produce this effect by getting the hearer to recognize this intention to produce it, and he intends to get the hearer to recognize this intention in virtue of the hearer’s knowledge of the rules that govern the utterance of the sentence (Searle 1975: 161).

There is, however, another category of speech acts, which are described as indirect speech acts by Searle. In indirect speech acts, the speaker may utter a sentence and mean what he says, but also mean something more. A request, for example, may be made with a statement, as in “I want you to shut the door”, or by using an imperative construction, as in “Shut the door!”, but these can be interpreted as impolite and inappropriate in many situations. The same request can be made with a question, as in “Could you shut the door?” or “Would you mind shutting the door?”, or simply by using a statement, as in “The door is open”. In such utterances the speaker intends to produce a certain perlocutionary effect in the hearer and cause the hearer to feel persuaded, ashamed, respected, warned, etc. A successful indirect speech act, also called a perlocutionary act, requires a set of devices indicating linguistic function and a sufficient non-linguistic condition to make it possible for the hearer to understand the indirect speech act when the sentence he hears and understands means something more than what is said. Searle maintains:

In indirect speech acts the speaker communicates to the hearer more than he actually says by way of relying on their mutually shared background information, both linguistic and nonlinguistic, together with the general powers of rationality and inference on the part of the hearer (Searle 1975: 162).

Our preliminary observations of the use of *magar* in *Tārikh-e Beyhaqi* indicate that in some instances *magar* has been used in order to produce a perlocutionary effect in the hearer. A correct interpretation of such instances of *magar* requires an analysis of the shared background information of the language users, the inferences they make, and the assumptions they hold.

To conduct our study, we have extracted all instances of *magar* in *Tārikh-e Beyhaqi*, using the digital version of the Yahqqi and Sayyedi edition. These instances are first classified into different patterns in accordance with their syntactic functions. Those instances of *magar* that allow for a pragmatic analysis have been followed by an analysis using the theory of indirect speech acts. This classification will be preceded by a presentation of previous writings on *magar* in Persian and in other languages, and followed by some concluding remarks.

3. *Magar* as presented in lexicographical works, grammars, and textbooks in Persian and other languages

In Persian grammatical and lexicographical writings, *magar* is usually described in semantic terms using examples from Classical Persian texts. One typical example is Dehkhoda's dictionary *Loghatnāme* (1325–1354). In Dehkhoda's account, different uses of *magar* have been described and organized in eleven categories based on instances retrieved from different types of Classical Persian texts from the 10th to the 16th centuries.⁵ As a part of speech, *magar* has been labelled as either a preposition or an adverb:

(1) as a preposition equivalent to the Arabic *illā* 'except' to express exception (1.1), and as conditional conjunction 'unless' (1.2):

(1.1)

<i>va</i>	<i>az ān</i>	<i>haftād</i>	<i>hezār</i>	<i>zangi</i>	<i>kas</i>	<i>jān</i>	<i>na-bord</i>
and	of that	seventy	thousand	slave	person	life	NEG-take.PAST.3SG

magar *andak=i*
except few=IND

'Of those seventy thousands slaves⁶ no one survived, except for a few.'

(1.2)

<i>va</i>	<i>hajjāj</i>	<i>sowgand</i>	<i>khvord</i>	<i>ke</i>	<i>u</i>	<i>rā</i>	<i>az</i>
and	Hajjaj	oath	eat.PAST.3SG	CLM	PN.3SG	OM	from

<i>dār</i>	<i>foru-na-gir-ad</i>	<i>magar</i>	<i>mādar=ash</i>
gallows	PREV-NEG-take.PRES-3SG	unless	mother=PC.3SG

shafā'at *kon-ad*
intercession do.PRES-3SG

'And Hajjaj took an oath (swore) that he would not take him down from the gallows unless his mother interceded for him.'

(2) either optatively to express a wish or hope (2.1), or as the adverbial of possibility (2.2):⁷

(2.1)

<i>goft</i>	<i>magar</i>	<i>khodāy</i>	<i>ta'ālā</i>	<i>in</i>	<i>doshman=e</i>	<i>shomā</i>
say.PAST.3SG	may	God	high.exalted	this	enemy=EZ	PN.2PL

halāk *kon-ad*
destroyed do.PRES-SG

'He said: may God, high exalted, destroy this enemy of yours.'

⁵ A number of these examples are retrieved from other Persian lexicographical works such as *Borhān-e qāte'*, *Farhang-e nafisi* (*Nāẓem al-aṭebbā*), *Farhang-e ānendrāj*.

⁶ Black slaves from Zanzibar.

⁷ There is a vague and shifting boundary between these two uses of *magar*. It is not always easy to determine if *magar* is being used in an optative sense or to express modality. To determine the exact sense of a particular instance of *magar*, one must rely heavily on the context.

(2.2)

andishid *ke* ***magar*** *hanuz* *gabr* *bāsh-ad*
 think.PAST.3SG CLM **might** still Zoroastrian be.PRES-3SG
 ‘He thought that he might still be a Zoroastrian.’

(3) in the sense of the disjunctive *yā* ‘or’

majles *ast* *in* ***magar*** *behesht=e* *barin*
 gathering be.PRES.3SG this **or** paradise=EZ high
 ‘Is this a gathering or the high paradise?’

(4) interrogatively with an affirmative verb when a negative answer is expected (4.1), or with a negative verb when an affirmative answer is expected (4.2):

(4.1)

magar *āsudegi* *bar* *mā* *ḥarām* *ast*
IP tranquility on PN.1PL forbidden be.PRES.3SG
 ‘Tranquility is not forbidden for us, is it?’

(4.2)

chun *gerd* *āmadan=e* *khalq* *mowjeb=e* *pādshāhi*
 as round come.INF=EZ people prerequisite=EZ kingdom

ast *to* *khalq* *rā* *cherā* *parishān* *mi-kon-i*
 be.PRES.3SG PN.2SG people OM why split IPFV-do.PRES-2SG

magar *sar=e* *pādshāhi* *na-dār-i*
IP head=EZ kingdom NEG-have.PRES-SG
 ‘As the union of a people is the prerequisite for a kingdom, why do you split them?
 Don’t you have the intention of becoming a king?’

(5) in the sense of the intensifier *hamānā* ‘certainly, surely’⁸

solṭān *rā* *khāṭer* *oftād* *ke* ***magar*** *ḥilat=i*
 Sultan OM mind fall.PAST.3SG CLM **INTS** trick=IND

ast *tā* *chiz=i* *be-st-ān-ad*
 be.PRES.3SG so.that thing=IND SBJN-get-CAUS.PRES-3SG
 ‘It fell into the mind of the Sultan that there must [certainly] be a trick to getting something from him.’

(6) in the sense of *guyā/guyi*, *pendāri* ‘as if’, *zāheran* ‘apparently’

magar *modām* *dar* *in* *faṣl* *khāk* *mast* *bov-ad*
apparently constantly in this season soil drunk be.PRES-3SG

as bas *ke* *bar* *vei* *riz-and* *jor’e-hā=ye* *modām*
 so much CLM on it pour.PRES-3PL gulp-PL=EZ wine
 ‘Apparently the soil is constantly drunk in this season, because of all the wine poured on it’

⁸ This type of *magar* is usually preceded by the clause linkage marker *ke* or the subordinate conjunction *tā*. In both cases either *magar* or *ke/tā* can be omitted without any problem.

(7) in the sense of ‘it happened/by chance/once’⁹

be herāt pādeshāh=i bud nām=e u shemirān...
in Herat king=IND be.PAST.3SG name=EZ PN.3SG Shemiran...

magar *ruz=i shāh shemirān bar manzare neshaste*
by.chance day=IND king Shemiran on veranda sit.PSTP

bud va bozorg-ān pish=e u
be.AUX.PAST.3SG and noble-PL before=EZ PN.3SG

‘There was a king in Herat, called Shemiran... Once, King Shemiran sat on the veranda; with the nobles before him.’

(8) In the sense of the noun ‘doubt’

pir=e tariqat goft niāzmand rā radd
sheikh=EZ order say.PAST.3SG needy OM rejection

nist va dar pas=e divār=e niāz magar
NEG.be.PRES.3SG and in behind=e wall=EZ need **doubt**

nist

NEG.be.PRES.3SG

‘The Sheikh of the order said: there is no rejection from the needy and there is no doubt behind the wall of need.’

(9) in the sense of ‘only’

chehel sāl sar bar bālin na-nehād va andar ferāsh
forty year head on pillow NEG-put.PAST.3SG and in bed

na-khoft magar be ta'abbod=e izad=e ta'ālā
NEG-sleep.PAST.3SG **only** to praying=e God=EZ high.exalted

mashghul bud

occupied be.PAST.3SG

‘For forty years he didn’t put his head on a pillow and didn’t sleep in a bed and was only occupied with praying to God high exalted.’

(10) in the sense of ‘only if’

magar *to ruy be-push-i o fetne*
only.if PN.2SG face SBJN-cover.PRES-2SG and temptation

bāz-nesh-ān-i ke man qarār
PREV-quench-CAUS.PRES-2SG CLM PN.1SG intention

na-dār-am ke dide az to be-push-am
NEG-have.PRES-1SG CLM eye from you SBJN-cover.PRES.1SG

‘It is only up to you to cover your face and quench the temptation, as I have no intention of turning my eyes away from you.’

⁹ According to Natel Khanlari, for opening a narrative, see below.

(11) in the sense of ‘it will be good if...’

<i>loqmān=e</i>	<i>ḥakim</i>	<i>andar</i>	<i>ān</i>	<i>qāfele</i>	<i>bud</i>	<i>yek=i</i>	<i>az</i>
Loqman=EZ	wise	in	that	caravan	be.PAST.3SG	one=IND	of

<i>kārvāni-ān</i>	<i>goft</i>	<i>magar</i>	<i>in-ān</i>	<i>rā</i>
traveller-PL	say.PAST.3SG	it.will.be.good.if	this-PL	OM

<i>naṣiḥat</i>	<i>kon-i</i>	<i>goft</i>	<i>darigh</i>	<i>bāsh-ad</i>	<i>kalame=i</i>
advice	do.PRES-2SG	say.PAST.3SG	waste	be.PRES-3SG	word=IND

<i>ḥekmat</i>	<i>bā</i>	<i>ishān</i>	<i>goftan</i>
wisdom	with	PN.3PL	say.INF

‘Loqman, the wise, was with the caravan. One of the travellers said: it will be good if you could give them some pieces of advice. He said: it would be a waste to give them any word of wisdom.’

Having studied more than 150 works that cover Early Classical Persian writings up to the early 13th century, Natel Khanlari recognizes only four uses of *magar*, and these are already covered by Dehkhoda: to express exception, to express possibility, as an interrogative word with a negative verb when an affirmative answer is expected, and for opening a narrative to emphasize evidentiality (Natel Khanlari 1366: 418–20; 1373: 256–257). According to Dehkhoda and Natel Khanlari, *magar* is primarily used in Classical Persian to express exception.¹⁰

In most writings on New Persian in other languages, *magar* is presented very briefly, and mainly with examples from Modern Persian (Jensen 1931: 254; Lazard 1992: 212; Thackston 1993: 174–175; and Roberts 2009: 225–226). There are some exceptions, such as *La langue des plus anciens monuments de la prose persane* by Gilbert Lazard (1963), which is based on Early Classical Persian texts up to the 12th century, including *Tārikh-e Beyhaqi*. In this work, for example, Lazard (1963: 487) mentions *magar* only once and in connection with combined conjunctions; as *magar ki* ‘unless’. Another work in this category is *Persische Grammatik mit Litteratur, Chrestomathie und Glossar* by Carl Salemann and Valentin Shukovski (1889). In this work, several functions of *magar*, as interrogative particle, preposition of exception, conditional conjunction, and modal particle of probability, have been mentioned (Salemann and Shukovski 1889: 85, 127).

As seen above, Dehkhoda’s dictionary presents a wide variety of uses of *magar* in Classical Persian texts belonging to the 10th–16th centuries. The problem with Dehkhoda’s description is that some different uses of *magar* have been combined within one and the same category, while some other instances can easily be moved from one category to another. But despite sometimes being eclectic, it provides us with a rough frame of reference. No other description of *magar* includes so many different uses of this particle.

¹⁰ *magar* is mentioned in several other writings in Persian. None of these works presents any new semantic field for *magar* than what is presented by Dehkhoda and Natel Khanlari. See for instance, Moin (1347: 4324–4325); Nafisi (2535: 3482); Mashkur (2535: 57, 157, 172); Bahar (1337: 304); and Farshidvard (1375: 36).

In our study, we have tried to avoid classifying different uses of *magar* within one and the same pattern as far as possible. Each pattern is followed by a pragmatic analysis using the theory of indirect speech acts (perlocutionary acts), whenever such an analysis has been possible.

However, we must admit that it has not been an easy task to determine the exact semantic scope of all the extracted instances of *magar* that function as modal particles. There are instances that can be interpreted in several senses. This ambiguity is partly due to the fact that we are dealing with a written context where some devices indicating linguistic function, like intonation, are absent, and partly due to the elusive nature of *magar*, which functions differently in different contexts. In such cases, we have chosen to mention alternative possible interpretations in footnotes.

4. Results of the investigation

As mentioned above, all instances of *magar* have been extracted. This has resulted in 82 instances, four of which occur in verses. These latter instances are not included in our analysis as the verses in question were not composed by Beyhaqi. Moreover, Classical Persian verses are not especially reliable objects of linguistic analysis because of metrical considerations. The remaining 78 instances have been classified into different patterns in accordance with the structural-functional approach as follows:

Pattern 1: as a preposition to express exception with affirmative and negative verbs (21 sentences):

(12) P. 141

<i>man</i>	<i>hame=ye</i>	<i>shoghl-hā</i>	<i>bedu</i>	<i>khvāh-am</i>	<i>sepord</i>
PN.1SG	all=EZ	affair-PL	to.PN.3SG	shall.PRES-1SG	entrust.SHIF

magar *neshāt* *o* *sharāb* *o* *chugān* *o* *jang*
except convivial.session and wine and polo and contest
 ‘I shall entrust all affairs to him except for arranging convivial sessions of wine drinking, polo-playing, and equestrian contests.’

(13) P. 560

<i>va</i>	<i>chonin</i> ¹¹	<i>aḥvāl</i>	<i>kas</i>	<i>az</i>	<i>dabir-ān</i>	<i>vāqef</i>
and	such	event.PL	person	of	secretary-PL	aware

<i>na-bud-i</i>	<i>magar</i>	<i>ostād=am</i>	<i>bu naṣr</i>
NEG-be.PAST.3SG-IPFV	except	master=PC.1SG	Bu Nasr

‘And none of the secretaries knew about such events, except my master Bu Nasr.’

(14) P. 351

<i>aknun</i>	<i>hame</i>	<i>bar</i>	<i>jāy</i>	<i>and</i>	<i>magar</i>	<i>ḥasanak</i>
now	all	on	place	be.PRES.3.PL	except	Hasanak

‘Now, all of these persons are still alive, except Hasanak.’

¹¹ Fayyaz, p. 734: *bar chonin*.

Pattern 2: as a conditional conjunction to express a prerequisite for the fulfilment of another action expressed in the main clause either with a negative verb (15) or with an affirmative verb (16) (2 sentences):

(15) P. 130

be hich ḥāl in rāst na-yā-yad magar u
to no circumstance this right NEG-come.PRES-3SG unless PN.3SG

rā bedān daraje bar-i ke az
OM to.PN.3SG level take.PRES-2SG CLM from

avval bud
first be.PAST.3SG

‘Under no circumstance can this ever come about, unless you take him down to the level he was at before.’

(16) P. 47

magar ān pādshāh rā sharm ā-yad vagarna shomā
unless that king OM remorse come.PRES-3SG otherwise you

bar shorof=e halāk id
on edge=EZ perdition be.PRES.2PL

‘Unless that king experiences some feelings of remorse, you will be on the edge of perdition.’

Pattern 3: As a modal particle to express a hope/wish (7 sentences):

(17) P. 187

kh^vānande-gān¹² magar ma ‘zur dār-and va ‘ozr=e
reader-PL may forgiven hold.PRES-3PL and excuse=EZ

man be-pāzir-and
PN.1SG SBJN-accept.PRES-3PL

‘And may the readers forgive me and accept my excuses’¹³

(18) P. 55

va man in foṣul az ān jahat rānd-am ke
and PN.1SG this chapter.PL of that reason write.PAST-1SG CLM

magar kas=i rā be kār ā-yad¹⁴
hoping person=IND OM to use come.PRES-3SG

‘And I wrote these chapters hoping that someone might find it useful’

¹² Fayyaz, p. 246: *va kh^vānandegān*

¹³ In Beyhaqi (2011 vol. I: 290), this instance of *magar* has been translated with a combination of ‘hope’, ‘may’, and ‘perhaps’ as follows: “I hope that readers may perhaps forgive this and accept my excuses”. This is a good example of the elusive nature of *magar*.

¹⁴ *magar* here can even be interpreted as ‘perhaps, probably’.

Pattern 4a: As a modal particle to expresses a probability/possibility (34 sentences):¹⁵
(19) P. 151

ruz=e *ādine* *bude* *ast* *va* *dāneste* *bude*
day=EZ Friday be.PSTP be.AUX.PRES.3SG and know.PSTP be.PSTP

ast *ke* *khodāvand* *rāy=e* *shekār* *karde*
be.AUX.PRES.3SG CLM lord intention=EZ hunting make.PSTP

ast *magar* *bedān* *sabab* *na-yāmade*
be.AUX.PRES.3SG **probably** to.that reason NEG-come.PSTP

ast
be.AUX.PRES.3SG¹⁶

‘It was Friday and he might have known that the lord intended to go hunting; he had probably not come for that reason.’

(20) P. 142

javāb *dād* *ke* *kh’āje* *to* *rā* *dar-kh’āst*
reply give.PAST.3SG CLM kh’āje PN.2SG OM PREV-ask.PAST.3SG

ke *magar* *bar* *man* *e’temād* *na-dāsht*
CLM **perhaps** on PN.1SG faith NEG-have.PAST.3SG
‘He replied: Kh’āje asked for you, as he perhaps had no faith in me.’

Pattern 4b: As a modal particle literally meaning ‘perhaps/may be’ with a pragmatic function (8 sentences):

(21) P. 701

bu sahl *goft* *magar* *ṣavāb* *bāsh-ad* *rekābdār*
Bu Sahl say.PAST.3SG **perhaps** right be.PRES-3SG courier

nazdik=e *vazir* *rav-ad*
near=EZ vizier go.PRES.3SG

‘Bu Sahl said: the best course perhaps might be for the courier to go to the Vizier.’

¹⁵ In Contemporary Persian, the most common adverb of possibility is *shāyad*, which is derived from the modal verb of *shāyestan* ‘to be proper/to be worthy’. The instances of *shāyad* in *Tārikh-e Beyhaqi* hardly amount to five in total (according to the word-index of the Yahaqqi and Sayyedi edition). All of these instances are also used in the sense of ‘it is proper/it is worthy’, not as an adverb of possibility/probability.

¹⁶ *magar* here can also be interpreted as ‘certainly, surely’. This indicates that *magar*, as a modal particle, can express both alethic and epistemic modality. Crystal (1997: 109) describes these two types of modality by referring to the use of modals in sentences such as *The car must be ready*. The alethic modality would, according to him, interpret this sentence as ‘It follows that the car is ready’ and the epistemic modality would interpret it as ‘It is surely the case that the car is ready’.

(22) P. 126

goftam *şavāb* *bāsh-ad* *ke* ***magar*** *chiz=i*
 say.PAST.1SG right be.PRES-3SG CLM **perhaps** ting=INDF

nebeshte *ā-yad* *ke* *bar* *khodāvand* *hojjat*
 write.PSTP come.PRES-3SG CLM on lord evidence

na-kon-ad

NEG-do.PRES-3SG

‘I said: the best course perhaps would be to write something which cannot be used as evidence against the lord.’

(23) P. 565

vazir *solţān* *rā* *goft* ***magar*** *şavāb* *bāsh-ad*
 vizier Sultan OM say.PAST.3SG **perhaps** right be.PRES-3SG

ke *khodāvand* *in* *tākhtan* *na-kon-ad* *va* *injā* *be*
 CLM lord this attack.INF NEG-kon.PRES-3SG and here to

*rāvan*¹⁷ *moqām* *kon-ad*

Rāvan staying do.PRES.3SG

‘The Vizier said to the Sultan: Perhaps it would be wiser for the lord not to embark upon this attack but to remain here at Rāvan.’

(24) P. 97

ishān *goft-and* ***magar*** *şavāb* *ān* *ast* *ke*
 they say.PAST-3PL **perhaps** right that be.PRES.3SG CLM

khodāvand *nadim-ān=e* *kheradmand-tar*
 lord boon.companion-PL=EZ wise-CMPR

*istād-ān-ad*¹⁸ *pish=e* *kh'ish*

station-CAUS.PRES-3SG before=EZ self

‘They replied: Perhaps it may be the right course for the lord to station before himself the wisest of his boon-companions.’

As can be seen, this type of *magar* is always accompanied by the phrase *şavāb bāshad/şavāb ast* ‘It would be best/it may be the right course’. Examining these instances of *magar* in their specific contexts, we have found that these instances fulfil a pragmatic function. None of these sentences convey uncertainty about the best course of action. The speaker is sure about the best thing to do and tries to persuade the listener to perform that action. By using *magar*, the speaker is just trying to produce a perlocutionary effect, to make the hearer feel respected and avoid sounding

¹⁷ Fayyaz, p. 739: *parvān*.

¹⁸ Fayyaz, p. 127: *istānad*.

like a know-it-all. He relies on his own and the hearer's mutually shared background information and, as Searle puts it, on "the general powers of rationality and inference on the part of the hearer." This use of *magar* is especially important in hierarchic structures. It is governed by who is talking to whom in the hierarchy. In all these instances, people lower in the hierarchy, the subordinates, are talking to someone higher in the hierarchy, the superordinate. This is clearly indicated in three of the four above instances. We believe that all these sentences could have been uttered without *magar* and even in imperative form if they were not addressed to a person of higher status in the hierarchy.¹⁹ We have not been able to find any such sentences uttered by the Sultan directed to his Vizier or to others in his service or by someone higher in position addressed to someone lower in position. There are, on the other hand, similar sentences uttered by someone in almost the same or higher position than the hearer. In such sentences *magar* is absent. In one scene we read, for example, about kh'āja Ahmad b. Hasan (later Grand Vizier) trying to get General Eryāruq to follow him to the Royal court:

u [kh'āja aḥmad] eryāruq-e ḥājeb, sālār-e hendustān, rā gofte bud ke: nāmi zeshtgune bar to neshaste ast, šavāb ān ast ke bā man beravi va ān khodāvand rā bebini (Beyhaqi 1388: 138).

Kh'āja Aḥmad had then said to the General Eryāruq, the Commander of the Indian Troops [...]: you have been placed in a bad light, and it would be therefore advisable for you to accompany me to the court and have an audience with the Amir (Beyhaqi 2011 vol. I: 236–237).

Examples from this category can be compared to using "It'd be good to have the door shut" instead of "Shut the door!" in terms of the theory of indirect speech acts. The speaker persuades the listener to perform what he considers to be the right action by using an adverb of uncertainty.

Pattern 5: as a substitute for the subordinate conjunction *tā* 'so that' in final subordinate clauses (2 sentences):

(25) P. 704

<i>omid=e</i>	<i>hame-gān</i>	<i>be</i>	<i>kh'āje=ye</i>	<i>bozorg</i>	<i>ast...</i>	<i>tā</i>
hope=EZ	everybody-PL	to	vizier=EZ	grand	be.PRES.3SG...	so.that

<i>in</i>	<i>tadbir=e</i>	<i>khaṭā</i>	<i>rā</i>	<i>bezudi</i>	<i>dar-yāb-ad ...</i>
this	plan=EZ	wrong	OM	soon	PREV-encounter.PRES-3SG...

<i>magar</i> ²⁰	<i>in</i>	<i>tadbir=e</i>	<i>nāšavāb</i>	<i>be-gard-ad</i>
so.that	this	plan=EZ	incorrect	SBJN-renounce.PRES-3SG

'All their hopes are concentrated on the Grand Vizier ... encountering this wrong plan soon... so that this misguided plan will be renounced.'

¹⁹ In all the above sentences in this group, the entire combination of *magar* and *šavāb bāshad/šavāb ast* could also have been replaced by the short infinitives *bāyad* 'must' or *shāyad* 'it is appropriate', if they were not addressed to a person higher in the hierarchy.

²⁰ In this example, *magar* can also be interpreted as modal particle expressing either possibility or hope.

(26) P. 680

lashkar=i²¹ *bāyad* *ferestād* **magar** *balkh* *be* *dast=e* *mā*
 army=IND must.be send.SHIF so.that Balkh in hand=EZ PN.1PL

be-mān-ad

SBJN-stay.PRES-3SG

‘And an army must be sent so that Balkh remains in our hands.’

In those instances where *magar* functions as a substitute for the conjunction *tā*, one can insert *tā* before *magar* without causing any syntactic problem. The semantic effect will be a slightly higher degree of uncertainty, as *magar* falls into its ordinary slot as an adverb of possibility/probability. There are examples (28) in which *magar* is preceded by the subordinate conjunction *tā* and functions either as an adverb of possibility/probability or as an intensifier.

Pattern 6: as the intensifier *hamānā* ‘certainly, surely’ (2 sentences):

(27) P. 78

va *javāb* *yāft* *ke* *chun* *be-raft* **magar**
 and reply receive.PAST.3SG CLM when PFTV-go.PAST.3SG INTS

zesht *bash-ad* *bāz-gashtan*
 rude be.PRES-3SG PREV-turn.INF

‘He received the reply that, since he had set off, it would be rude of him to return.’

(28) P. 183

hīlat *bāyad* *kard* *tā* **magar**²² *vei* *rā* *bar*
 stratagem must do.SHIF so.that INTS PN.3SG OM on

pesar=ash *be-tavān-id* *gozar-ān-id*
 son=PC.3SG SBJN-can.PRES-2PL pass-CAUS.PRES-2PL

‘You must devise a stratagem to get her to pass by her son.’

As seen, no equivalent is considered for *magar* in the translations. The sentences could have been written without *magar* as well. In (28), one can easily remove either *magar* or *tā*.

²¹ Fayyaz, p. 780: *va lashkar*.²² In Beyhaqi (2011 vol. I: 286), *magar* has been translated as ‘perhaps’.

Pattern 7: as an interrogative particle (1 sentence):

(29) P. 310

<i>goft</i>	<i>magar</i>	<i>gusht</i>	<i>na-yāfte</i>	<i>bud-i</i>	<i>va</i>
say.PAST.3SG	IP	meat	NEG-find.PSTP	be.AUX.PAST-2SG	and

<i>noql</i>	<i>ke</i>	<i>marā</i>	<i>va</i>	<i>kadkhodā-yam</i>	<i>rā</i>
delicacy	CLM	PN.1SG.OM	and	counsellor=PC.1SG	OM

be-kh'ord-i?

PFTV-eat.PAST-2SG

‘He said: Couldn’t you find any meat or delicacies, that you had to eat me and my counsellor?’²³

The speaker is criticizing the hearer for having spoken ill of him and his chamberlain. It has nothing to do with meat and delicacies. The speaker neither expects any direct answer to that question nor has any doubt about the hearer’s access to meat and snacks. This is another example that fits well with the theory of indirect speech acts. A question is used to express a criticism that normally is expressed in the form of a statement. The speaker could have expressed his intention by saying: “I am very disappointed about your having spoken ill of me and my counsellor.” Using the question form and the verb *kh'ordan* ‘to eat’ metaphorically, the speaker tries to produce a perlocutionary effect in the hearer and cause him to feel ashamed.

Pattern 8: in the sense of ‘it happened/once’ for opening a narrative (1 sentence):

(30) P. 403

<i>chonān</i>	<i>oftād</i>	<i>az</i>	<i>qazā</i>	<i>ke</i>	<i>nadim</i>	<i>bu no'aim=e</i>
so	happen.PAST.3SG	of	chance	that	counsellor	Bu No'aim=EZ

<i>magar</i>	<i>be</i>	<i>ḥadith=e</i>	<i>in</i>	<i>tork</i>	<i>del</i>	<i>be</i>	<i>bād</i>	<i>dāde</i>
NOIP	to	thought=EZ	this	Turk	heart	to	wind	give.PSTP

bud

be.AUX.PAST.3SG

‘It happened (by chance) that Bu No’aim, the counsellor, had become besotted by the beauty of this Turk.’²⁴

As is seen, *magar* here completes what has already been expressed by *az qazā* ‘by accident’. It contributes to opening a narrative, even if it is placed in the middle of the sentence. It could have been placed at the beginning of the narrative as well.

²³ In Beyhaqi (2011 vol. I: 432), an expressive explanation is given within the parentheses as follows: (i.e. slander us behind our backs?).

²⁴ In Beyhaqi (2011 vol. II: 61), *magar* has here been translated as ‘perhaps’.

	+	–	
	Pragmatic function	Pragmatic function	TOTAL
Preposition of exception		21	21
Conditional conjunction		2	2
Modal particle: probability	8	34	41
Modal particle: wish/hope		7	7
Intensifier		2	3
Interrogative particle	1		1
Substitute for <i>tā</i>		2	2
Narrative opener		1	1
TOTAL	9	69	78

Table 1. Frequencies of different uses of *magar* in *Tārikh-e Beyhaqi*

5. Concluding Remarks

As Table 1 shows, *magar* is most frequently used as the modal particle of possibility, followed by being used as the preposition of exception.²⁵ The function and meaning of *magar* are much clearer when it is used as a preposition, ‘except’; as a conditional conjunction, ‘unless’/‘only if’; and as an interrogative particle. As a preposition and as an interrogative particle it is used in sentences referring to the present, past, or future. However, the number of sentences with reference to the past is much higher, probably due to the historiographical nature of the text. We have not been able to find any instances of *magar* as a combined conjunction as mentioned by Lazard.

There is no sharp line between *magar* as a subordinate conjunction *tā*, *magar* as ‘perhaps’, the optative *magar* as ‘hoping’, and *magar* as an intensifier. It is only in relation to the context that one can determine in what sense a certain *magar* has been used. This being the case, there is room for slightly different interpretations of these instances of *magar*. As a modal particle to express probability/possibility or hope/wish, and as a conditional and subordinate conjunction, *magar*, as expected, is usually used in sentences referring either to the future or to the present. As a modal particle of probability, it has also been used in sentences referring to the past. As an intensifier it is used with verbs referring to the present or future.

Some uses of *magar* mentioned by Dehkhoda, such as its use as the disjunctive ‘or’, as the noun ‘doubt’, as ‘only’, and as ‘as if’, do not occur in *Tārikh-e Beyhaqi*. There are two different uses of *magar* (patterns 5 and 8) in *Tārikh-e Beyhaqi* that fit well with the theory of indirect speech acts.

²⁵ In a study of the frequency of the words in Sa’di’s *ghazals*, Sadiqian (1378: 1682–1688) recognizes 194 occurrences of *magar*: a) 72 times as a preposition/conjunction meaning except/unless; b) 49 times either optatively or as the adverb of possibility; c) 41 times as sentence modifier like *guyā/guyi* ‘as if’, *pendāri* ‘as if’; e) 20 times as an interrogative particle with a negative verb when an affirmative answer is expected; and f) 12 times in the meaning ‘it will be possible only if’. The results of the present paper are somewhat different from the results obtained in Sadiqian’s study. However, as Sadiqian’s study is based on poetry, one should be cautious about drawing any conclusions, given the limitations that the defined metres of Classical Persian poetry impose on the use of language.

A preliminary, not yet published, study of Persian classical texts carried out by us indicates that both the frequency and semantic variation of this particle began decreasing in the 14th century, but then began increasing again in the 18th century. Some uses of *magar* mentioned by Dehkhoda do not occur in Contemporary Persian either. There are, however, several other uses of this particle in Contemporary Persian, which are particularly interesting from the point of view of pragmatics and the theory of indirect speech acts. We hope to be able to follow up this paper in the near future with a synchronic study of the uses of *magar* in Contemporary Persian.

Other interesting topics for further research are a diachronic study of this particle, a comparative study of its use in other Iranian languages, and a study of the use of a similar particle in some other Indo-European language, like Italian.

Abbreviations

1	first person	NEG	negation
2	second person	NOP	narrative opener
3	third person	OM	object marker
-	affix boundary	PAST	past tense
=	clitic boundary	PC	pronominal clitic
AUX	auxiliary verb	PFTV	perfective aspect
CAUS	causative	PL	plural
CLM	clause linkage marker	PN	pronoun
CMPR	comparative degree	PRES	present tense
CONJ	conjunction	PREV	preverbal particle
DISJ	disjunctive	PSPT	past participle
EZ	ezafe	IP	interrogative particle
IND	indefinite	SG	singular
INTS	intensifier	SBJN	subjunctive
IPFV	imperfective aspect	SHIF	short infinitive

Bibliography

- Austin, John L. (1961). "Performative Utterances". In: Martinich, A. P. (ed.), *The Philosophy of Language*. New York and Oxford: Oxford University Press, pp. 105–114.
- Bahar, Mohammad T. (1369 [1990/91]). *Sabkshenāsi: yā taṭavvor-e nathr-e fārsi, jeld-e dovvom*. Tehran: Amir Kabir.
- Beyhaqi, Abu al-Fazl (2536 [1977/78]). *Tārikh-e Beyhaqi*, edited by Ali A. Fayyaz. Mashhad: Dāneshgāh-e Ferdowsi. (Referred to as Fayyaz.)
- Beyhaqi, Abu al-Fazl (1388 [2009/10]). *Tārikh-e Beyhaqi*, edited by Mohammad J. Yahaqqi and Mehdi Sayyedi, Tehran: Sokhan.
- Beyhaqi, Abu'l-Fazl (2011). *The History of Beyhaqi (The History of Sultan Mas'ud of Ghazna 1030–1041)*, vol. I-II, translated by C. E. Bosworth and revised by Mohsen Ashtiany. Cambridge, etc.: Ilex Foundation & Center for Hellenic Studies.
- Crystal, David (1997). *A Dictionary of Linguistics and Phonetics*. Oxford: Blackwell.
- Daniel, Elton L. (2012). "The Rise and Development of Persian Historiography." In: Melville, Charles (ed.), *Persian Historiography*. London and New York: I. B. Tauris, pp. 101–154.
- Dehkhoda, Ali A. (1325 [1946/47]–1354 [1975/76]). *Loghatnāme*. Tehran: Enteshārāt-e majles-e showrā-ye melli.

- Farshidvard, Khosrow (1375 [1996/97]). *Goftārḥā-yi dar bāre-ye dastur-e zabān-e fārsi*. Tehran: Amir Kabir.
- Jensen, Hans. (1931). *Neupersische Grammatik mit Berücksichtigung der historischen Entwicklung*. Heidelberg: C. Winter.
- Khatibi, Hoseyn (1366 [1987/88]). *Fann-e nathr dar adab-e pārsi*. Tehran: Zavvār.
- Langarudi, Shams. (1377 [1998/99]). *Tārikh-e taḥlīli-ye she'r-e now: jeld-e sevvom, 1341–1349*. Tehran: Nashr-e markaz.
- Lazard, Gilbert (1963). *La langue des plus anciens monuments de la prose persane*. Paris: Librairie C. Klincksieck.
- Lazard, Gilbert (1992). *A Grammar of Contemporary Persian*, translated into English by Shirley A. Lyon. Costa Mesa: Mazda Publishers.
- MacKenzie, David N. (1971). *A Concise Pahlavi Dictionary*. London: Oxford University Press.
- Mashkur, Mohamm J. (2535 [1976/77]). *Šarf va naḥv-e zabān-e fārsi*. Tehran: Mo'asese-ye maṭbu'āti-ye sharq.
- Moin, Mohammad (1374 [1995/96]). *Farhang-e fārsi*. Tehran: Mo'asese-ye chāp va enteshārāt-e Amir Kabir.
- Nafisi, Ali A. (2535 [1976/77]). *Farhang-e nafisi*. Tehran: Ketābforushi-ye Khayyām.
- Natel Khanlari, Parviz (1366 [1987/88]). *Tārikh-e zabān-e fārsi: jeld-e sevvom*. Tehran: Nashr-e now.
- Natel Khanlari, Parviz (1373 [1994/95]). *Dastur-e tārikhi-ye zabān-e fārsi*. Tehran: Tus.
- Roberts, John R. (2009). *A Study of Persian Discourse Structure* [Studia Iranica Upsaliensia 12]. Uppsala: Acta Universitatis Upsaliensis.
- Sadiqian, Mahindokht (1378 [1999/2000]). *Farhang-e vāzhenāme-ye ghazaliāt-e sa'di be enẓemām-e farhang-e basāmadi*. Tehran: Pazhuheshgāh-e'olum-e ensāni va moṭāle'āt-e farhangi.
- Salemann, Carl, and Shukovski, Valentin (1889). *Persische Grammatik mit Litteratur, Chrestomathie und Glossar*. Berlin: Reuther & Richard.
- Searle, John R. (1965). "What is a Speech Act?" In: Martinich, A. P. (ed.), *The Philosophy of Language*. New York and Oxford: Oxford University Press, pp. 115-125.
- Searle, John R. (1975). "Indirect Speech Act?" In: Martinich, A. P. (ed.), *The Philosophy of Language*. New York and Oxford: Oxford University Press, pp. 161-175.
- Thackston, Wheeler M. (1993). *An Introduction to Persian*. Bethesda, Maryland: IBEX Publishers.

Diphthongization in Five Iranian Balochi Dialects

Farideh Okati, University of Zabol and Uppsala University

farideh.okati@uoz.ac.ir

Pétur Helgason, Uppsala University

petur.helgason@lingfil.uu.se

Carina Jahani, Uppsala University

carina.jahani@lingfil.uu.se

Abstract

This paper deals with the phenomenon of diphthongization, *e:* > *ie* and *o:* > *ue*, in different varieties of Iranian Balochi dialects spoken in the five regions of Sistan, Saravan, Khash, Iranshahr and Chabahar in the southeast corner of Iran. The study reveals that diphthongized production of these vowels is predominant in the Khash dialect, suggesting that they should be represented as the diphthongs *ie* and *ue* in the vowel inventory of this dialect. In the Iranshahr and Chabahar dialects, which show the second and third highest degrees of diphthongization among the dialects under study, the data indicate a diphthongization tendency rather than a shift to predominantly diphthongized productions. Sistan and Saravan show only sporadic tendencies toward diphthongization. Balochi, in general, has eight vowels *i*, *i:*, *u*, *u:*, *a*, *a:*, *e:*, *o:* (also called the Common Balochi vowel system), as well as the speech sounds *ay* and *aw*, which are referred to as diphthongs by some scholars and are believed to be sequences of V+C by others. The occurrence of diphthongization in the dialects under study seems to be system-internal rather than due to external influence. Factors such as age, education, or language contact with surrounding languages such as Persian, do not appear to contribute to the occurrence of diphthongized vowel production.

Keywords: Balochi, Iranian Balochi dialects, vowels, diphthongization, Persian

1. Introduction

The Balochi language is mainly spoken in southwestern Pakistan, southern Afghanistan, and southeastern Iran. This language is considered a Northwestern Iranian language, and is most closely related to “Kurdish, Tati, Talyshi and other Northwestern Iranian languages” (Jahani 2003: 114). Jahani and Korn (2009: 636) divide Balochi into the three major dialects of Western, Southern, and Eastern Balochi, among which there are diverse varieties. The Balochi varieties spoken in Iran belong to Western and Southern Balochi.

Factors such as geographical distribution, contact with surrounding languages such as Persian, Pashto, Urdu, Panjabi, and Sindhi, as well as borrowings from these languages account for the existence of many dialect variants of Balochi. For the Iranian Balochi dialects, Mahmoodzahi (2003: 148) points out that “Balochi and Persian must have been to at least a certain degree in constant contact with each other for centuries.” The existence of a younger generation who receive education (in the dominant language) further increases dialect divergences.

This study investigates the status of diphthongization of *e:* to *ie* and *o:* to *ue* in the Iranian Balochi dialects spoken in five selected locations, Sistan (SI), Khash (KH), Saravan¹ (SA), Iranshahr (IR), and Chabahar (CH) in Sistan and Baluchestan² prov-

¹ Spelled *Sarawan* in some sources.

² This is the official English spelling used in Iran; also spelled *Sistan va Baluchestan*.

ince in the southeast of Iran.³ The language data were gathered in the form of free speech and verbal elicitation from more than 20 literate and non-literate male and female language consultants. In most cases, the sentences and words were elicited from the consultants by asking them to translate the Persian sentences or words into Balochi and repeat them four times (in Balochi). The set of targeted words contained the vowels /e:/ and /o:/, which tend to be diphthongized in all or some of the dialects under investigation.

The elicited data for this study were in the form of unscripted speech. The age of the consultants ranged from 25 to 82, though the majority of them were in the age range 40 to 70. For each dialect, 2 males and 2 females were used for the spectral (i.e. formant) measurements of vowels. The data were recorded in a lossless audio format using a handheld, digital recording device with built-in microphones. The data were labeled using WaveSurfer (Sjölander and Beskow 2000) and the labeling was used as the basis for automatic extraction of spectral and durational information. For the formant analysis, the program Praat⁴ was used. Numerical analysis and graphing were done using MS Excel as well as the statistical program Minitab.

Since not all the Balochi dialects in Iran can be described using the same vowel inventory, Jahani and Korn (2009: 642) present two different inventories for the vowel systems in these dialects. Thus, some dialects show a system with the monophthongs *a*, *a:*, *e*, *e:*, *i:*, *u:*, *o*, *o:*, as well as the vowel + glide sequences *ey* and *ow*, while other dialects show a system with the monophthongs *a*, *a:*, *e*, *i:*, *u:*, *o*, the diphthongs *ie* and *ue*, as well as the vowel + glide sequences *ey* and *ow*.⁵ There is no mention of the distribution of these two systems across Balochi dialects spoken in Iran.

The following table presents a sample of words used for the diphthongization analysis.

/e:/ >	[ie]	/o:/ >	[oe]
ke:p >	kɛp 'bag'	bo: >	bœ 'smell'
se:b >	sɛb 'apple'	go:š >	goeš 'ear'
re:k >	ɾɛk 'sand'	do:g >	doeg 'buttermilk'
je:b >	ʝɛb 'pocket'	go:k >	goek 'cow'
de:r >	dɛr 'late'	do:st >	doest 'like'
de:m >	dɛm 'face'	go:št >	goešt 'meat'
sre:n >	sɾɛn 'waist'	čo:ne >	čœne 'how are you'
če:ra: >	čɛra: 'under'	bšo:d >	bšœd 'wash!'
be:ga: >	bɛga: 'afternoon'	lo:ti:n >	loeti:n 'I want'
ge:štir >	gɛštɪr 'more'	na:ko: >	na:kœ 'uncle'
gre:tun >	gɾɛtun 'I cried'	o:šta:t >	œšta:t 'he/she stood'
be:ho:š >	bɛhoeš 'forget'	balo:č >	balœč 'Baloch'
ne:mro:č >	nɛmroeč 'noon'	do:či:n >	doeči:n 'I sew'

Table 1. Examples of words with diphthongization

³ The data for this study were gathered by Farideh Okati and the phonetic analysis was carried out by Farideh Okati and Pétur Helgason. All three authors contributed to the final editing of the text.

⁴ <http://www.fon.hum.uva.nl/praat/>

⁵ Based on the text in Jahani and Korn (2009: 642). Table 11: 3 in Jahani and Korn (*ibid.*) gives *u* instead of *o*, which according to personal communication with Jahani is a *lapsus calami*.

2. Previous studies

Previous studies on the Balochi language mention the existence of different speech sounds in the vowel systems of the Balochi dialects, which are referred to as diphthongs by some scholars, while others regard them as sequences of V+C (glide) in view of the existing syllable patterns (Jahani and Korn 2009: 641). For example, *ai* and *au* are reported in Eastern Balochi by Dames (1891) and in Eastern and Western Balochi by Grierson (1921); *ay* and *aw* in the Rakhshani dialect by Barker and Mengal (1969) and in Balochi in general by Elfenbein (1989); *ay*, *ey*, and *aw* in the Saravani dialect by Spooner (1967) and by Baranzehi (2003); *ai*, *ao*, and *ei* in Karachi Balochi by Farrell (1990); and *ey*, *ow*, *ie*, *ue* in Balochi dialects in Iran by Jahani and Korn (2009).

Dames (1891: 6) divides his vowel list into “Short”, “Long”, and “Diphthongs” and displays “e, ai, o, au” under the name of diphthongs; he does not explain why he categorizes *e* and *o* under this group. Grierson (1921: 336) presents the *ai* and *au* in the vowel list and does not refer to them separately as diphthongs or sequences of sounds. Gilbertson (1923: 2–3) refers to *au* and *ai* as “diphthongs” and explains that they sound like “‘ou’ in the English word ‘house’” and “‘ai’ in the word ‘aisle’” respectively. Barker and Mengal (1969: xliv) refer to *ay* and *aw* as sequences of a vowel and an off-glide; they emphasize that “[t]hese sequences must be carefully distinguished, since Baluchi has some rather similar vowel clusters: e.g. /əi/, /əe/, /əo/.” Elfenbein (1989: 352) also divides the Balochi vowels into “short”, “long”, and “diphthongs” and places the *ay* and *aw* under the diphthongs division. Farrell (1990: 11) lists *æe*, *əo*, *ao*, *ei*, *ai* in a phonemic representation (in / /), under the name of diphthongs, but he does not include them in the vowel chart he presents later on the following page. Jahani and Korn (2009: 642) categorize *ie* and *ue* under the diphthongs for Iranian Balochi, but they believe that *ey* and *ow* are sequences of a vowel plus a glide (V+C).

3. Diphthongs in neighbouring languages

There are different views about the diphthongs in Persian, which belongs to the Southwestern group of Iranian languages and is the language most in contact with the Balochi dialects spoken in Iran. Samareh (1368) posits six diphthongs for the Persian language varieties spoken in Iran, but points out that they are only diphthongs from a phonetic point of view and can also be described as sequences of V+C. Ganjavi *et al.* (2003) believe there are no diphthongs in Iranian Persian, while Yaesoubi (2010) holds that it has two diphthongs, *ej* and *ow*, although he does not mention if these diphthongs have phonemic or phonetic status. Hakimi (2012: 9) finds that the “digraphs or diphthongs” in Iranian Persian, *ay*, *ey*, *uy*, *oy*, *ay*, and *ow*, are combined sequences of a vowel plus a consonant from a phonemic viewpoint. Hakimi (*ibid.*) also points out that Tajiki has “four diphthongs”, *ay*, *aw*, *uy*, and *o:y*, which have the same phonemic status (V+C) as those in Iranian Persian.

Hatami (2011), studying Ruini, a dialect of Khorasani Persian, uses the term diphthongs to refer to *aw*, *ao*, *oe* in this dialect, but declares that the second components in these diphthongs are not vowels at all, but are semivowels and change to /v/ and /

y/ before vowels (e.g. *jelaw* > *je.la.ve*). Diphthongs are also reported for other languages in Iran; e.g. the diphthongs /*aw*, *ew*, *ow*, *øw*, *əw*, *ey*/ are reported for Laki spoken in Kermanshah, by Mirdehghan and Moradkhani (2010). Apparently, these so-called diphthongs are sequences of V+C in this dialect, because Mirdehghan and Moradkhani (*ibid.*: 515) point out that “an extra semi-vowel /w/ is observed in the dialect comparatively to Persian”.⁶ Aldaghi and Tavakoli (2011) mention eight diphthongs, as they call them, *ay*, *ey*, *oy*, *uy*, *iy*, *ay*, *aw*, *ow*, e.g. in *tow* ‘fever’ and *kawsh* ‘shoes’, for the Sabzevari dialect, a Persian dialect spoken in Khorasan in Iran, and Ebrahimi (2012) indicates *ey*, *oy*, *a:y*, *ow* for Bayzai dialect, a Persian dialect spoken in Fars in Iran. Although Ebrahimi (*ibid.*) refers to these sounds as diphthongs, he describes them as sequences of a vowel plus the semi-vowels *w* and *j*.

Urdu, which belongs to the Indo-Aryan branch of the Indo-Iranian family, is one of the languages in contact with the Balochi dialects. Hussain *et al.* (2011) point out that, unlike English, Urdu lacks diphthongs, but sometimes English diphthongs such as *ai*, which have entered Urdu through loanwords, are not modified and therefore are preserved as diphthongs in this language.

4. Diphthongization analysis

The diphthongs *ie* and *ue*, derived from the long vowels *e:* and *o:*, are observed in the five Iranian Balochi dialects investigated in this study. The data analysis shows that the diphthongization is not present to the same extent in all five dialects. It is the predominant production type in the Khash dialect, but is less prominent in Iranshahr and even weaker in Chabahar. No consistent diphthongization is seen in Sistan and Saravan. Only a few cases, with a low degree of diphthongization, are found in the speech of one consultant from the town of Gosht, close to the border of Khash region, where diphthongization is predominant and a very weak trace of diphthongization is also observed in the speech of one speaker of the SI dialect. These very small tendencies toward diphthongization in SA and SI can, of course, be a starting point from which this phenomenon will grow in these dialects in the future.

The following spectrograms show examples of the diphthongization of *o:* to *ue* in the word *go:k* ‘cow’ (Figures 1–5, female speakers) and *e:* to *ie* in the word *se:b* ‘apple’ (Figures 6–10, male speakers), in the different dialects. For *go:k* ‘cow’, in Figures 1 through 5, the rise in the frequency of F2, the second line from the bottom (the green line) in the spectrograms indicates the degree of diphthongization of *o:* to *ue* in the different dialects under study. A back high vowel [o] should have a low F1 and a low F2 throughout its production. A rise in F2 indicates that the vowel becomes fronted (and possibly unrounded). The greater the rise, the greater the diphthongization tendency in the *o:*.

⁶ [w] is also observed as allophone of /v/ in Sistani Persian spoken in the Sistan region in the southeast of Iran (Okati, *The phonology of the Iranian Sistani dialect*, forthcoming).

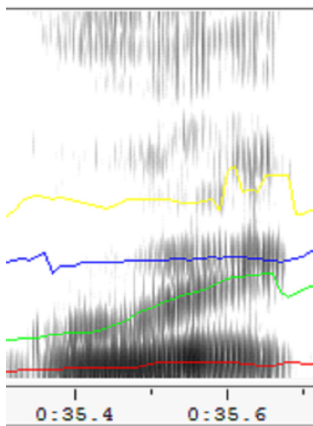


Figure 1. go:k 'cow'- KH

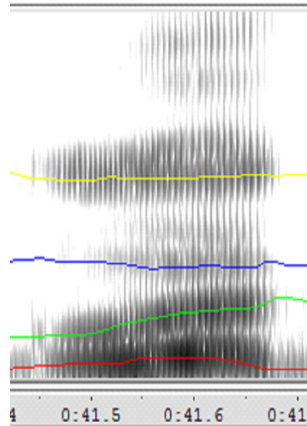


Figure 2. go:k 'cow'- IR

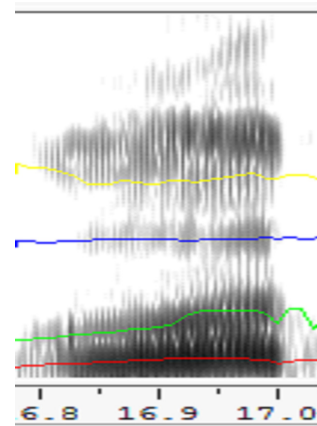


Figure 3. go:k 'cow'- CH

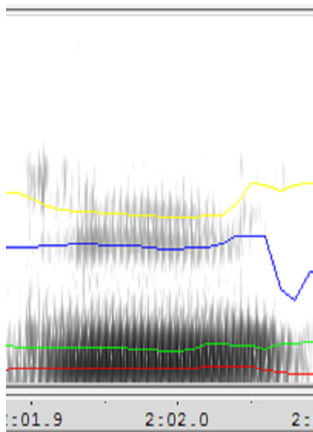


Figure 4. go:k 'cow'- SA

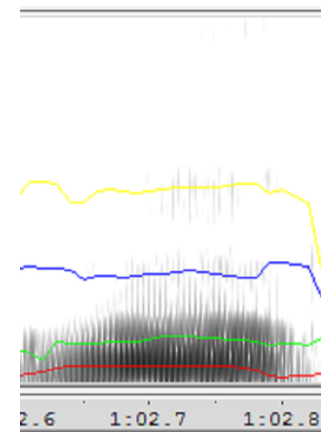


Figure 5. go:k 'cow'- SI

For *se:b* 'apple', in Figures 6 through 10, the fall in the frequency of F2 indicates the degree of diphthongization for *e:* to *ie:*; a front high vowel [e] should have a high F2 throughout its production. A fall in the frequency of F2 indicates that the vowel becomes lower and possibly more centralized. The sharper the fall, the greater the diphthongization tendency in the *e:*.

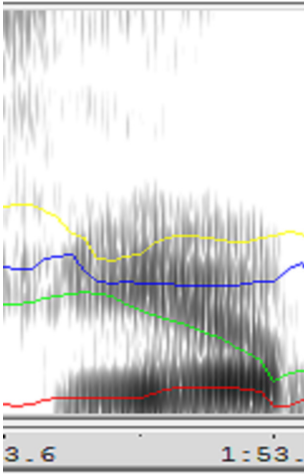


Figure 6. se:b ‘apple’- KH

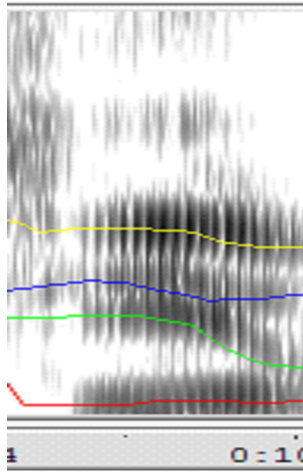


Figure 7. se:b ‘apple’- IR

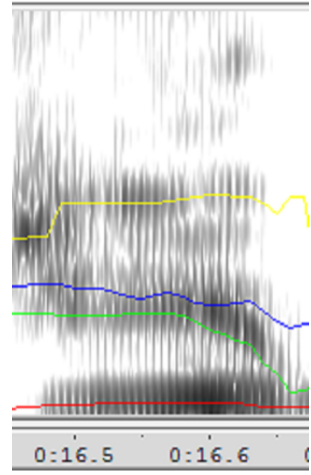


Figure 8. se:b ‘apple’- CH

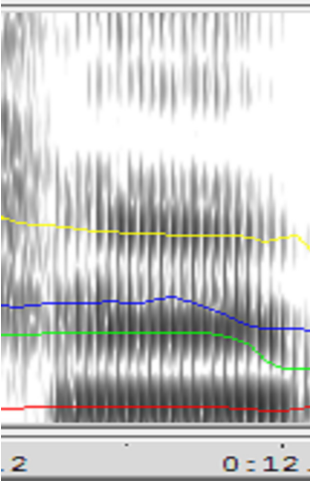


Figure 9. se:b ‘apple’- SA

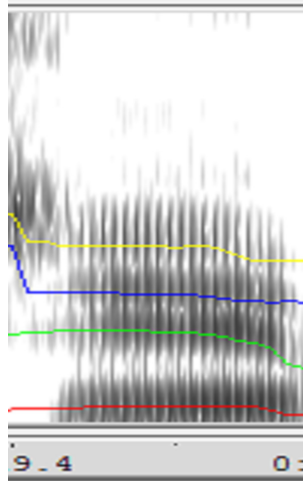


Figure 10. se:b ‘apple’- SI

In order to compare the degree of diphthongization of /o:/ and /e:/ in the different dialects, the changes in the frequency of the second vowel formant (F2) during the production of the vowel can be estimated. For a diphthongized /o:/, F2 rises, and the degree of diphthongization is directly correlated with the degree of rise in the formant. Conversely, when /e:/ is diphthongized, F2 sinks, and the degree of diphthongization is correlated with the degree of lowering. Therefore, formant frequency measurements are made at three different time points in the vowel. The first time point is determined by multiplying the total vowel duration by 0.1, yielding a time point 10% into the vowel duration, i.e. near the onset of the vowel. This measure-

ment is henceforth referred to as the *initial* formant measurement, or *ini* for short. The second time point is determined by multiplying the total vowel duration by 0.5, yielding a time point 50% into the vowel duration, i.e. at the centre of the vowel. This measurement is henceforth referred to as the *medial* measurement, or *mid* for short. Lastly, the third time point is determined by multiplying the total vowel duration by 0.9, yielding a time point 90% into the vowel duration, i.e. near the end of the vowel. This measurement is henceforth referred to as the *final* measurement, or *fin* for short.

To be able to use the measured changes in F2 at these three time points to make comparisons across dialects and speakers, two obstacles have to be overcome. First, for cross-dialect comparisons, the variability in phonetic contexts does not allow the pooling of formant measurements for a specific dialect. This is because the various places of articulation of consonants before and after the vowel affect the formant trajectories in differing ways, which may obscure the F2 differences. Pooling formant measurement values obtained near vowel onset and offset from different words would therefore not give representative results. Instead, the approach adopted here is to select specific words for which several tokens for each dialect have been obtained and to compare the degree of diphthongization in the target vowel in these words in the different dialects. The second obstacle is that the speakers have vocal tracts of different sizes, which results in differences in formant frequencies: long vocal tracts yield lower formant frequencies than shorter ones. This means that the frequency measurements should be normalized for vocal tract size.

For the present comparison of the degree of diphthongization, normalization is achieved by estimating the relative deviation of the initial and final F2 measurements from the medial measurement. First, the initial F2 measurement is divided by the medial measurement, which yields a quotient that indicates how much the formant has changed from vowel onset to the middle of the vowel. This quotient will be referred to here as the initial $F2_Q$. Then the final F2 measurement is divided by the medial measurement, which yields a quotient that indicates how much the formant has changed from the middle of the vowel to the end of the vowel. This quotient will be referred to as the final $F2_Q$. The medial measurement divided by itself always yields the value 1, which therefore can serve as a baseline for F2 comparisons across speakers and dialects. This constant value will be referred to as the baseline value.

Let us illustrate this using a hypothetical example. Suppose that, for a specific word and one speaker, the initial, medial, and final F2 measurements for a token of /o:/ yield the values 1500 Hz, 2000 Hz, and 2500 Hz respectively. By dividing the initial value by the medial value (1500/2000) we get an $F2_Q$ of 0.75. Thus we have effectively obtained a measure of the relative deviation between the initial and the medial measurements; the initial F2 value is 75% of the medial one. By dividing the final value by the medial value (2500/2000) we get an $F2_Q$ of 1.25; the final F2 value is 125% of the medial one. Measuring the deviation of the start and end point values from the medial value for every vowel token for a word allows us to compare all tokens of a particular word in a way that is effectively normalized for vocal tract length. Suppose now that, for an /o:/ vowel token from the same word as the previous example, but for another speaker, we have obtained the values 1200 Hz, 1600

Hz, and 2000 Hz. These values may at first seem very different from the first set of values but, in fact, the initial $F2_Q$ is 0.75 and the final $F2_Q$ is 1.25, just as in the first example. Thus we have established that these two hypothetical vowels are very similar in terms of how $F2$ changes in the vowel during its production. Lastly, we can plot a trajectory for the relative formant changes in the vowel from the initial $F2_Q$ to the final $F2_Q$. For the medial measurement, a quotient value is obtained by dividing the medial formant frequency by itself. This, of course, always yields 1, a constant that serves as the baseline for comparison across speakers and dialects.

Several words, such as *go:k* ‘cow’, *go:š* ‘ear’, *balo:č* ‘Baloch’, *se:b* ‘apple’, *be:ga:* ‘afternoon’, and *če:ra:* ‘under’, have been used to compare the degree of diphthongization of /o:/ and /e:/ in the five dialects under study. Among these, the words *go:k* ‘cow’ and *se:b* ‘apple’ have been chosen to be presented here as examples. The graphs in Figures 11 and 12 show the degree of diphthongization in these words in the different dialects. The graphs show the initial $F2_Q$ (ini), the baseline for the vowel (mid, which is always 1), and final $F2_Q$ (fin). The related tables following each figure show the $F2_Q$ for the initial and final $F2$ measurements in the vowel, as well as the difference between the initial and final $F2_Q$ values.

Figure 11 compares the degree of diphthongization of /o:/ in the word *go:k* ‘cow’ in different dialects. If *go:k* is produced with little or no diphthongization, one would expect the initial and final $F2_Q$ values to be similar without much deviation from the medial (baseline) value, since the preceding and following velar stops are highly compatible with the production of a back, high vowel. However, if the /o:/ in *go:k* is diphthongized and produced as [œ], one can expect a sharp rise from the initial $F2_Q$ to the final $F2_Q$; i.e. $F2$ increases considerably. Table 2 gives the initial and final quotients (rows 1 and 3) and the baseline (row 2). In addition, Table 2 shows the difference between the initial and final $F2_Q$ values (bottom row), which can be used as a rough indicator of degree of diphthongization – the greater the increase from initial to final $F2_Q$, the more diphthongization there is in the /o:/ vowel.

In Figure 11, the initial and final $F2_Q$ values in both SA and SI are similar and do not deviate much from the baseline, which means that these dialects have no tendency toward diphthongization in *go:k*. CH appears to have similar $F2_Q$ values as SA and SI, but it should be noted that the initial value is below the baseline and the fin value is above the baseline, which indicates a greater rise in $F2$ from the start to the end of the vowel than was found for SA and SI. In terms of the difference between the initial and final $F2_Q$, the KH and IR dialects have by far the greatest degree of diphthongization in *go:k*, with a difference in initial and final $F2_Q$ of 0.38 and 0.30 respectively. CH shows a slight tendency toward diphthongization (0.08), but SA and SI do not show any such tendency at all (−0.01 and 0.00 respectively).

In the diphthongization of /e:/, the movement of the tongue is primarily from a front mid-high position to a more back and slightly lower position. Acoustically, this movement means that $F2$ is lowered. Thus, for /e:/ one can compare the degree of diphthongization in the different dialects by calculating the initial and final $F2_Q$ to obtain a normalized estimate of how much $F2$ is lowered. Figure 12 and Table 3 compare the degree of diphthongization of the word *se:b* ‘apple’ in the five dialects. If the vowel is produced as a monophthongal [e] vowel, neither the coronal /s/ pre-

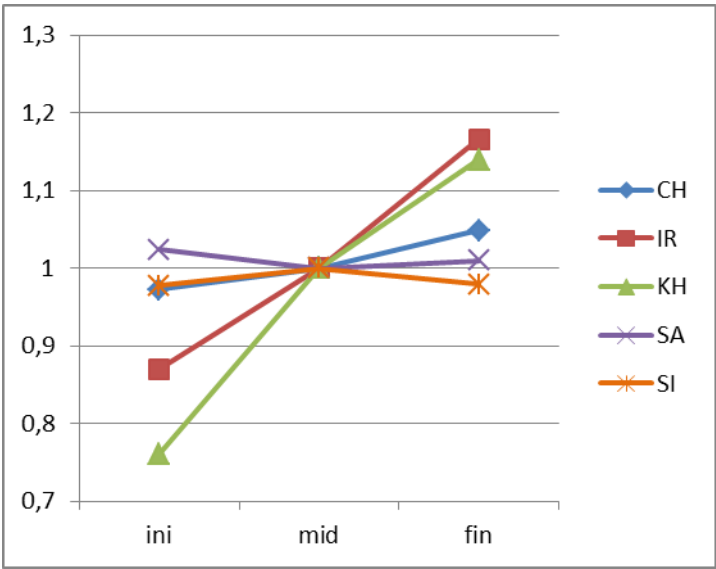


Figure 11. Comparison of the degree of diphthongization of the vowel /o:/ in go:k ‘cow’ in different dialects. The y-axis shows a normalized F2 deviation quotient (F2_Q) calculated at three time points in the vowel: an initial F2_Q value (ini) is obtained by measuring the F2 frequency near the onset of the vowel and dividing it by the F2 measurement made at the centre of the vowel; a final F2_Q value (fin) is calculated by measuring F2 near the end of the vowel and dividing it by the F2 measurement at the centre. The mid value is a baseline value obtained by dividing the measurement at the centre of the vowel by itself, which always yields the value 1. Further explanation in text.

Dialect	CH	IR	KH	SA	SI
F2					
F2 _Q -ini	0.972253	0.869918	0.76121	1.023486	0.977694
Baseline	1	1	1	1	1
F2 _Q -fin	1.048096	1.165517	1.139368	1.009649	0.979231
F2 _Q -ini – F2 _Q -fin	0,08	0,30	0,38	-0,01	0,00

Table 2. The normalized F2 deviation quotient (F2_Q) for the word go:k ‘cow’ in different dialects. Row one shows the initial F2_Q, row two shows the baseline, and row three shows the final F2_Q. The bottom row shows the difference between the initial and final F2_Q.

ceding the vowel nor the bilabial /b/ following it should affect the F2_Q values to any great extent; i.e. the initial and final F2_Q values should be similar. However, if the vowel is diphthongized and produced as [ɪɛ], one would expect to see a sharp fall from the initial F2_Q to the final F2_Q.

A similar pattern as for /o:/ emerges for /e:/ in se:b. Figure 12 shows that in KH, IR, and CH there is a fall in F2_Q, whereas for SI there is only a slight decrease and in SA the initial and final F2_Q values are almost equal. Table 3 shows that the difference between the initial and final F2_Q in KH is –0.37, indicating a far greater degree

of diphthongization than in CH and IR, for which the differences are -0.18 and -0.15 respectively. For SI the difference is -0.05 and for SA it is 0.00 , which indicates that these dialects have little or no tendency toward diphthongization.

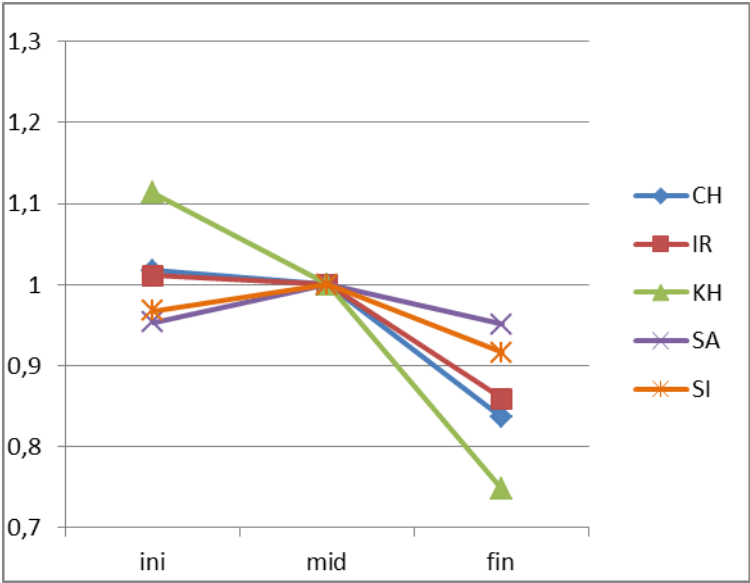


Figure 12. Comparison of the degree of diphthongization of *se:b* ‘apple’ in different dialects. Further explanation in text as well as in the caption to Figure 11.

Dialect	CH	IR	KH	SA	SI
F2					
F2 _Q -ini	1,017641	1,011199	1,113282	0,953745	0,967414
Baseline	1	1	1	1	1
F2 _Q -fin	0,836703	0,85938	0,747476	0,951256	0,916288
F2 _Q -ini – F2 _Q -fin	-0,18	-0,15	-0,37	0,00	-0,05

Table 3. F2_Q values for the word *se:b* ‘apple’ in different dialects; initial F2_Q is given in row one, the baseline in row two, final F2_Q in row three, and the difference between initial and final F2_Q in the bottom row.

5. Discussion and conclusions

The data show that the speakers in these five dialects have varying tendencies toward diphthongization, with some of the speakers displaying diphthongization more strongly in their speech than others. Generic factors such as speech rate and idiolect cause some of the variability in the degree of diphthongization, but overall, the primary source of this variability in diphthongization is greater tendency of the speakers of the KH dialect to produced diphthongized variants.

The speakers of IR, which has the second highest rate of diphthongization among the dialects investigated, do not show the same degree of diphthongization in their speech as do the KH speakers. This means that there are eligible words which are not always pronounced with a diphthong by the speakers of this dialect. It seems that the females use the diphthongs more markedly in their utterances than the males, but still diphthongization is seen in the speech of all speakers in the IR dialect. The speakers of the CH dialect show diphthongization to a less extent than the speakers of KH and IR, and diphthongization does not appear in all eligible words. The female speakers show more diphthongization than the males, yet they do not utter all target words with a diphthong. Finally, in the SA and SI dialects, there is no strong evidence of a diphthongization process in the speech of any of the speakers investigated.

It was observed that in the word-final position, regardless of the environment, the degree of diphthongization of /o:/ is weaker and it is less salient in impressionistic analysis, e.g. in *bo*: 'smell' as *bue*, *na:ko*: 'uncle' as *na:kue*. In addition, it seems that in retroflex contexts, the degree of diphthongization of /o:/ is lesser than in the other contexts, e.g. before the retroflex /ɾ/ in the word *na:jo:ɾ* 'ill'. However, the data show no apparent restriction on the diphthongization of /e:/ in different contexts.

Factors such as age, gender, and education do not seem to correlate with the diphthongization in the Balochi dialects spoken in Iran. Because various diphthongs have been reported for different Balochi dialects in general, and because there are no diphthongs in the systems of the surrounding languages and dialects, the appearance of diphthongs in the Iranian Balochi dialects can be attributed to an inherent tendency. A similar process of monophthongization, the opposite of diphthongization, has been taking place in Standard Persian which is the dominant surrounding language of these dialects. In this process *ow* shifts to *o:*.⁷

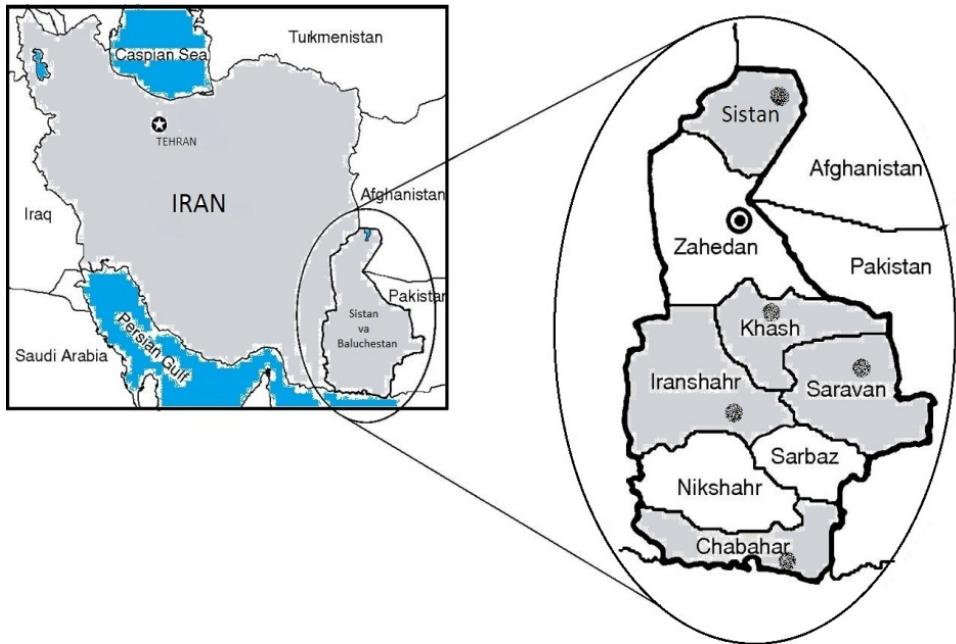
To conclude, diphthongization is a predominant feature in the production of the Common Balochi *e:* and *o:* in the Khash dialect, and is observed to occur to a lesser degree in the Iranshahr dialect and only occasionally in the Chabahar dialect. Diphthongization is practically absent in the Sistan and Saravan dialects, as only in a few cases was a weak tendency toward diphthongization observed. The available evidence shows that neither factors such as age, sex, or education, nor the influence of Persian, the surrounding dominant language with no diphthongs in its inventory, can account for the diphthongization in the Balochi dialects spoken in Iran. Instead, this phenomenon seems to be a system-internal development constituting a vowel change in the KH dialect, and, possibly, the start of a process of vowel change in the IR and CH dialects.

Seeing phonemic representation as an abstract representation of relations in the vowel system, the traditional phonemic representation of the Common Balochi *e:* and *o:* as /e:/ and /o:/ can be used for descriptive purposes in the dialects discussed here, although for the KH dialect, it seems more straightforward to describe these vowels phonemically as the diphthongs /ie/ and /ue/.

⁷ An example of this process is *xosrow* pronounced as *xosro*: 'Khosro'; of course the form *ow* is still used in poetry, and also exists in some Persian dialects.

References

- Aldaghi, Anita, and Tavakoli, Nastaran (2011). "Major Phonetic Processes in Sabzevari Dialect." *Theory and Practice in Language Studies* 1(3), pp. 292–301.
- Baranzehi, Adam Nader (2003). "The Sarawani Dialect of Balochi and Persian Influence on It." In: Jahani, Carina, and Korn, Agnes (eds), *The Baloch and Their Neighbours: Ethnic and Linguistic Contact in Balochistan in Historical and Modern Times*. Wiesbaden: Reichert, pp. 75–111.
- Barker, Muhammad Abd-al-Rahman, and Mengal, Aqil Khan (1969). *A Course in Baluchi*, I. Montreal: McGill University.
- Dames, M. Longworth (1891). *A Text Book of the Balochi Language: Consisting of Miscellaneous Stories, Legends, Poems, and a Balochi-English Vocabulary*. Lahore: Punjab Government Press.
- Ebrahimi, Ghorbanali (2012). "Investigating the Linguistics Features of Bayza Dialect." *Journal of Language Teaching and Research* 3(3), pp. 369–378.
- Elfenbein, Josef (1989). "Balochi." In: Schmitt, Rüdiger (ed.), *Compendium Linguarum Iranicarum*. Wiesbaden: Reichert, pp. 350–362.
- Farrell, Tim (1990). *Basic Balochi, An Introductory Course* [Baluchistan Monograph Series, 1]. Naples: Istituto Universitario Orientale.
- Ganjavi, Shadi, Georgiou, Panayiotis G., and Narayanan, Shrikanth (2003). "ASCII Based Transcription Systems for Languages with the Arabic Script: The Case of Persian." In: *Proceeding of Automatic Speech Recognition and Understanding (ASRU) 2003*, pp. 595–600.
- Gilbertson, George Waters (1923). *The Balochi Language. A Grammar and Manual*. Hertford.
- Grierson, George A. (1921). "Balōchī." In: Grierson, George A. (ed.), *Linguistic Survey of India (LSI)* X, pp. 327–451.
- Hakimi, Abdolazim (2012). "Comparative Phonetic Study of Frequently Used Words in Iranian Farsi versus Tajik Farsi." *Journal of American Science* 8(4), pp. 6–16.
- Hatami, Mohammad-Hasan (2011). "Phonetic Processes in Ruini Dialect." *Journal of American Science* 7(8), pp. 679–682.
- Hussain, Qandeel, Mahmood, Rashid, and Mahmood, Muhammad Asim (2011). "Vowel Substitution: A Comparative Study of English Loans in Punjabi and Urdu." *International Journal of Linguistics* 3(1), pp. 1–13.
- Jahani, Carina (2003). "The Case System in Iranian Balochi in a Contact Linguistic Perspective." In: Jahani, Carina, and Korn, Agnes (eds), *The Baloch and Their Neighbours: Ethnic and Linguistic Contact in Balochistan in Historical and Modern Times*. Wiesbaden: Reichert Verlag, pp. 113–132.
- Jahani, Carina, and Korn, Agnes (2009). "Balochi." In: Windfuhr, Gernot (ed.), *The Iranian Languages*. London and New York: Routledge, pp. 634–692.
- Mahmoodzahi, Moosa (2003). "Linguistic Contact in Iranian Balochistan in Historical and Modern Times." In: Jahani, Carina and Korn, Agnes (eds), *The Baloch and Their Neighbours: Ethnic and Linguistic Contact in Balochistan in Historical and Modern Times*. Wiesbaden: Reichert, pp. 147–156.
- Mirdehghan, Mahinnaz, and Moradkhani, Simin (2010). "Personal Pronouns in the Kakavandi Laki Dialect of Harsin (Kermanshah, Iran)." *Iranian Studies* 43(4), pp. 513–531.
- Samare, Yadollah (Samare, Yadollāh) (1368 [1989–90]). *Āvāšenāsi-ye zabān-e fārsi, āvāhā va sāxt-e āvāyi-ye hejā*, Tehran: Markaz-e Našr-e Dānešgāhi.
- Sjolander, K., and Beskow, J. (2000). "WaveSurfer – an Open Source Speech Tool." In: Yuan, B., Huang, T., and Tang, X. (eds), *Proceedings of ICSLP 2000, 6th International Conference on Spoken Language Processing*. Beijing, pp. 464–467.
- Spooner, Brian (1967). "Note on the Balochī Spoken in Persian Baluchistan." *Iran, Journal of the British Institute of Persian Studies* 5, pp. 51–71.
- Yaesoubi, Maziar (2010). *Machine Transcription Conversion between Perso-Arabic and Romanized Writing Systems*. Master thesis, Linköping University. Online at: <http://liu.diva-portal.org/smash/record.jsf?pid=diva2:360341> [Retrieved November 10, 2012].



Map 1. Map of Iran with Sistan and Baluchestan province enlarged

Iranian Linguistics

Eds. Geoffrey Haig, Otto-Friedrich-Universität, Bamberg
Carina Jahani, Uppsala University

Introduction

Geoffrey Haig, Otto-Friedrich-Universität, Bamberg

geoffrey.haig@uni-bamberg.de

Carina Jahani, Uppsala University

carina.jahani@lingfil.uu.se

This special section in *Orientalia Suecana* 61 is a selection of papers presented at the Fourth International Conference on Iranian Linguistics (ICIL4) held in Uppsala, June 17–19, 2011. The conference was organized by Carina Jahani, Uppsala University, Simin Karimi, University of Arizona, Tucson, and Agnes Korn, Goethe University, Frankfurt am Main. Originally more than 100 abstracts were submitted to the conference, from which 36 were selected for oral presentations and 19 for poster presentations. A number of contributors to the conference subsequently submitted their papers for publication in this issue of *Orientalia Suecana*, and after the reviewing process, the seven papers below were finally selected for publication.

The aim of the International Conferences on Iranian Linguistics is to provide a common venue for scholars from different disciplines, such as general linguistics, Iranian studies, and comparative Indo-European studies, who share a research focus on the Iranian languages. The first ICIL conference was held in Leipzig in 2005, and thereafter it has become a bi-annual event, hosted in turn by the University of Hamburg (2007), the University of Sorbonne Nouvelle, Paris (2009), and Uppsala University (2011). Selections from two previous conferences have been published, Karimi *et al.* (2008) and Korn *et al.* (2011). This section in the 2012 issue of *Orientalia Suecana* thus continues a now well-established tradition of publishing a selection of state-of-the-art contributions covering a broad spectrum of topics in current Iranian linguistics.

Iranian languages are spoken over a vast geographic area, stretching from the Pamir region in the northeast and across Central Asia to the Caucasus, southern Russia, and eastern Turkey in the northwest. On the Iranian plateau, different languages belonging to the Iranian family are spoken all the way to the southern shores of the Persian Gulf and even across the Gulf on the Arabian Peninsula. In the southeast there are speakers of Iranian languages all the way to Sindh province in Pakistan. Studies of linguistic contact between Iranian languages and particularly Turkic, Semitic, and other Indo-European languages, e.g. languages from the Indo-Aryan family, Armenian, and Tocharian, are therefore highly rewarding.

Persian has for centuries enjoyed a special position among the Iranian languages. For more than a thousand years it has been used at courts on the Iranian plateau, in Turkey, Central Asia, and India, where it developed as the vehicle of an elevated written literature. Long before that, in its Middle Persian form, it was the medium of a rich oral literary tradition as well, and oral literature has continued to play an important role in Persian in modern times. Persian was also an important *lingua franca* among traders along the Silk Road. There are today more than 100 million speakers

of Persian (in its three dialect variants Farsi, Dari, and Tajik), either as first or second language.

Persian is, however, not the only Iranian language with several million speakers. Kurdish, Pashto, and Balochi each count around 10 million speakers or more, and in historical times Sogdian was a large Iranian language spoken over vast areas in Central Asia. But there are also very small languages belonging to the Iranian family, among others the so-called “Pamir languages” spoken in southern Tajikistan, eastern Afghanistan, and northern Pakistan, some of which only have a few thousand speakers. The area of Iranian linguistics therefore offers rich potential for a highly diversified range of diachronic and synchronic studies, covering a considerable number of languages in various geographic and socio-cultural settings.

The individual articles in this section of *Orientalia Suecana* are arranged alphabetically, based on the surname of the author. It would, of course, have been possible to arrange the articles based on content as well, but we feel that the themes of the articles are rather intertwined, and that any thematic split would be rather arbitrary. There are three articles that treat phonological subjects. Kümmel and Miller deal with phonological changes over time and Hosseini with prosody and syllabification in Persian in connection with clitics. Cliticization is also discussed by Nourzaei and Jahani, though this time in Balochi. Levinsohn also deals with Balochi, but investigates the discourse level. Shokri and Öpengin both study morphosyntax, but there are clear links between the articles by Shokri and by Nourzaei and Jahani as well, since both deal with the verb phrase. There is also a good range of specific languages under study: Balochi, Kurdish, Mazandarani, and Persian. Kümmel’s article is a general discussion of phonological development and not limited to any one Iranian language. Kümmel and Miller have a clear diachronic perspective in their articles, while the rest are of a mainly synchronic nature.

The first article, written by Ayat Hosseini, is entitled “The Prosodization of Function Words in Persian” and investigates the prosodic structure of stressless function words in Persian (enclitics and proclitics) within the framework of Prosodic Phonology. Hosseini argues for an analysis that differentiates enclitics from proclitics. Enclitics group with their hosts in a recursive structure, forming a single PWord, while proclitics exhibit a PWord boundary between themselves and their hosts.

In the next article, Martin Joachim Kümmel studies “The Iranian Reflexes of Proto-Iranian **ns*” and proposes a solution uniting the apparently diverging Avestan *tqθra-* ‘darkness’ with corresponding words in other Iranian languages under a common preform. He also gives a similar explanation for a parallel case of apparent variation, namely Avestan *pqsnu-* ‘ashes’ and its cognates, and concludes by discussing the development of Proto-Iranian **ns* in Iranian languages and its relative chronology.

In his article entitled “Introducing Reported Speeches in Balochi of Sistan with *ki*”, Stephen H. Levinsohn investigates the hypothesis that *ki* in Balochi is used as an introducer of a reported speech when the consequences of the speech are highlighted, rather than the speech itself. He finds that the main function of *ki* is to indicate that the following speech is to be understood as a representation of an utterance or thought, rather than its verbatim replication.

In “Variation in Persian Vowel Systems”, Corey Miller addresses the diachronic development of the Early New Persian vowel system in the three contemporary variants of Persian – Farsi, Dari, and Tajik – with respect to Labov’s principles of vowel shifting. He concludes that the changes that have taken place in the Persian vowel system sometimes corroborate Labov’s findings and sometimes provide alternative directions. The notion of a pan-dialectal vowel system is explored both as a theoretical construct in understanding the diachrony of vowel systems, but also for possible pedagogical applications.

Maryam Nourzaei and Carina Jahani investigate “The Distribution and Role of the Verb Clitic =a/a= in Different Balochi Dialects”, including four varieties spoken west of Balochistan proper. They discuss the enclitic versus proclitic status of the verb clitic in altogether eleven Balochi dialects and find that the verb clitic has proclitic status in three of the four previously undescribed westernmost dialects, whereas it occurs as an enclitic, attaching to the word preceding the verb with certain restrictions to what word can host it, in a majority of the Balochi dialects for which there are descriptions available, and in one of the dialects under study it is totally absent.

In his article entitled “Adpositions and Argument Indexing in the Mukri Variety of Central Kurdish: Focus on Ditransitive Constructions” Ergin Öpengin describes how adpositions in this dialect may have the syntactic functions of introducing a third participant into the speech event, a causee in an indirect causation construction, and a “weakened actor” in the passive construction. Adpositions combine with pronominal clitics to yield a system of argument indexing that is of considerable complexity and has hitherto not been adequately analysed.

Guiti Shokri studies “Past and Non-past Structures in the Mazandarani Dialect Spoken by the Galesh of Ziarat” and finds that the non-past indicative verb forms in this dialect combine Persian indicative prefixes and personal endings with a Mazandarani stem. In the simple past, the dialect of Ziarat follows a Mazandarani structure, apart from the personal endings, which are Persian, except in the third person singular. This is an example of the comparatively rare instance of a language borrowing entire inflectional paradigms in a language contact setting.

As is evident, the contributions to this section cover a broad range of topics and languages, and illustrate several distinct theoretical traditions. We are very grateful to the authors for submitting their articles to *Orientalia Suecana* and for their co-operative attitude during the revision process. Sincere thanks also to the anonymous reviewers of the submitted papers. We are indebted to Everett Thiele for a highly professional checking of the English and stylistic accuracy of the articles, and to John Wilkinson, Textgruppen i Uppsala AB, for typesetting.

Publications from previous ICIL

Karimi, Simin, Samiiian, Vida, and Stilo, Donald (eds) (2008). *Aspects of Iranian Linguistics*, Newcastle upon Tyne: Cambridge Scholars Publishing. ICIL1.

Korn,ANGES, Haig, Geoffrey, Karimi, Simin, and Samvelian, Pollet (eds) (2011). *Topics in Iranian Linguistics*. Beiträge zur Iranistik, 34. Wiesbaden: Reichert. ICIL3, reviewed in this volume.

The Prosodization of Function Words in Persian

Ayat Hosseini, The University of Tokyo
hosseini@phiz.c.u-tokyo.ac.jp

Abstract

This paper investigates the prosodic structure of weak (stressless) function words in Persian within the framework of Prosodic Phonology. Weak function words in Persian are prosodic clitics that form phrases with the material following them (proclitics) or preceding them (enclitics). So far, only the prosody of enclitics has been studied in the literature. The present study proposes that, in Persian, enclitics are affixal clitics, while proclitics are free clitics. These proposals contribute to the findings in Prosodic Phonology that prosodic categories can in fact be recursive and non-exhaustive.

Keywords: Persian, stress, prosodic structure, function words, recursivity, exhaustivity

1. Introduction

Following the standard prosodic distinction between lexical and function words set forth in Selkirk (1995) and subsequent work, we assume that morpho-syntactic words in Persian can be divided into two groups. The first group consists of words which can have stress, and since most of these belong to the category of lexical or content words, following the literature, we refer to them as lexical (*lex*) words. The second group comprises weak function words (*func*) which are always stressless, except when focused or used metalinguistically. In Persian, demonstrative and personal pronouns such as *un* ‘that, he/she’ or *to* ‘you’, although being semantically function words, are stressed and thus belong to the first group.

Weak function words are prosodic clitics; i.e., they combine with other material into a single prosodic unit rather than being autonomous units themselves (Anderson 2005: 42). In this sense, Persian has both enclitics and proclitics. This study proposes that, in Persian, enclitics are affixal clitics, while proclitics are free clitics. Following the definitions in Selkirk (1995), in affixal clitics the lexical host comprises its own prosodic word, and the clitic is adjoined to the host in a recursive structure. Free clitics on the other hand are incorporated directly into the phonological phrase, as a sister of the prosodic word containing the lexical word, in a non-exhaustive structure.

Some of the most frequently used enclitics in Persian are the indefinite article and relative particle *-i*, the direct object particles *-o* and *-râ*, the conjunction *-o*, the Ezafe particle *-e*, the verbal copula *-e*, the particle meaning ‘also, even’ *-am / -ham*, pronominal suffixes such as *-et* ‘your’, and personal agreement suffixes when attached to past tense verbs. Some examples of proclitics are class1 (true) prepositions; class2a prepositions, which are proclitics only when used without Ezafe linker (Pantcheva 2006: 1); the conjunctions *va* and *yâ*; and the particle *ke* in some of its usages.

The phonological theory applied in this study is known as Prosodic Phonology, Phrasal Phonology or Prosodic Hierarchy Theory, and is based on the work of such

researchers as Selkirk (1978, 1981, 1995, 2011), Nespor and Vogel (1986), Ito and Mester (2009) among others. The main idea of Prosodic Phonology, as shown in (1), is that utterances are phrased in prosodic constituents which are themselves organized into a hierarchy.

- | | | |
|-----|-------------------------------|---|
| (1) | Intonational Phrase (IPhrase) | ι |
| | Phonological Phrase (PPhrase) | φ |
| | Prosodic Word (PWord) | ω |
| | Foot | π |
| | Syllable | σ |

2. Background

The prosodic structure of proclitics in Persian has not been specifically explored in previous literature, though there are a number of works that study enclitics. In this section we briefly present Amini (1997), Kahnemuyipour (2003), and two other approaches, and point out their theoretical and empirical problems.

2.1. Amini (1997)

Amini proposes for Persian enclitics a non-recursive exhaustively parsed structure in which enclitics form a single Phonological Word (PWord) with their hosts. She then explains the lack of stress of these enclitics using Extrametricality Theory. The examples in (2) demonstrate Amini's approach. The last syllable in (2i) and the last two ones in (2ii) are considered extrametrical.

- | | | | | | |
|-----|-----|--------------|---------|---|-----------------|
| (2) | i. | ketâb | -i | → | ke.tấ <bi> |
| | | book | IND | | |
| | | 'a book' | | | |
| | ii. | ketâb | -ešun | → | ke.tấ <be.šun> |
| | | book | POSS3pl | | |
| | | 'their book' | | | |

Firstly, Amini's OT analysis lacks an active constraint which triggers extrametricality in cliticized words exclusively. More importantly, according to Hayes (1995: 57) only constituents (segment, syllable, foot, phonological word, etc.) may be marked as extrametrical, but in Persian, we frequently encounter cases in which more than one enclitic is attached to a host resulting in a clitic cluster. In most cases, such clusters, which are considered to be "extrametrical" in Amini's analysis, do not form a constituent in either a prosodic or morphological sense. Example (3) illustrates such cases:

- | | | | | | | |
|-----|-----|------------------|---------|-----|---|----------------------|
| (3) | i. | ketâb | -ešun | -am | → | * ke.tấ <be.šu.nam> |
| | | book | POSS3pl | too | | |
| | | 'their book too' | | | | |
| | ii. | ketâb | -etun | -o | → | * ke.tấ <be.tu.no> |
| | | book | POSS2pl | ACC | | |
| | | 'your book-ACC' | | | | |

If we assume with Amini (1997) that stresslessness of enclitics in Persian is due to extrametricality, then the entire strings within <> in example (3) must be considered as extrametrical constituents. However, these strings are not a single prosodic or morphological constituent, and thus cannot be subject to extrametricality. Note that in cases like (3) an approach assuming a cyclic process in which each morpheme is marked as extrametrical in a different cycle is also unacceptable, because the extrametricality rule/constraint has access only to the final constituent of a prosodic category.

In sum, explaining the lack of stress on Persian enclitics by extrametricality is problematic, since it fails to account for a wide range of cases. Therefore, it is impossible to take enclitics and their hosts as a single non-branching PWord as suggested in Amini (1997).

2.2. Kahnemuyipour (2003)

In this comprehensive study, prosodic parsing is exhaustive in all levels and no recursive structure is allowed. All morpho-syntactic words including weak function words such as enclitics form their own PWords, and due to the phrasal stress rule, the leftmost PWord in the phonological phrase (PPhrase) takes stress. The reason enclitics never receive stress is that these functional morphemes happen to follow lexical words, and the phrasal stress rule puts the stress on the leftmost lexical word: $((ketáb)_\omega (i)_\omega)_\phi$ ‘a book’.

As discussed in the literature (Selkirk 1995; Anderson 2005; Ito and Mester 2009) weak (stressless) function words cannot constitute PWords by themselves, and essentially tend to phrase with their adjacent lexical words. So theoretically it is inappropriate to take Persian enclitics as autonomous PWords in the absence of any phonological or phonetic evidence.

Taking Persian function words as independent PWords also suffers from a major empirical problem. Kahnemuyipour (2003) does not deal with proclitics, but if all morpho-syntactic words, including enclitics, can constitute PWords, then proclitics such as class1 prepositions should also have the right to form PWords on their own. However, in the case of most proclitics, this analysis will wrongly predict the PPhrase stress on the phrase-initial proclitic as shown in (4).

- (4) i. [az Tehrân]_{pp} → * ((áz)_ω (Tehrân)_ω)_φ
 from Tehran
- ii. [be un xune]_{pp} → * ((bé)_ω (un)_ω (xune)_ω)_φ
 to that house

In brief, treating weak function words in Persian as independent PWords encounters theoretical and empirical problems and should therefore be avoided.

2.3. Other Proposals

Clitic Group (CG) and Accentual Phrase (AP) categories have also been proposed for Persian.¹ Both CG and AP are composed of a lexical word and its enclitics. One

major problem with these approaches is that they do not take proclitics into consideration, and thus fail to account for the asymmetry between proclitics and enclitics. More importantly, the necessity of such additional and non-universal categories as AP and CG has been questioned and rejected frequently in the literature (Selkirk 1995, 2011; Anderson 2005; Ito and Mester 2009; Féry 2010). As pointed out in Ito and Mester (2009: 168), Occam's razor militates against any additional category such as Clitic Group, as long as the existing ones are sufficient to represent the prosodic phrasings.

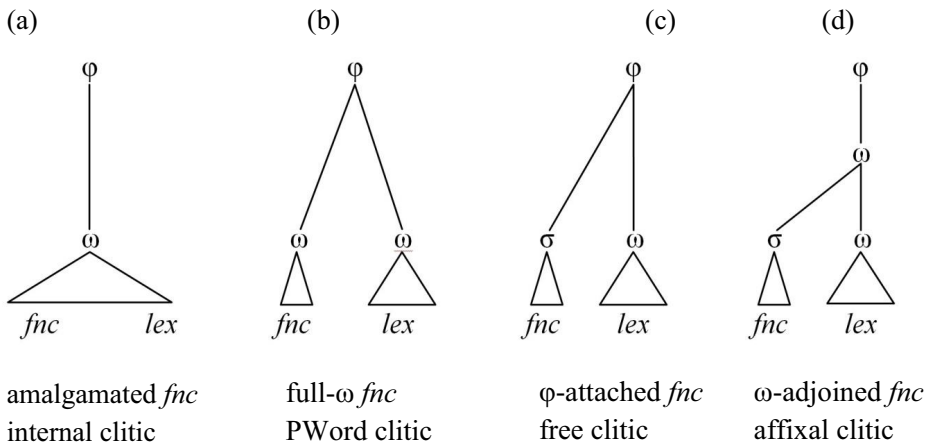
3. Discussions and proposals

This section examines the prosodization of Persian function words, using phonetic and phonological evidence.

3.1. Proclitics

As Selkirk (1995) points out, for any given host-plus-clitic combination, irrespective of their order, there are essentially four formal possibilities. These four possibilities are illustrated for proclitics in (5).

(5)



A crucial point in (5) is that the proclitic *fnc* is ω -initial (in the initial position of a PWord) in (a), (b), and (d), but not in (c). The *fnc* in (c) is the only one which is not parsed into ω , and is immediately dominated by the ϕ (PPhrase) node, although in all four structures the proclitic *fnc* is ϕ -initial.

Selkirk (1995) reports that the initial position in a PWord is often associated with effects involving the phonetic realization of segments. In English, for example, a word-initial voiceless stop is aspirated, even when the syllable to which it belongs is stressless.

¹ For a detailed discussion of AP see Sadat Tehrani (2007), and for CG, refer to Abolhasanizadeh, Gusenhoven and Bijankhan (2012).

Cooper (1991, 1994), as cited in Selkirk (1995), shows that there is a distinct word-initial aspiration effect which cannot be reduced to a simple syllable-initial effect. Prosodic structure theory takes such “word-initial” effects to be PWord-initial effects.

Aspiration in Persian is studied in Samare (1985: 27) according to which all voiceless stops are aspirated, but the degree of aspiration depends on the position of the stop. In the onset position of stressed syllables and word initially, stops are more aspirated. For example /p/ in *parváz* ‘flight’ and *sepáh* ‘army’ are more aspirated than in *sepâye* ‘tripod’. If we accept the claim in Samare (1985), then it becomes clear that Persian is similar to English in having a word-initial aspiration effect.

Therefore, if we can demonstrate that proclitic-initial stops are significantly less aspirated than word-initial stops, then we can claim that proclitics are not located in the initial position of a PWord, and this will support the structure in (5c).

An experiment was conducted to compare the degree of aspiration between proclitic-initial stops and word-initial stops.

3.1.1. PWord-initial aspiration, an experiment

In an experiment, the Persian class 1 preposition *tâ* and class 2a preposition *tu* were used in two sentences before polysyllabic nouns starting with the syllables /ta-/ and /tu-/, in non-t-initial positions.

Six Persian native speakers read each of the two sentences (6) and (7) twice, and the utterances were recorded at a sampling frequency of 44,100 Hz and analysed using the phonetics software Praat.

(6) *dišab* *tu* *tunél* *tasâdóf* *kard-am*
 last night in tunnel accident do.PAST-1SG
 ‘I had an accident in the tunnel last night.’

(7) *behrúz* *tâ* *tâbestún* *kár* *kard-Ø*
 Behruz until summer work do.PAST-3SG
 ‘Behruz worked till summer.’

The Voice Onset Time (VOT) in proclitics and word-initial syllables was measured and the mean values and standard deviations were calculated. A paired Student’s t-test was conducted to determine the significance of differences between VOT values in the two positions. The significance threshold of the t-test was 0.01, the degree of freedom being the number of participants minus one. Table 1 presents a summary of the statistical analyses.

	<i>ta</i> (proclitic)	<i>ta-</i> (word-initial)	<i>tu</i> (proclitic)	<i>tu-</i> (word-initial)
Mean VOT (ms)	22.35	51.35	26.55	62.99
Standard deviation	3.597	3.68	4.13	6.47
t-test results	t(5) = 131.13, p < 0.01		t(5) = 106.2, p < 0.01	

Table 1. Summary of VOT analyses

As is apparent from the t-test results, the VOT value (degree of aspiration) in proclitics is significantly less than that in word-initial position, showing that there is less aspiration on proclitics than on word-initial syllables. In Figure 1, two examples of the recorded data in this experiment are presented, in which the substantial difference between the aspiration of initial voiceless stop on proclitics and the first syllable of lexical words is observable.

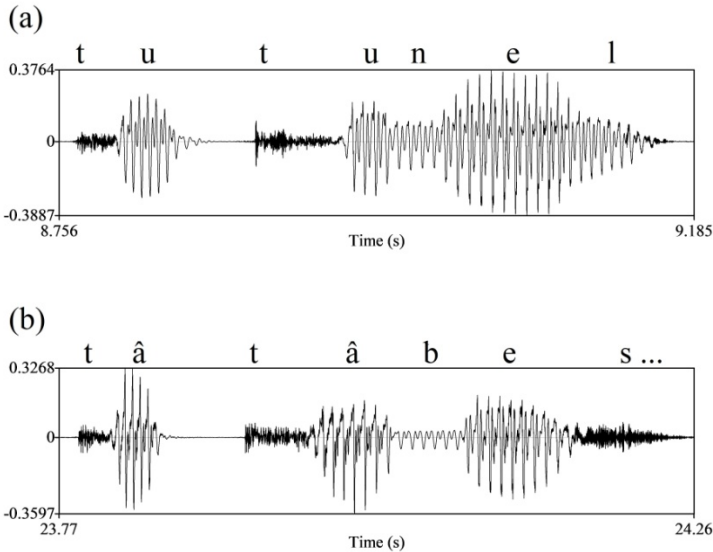


Figure 1. VOT differences observable in recorded data

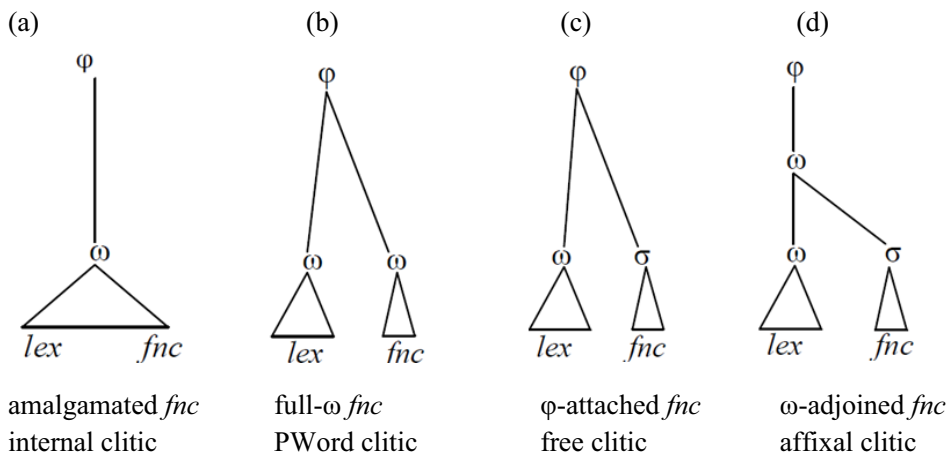
Based on these findings, we can conclude that proclitics in Persian are not in PWord-initial position; thus they are not ω -adjoined, and rather are ϕ -attached free clitics as shown in (5c). Section 4 provides more evidence for this claim.

3.2. Enclitics

The structures in (8) demonstrate the four possibilities of host-plus-clitic combinations for enclitics.

The structure in (8a) is the one proposed in Amini (1997), and as we discussed in section 2.1 it is not acceptable because if we take Persian enclitics as amalgamated function words, their lack of stress cannot be explained by extrametricality. The structure in (8b), in which weak function words are independent phonological words, is the one proposed in Kahnemuyipour (2003), and as we saw in section 2.2, this too is not acceptable due to both theoretical and empirical problems. Consequently, there are only two remaining possibilities for the structure of enclitics in Persian, namely (8c) and (8d). The main difference between (8c) and (8d) is that in the latter, the clitic and its host are parsed into the same PWord, but in the former, the clitic is not parsed into the PWord. The evidence for our argumentation comes from syllabification.

(8)



3.2.1. Syllabification

Syllabification is carried out on different levels. At the lexical level, syllabification rules/constraints require all segments in the domain to syllabify with each other, but at post-lexical levels, factors such as high rate of speed can cause syllabification to optionally take place at domain boundaries. Laeufer (1995: 118–119) states that syllabification applies cyclically to roots and is reapplied after each word formation process to incorporate newly added elements. In some languages however it is sensitive to morphological structure, while in others it is not. In other words, in some languages the domain of lexical syllabification is the morpheme, and in others the word.

The canonical syllable type in Persian is (C)V(C)(C) and the prosodic requirements of the syllable, such as obligatory onset, allowing no consonant clusters in onset position, and allowing at most two segments in coda position, determine the syllable structure in the word domain, ignoring the morpheme boundaries (Mahootian 1997: 303–305). In (9i and ii) morpheme boundaries are shown by ‘-’ on the left, while a ‘.’ marks syllable boundaries on the right.²

(9)

- i. dân-eš-mand-ân → *dâ.neš.man.dân* / **dân.(?)eš.mand.(?)ân* ‘scientists’
- ii. ham-âhang-i → *ha.mâ.han.gi* / **ham.(?)â.hang.(?)i* ‘harmony’

As can be seen, the syllabification is not sensitive to morphological structure and takes the whole word as its domain. Since in Prosodic Phonology PWords are the equivalents of words in syntactic constituent structure (Selkirk 2011), we can conclude that in Persian lexical syllabification takes place in PWord domains. However at the post-lexical level, resyllabification is also possible at higher levels, ignoring PWord boundaries. This is shown in (10):

² Word-initial glottal stops are shown in parentheses, since their existence in phonological representation is controversial and is beyond the scope of this study.

(10)

- i. $((\text{har})_{\omega} (\text{âdam})_{\omega})_{\phi} \rightarrow \text{har. } (\text{ʔ})\text{â.dam} / \text{ha.ra.dam}$ ‘each human’
 ii. $((\text{panj})_{\omega} (\text{angošt})_{\omega})_{\phi} \rightarrow \text{panj. } (\text{ʔ})\text{an.gošt} / \text{pan.jan.gošt}$ ‘five fingers’

The syllabic structures to the left of the slash (/) show the lexical level syllabification, while the structures to the right show the post-lexical resyllabification, which ignores PWord boundaries. The existence of such resyllabifications proves that the obligatory syllabification is only a characteristic of the PWord domain and does not necessarily apply to higher domains such as PPhrase.

A close investigation of syllabification in cliticized words reveals that content words and their enclitics are necessarily syllabified with each other and it is not possible to syllabify them independently, as shown in (11).

(11)

- i. $[\text{mârd}]_{\text{lex}} \quad [\text{i}]_{\text{fnc}} \rightarrow \text{mâr.di} / * \text{mârd.}(\text{ʔ})\text{i}$ ‘a man’
 ii. $[\text{pedâr}]_{\text{lex}} \quad [\text{am}]_{\text{fnc}} \rightarrow \text{pe.dâ.ram} / * \text{pe.dâr.}(\text{ʔ})\text{am}$ ‘my father’
 iii. $[\text{ketâb}]_{\text{lex}} \quad [\text{o}]_{\text{fnc}} \rightarrow \text{ke.tâ.bo} / * \text{ke.tâb.}(\text{ʔ})\text{o}$ ‘book-ACC’
 iv. $[\text{gârm}]_{\text{lex}} \quad [\text{e}]_{\text{fnc}} \rightarrow \text{gâr.me} / * \text{gârm.}(\text{ʔ})\text{e}$ ‘It’s hot’

This shows that host+enclitic combinations are more similar to the structures in (9) than those in (10), meaning that there are no PWord boundaries between a host and its enclitic. If we return to the two possible structures for enclitics mentioned in 3.2, we now have enough evidence to claim that the only acceptable structure for host+enclitic combinations is the one shown in (8d), which is a ω -adjoined affixal clitic.

There are two PWord nodes (shown by ω) in (8d). The lower PWord, which does not dominate any other PWord, is called a Minimal PWord, and the higher PWord, which is not dominated by any other PWord, is called a Maximal PWord. In Persian, the domain of obligatory syllabification is Maximal PWord, and stress is culminative and obligatory in this domain. Minimal PWord is right-aligned with a *lex*, and stress is rightmost in this domain.

4. Further evidence

In this section, three additional pieces of evidence are provided, that support the claim that proclitics and enclitics behave asymmetrically in Persian, and that there is a PWord boundary between a proclitic and its host, while enclitics and their hosts are parsed into a single PWord.

4.1. Proclitics and Syllabification

In 3.2.1. we argued that the obligatory syllabification of host+enclitic combinations in Persian suggests that there is no PWord boundary between a host and its enclitic. Proclitics, on the other hand, behave like the structures in (10), suggesting that they are not parsed into PWords together with their hosts and that there is a PWord boundary between them and their lexical hosts. This is exemplified in (12):

(12)

- | | | | | | |
|------|---------------------|-------------------------|---|--------------------------------|---------------|
| i. | [az] _{fn} | [injá] _{lex} | → | (ʔ)a.zin.já / (ʔ)az.(ʔ)in.já | ‘from here’ |
| ii. | [dar] _{fn} | [âsemân] _{lex} | → | da.râ.se.mân / dar.(ʔ)â.se.mân | ‘in the sky’ |
| iii. | [joz] _{fn} | [Írán] _{lex} | → | jo.zi.rân / joz.(ʔ)i.rân | ‘except Iran’ |

4.2. Uninterruptability

It is uncontroversial that it is prosodically unacceptable to insert a pause in the middle of a PWord: *[pe...dár]_ω. This principle is known as *uninterruptability* and is used as one of the major criteria of wordhood (Bauer 2003: 63–64). An examination of clitics in Persian shows that the insertion of a pause between a *lex* and its enclitic is not possible, however, one can make a pause between a *lex* and its proclitic. As shown in (13) pause insertion is not possible between a noun and the Ezafe particle (13a), between a verb and a personal agreement suffix (13b), or between a noun and the conjunction -o (13c). However a pause can be made between a proclitic and its host, for instance between the complementizer *ke* and the material following it (13d), a preposition and a noun (13e), and the conjunction *va* and the following noun (13f).

- | | | | | | |
|---------|----------------------------|----|-------------|----|-----------------|
| (13) a. | *mard...-e dānā | b. | *ráft...-im | c. | *zán...-o mard |
| | ‘wise man’ | | ‘we went’ | | ‘woman and man’ |
| d. | behrúz ke...dirúz | e. | az...šomá | f. | zan va...mard |
| | ‘Behruz that yesterday...’ | | ‘from you’ | | ‘woman and man’ |

4.3. Glide epenthesis

In Persian words no adjacent vowels are allowed (Windfuhr and Perry 2009: 429). Adjacent vowels are also avoided when the first of two vowels is the final vowel of a *lex* and the second one is the initial vowel of an enclitic. In this case, an epenthetic segment (mostly a glide) is inserted between the two vowels. In contrast, there is no such epenthetic segment insertion between a proclitic-final vowel and a *lex*-initial vowel. In (14i and ii) word-final vowels are followed by a clitic vowel, and to avoid vowel adjacency, a glide is epenthesized between the two vowels.

- | | | | | |
|---------|------------------------|-------------------|---|----------|
| (14) i. | [xodá] _{lex} | [i] _{fn} | → | xodá-yi |
| | God | IND | | |
| | ‘a God’ | | | |
| ii. | [pesté] _{lex} | [e] _{fn} | → | pesté-ye |
| | pistachio | EZ | | |
| | ‘pistachio-EZ’ | | | |

In (15i and ii) we have exactly the same vowels as in (14), but unlike (14), glide epenthesis between the two vowels is unacceptable.

- (15) i. $[bâ]_{fnc}$ $[în]_{lex}$ \rightarrow $*bâ-y-în$
 with this
 'with this'
- ii. $[be]_{fnc}$ $[eskâtland]_{lex}$ \rightarrow $*be-y-eskâtland$
 to Scotland
 'to Scotland'

In cases such as $bâ + -am \rightarrow bâhâm$ 'with me', it may seem that an epenthetic segment is inserted after the preposition $bâ$, suggesting that $bâ$ may not exclusively be a proclitic. However, as Naderi and van Oostendorp (2011: 164–165) observe, the insertion of /h/ in $bâhâm$ or $behem$ 'to me' is not productive, and happens in only a few frozen phrases, and should not be considered an epenthesis process in Persian. In these cases the second part of the combination (the nominal suffix $-am$) is a prosodically weak function word itself and cannot be a legitimate host for a proclitic. These fossilized combinations of two clitics form a single PWord in these examples, while $bâ$ and be , when used as true prepositions before lexical words, are always proclitics.

5. Multiple function words

In Persian, more than one weak function word can occur before or after a lexical word. In cases with multiple proclitics, e.g., $az\ tu\ xune$ 'from inside the house', we observe that none of the proclitics are stressed, suggesting that they do not form PWords on their own. Even in literary forms such as $vazân$ and $kazîn$, which are contracted forms of $va\ az\ ân$ 'and from that' and $ke\ az\ in$ 'that from this' respectively, the whole contracted form is cliticized to the following lex without itself being stressed.³ Also, in multiple proclitics there is more than one possible syllabification, similar to what we saw in (10) and (12). For example $joz\ az\ irân$ 'except from Iran', when produced without any focus, is a sequence of two proclitics and a host and can be syllabified in four ways:

$jo.za.zi.rân$ $joz.(?)a.zi.rân$ $jo.zaz.(?)i.rân$ $joz.(?)az.(?)i.rân$

The possibility of syllabifying each proclitic independently from the host and other proclitics, suggests that neither of the proclitics is parsed into the PWord, a domain in which, only one specific syllabification is allowed. Thus the structure for multiple proclitics would be multiple free clitics $(fnc_2 fnc_1 (lex)_\omega)_\phi$.

So far we have argued that enclitics in Persian are affixal clitics that adjoin to PWords in a recursive manner. In the case of multiple enclitics, e.g., $pedâreto$ 'father-POSS2sg-ACC', there are two possible structures:

- (16) i. $((lex)_\omega fnc_1 fnc_2)_\omega)_\phi$
 ii. $((lex)_\omega fnc_1)_\omega fnc_2)_\omega)_\phi$

³ Please also note that contracted forms like $kazîn$ or $kîn$ also support the proclitic nature of ke in such cases.

In the relatively plain structure in (16i) there are only two levels of PWord, namely a minimal and a maximal one, while in the more nested structure in (16ii) there is an intermediate level PWord between the minimal and the maximal one. In fact there seems to be neither a phonological process nor phonetic evidence in Persian supporting either of the structures in (16).

From a theoretical point of view, it is assumed in Prosodic Phonology that recursive morpho-syntactic structures and recursive prosodic structures should correspond to each other. As is asserted in Selkirk (2011) among others, prosodic structure reflects morphological structure as closely as possible. Kabak and Revithaidou (2009: 112–113) argue that in inherently recursive morpho-syntactic structures, the category of the whole construction is the same as its head. This category includes constructions that contain function words that are adjoined to syntactic lexical heads such as clitics. Based on these theoretical considerations, in the case of multiple enclitics above, the structure in (16ii) would be more appropriate because it mirrors the recursive morphological structure of the construction more accurately.

6. Conclusion

We have examined the existing proposals for the prosodic structure of function words in Persian and shown that taking function words as amalgamated clitics and independent PWords is theoretically and empirically problematic. The proposal made in this paper is that enclitics are grouped with their hosts in a recursive structure, while proclitics are not parsed into PWords, but rather are immediately dominated by a PPhrase in a non-exhaustive structure. In recursive PWords, syllabification is obligatory in the Maximal PWord domain, and stress is also obligatory and culminative in this domain. Minimal PWords are always right-aligned with a *lex* and stress is rightmost in their domain.

Abbreviations

ACC	accusative	PPhrase	phonological phrase
EZ	ezafe	PWord	phonological word
fnc	function word	SG	singular
IND	indefinite	ι	intonational phrase
IPhrase	intonational phrase	π	foot
lex	lexical word	σ	syllable
pl	plural	φ	phonological phrase
POSS	possessive	ω	phonological word

References

Abolhasanizadeh, Vahideh, Gussenhoven, Carlos, and Bijankhan, Mahmood (2012). “The Persian Pitch Accent and its Retention after the Focus.” *Lingua*, vol. 122, issue 13, pp. 1380–1394.

Amini, Afsaneh (1997). “On Stress in Persian.” *Toronto Working Papers in Linguistics* 16(1), pp. 1–20.

Anderson, Stephen (2005). *Aspects of the Theory of Clitics*. Oxford: Oxford University Press.

Bauer, Laurie (2003). *Introducing Linguistic Morphology*, 2nd ed. Washington, DC: Georgetown University Press.

- Cooper, André (1991). *An Articulatory Account of Aspiration in English*. Ph.D. dissertation, Yale University.
- Cooper, André (1994). "Aspiration in English: How Should it be Defined and Where Should it be Described?" (unpublished manuscript).
- Féry, Caroline (2010). "Recursion in Prosodic Structure." *Phonological Studies* 13, pp. 51–60.
- Hayes, Bruce (1995). *Metrical Stress Theory*. Chicago: University of Chicago Press.
- Ito, Junko, and Mester, Armin (2009). "The Extended Prosodic Word." In: Grijzenhout, J., and Kabak, B. (eds), *Phonological Domains: Universals and Deviations*. Berlin: Mouton de Gruyter, pp. 135–194.
- Kabak, Barış and Reviathiadou, Anthi (2009). "An Interface Approach to Prosodic and Word Recursion." In: Grijzenhout, J., and Kabak, B. (eds), *Phonological Domains: Universals and Deviations*. Berlin: Mouton de Gruyter, pp. 105–134.
- Kahnemuyipour, Arsalan (2003). "Syntactic Categories and Persian Stress." *Natural Language and Linguistic Theory* 21, pp. 333–379.
- Laeuffer, Christiane (1995). Morphology and Syllabification Domains. *Lingua* 97, pp. 101–121.
- Mahootian, Shahrzad (1997). *Persian*. Routledge, London.
- Naderi, Navid, and van Oostendorp, Marc (2011). "Reducing the Number of Farsi Epenthetic Consonants." In: Korn, Agnes, Haig, Geoffrey, Karimi, Simin, and Samvelian, Pollet (eds), *Topics in Iranian Linguistics*. Wiesbaden: Reichert pp. 153–166.
- Nespor, Marina, and Vogel, Irene (1986). *Prosodic Phonology*. Dordrecht: Foris Publications.
- Pantcheva, Marina (2006). "Persian Preposition Classes." In: Svenonius, Peter (ed.), *Nordlyd: Tromsø Working Papers in Linguistics* 33(1), pp. 1–25.
- Sadat Tehrani, Nima (2007). *The Intonational Grammar of Persian*. PhD dissertation, University of Manitoba.
- Samareh, Yadollah (Samare, Yadollâh) (1378 [1999–2000]). *Âvâšenâsi-ye Zabân-e Fârsi* [Phonetics of the Persian Language], 5th ed. Markaz-e našr-e dânešgâhi, Tehrân.
- Selkirk, Elisabeth (1978). "On Prosodic Structure and its Relation to Syntactic Structure." In: Fretheim, Thorstein (ed.), *Nordic Prosody II: Papers from a symposium*. Trondheim: TAPIR, pp. 111–140.
- Selkirk, Elisabeth (1981). "The Nature of Phonological Representation." In: Anderson, John, Laver, John, and Myers, Terry (eds), *The Cognitive Representation of Speech*. Dordrecht: North Holland Publishing Company, pp. 379–388.
- Selkirk, Elisabeth (1995). "The Prosodic Structure of Function Words." In: Beckman, Jill, Dickey, Laura W., and Urbanczyk, Suzanne (eds), *Papers in Optimality Theory. University of Massachusetts Occasional Papers* 18. Amherst, Mass.: GLSA, pp. 439–469.
- Selkirk, Elisabeth (2011). "The Syntax-phonology Interface." In: Goldsmith, John A., Riggle, Jason, and Yu, Alan (eds), *The Handbook of Phonological Theory*. Blackwell Publishing, 2nd edition, pp. 435–484.
- Windfuhr, Gernot, and Perry, John R. (2009). "Persian and Tajik." In: Windfuhr, Gernot (ed.), *The Iranian Languages*. London and New York: Routledge, pp. 416–544.

The Iranian Reflexes of Proto-Iranian **ns*

Martin Joachim Kümmel, Friedrich-Schiller-Universität Jena
martin-joachim.kuemmel@uni-jena.de

Abstract

The obvious cognates of Avestan *tqθra-* ‘darkness’ in the other Iranian languages generally show no trace of the consonant *θ*; they all look like reflexes of **tār°*. Instead of assuming a different word formation for the non-Avestan words, I propose a solution uniting the obviously corresponding words under a common preform, starting from Proto-Iranian **taNsra-*: Before a sonorant **ns* was preserved as *ns* in Avestan (feeding the change of tautosyllabic **sr > *θr*) but changed to **nh* elsewhere, followed by **anhr > *ā(h)r*. A parallel case of apparent variation can be explained similarly, namely Avestan *pqsnu-* ‘ashes’ and its cognates. Finally, the general development of Proto-Indo-Iranian **ns* in Iranian and its relative chronology is discussed, including word-final **ns*, where it is argued that the Avestan accusative plural of *a*-stems can be derived from **-āns*.

Keywords: Proto-Iranian, nasals, sibilant, sound change, variation, chronology

1. Introduction

The aim of this paper is to discuss some details of the development of the Proto-Iranian (PIr) cluster **ns* in the Iranian languages. Before we proceed to do so, it will be useful to recall the most important facts concerning the history of dental-alveolar sibilants in Iranian.

1) PIr had inherited a sibilant **s* identical to Old Indo-Aryan (Sanskrit/Vedic) *s* from Proto-Indo-Iranian (PIIr) **s*. This sibilant changed to Common Iranian (CIr) *h* in most environments, while its voiced allophone *z* remained stable all the time.

2) There was a different sound that yielded *s* in most of Iranian but *θ* in Old Persian (OP); this is in my view derived from a PIr affricate **c* [ts], that corresponds to Old Indo-Aryan (OIA) “palatal” *ś* [ʃ] from a PIIr palatoalveolar affricate **č* [tʃ] (cf. Lipp 2009: I: 139–149; 333–334), parallel to the corresponding voiced sound Av. *z* = OP *d* < PIr **j* [dz] = OIA *j* [dʒ], *h* [ɦ] < PIIr **j* [dʒ], **j^h* [dʒ^h]. Thus, PIr **c*, **j* as used here correspond to **ś*, **z* or **č*, **j* in other approaches (e.g., Hoffmann & Forssman 1996; Tremblay 2005: 677–679).

2. The words for ‘dark’, ‘darkness’

a. In Younger Avestan (YAv), there is a noun *tqθra-* that is only attested in the plural which has a collective sense ‘darkness’. An adjective *tqθriia-* ‘dark’ was derived from that stem. In most if not all other Iranian languages similar words with

¹ For valuable discussion and suggestions I wish to thank Agnes Korn, Reiner Lipp, Eva Tichy, and the reviewers and editors. Of course, I remain responsible for all shortcomings of this article. My research was made possible by a Heisenberg grant from the Deutsche Forschungsgemeinschaft (DFG).

identical meaning exist. However, they seem to presuppose slightly different pre-forms **tāhra-/tāra-*, **tāriya-*, cf. Middle Persian (MP) and Parthian *t'r /tār/* 'darkness', Khotanese (Khot.) *ttāra-* 'dark', Ossetic (Oss.) *tar* 'dark(ness)', Pashto *tor* 'black', *tyār ə* 'darkness', Balochi (Bal.) *tahār* 'dark' etc.²

b. The difference between **tanθr°* and **tahr°/tār°* has been explained as a regular change of **tanθr°* to **tāθr°* > **tāhr°* > **tār°* in later Iranian (cf. Morgenstierne 1974: 81a; 2003: 82–83). Indeed, CIr **θ* is frequently lost before *r*, often via **h*. This is the regular development in most of Western Iranian (though not of Persian proper and Balochi, see below), cf. especially New Persian (NP) *tār* 'string' < **tanθra-*, and PIr **manθra-* > *mahr* in Parthian = MP *mhr* 'mantra' < **manθra-* (Av. *māθra-*) with a variant *mār-*: **θr* first became **hr*; if **h* was lost, the vowel was lengthened; before preserved *hr*, long vowels were shortened (cf. Hübschmann 1895: 204–207; 1899: 20; Cantera 1998).

c. However, in some of the languages reflecting **tahr°/tār°*, the CIr cluster **θr* did not normally develop to *hr* or to simple *r*. E.g., **dāθra/i-* 'sickle' developed to MP and NP *dās*,³ Yaghnobi **dāθr* > **drāθ(r)* > *dirot/diros*, Pam. Bartangi *ḏōc*, Wakhi *ḏitr* etc. (Rastorgueva & Édél'man 2003: 440–441), **māθr-* 'mother' > Bal. *mās* (cf. Korn 2005: 89). In such words, **θ* was obviously preserved for quite a long time, and it would be difficult to understand why a word like **tāθra-* should have developed differently.

d. To explain **tāra-* in these languages, some scholars have therefore proposed⁴ assuming a different word formation outside of Avestan, namely **tmH-ra-* > **tāra-* (cf. Vedic *tāmrā-*). But of course, a solution with a unitary PIr preform should be preferred, if possible; and this solution does not work for Bal. *tahār* anyway.

3. PIr **nsr* and **sr*

a. Now it should be remembered that YAv. *tqθra-* must be derived from a PIr and PIr predecessor containing a cluster **sr* identical to OIA *tāmisrā-* and inherited from PIE: < PIr **taNsra-* < PIr **tamH(ī)sra-* < PIE **témHs-ro-* (where *sr* may have been tautosyllabic, cf. Lipp 2009: II 405). This means we must assume a secondary development **sr* > *θr* in Avestan.

b. Such a development is clearly attested for **sr-* in initial position, cf.

- 1) *θraotah-* 'stream' in *θraotō.st°*, cf. OIA *srótas-* 'id.'
- 2) *θraxti-* 'corner, side' (varia lectio *sr°* influenced by Pahlavi) = OIA *srakti-* 'id.'
- 3) perhaps also in *θrañhi-* 'corner of the mouth' from **sra(n)s-*, cf. OIA *srams-* 'fall down'?

² Cf. Hübschmann (1895: 207); Bailey (1979: 126); Morgenstierne (1974: 81a; 2003: 82–83). Thanks to Carina Jahani for bringing the Bal. word to my attention; however, it is not clear to me how the long vowel in this form can be explained if it continues **tāhra-*.

³ Likewise **pāθra-* > *pās* 'watch'. This is apparently the development of Southwest Iranian proper, cf. **puθra-* > OP *puça-* 'son' > MP, NP *pus*, while **θr* > *hr* reflects Northwestern dialects (cf. Hübschmann 1895: 204; Klingenschmitt 2000: 202).

⁴ See Cheung (2002: 229–230); Kümmel (2007: 153 n. 93).

c. After a vowel, however, **sr* obviously yielded *hr* > Av. **ḡhr* > *ḡr*, cf.

1) YAv. *hazaḡra-* ‘1000’ < pre-Av. **hazahra-* < Plr **sajasra-* = OIA *sahásra-*

2) Av. *aḡra-* ‘evil’ < **ahra-* < **asra-* = OIA *asrá-*

3) Av. *daḡra-* ‘skilful’ < **dahra-* < **dasra-* = OIA *dasrá-*

d. This development **sr* > (**h*)*r*- has also been claimed as a variant for initial position (e.g., by Bartholomae 1894: 36), but the evidence is far from certain.

OAv. *rāma-* ~ *rāma-* ‘sprain?/brutality?’ has been compared to OIA *srāmá-* ‘lame’ (Mayrhofer 1996: 785), but as the meaning is unclear, other etymologies are possible. Younger Avestan cases are even less certain.⁵ Even the best one is dubious. The genitive plural *raonqm* was traditionally taken to mean ‘river beds’ and compared to *θraotah-*, but a meaning ‘lowland, plains’ fits the context better, as the word contrasts with *gairinqm* ‘mountains’, and this meaning is also attested for other Iranian cognates, cf. Khwar. *r’wyn* ‘land, place’, MP *rōn* ‘region, direction’ (see Schwartz 1971: 301), so it would better be derived from a root beginning with simple **r*, cf. also Av. *rauuah-* ‘free space’.

Thus there seems to be no counterevidence to the following Avestan rules: while Plr **sr* it became *hr* between vowels, it yielded *θr* in word-initial position and after **n*, which may be lumped together as syllable-initial position (cf. Lipp 2009: II 405 fn. 112).

e. Elsewhere in Iranian, however, the evidence points to a development **sr* > *hr* in all contexts rather than **sr* > **θr*, and thus we can easily assume that **tansra-* yielded *tanhra-* > **tāhra-* > **tahra-* which then developed quite parallel to **hazahra-* ‘thousand’ > MP, Parthian *hz’r*, Sogd. *z’r*, Khot. *ysāru* etc. For initial position, compare the following cases:

1) **srautah-* > **hrautah-* > OP *rauta* ‘river’ > MP, Parth. *rwd*, Sogd. *rwt*, NP *rōd*.

2) **sraxwa-* > **hraxwa-* ‘corner of mouth?’ (= OIA *srákva-*) > Armenian LW *erax* ‘mouth’, Persian *rox*, Pashto *rāx/-rxo* ‘cheek’, Bal. *rakk* ‘lip’ (cf. *pakkag* ‘ripe’ < **paxwaka-*⁶).

3) **sraba-* > **hraba-* > Pashto *raw-*, Shughni *rāv-* ‘to suck’ ... < IE **srebʰ-* ‘to sip, to slurp’ (see Cheung 2007: 140).

f. The special development in Avestan found here is an archaism insofar as *sr* must have been retained until it could change to *θr*, but this very change must be considered an early innovation of Avestan in contrast to the other languages, where *sr* must have been preserved until it yielded *hr* in all positions.⁷ All these changes must have preceded the later dialectal development of **cr* > *sr* / *θr*.

⁵ On *urūdi* see Kellens (1974: 80–84); *raḡha-* is semantically unclear (the meaning ‘epileptic’ is merely a guess based on etymological connection with IA *srams-* ‘to fall down’).

⁶ I wish to thank Agnes Korn for discussing these Balochi words with me. As discussed by Korn (2005: 112), the possibility cannot be excluded that *pakkag* derives from *paxtaka-*, cf. *sak(k)* ‘hard, very’, if from < **saxta-* (cf. NP *saxt* ‘difficult’) and not from Urdu *sak*. Less sure possible cognates are Khot. *rahā* ‘notch’ Bailey (1979: 360b) and Wakhi *ruk* ‘forehead’ (→ Sargoli *rak*, cf. Morgenstierne 1974: 67a).

⁷ A problematic word in this context is **strī-* ‘woman’, if originally < **srī-/sryā-*, as proposed by Kim (2005: 132–133). But nothing speaks against deriving attested Iranian *sr-* from **str-*, so **str-* must have been Proto-Indo-Iranian, however this can be explained.

4. PIIr **nsn* and **sn*

This one case of *ns* preceding a sonorant can be strengthened by parallel phenomena found with the cluster **sn*, where we can observe a difference between word-initial and intervocalic position, too.

a. YAv. *pqsnu-* ‘ashes, dust’ clearly corresponds to Khot. *phāna-* ‘dust, med’, Oss. Digor *funuk*, Iron *fənyk* ‘ashes’.⁸ As shown by OIA *pāmsú-* ‘dust’, Av. *s* must be old here.⁹

b. This means that in **pānsnu-*, **sn* was preserved in Avestan, like initial *sn-*.¹⁰ In the other languages there apparently was a development **sn* > **hn* parallel to **sr* > *hr* above. Thus, Khot. **pāhnu-* > **pāhn-* > *phān-*, and Oss. **panhnu-* > **fanu-* + *-k*.

c. In this case, preservation of initial **sn-* is the most widespread development in Iranian and is found in most languages (including some where further changes applied, like *sn-* > *st-* in Wakhi; *sn-* > *zn-* in Khotanese, Pamiri, and Yidgha-Munji). In some languages, however, loss of the sibilant seems to point to **hn-*, namely in Ossetic, partly in Khwarezmian, in Pashto, and in Balochi (see Hübschmann 1895: 253; Korn 2005: 128). Cf. the following words:

- 1) **snāwar/snāwan-* ‘sinew’ (Skt. *snāvan-*) > YAv. *snāuuarə*, Sogd. *sn* ‘w’ > **hnāwar-* > Oss. Digor *nawær* (cf. Cheung 2002: 209)
- 2) **snuš-* ‘daughter-in-law’ > Sogd. *šwnšh* (with assimilation), Wakhi *stāx*, Shughni *zināy*
**snaušā-* > Bactrian ασνωο
> **hnuš-* > Khwar. ‘*nh* and (with analogical extension¹¹) Pashto *nzor*, Bal. *nišār*
> **hnauš-tā-* > Oss. Digor *nostaē* (cf. Morgenstierne 1974: 108b; Cheung 2002: 208)
- 3) **snauda-* ‘cloud’ > YAv. *snaoḍa-*, MP pahl. *snwd*
> **hnauda-* > Bal. *nōd*
- 4) **snā-* ‘to bathe’ > YAv. *snā°*, Khwar. *sn* ‘d’, Parth. *sn* ‘c-’, MP man. ‘*sn* ‘y-’,¹² Khot. *ysānā°* etc.
> **hnā-* > Oss. *naj-/nad-* (cf. Bailey 1979: 351; Cheung 2002: 30)
- 5) **snaij-*, **snaig-* ‘(to) snow’ > YAv. *snaēža-*, Sogd. *šnyš-* (with assimilation), MP Pahl. *snyck*, Pam. Yazghulami *zənay* (cf. Morgenstierne 1974: 108b, 110b)

⁸ Cf. Bailey 1979: 261; Abaev 1958: 449; Cheung 2002: 125, 186; possibly also in Parachi *yā-phunē* ‘dusty wind’. Sogdian *spn* ‘k/spnyy’ ‘dirt, filth’ differs semantically and need not be related.

⁹ Contra Turner (1966: No. 8019) there is no good evidence for an original palatal **ś*: The late Sanskrit variant *pāmsū-* has no value; affricate reflexes in Nuristani may as well result from **ns* > **nts*; and Romani *poši(k)* ‘sand, dust’ must be a loanword from Armenian *p’oši* rather than inherited (cf. Wolf 1960: No. 2541).

¹⁰ Unfortunately, there seems to be no evidence for **sn* after a vowel.

¹¹ < **(h)nušā-r* with *-r* taken from other kinship terms, as in NP *sunār* which seems to result from dissimilation of a **sun(u)šār* < **snušār*; cf. Korn 2005: 128.

¹² In MP we also find Pahl. *šn* ‘c-’, Man. *šn* ‘z-’ ‘to wash’, cf. NP *šenā-* (Horn 1893: Nr. 792), which may be influenced by ruki variants in compounds (thus Hübschmann 1895: 81). Alternatively, one might suspect that *šn-* was the genuine Persian outcome of **sn-*, and all the words with *sn-* are northwestern loans.

d. This seems to imply the following dialectal distribution concerning PIr **sn*

- 1) In Avestan and probably some other languages, **sn* developed to **hn* only after a vowel, but **sn* was preserved (or perhaps first affricated **s > ts?*) when it was word-initial or followed another *n*, where *sn* might have been in syllable-initial position (although *ns.n* is thinkable, too).
- 2) In Khotanese and possibly other languages, only word-initial **sn* was preserved (as **zn?*), but internal yielded **hn* also after **n*.
- 3) In some languages spoken in the periphery of Iranian (Ossetic, partially Khwarezmian, Pashto, and Balochi) **sn* yielded **hn* everywhere.

Again the relative chronology implies that **sn > hn* was earlier than **cn > sn/šn*. The development was largely parallel to that of **sr*.

5. **ns* elsewhere

Let us now take a brief look at what happened to **ns* in other environments, i.e., when not followed by a sonorant.

a. **s* was preserved before voiceless stops (as generally in this position). The nasal was then lost before the sibilant with nasalization of the preceding vowel, cf. **ansta-* > Av. *qsta-* ‘envy’, **ansti-* > Khwar. ‘*sc*’ (Benzing 1983: 83).

b. Elsewhere (i.e., before vowels and in final position) Clr **Vns* yielded **Vnh*, as **Vs* yielded **Vh* in these positions. In Ossetic, *n* seems to have been preserved before *h* as well as before other fricatives, thus we get **anha-* > (*i-*)*on(æ)* ‘shoulder’ (cf. Cheung 2002: 211–212), as in **panhnu-* > **fan(n)u-* ‘ashes’. Elsewhere nasalization occurred in this context, too (possibly preceding **s > h*): *Vn > Vā/Vĩ*: As the nasal vowel was long in any case, original quantity distinctions of the vowels disappeared. In Avestan, the peculiar development of *h > ŋh* followed and led to a reoralization of the nasal following the raised nasal vowel. The original quantity of this vowel itself could not be restored, however: Both **āNh* and **ǎNh* must have merged in Pre-Avestan **āh* [ǎ:h] which then developed to **ǎŋh* > OAv. *ǎŋgh* and YAv. *aŋh* preceding *ǎ* and *qh* preceding *ĩ*.

c. This implies that word-final **-āh > *-ǎŋ(h) > OAv. -ǎŋg*, YAv. *-ǎ/-q* may also continue PIr < PIIr **-āns*, and the Av. *a*-stem accusative plural ending might thus be identical to Indo-Aryan *-āms#* rather than continuing a different ending **-ǎns* with a short vowel, as is normally thought (cf. Hoffmann & Forssman 1996: 120). Older **-ānh* is also presupposed by the Av. accusative plural *maqθrā* < **manθrāh*, apparently showing early dissimilation of **manθrānh > *manθrāh* (see Hoffmann 1955: 35 = 1976: 378 n. 1), while an older **manθranh* would have yielded †*maqθrā* by the same kind of dissimilation.

It has to be conceded that no other secure examples of Clr **āNh* are known. The only possible case would be *māŋhəm* in Yašt 8,1 if it originally meant ‘food’ < ‘flesh, meat’ (Gershevitch 1976) and were to be compared to OIA *māmsām* ‘flesh, meat’. But the exact interpretation of this passage is difficult (cf. Panaino 1990:

87–88), and the word might also have been influenced by **māh* = OIA *mās* ‘meat’ probably attested by OAv. **mā* (cf. Tichy 1997: 98–99) and/or by much more frequent *māṇhəm* ‘moon’ < **māham* appearing in the near context. Therefore, *māṇhəm* cannot be considered to be valid evidence for preservation of **ānh* (> *āṇh*) in Avestan rather than a development to **əṇh*.

6. Conclusions

a. We have seen that **s* yielded *h* / *._n,r* (= before tautosyllabic¹³ coronal sonorants) later than elsewhere and this development was preceded by some dialectal developments precluding that change in some Iranian languages (especially in Avestan).

It follows that the “aspiration” of **s* > *h* thus seems to have had different stages depending on the phonological environment. This strengthens the case for a CIr rather than PIr change, or to be more precise, for a change that may have started in PIr times but was not fully carried through before the separation of some branches (cf. Szemerényi 1966: 192–193; Hintze 1998; Schmitt 2000: 19; Tremblay 2005: 682–683).

b. The preservation of tautosyllabic **sr* followed by an early specific change *sr* > *θr* / *\$_sets* (Younger?) Avestan apart from all or at least most other Iranian languages. If we could find this change in another Iranian language, it might be highly relevant for early dialect grouping.

Abbreviations

Av.	Avestan
Bal.	Balochi
Khot.	Khotanese
Khwar.	Khwarezmian
LW	loanword
Man.	Manichaean
MP	Middle Persian
N	unspecified nasal
NP	New (= Modern) Persian
OAv.	Old Avestan
OIA	Old Indo-Aryan (Vedic)
OP	Old Persian
Oss.	Ossetic
Pahl.	Pahlavi
Pam.	Pamiri
PIE	Proto-Indo-European
PIr	Proto-Iranian
PIIr	Proto-Indo-Iranian
Sogd.	Sogdian
V	unspecified vowel
YAv.	Younger Avestan

¹³ As noted above, the most plausible way to combine word-initial and postnasal position into one conditioning is to assume that the cluster was tautosyllabic after sonorants.

References

- Abaev, V. I. (1958, 1973, 1989). *Istoriko-Etimologičeskij Slovar' Osetinskogo Jazyka*. 4 Vols. Leningrad.
- Bailey, H. Walter (1979). *Dictionary of Khotan Saka*. Cambridge: Cambridge University Press.
- Bartholomae, Christian (1894). "I. Vorgeschichte der Iranischen Sprachen." In: Geiger, Wilhelm an Kuhn, Ernst (eds): *Grundriss der iranischen Philologie*, 1:1. Strassburg: Trübner, pp. 1–151.
- Benzing, Johannes (1983). *Chwaresmischer Wortindex*. Mit einer Einleitung von Helmut Humbach, herausgegeben von Zahra Taraf. Wiesbaden: Harrassowitz.
- Cantera, Alberto (1998). "Phl. *mānsarspan/māraspand* und die Entwicklung der Gruppe *-θr-* im Pahlavi." *Indo-Iranian Journal* 41, pp. 351–369.
- Cheung, Johnny (2002). *Studies in the Historical Development of the Ossetic Vocalism*. Wiesbaden: Reichert.
- Cheung, Johnny (2007). *Etymological Dictionary of the Iranian Verb*. Leiden and Boston: Brill.
- Gershevitch, Ilya (1954). *A Grammar of Manichaean Sogdian*. Oxford: Blackwell.
- Gershevitch, Ilya (1976). "More Meat in Iranian." In: Morpurgo Davies, Anna, and Meid, Wolfgang (eds), *Studies in Greek, Italic, and Indo-European Linguistics offered to Leonard R. Palmer on the occasion of his seventieth birthday*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck, pp. 63–64.
- Hintze, Almut (1998). "The Migrations of the Indo-Iranians and the Iranian Sound-Change *s > h*." In: Meid, Wolfgang (ed.), *Sprache und Kultur der Indogermanen. Akten der X. Fachtagung der Indogermanischen Gesellschaft, Innsbruck, 22.–28. September 1996*. Innsbruck: Institut für Sprachwissenschaft der Universität, pp. 139–153.
- Hoffmann, Karl (1955). "Ein grundsprachliches Possessivsuffix." *Münchener Studien zur Sprachwissenschaft* 6, pp. 35–40. [= 1976, 378–383].
- Hoffmann, Karl (1976). *Aufsätze zur Indoiranistik*, edited by Johanna Narten. Vol. 2. Wiesbaden: Reichert.
- Hoffmann, Karl, and Forssman, Bernhard (1996). *Avestische Laut- und Flexionslehre*. Innsbruck: Institut für Sprachwissenschaft der Universität.
- Horn, Paul (1893). *Grundriß der neupersischen Etymologie*. Strassburg: Trübner.
- Horn, Paul (1901). "Neupersische Schriftsprache." In: Geiger, Wilhelm and Kuhn, Ernst (eds), *Grundriss der iranischen Philologie*, 1:2. Strassburg: Trübner, pp. 1–200.
- Hübschmann, Heinrich (1895). *Neupersische Studien*. Strassburg: Trübner.
- Hübschmann, Heinrich (1899). Review of Horn 1898. *Indogermanische Forschungen Anzeiger*, pp. 18–41.
- Kellens, Jean (1974). *Les noms-racines de l'Avesta*. Wiesbaden: Reichert.
- Kim, Ronald I. (2005). "Ossetic *silə/syl* and the Indo-Iranian Word for 'Female'." *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 2, pp. 123–168.
- Klingenschmitt, Gert (2000). "Mittelpersisch." In: Forssman, Bernhard and Plath, Robert (eds), *Indo-arisch, Iranisch und die Indogermanistik. Arbeitstagung der Indogermanischen Gesellschaft vom 2. bis 5. Oktober 1997 in Erlangen*. Wiesbaden: Reichert, pp. 191–229.
- Korn, Agnes (2005). *Towards a Historical Grammar of Balochi. Studies in Balochi Historical Phonology and Vocabulary*. Wiesbaden: Reichert.
- Kümmel, Martin Joachim (2007). *Konsonantenwandel. Bausteine zu einer Typologie des Lautwandels und ihre Konsequenzen für die vergleichende Rekonstruktion*. Wiesbaden: Reichert.
- Lipp, Reiner (2009). *Die indogermanischen und einzelsprachlichen Palatale im Indoiranischen*. Vol. I: Neurekonstruktion, Nuristan-Sprachen, Genese der indoarischen Retroflexe, Indoarisch von Mitanni. Vol. II: Thorn-Problem, indoiranische Laryngalkokalisation. Heidelberg: Winter.
- Mayrhofer, Manfred (1996). *Etymologisches Wörterbuch des Altindoarischen*. Vol. 2. Heidelberg: Winter.
- Morgenstierne, Georg (1974). *Etymological Vocabulary of the Shughni Group*. Wiesbaden: Reichert.
- Morgenstierne, Georg (2003). *A New Etymological Vocabulary of Pashto*, compiled and edited by Josef Elfenbein, David Neil MacKenzie, and Nicholas Sims-Williams. Wiesbaden: Reichert.
- Panaino, Antonio (1990). *Tištīrya. Part I, The Avestan Hymn to Sirius*. Roma: Istituto Italiano per il Medio ed Estremo Oriente.

- Rastorgueva, Vera S., and Ėdel'man, Džoj Iosifovna (2003). *Ėtimologičeskij slovar' iranskix jazykov*. Tom 2, b-d. Moscow: Vostočnaja Literatura.
- Schmitt, Rüdiger (2000). *Die iranischen Sprachen in Geschichte und Gegenwart*. Wiesbaden: Reichert.
- Schwartz, Martin (1971). "On the Vocabulary of the Khwarezmian Muqaddimatu l-Adab, as Edited by J. Benzing." *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 120, pp. 288–304.
- Szemerényi, Oswald (1966). "Iranica II." *Die Sprache* 12, pp. 190–226.
- Tichy, Eva (1997). "Fleisch und Haoma. Zur Interpretation von Yasna 48,10." *Die Sprache* 39, pp. 89–104.
- Tremblay, Xavier (2005). "Bildeten die iranischen Sprachen ursprünglich eine genetische Familie oder einen Sprachbund innerhalb des indo-iranischen Zweiges? Beiträge zur vergleichenden Grammatik der iranischen Sprachen V." In: Meiser, Gerhard, and Hackstein, Olav (eds), *Sprachkontakt und Sprachwandel, Akten der XI. Fachtagung der Indogermanischen Gesellschaft, Halle an der Saale, 17.–23. September 2000*. Wiesbaden: Reichert, pp. 673–688.
- Turner, Ralph Lilly (1966). *A Comparative Dictionary of the Indo-Aryan Languages*. London: Oxford University Press.
- Wolf, Siegmund A. (1960). *Großes Wörterbuch der Zigeunersprache (romani tšiw): Wortschatz deutscher und anderer europäischer Zigeunerndialekte*. Mannheim: Bibliographisches Institut.

Introducing Reported Speeches in Balochi of Sistan with *ki*

Stephen H. Levinsohn, SIL International
stephen_levinsohn@sil.org

Abstract

The Balochi conjunction of general subordination, *ki*, sometimes introduces complement clauses that report speeches. Barjasteh Delforooz (2010a: 224) suggested that, when so used, *ki* “has a highlighting function. ... The marked speeches push the story forward to its goal.” The current paper argues that, in fact, *ki* highlights not the speech itself, but its consequences. When *ki* precedes a reported question, for example, the answer is more important than the question. The presence of *ki* may also indicate that the words in a speech were not said on a particular occasion, but represent the substance of what someone else said or used to say, is to say or could have said. In other words, *ki* is a “linguistic indicator of interpretive use” (Blass 1990: 104; Farrell 2005: 1).¹ The paper ends by discussing how the above motivations for using *ki* before a reported speech in Balochi can be reconciled.²

Keywords: Balochi, interpretive use marker, reported speech

1. Introduction³

Ki in Balochi is a “conjunction of general subordination” (Barjasteh Delforooz 2010a: 16) found in adverbial, relative and complement clauses (sec. 2). This paper concentrates on its presence before complement clauses and, in particular, those that report a speech.

Roberts (2009: 295, 300) claims for Persian that the “clause linkage marker” *ke* “is used primarily in spoken texts to give prominence to speeches that the author considers are important to the story”. In the pre-defence version of his dissertation, Barjasteh Delforooz (*ibid.*: 224) likewise suggested that, when introducing a direct speech, Balochi *ki* “has a highlighting function. ... The marked speeches push the story forward to its goal.”

¹ Sadly, when I presented this paper, I was unaware that Farrell had already asserted that “*ki* may be broadly analysed in all its uses as introducing a representation of another representation – a thought, utterance or state of affairs that could possibly be entertained” (2005: 17). The uses Farrell considered include *ki* as a relative and subordination marker; in compounds; as a comparative; introducing purposes, results and reasons; as a temporal marker; in both positive and negative disjunctions; and in connection with indirect and direct quotations (*ibid.*: 3–15).

² The conclusions of this paper are based on the texts found in Appendix 2 of Barjasteh Delforooz 2010b (pp. 287–392). To ensure that the paper covers all the major uses of *ki*, every instance in texts XM (*Xarmizza* ‘Melon’ – pp. 288–295) and MG (*Hazrat-i Mūsā u gušnagēn bandag* ‘Moses and the Starving Man’ – pp. 296–304) has been cited. Reference is also made to texts BP (*Baxtay padā* ‘Seeking the Fortune’ – pp. 305–320), BU (*Pirēn balōč u uštir* ‘The Old Baloch and the camel’ – pp. 359–367), KH (*Khudānizar Khān* ‘Khudanizar Khan’ – pp. 328–337) and PJ (*Pir jāngī* 3 ‘Pir Jangi’ – pp. 321–327).

³ A shorter version of this paper was presented at the 4th International Conference on Iranian Linguistics (ICIL 4), Uppsala University, Sweden in June 2011. Barjasteh Delforooz classifies Balochi of Sistan “as belonging to the Sarhaddī subdialect of the Raxšānī or Western group of Balochi dialects”. It “belongs to the so-called north-western group of Iranian languages” and is mainly spoken in Iranian and Afghani Sistan (2010a: 18–20).

However, further research suggested that, when *ki* introduces a reported speech in Balochi, it is the consequences of the speech that are highlighted, rather than the speech itself. In particular, “when a question is answered, the answer is usually more important than the question, which is why the effect of marking a question with *ki* is to highlight the answer” (Barjasteh Delforooz 2010b: 227 – see sec. 4 for examples). Furthermore, *ki* may introduce a speech, not to highlight anything, but to indicate that its words represent the substance of what one or more people said, are to say or could have said, as opposed to having been uttered on a particular occasion (*ibid.*: 229).

Like Farrell (2005), this paper argues that the reason that *ki* can have such diverse effects is that it is an “indicator of interpretive use” (Blass 1990: 104). This means that, when it introduces a reported speech, the speech is to be understood not as a description of what was said on a particular occasion, but rather as a representation of an utterance or thought (see sec. 3).

2. *Ki* as a conjunction of general subordination

Subordinate clauses can be divided into three basic types, adverbial, relative, and complement (see Whaley 1997: 247), and *ki* is found in all three in Balochi. This section briefly illustrates its typical use in the three types.

2.1. *Adverbial clauses subordinated by ki*

In adverbial clauses, *ki* most often occurs at the beginning of the clause, unless it is of time, in which case it immediately precedes the subordinated verb or verb phrase, following the subject, if present.

When the clause is post-nuclear, it is usually of reason or result (if realis) or purpose (if irrealis) (Barjasteh Delforooz 2010b: 200), though other relations are also found. In (1), the subordinated clause (1b) gives a reason for the event of the main clause (1a).

- | | | | | | |
|------|-----------------------------|---------------------------------|---------------------------|--------------------------------|-----------------------------|
| (1)a | <i>nām=ay</i>
name=PC.3s | <i>guṛā</i>
then | <i>galaw-ā</i>
melon-O | <i>išt-ant</i>
leave.PST-3p | <i>xarmizza</i>
xarmizza |
| b | <i>ki</i>
SUB | <i>mizzag=ay</i>
taste=PC.3s | <i>awal</i>
first | <i>xar</i>
donkey | <i>burt</i>
take.PST.3s |
- Then they named the melon ‘xarmizza’, since it was a donkey that tasted it first. (XM 109–110)⁴

When the clause is pre-nuclear, it is usually of time, though other relations are also found. In (2), the subordinated clause (2a) gives the time for the event of the main clause (2b–c).

- | | | | | |
|------|----------------------------------|-------------------------|------------------------|---------------------------|
| (2)a | <i>bādišā</i>
king | <i>ki</i>
SUB | <i>ēši-rā</i>
DEM-O | <i>dīst</i>
see.PST.3s |
| b | <i>zānt</i>
understand.PST.3s | | | |

⁴ See also XM 74–75 (reason), MG 61–62 (result) and KH 111 ((8b) below) (purpose).

- c *ki* *diga* *kišwar-ay* *mardum=ē*
 SUB other country-GEN person=IND
 When the king saw him, he understood that (he was) a person from another country. (BP 27–29)⁵

What is noteworthy for the present paper is that, whether pre-nuclear or post-nuclear, the information in adverbial clauses that are subordinated with *ki* can easily be related to information that has recently been stated in the discourse. Thus, in (1), the hearers already know that it was a donkey that first ate the melons (XM 83–87). Likewise, in (2), the hearers already know that the traveller has arrived in a country whose king is out hunting (BP 24–26).

2.2. Relative clauses subordinated by *ki*

Relative clauses are commonly divided into two types: restrictive and non-restrictive. Restrictive (identifying) relative clauses serve to delimit the potential referents (Comrie 1989: 138). In such clauses in Balochi, *ki* occurs immediately after the head NP that it modifies. In (3), the head NP is *har yazāu har mēwagē* ‘every kind of food and fruit’, and the relative clause *ki dīst* limits the referent to that ‘which he saw’.

- (3) *am=ē* *har* *yazā=u* *har* *mēwag=ē*
 EMP=DEM each food=and each fruit=IND
 ki *dīst* *zīt=u* *wārt*
 SUB see.PST.3s buy.PST.3s=and eat.PST.3s
 He bought and ate every kind of food and fruit which he saw. (MG 58–60)⁶

Non-restrictive relative clauses serve “merely to give the hearer an added piece of information about an already identified entity, but not to identify that entity” (*ibid.*). Such clauses in Balochi begin with *ki*, rather than placing it after the head NP. In (4), for instance, the head NP is *yakk tīri baryē* ‘a light post’ (4a), and the relative clause of (4b) gives additional information about it.

- (4)a *ē* *bādišā* *bi=m=ē* *wat-ī* *šār-ay*
 DEM king in=EMP=DEM RFL-GEN town-GEN
 wasat-(t)ā *yakk* *tīr=i* *bary=ē* *dāšt*
 middle-OBL one pole=IZ electricity=IND have.PST.3s
 b *ki* *harčī* *am=ē* *tīlīpunān-ī* *sīm=at-ant*
 SUB whatever EMP=DEM telephones-GEN wire=COP.PST-3p
 bi *am=ēšī* *wasl=at-ant*
 to EMP=DEM.OBL connected=COP.PST-3p

This king had a light post in the centre of his town, to which were connected whatever phone wires there were. (XM 3–5)

⁵ See also XM 7, 31, 36, 50, 55, 92, 106; MG 9, 45 (reason). See BP 106–107 for a complement following *zānt* ‘understand, know’ that is not introduced with *ki*.

⁶ See also XM 6, MG 90.

(5b) is different, as *ki* introduces the presupposition (established information – see (5a)) in the equivalent of an “it-cleft structure” (Levinsohn 2011a: 65).

- (5)a ‘Finally, he became the (biggest) trader of the world, the trader of the entire world.’
- (5)b
- | | | |
|--|------------------|---|
| FOCUS
<i>marg na(y)-āt</i>
death NEG-come.PST.3s
(It was through) death not coming (contrary to expectation) ⁷ that he became a trader. (MG 94–97) | <i>ki</i>
SUB | presupposition
<i>tajjār būt</i>
trader become.PST.3s |
|--|------------------|---|

2.3. Complement clauses subordinated by *ki*

In complement clauses, *ki* typically introduces the complement, as in (2c) above and (6).

- (6)
- | | | | | | | |
|---------------|--------------------------|----------|-------------|-------------|------------|----------|
| <i>b(y)-ā</i> | <i>ki</i> | <i>ē</i> | <i>rang</i> | <i>ēš-ī</i> | <i>mās</i> | <i>ē</i> |
| SBJ-come.PRS | SUB | DEM | manner | DEM-GEN | mother | DEM |
| <i>rang</i> | <i>guṭṭō=at=ō</i> | | | | | |
| manner | strangled=COP.PST.3s=and | | | | | |
- It happened that its mother was strangled in this way and... (XM 59)⁸

When the complement is a reported speech, however, the default is for *ki* to be absent (see (7) below), so its presence before a reported speech is noteworthy. The following sections discuss the pragmatic effects of introducing a reported speech with *ki*.

3. *Ki* as an indicator of interpretive use

In Levinsohn (2011a: 115), it is argued that a number of languages exhibit what may be termed a marker of interpretive use:

When a speech is reported directly, it usually purports to **describe** a state of affairs – what was said on a particular occasion (a “descriptive use” – Sperber and Wilson 1986: 224–31). However, some reported speeches do not inform the reader of what was said so much as **represent** an utterance or thought that resembles it.

Some languages have “an explicit linguistic indicator of interpretive use” (Blass 1990: 104) whose function is to indicate that the speech concerned is not describing what was said on a particular occasion, but rather represents an utterance or thought [as illustrated below]. Such markers are often found in speech orienters. A variety of circumstances motivate the presence of interpretive use markers.

I now argue that Balochi *ki* is such an interpretive use marker and discuss some factors that motivate its presence.

When *ki* does not introduce a reported speech, then the sentence concerned pur-

⁷ This *ki* may be a “proclitic” (Hosseini, this volume) that indicates contrastive emphasis.

⁸ See also MG 107, 114. In MG 72 (*man šapī ki mir-īn* [I tonight SUB die.PRS-1s] ‘(It is sure) that I will die tonight’), *ki* appears to introduce a complement (*mir-īn*), even though no orienter is present.

ports to describe what was said on a particular occasion. So, in (7), by not using *ki*, the storyteller implies that, on a particular occasion, someone said to the king, “Lord king, it is a dragon”.

- (7) *gušt* *bādšā* *sāib* *aždiyā=(y)ē*
 say.PST.3s king master dragon=IND
 He said, “Lord king, it is a dragon”. (XM 15–16)

One of the reasons for introducing a reported speech with *ki* is to indicate that the words that follow were NOT said on a particular occasion.⁹ In (8c), for instance, the words that follow *ki* are not a report of what Sabzo said on some occasion. Rather, they are hypothetical, representing what she might have said, had Khudanizar Khan not given her a generous dowry.¹⁰

- (8)a *jam=ē* *ku*
 collection=PC.3s do.PST.3s
 b *ki*¹¹ *yānē* *sabzō* *ma-guš-īt*
 SUB means Sabzo PROH-say.PRS-3s
wat-ī *dil-ay* *tā* *pikr* *ma-kan-t*
 RFL-GEN heart-GEN in thought PROH-do-3s
 c *ki* *xudānizar* *xān* *mnā* *bi xwār-ēn*
 SUB Khudanizar Khan me to lowly-ATTR
mard=ē *dāt*
 man=IND give.PST.3s
 He collected them (almost 200 sheep and 100 camels) in order for Sabzo not to say, not to think in her heart, “Khudanizar Khan married me off to a lowly man”. (KH 110–113)

Similarly, what follows *ki* in (9d, f) is not a report of what Pirakk said on a specific occasion. Rather, it gives the substance of what he said a number of times, as the imperfectives (9c, e) imply.

- (9)a *būt=u* *mardum=ē*
 become.PST.3s=and person=IND
 b *ki* *b(y)-āt-ēn* *bi pīrakk-ay* *gis-ā*
 SUB SBJ-come.PST-SBJ.3s to Pirakk-GEN house-OBL
 c *bass* *ša* *xudānizar* *kissa=a* *kurt*
 just from Khudanizar story=IMPF do.PST.3s
 d *ki* *xudānizar* *pa(m)=man* *ē* *rang* *kurt=u*
 SUB Khudanizar for=me DEM manner do.PST.3s=and

⁹ Although this subset of examples of *ki* can be viewed as functioning as a “non-first hand” or “reported” evidential whose information has been “inferred from indirect evidence” (Aikhenvald 2004: 28, 1), the Relevance-Theoretic approach of this paper enables it to be described in terms of a single function (see Blass 1990: 259 on treating “hearsay” particles as “interpretive use markers”). As Farrell (2005: 17) observes, “Balochi *ki* is used in a wide variety of constructions. What is needed is an analysis that, if possible, will capture all of the varied uses under one function.”

¹⁰ *Ki* also tends to be used to introduce a message that the addressee is asked to pass on to someone else. This is probably because, at this stage, it is hypothetical. See MG 5–6.

¹¹ The *ki* in (8b) introduces a subordinate clause of purpose (see footnote 4).

e	<i>gurā</i>	<i>dēm-ā=a</i>	<i>gardēnt</i>			
	then	face-O=IMPF	turn.PST.3s			
f	<i>ki</i>	<i>sabzō</i>	<i>xudānizar</i>	<i>am=ē</i>	<i>rang</i>	<i>kurt</i>
	SUB	Sabzo	Khudanizar	EMP=DE M	manner	do.PST.3s
	<i>yā</i>	<i>na-kurt</i>				
	or	NEG-do.PST.3s				

It so happened that, whenever someone came to Pirakk's house, he used to talk so much about Khudanizar: "Khudanizar did this kind of thing for me"; and then he would turn (to his wife): "Sabzo, Khudanizar did this kind of thing, didn't he?" (KH 125–131)

When a well-known folktale is related, it is not usually thought of as something that was told on one particular occasion, so it is natural that *ki* should introduce it, as in (10b) (the present is used in (10a)):

- (10)a *guš-īt*
say.PRS-3s
- b ***ki*** *yag bādišā=(y)ē=at*
SUB one king=IND=COP.PST.3s
- They say that there was a king. (XM 1–2)

Ki may also introduce the substance of a whole conversation, as in (11b). It is most unlikely that three thieves would chorus together, "Let's go and steal from the king's treasury!" Rather, this speech represents what they decided after a discussion.

- (11)a *say duzz irāda kurt-at-ant*
three thief desire do.PST-COP.PST-3p
- b ***ki*** *b-raw-an bādišā-ay xazānag-ā b-ĵan-an*
SUB SBJ-go.PRS-1p king-GEN treasury-O SBJ-hit.PRS-1p
- Three thieves had taken a decision: "Let's go and steal from the king's treasury". (PJ 60–62)

4. *Ki* and grounding

I argued in section 3 that one of the reasons for introducing a reported speech with *ki* was to indicate that the words that follow were NOT said on a particular occasion. I now turn to passages in which speeches introduced with *ki* are backgrounded in relation to what follows.

I first return to extracts (10) and (11) above, as they provide evidence that the use of *ki* to introduce a reported speech does not highlight the speech concerned (contrast the claim by Roberts (2009: 300) for Persian that the comparable marker "*ke*" is used primarily in spoken texts to give prominence to speeches that the author considers are important to the story").

In (10b), the copula indicates that the information concerned is of a background nature, as far as the theme line of the narrative is concerned (Levinsohn 2011a: 68), even though it is introduced with *ki*.

In (11a), a pluperfect (translated ‘had taken a decision’) introduces the speech. Pluperfects are associated with backgrounding in narrative, as they are used for events that take place prior to the theme-line events (*ibid.*: 70). So the reported speech of (11b) is backgrounded with respect to what follows (the later decision of the thieves to pour a lap of gold into the very grave where the hero of the story is hiding – PJ 64–67).

Further evidence that *ki* does not highlight the speech that it introduces is that, when it precedes a reported question, the ANSWER (which is not introduced with *ki*) is more important than the question. Such is the case in (12).¹²

- (12) question *gušt=ī* *ki* *pīramard bābā*
 say.PST.3s=PC.3s SUB old.man father
ē *čē=(w)ant*
 DEM what=COP.PRS.3p
 ANSWER *gušt=ī* *ē* *bēxī* *ǰwān-ēn* *čīz=ant*
 say.PST.3s=PC.3s DEM entirely good-ATTR thing=COP.PRS.3p
man=um *wārt-a*
 I=also eat.PST-PP
ē *ar-ā=um* *dāt-a=un=ō*
 DEM donkey-O=also give.PST-PP=COP.PRS.1s=and
šumā=um *bōr-it*
 you=also SBJ.eat.PRS-2p

He said, “Dear old man, what are these?” He said, “These are very good things. I’ve eaten them and also given them to this donkey. You should eat them, too.” (XM 93–99)

A similar pattern is sometimes found when a reported “proposal” is followed by “its non-speech execution” (*ibid.*: 111). The effect of introducing the proposal with *ki* is to background it in relation to its non-speech EXECUTION. This is seen in (13).¹³

- (13) proposal *gušt=ī* *ki* *b-ra...*
 say.PST=PC.3s SUB SBJ-go.PRS
 (God) said: “Go (and tell that poor fellow...)”
 EXECUTION His Holiness Moses came and gave God’s message to the
 fellow... (MG 14-24)

¹² In the *Xarmizza* (XM) folktale, which concerns the discovery that melons are good to eat, three reported questions are introduced with *ki*, all of them asked by the king. The first leads to the discovery that a dragon has come to ask for help (12–16) and the second to the carpenter receiving a strange seed as a reward for helping the dragon (29–54). After someone sows the seeds, feeds the fruit to his donkey and, eventually, tries it himself (68–91), the third of the king’s questions (12) leads to him being persuaded to try the fruit, which results in it being called *Xar-mizza* (‘donkey-tasted’) (95–111 – see (1) above).

¹³ See also MG 29–38. In XM 12–13 (*bādšā dēm dāt yakk=ē=rā ki ē čī=(y)ē* [king face give.PST.3s one=IND=O SUB DEM what=IND] ‘The king sent someone [to check] who it is’), no speech verb is used to introduce the proposal, but the presence of *ki* again backgrounds this event in relation to its EXECUTION (14–16).

Cross-linguistically, indicators of interpretive use often introduce **indirect** speech (*ibid.*: 116), as it frequently communicates only the substance of the original communication (*ibid.*: 106), and indirect reporting is associated with backgrounding (Lowe and Hurlimann 2002: 75). Reported speeches in Balochi that are introduced with *ki* are not classified as indirect, since they retain first and second person pronominal references. Nevertheless, they tend to behave like indirect speeches as far as grounding is concerned. Given that one cross-linguistic way of highlighting an event is to background the one that immediately precedes it (Levinsohn 2011a: 79), I conclude that if the effect of using *ki* to introduce a reported question is to thereby highlight the answer, this is consistent with it being an indicator of interpretive use.

5. Further evidence that *ki* is an indicator of interpretive use

One reason for using an interpretive use marker is to signal that what follows relates back to and interprets something in the immediate context. So, for example, when a demonstrative is used cataphorically to point forward to and highlight what follows, it is cross-linguistically normal to introduce what follows with such a marker. Such is the case in (14): ‘that which I ate at once’ relates back to and interprets *amēš* ‘this’.

- (14) *mnī* *rōzī* *am=ēš=int* ***ki*** *man* *yakk-war-ā*
 my ration EMP=DEM=COP.PRS.3s SUB I one-time-ADVZ
 wārt-un
 eat.PST-1s
 My ration is this: that which I ate at one go. (MG 67–68)¹⁴

Similarly, in (15), the material following *ki* relates back to and interprets *allāay payyāmān* ‘God’s message’.

- (15)a *allā-ay* *payyāmān-ā* *pa* *bandag-ā* *dāt*
 God-GEN messages-O for servant-OBL give.PST.3s
 b ***ki*** *ay* *bandag* *tī* *rabb* *ē* *rang* *gušt*
 SUB oh servant your God DEM way say.PST.3s
 He gave God’s message to the fellow: “Oh fellow, your God said like this...” (MG 24–25)

See also KH 58–59, in which the material following *ki* (‘I swear by God...’) relates back to and interprets *kasam wā* ‘took an oath’.¹⁵

Extract (16) indicates the most significant points in the second half of a folktale about a camel that has become so exhausted from neglect and ill-treatment that it “laid down its neck to die” (BU 13–14). What is noteworthy is that *ki* introduces the

¹⁴ See also MG 77–78. In BU 55–56, the material following *ki* relates back to and interprets *anēō* ‘such’. PJ 3–8 (with *anēn sawt...* ‘such a voice’) is similar.

¹⁵ See also BP 21–22 (interpreting *sōj kan* ‘pose a question’). In (11) above, it could be argued that the material following *ki* relates back to and interprets *irāda kurt-at-ant* ‘had taken a decision’.

culminating speech by the camel (16b), which would be consistent with the claim by Roberts (2009: 300) for Persian that *ke* gives prominence to the speeches it introduces. However, it is clear from the content of the camel's speech that it is a direct response to its owner's request for forgiveness (16a). As such, the presence of *ki* can be understood as an overt indication that the reply relates to and interprets the owner's request (see further below).¹⁶

(16a) The Baloch nomad told the camel, "I have come to my senses now and now I know that [*ki*] I have been unjust to you. ... I want that [*ki*] you forgive me before you die, forgive me and do not take my negligence into consideration..." (BU 27–49)

b Then this camel, by the order of God the Almighty, turned its face and, in the very agony of death, began to speak and said to its owner *ki* "It doesn't matter. ... If you have loaded me up with heavy loads, I will also forgive you. ... I will forgive whatever you have done to me. But I will not forgive one thing, I will not forgive it until doomsday; that is this, that [*ki*] you didn't understand anything of my lawful and clean flesh... But one deed of yours I will not forgive, that [*ki*] ... you tied my rein to the tail of a crop-tailed donkey and made the donkey my leader and made the donkey my way-guide. This deed I will not forgive." (BU 50–105)

6. Conclusion

This paper has argued that, when introducing a reported speech or other complement, the subordinating conjunction *ki* functions as an "an explicit linguistic indicator of interpretive use" (Blass 1990: 104). In the majority of examples, its sole function is to indicate that the following speech is to be understood not as a description of what has been said on a particular occasion, but rather as a representation of an utterance or thought. In some passages, however, the pragmatic effect of introducing the speech with *ki* has been to background the question or proposal concerned, thereby highlighting the following answer or execution of the proposal. And, in one passage, *ki* has introduced the culminating speech of a folktale, to indicate explicitly that the speech relates to and interprets the previous speech.

This paper has not addressed the possibility that *ki* is a marker of interpretive use, not only when introducing a reported speech, but also in other constructions. Fortunately, Farrell has already addressed this issue. He concludes, "*ki* may be broadly analysed in all its uses as introducing a representation of another representation – a thought, utterance or state of affairs that could possibly be entertained" (2005: 17).

¹⁶ The following observation about the Koiné Greek interpretive use marker *hoti* applies equally to the way *ki* is used in (16b): "When *hoti* introduces direct speech in Luke-Acts, it not only indicates that the speech concerned "interprets" what has already been said ...; it also seems to mark that speech as the culmination of a narrative unit or sub-unit" (Levinsohn 2011b: 266).

Abbreviations¹⁷

1s/3s	1 st /3 rd person singular	IZ	<i>izafa</i> construction
1p/2p/3p	1 st /2 nd /3 rd person plural	NEG	negative
&	associative connective	NP	noun phrase
ADVZ	adverbializer	O	object
ATTR	attributive form of an adjective	OBL	oblique
COP	predicate copula	PC	pronominal clitic
DEM	demonstrative	PP	past participle
EMP	emphatic particle	PRS	present
GEN	genitive	PST	past
IMPF	imperfective	SBJ	subjunctive
IND	indefinite	SUB	conjunction of general subordination (<i>ki</i>)

References

- Aikhenvald, Alexandra Y. (2004). *Evidentiality*. Oxford: Oxford University Press.
- Blass, Regina (1990). *Relevance Relations in Discourse: A Study with Special Reference to Sissala*. Cambridge: Cambridge University Press.
- Comrie, Bernard (1989). *Language Universals and Linguistic Typology*, 2nd edition. Chicago: University of Chicago Press.
- Barjasteh Delforooz, Behrooz (2010a). *Discourse Features in Balochi of Sistan. Oral Narratives* [Studia Iranica Upsaliensia 15]. Uppsala: Acta Universitatis Upsaliensis.
- Barjasteh Delforooz, Behrooz (2010b). *Discourse Features in Balochi of Sistan. Oral Narratives* [Studia Iranica Upsaliensia 15]. Uppsala: Acta Universitatis Upsaliensis. Online at: <http://uu.diva-portal.org/smash/record.jsf?searchId=2&pid=diva2:345413>
- Farrell, Tim (2005). "That's that: *Ki* in Balochi as a Marker of Interpretive Use." Unpublished paper dated 10/8/05.
- Follingstad, Carl M. (2002). *Deictic Viewpoint in Biblical Hebrew Text: A Syntagmatic and Paradigmatic Analysis of the Particle Ki*. Dallas: SIL International.
- Hosseini, Ayat (2012). "Prosodization of Function Words in Persian." *Orientalia Suecana*, this volume.
- Levinsohn, Stephen H. (2011a). *Self-instruction Materials on Narrative Discourse Analysis*. Online at: <http://www.sil.org/~levinsohns>
- Levinsohn, Stephen H. (2011b). *Discourse Features of New Testament Greek: A Coursebook on the Information Structure of New Testament Greek*, 2nd edition, revised. Dallas: SIL International.
- Levinsohn, Stephen H. (2011c). *Some Notes on the Information Structure and Discourse Features of Exodus 1–14*. Online at: <http://www.sil.org/~levinsohns>
- Lowe, Ivan and Hurlimann, Ruth (2002). "Direct and Indirect speech in Cerma narrative." In: Güldemann, Tom, and von Roncador, Manfred (eds), *Reported Discourse: A Meeting Ground for Different Linguistic Domains*. Amsterdam and Philadelphia: John Benjamins, pp. 71–90.
- Roberts, John R. (2009). *A Study of Persian Discourse Structure* [Studia Iranica Upsaliensia 12]. Uppsala: Acta Universitatis Upsaliensis.
- Sperber, Dan and Wilson, Deirdre (1995). *Relevance: Communication and Cognition*, 2nd edition. Oxford: Blackwell.
- Whaley, Lindsay J. (1997). *Introduction to Typology: The Unity and Diversity of Language*. London: SAGE Publications.

¹⁷ The abbreviations are adapted from Barjasteh Delforooz (2010a: 15–16).

Variation in Persian Vowel Systems

Corey Miller, University of Maryland
cmiller6@umd.edu

Abstract

The vowel systems employed by different varieties of Persian across time and space exhibit a great deal of variation. This study attempts to describe that variation diachronically, in the spirit of Pisowicz (1985), and synchronically by examining the three major national varieties: Farsi, Dari and Tajik. We interpret the variation encountered through general principles of vowel shifting, as described by Labov (1994), from an Early New Persian baseline. We thus trace the historical development of the vowel systems in the major varieties of Persian spoken in Iran, Afghanistan and Tajikistan, clarifying the relationships among the varieties and the extent of their adherence to such principles.

Keywords: Persian, Farsi, Dari, Tajik, vowels, phonetics, phonology

1. Introduction: Early New Persian Vowel System¹

The history of Persian, a south-western Iranian language, is generally divided into three major periods: Old Persian, the language of the Achaemenids (558–330 BC), Middle Persian, the language of the Sassanids (224–651 AD), and New Persian, beginning in the seventh century AD, following the Arab Conquest (Windfuhr 2009: 445). The baseline vowel system from which we will be departing has been attributed both to Early New Persian (Windfuhr 2009: 457) and Classical Persian (Pisowicz 1985, Thiesen 1982). Both the dating and use of these terms is subject to divergence in the literature. Windfuhr and Perry (2009: 533) date Early New Persian (ENP) to 1100–1300 AD, while Windfuhr (2009: 447) assigns it to the 10th and 11th centuries. There is similar divergence in the dating of Classical Persian; Windfuhr provides several possibilities, such as between the 13th and 16th centuries (2009: 447), the 13th and 15th centuries (1979: 166), and along with Perry, between 1300 and 1600 (Windfuhr and Perry 2009: 533). However, according to Paul (2002: 21), Classical Persian “is not a well-defined linguistic term. It is a literary term only, but there is no agreement among Iranians or non-Iranians on what it actually denotes.” As we will see, the presence of divergences from this system in the periods associated with Classical Persian shows the necessity of ascribing this baseline to ENP. The vowel system of this variety, which we represent in **bold** in order to facilitate comparison with other varieties, is represented in Figure 1. The positions of the vowels are approximate and derived from Pisowicz (1985) and Horn (1901). In particular, Horn (1901: 19) presents evidence from Judeo-Persian that the short **a** had a fronted

¹ I would like to thank Erik Anonby, Youli Ioannesyan, Carina Jahani, and Geoffrey Haig for valuable comments and suggestions.

quality. In this and subsequent charts, arrows emanating from diphthongs are only intended to show the direction of the glides, rather than the location of their final targets. We transcribe diphthongs with offglides rather than vocalic targets because at this stage we have not investigated their phonetic details. Note that several researchers, including Jahani and Korn (2009: 648) in a discussion of similar phenomena in Balochi, prefer to label **aj** and **aw** as VC sequences due to their phonotactics. Okati (2012: 178–179) provides an excellent survey of the treatment of these sounds in various Iranian languages and dialects.

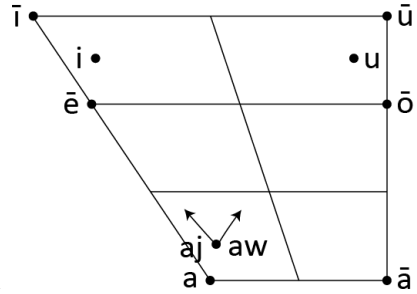


Figure 1. Vowel system of Early New Persian

The establishment of this baseline variety precedes the colonial Persian of India and the Ottoman Empire of the 15th–19th centuries (Windfuhr 1979: 166). Indeed, this vowel system is essentially reflected in the contemporary Persian of Afghanistan, also known as Dari (Farhadi 1955: 7), and the Persian of India in the Mogul period (16th–19th centuries), of which Phillott (1919) offers a comprehensive description; however, it differs markedly from contemporary Iranian Persian (CIP), or Farsi, as we will discuss below.

Since the Persian writing system, based on Arabic, does not provide any insight into changes in the pronunciation of these vowels, it is useful to consult sources which reference foreign languages to assist in establishing when this baseline was in effect. Meier (1981: 71–72) cites useful supporting evidence from Hamza al-Isbahani, who died prior to 970 AD. Writing in Arabic, al-Isbahani notes eight sounds found in Persian, but not Arabic. The only vowels included among these are **ō** as in **bō** (بو ‘smell’, CIP /bu/) and **ē** as in **sēr** (سير ‘full’, CIP /sir/).

Also useful in establishing prior pronunciation norms are early Latin transcriptions of Persian. One of these is the *Codex Cumanicus* (CC), a Latin-Persian-Turkish glossary compiled in the first half of the fourteenth century in Crimea. The Persian section of the manuscript was compiled by an unnamed Franciscan monk of Italian origin (Bodrogligeti 1971: 10). Bodrogligeti normalized the transcriptions found in the CC on the basis of evidence internal and external to the codex; we provide his normalized transcriptions here, with some modifications to represent consonants in the International Phonetic Alphabet (IPA). The CC offers plentiful examples of the existence of a distinction between the *majhul* (Persian/Arabic مجهول ‘unknown’) vowels, **ē** and **ō**, as distinct from their *ma’ruf* (Persian/Arabic معروف ‘known’) counterparts, **ī** and **ū**, so named to indicate their presence or absence in Arabic.

While Pisowicz (1985: 73), based on the presence of certain conversational elements of the material in the CC, believes that the informants were native Persians, he notes other research (Monchi-Zadeh 1969: 14) suggesting that the informants were Cumans, a Turkic people of the Eurasian steppe. Be that as it may, a sample of data from the CC indicating the distinctions among these vowels, contrasted with CIP, is shown in Table 1.

Vowel	Word	CC	CIP
ē	دیر 'late'	dēr	dir
	هیچ 'nothing'	hētʃ	hitʃ
ī	کیسه 'bag'	kīsa	kise
	تقویم 'calendar'	taqwīm	tayvim
ō	کوه 'mountain'	kō	kuh
	پوشش 'garment'	pōʃʃ	puʃeʃ
ū	موم 'wax'	mūm	mum
	گلو 'throat'	galū	gælu

Table 1. *Majhul* and *ma'ruf* vowels in the Codex Cumanicus and Contemporary Iranian Persian

As mentioned previously, areas where Persian spread early on retain the *majhul* vowels. According to Phillott (1919: 23), “The *majhul* sounds *o* and *e* are still preserved in the Persian spoken by Afghans and Indians, but they are now unknown in Persia.” Early Turkish borrowings from Persian indicate that the variety from which they were borrowed had maintained the *majhul* vowels. For example, Stein (2006: 147–148) cites *dost* ‘friend’ (cf. CIP دوست /dust/), *horoz* ‘rooster’ (cf. CIP خروس /xorus/), and *mešin* ‘sheepskin leather’ (cf. CIP میشن /miʃæn/). Kurmanji Kurdish is an example of a north-western Iranian language with a vowel system that maintains the *majhul* distinctions: e.g. *pêş* ‘front’ vs. *şîr* ‘milk’ and *roj* ‘day’ vs. *dûr* ‘far’ (Haig and Opengin 2012: 12–13).

2. Development of the Contemporary Iranian Vowel System

We will now explore the development of the contemporary vowels of Iranian Persian from the ENP model. Since ENP exhibits distinctions that are merged in different ways in other varieties, it is useful to view the ENP vowels as key vowels in a sense derived from Wells’s (1982) description of English. For example, Wells (1982: 120) defines several lexical sets, each identified by a keyword, which behave the same way “in respect of the incidence of vowels in different accents.” So words in the BATH set (consisting of words like *bath*, *path*, *staff* and *grass*) are generally pronounced with /æ/ in North American English and /a/ in Southern British English, whereas words in the TRAP set (e.g. *cat*, *back*, *mass*) are pronounced with /æ/ in both dialects. Returning to Persian, one can consider a lexical set consisting of words containing ENP ē and consider the transformations these words have undergone in other dialects. Labov (e.g. 1994: 164–165) has employed a similar notion called “word classes”, in order to facilitate comparison of dialects. Thiesen (1982: 9) developed a vowel notation that enables the reading of classical Persian poetry with

either a classical or a contemporary Iranian Persian pronunciation. As we will discuss below, such types of notation will be useful for developing a pandialectal pronunciation dictionary. Following Labov's notation for depicting vowels according to word classes, when comparing the ENP vowels to modern dialects, we place the ENP vowels in **bold**.

The ENP vowel system described above contrasts with that of CIP, as shown in Figure 2 (based on data from Majidi and Ternes 1999, Rees 2008 and Jahangiri 2000). As for the relative chronology of the changes between ENP and CIP, Windfuhr (1979: 144) states, "many questions actually have hardly been asked yet, such as those concerning the time of, and the conditions for, the lowering of the short high vowels to *e*, *o* ...". Pisowicz's (1985) study is a monumental effort in addressing this question, and we provide relevant examples from work cited in his study below. Note that we do not distinguish between long and short vowels in the modern system, since it appears that in contrast to the ENP system, the vowel system is currently based on quality rather than quantity (Lazard 1957, Toosarvandani 2004). Lazard (1957) characterizes the vowels derived from the ENP long vowels as "stable" and those derived from the ENP short vowels as "unstable". Compared to the stable vowels, the unstable vowels are more subject to fluctuations in quality and quantity.

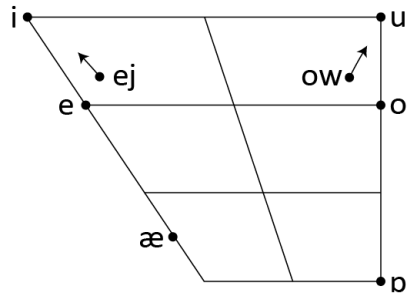


Figure 2. Vowel system of Contemporary Iranian Persian

The direction of changes between ENP long vowels and CIP is shown in Figure 3. As can be seen, all of these changes result in raising: $\bar{e} > i$, $\bar{o} > u^2$ and pre-nasally under certain linguistic, social, and stylistic conditions, $\bar{a} > u$ (Modaressi-Tehrani 1978: 74–109, Perry 1996: 274, Miller 2011). In Labov's (1994: 116) vowel shifting framework, these changes accord with Principle I: in chain shifts, long vowels rise. In fact, Labov (1994: 116) notes that this is the most robust of his principles of vowel shifting. According to Labov (1994: 118), a chain shift "is a change in the position of two phonemes in which one moves away from an original position that is then occupied by the other." Below we point out where the vowel shifts that are described result in this kind of situation, although at this stage we have not resolved the relative ordering of the shifts; e.g. whether these are push or pull chains (Labov 1994: 199–200).

² Haig and Opengin (2012: 42–43) describe the raising of /o:/ to /u:/ in Shemzinani Kurdish, a chain shift that was set in motion once the original /u:/ had been fronted to /y/.

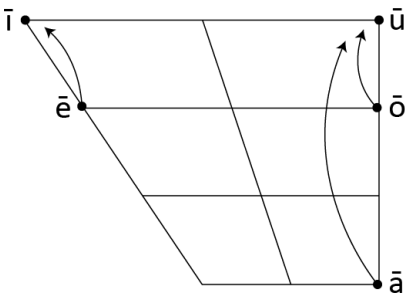


Figure 3. Long vowel raising between ENP and CIP

Note that the raising of *ē* > *i* and *ō* > *u* have resulted in mergers with CIP /i/ and /u/ which are descended from ENP *ī* and *ū*, respectively. These mergers have resulted in the following sets of words with distinct vowels in ENP and identical vowels in CIP, as shown in Table 2.

Word	ENP	CIP
سیر 'full'	sēr	sir
سیر 'garlic'	sīr	sir
شیر 'lion'	fēr	fīr
شیر 'milk'	fīr	fīr
رود 'river'	rōd	rud
بود 'was'	būd	bud

Table 2. Merger of ENP *ē*, *ī* and *ō*, *ū* in CIP

We will now explore the relative chronology and conditions for the shift from the ENP long vowel values to the contemporary ones. Meier’s (1981: 97) analysis of relevant rhymes in Persian poetry establishes that the merger of *ō/ū* preceded that of *ē/ī*, based on the persistence of failing to rhyme *ē/ī* after rhymes of *ō/ū* had become common. On the basis of such evidence, Perry (1996: 271) claims that *majhul* *ō* disappeared by the end of the 12th century and *majhul* *ē* disappeared between the late 15th and early 17th centuries. One example cited of a poet rhyming *ō/ū*, but not *ē/ī*, is Rumi, who lived in the thirteenth century AD. Here is one of several *abyāt* (sg. *beit*, analogous to a distich or couplet) provided by Nicholson (1926: 424) of rhyming words with *ō* and *ū*:

لطف حق این شیر را و گور را (gōrrā)
الف دادست این دو ضد دور را (durra)

The grace of God has given amity to this lion and wild-ass
these two far distant contraries³

³ Text: Nicholson 1925–1940, Vol. 1, p. 80, line 1294, translation: Nicholson 1925–1940, Vol. 2, p. 72, line 1294.

In contrast, the following *beit* alludes to the maintenance of the *ē/ī* distinction. While acknowledging that Rumi came from Balkh (in present-day Afghanistan), a region where the *majhul* distinctions most probably held sway, Meier (1981: 97) notes that he composed his poetry in the west (i.e. Anatolia), and took advantage of the full range of variational possibilities in his work:

کار پاکان را قیاس از خود مگیر (magīr)
گرچه باشد در نوشتن شیر، شیر (jīr)

Do not measure the actions of holy men by (the analogy of) yourself,
though *shér* (lion) and *shīr* (milk) are similar in writing⁴

As an indication that the status of *ē/ī* was in flux, however, Browne (1895: 239) offers the following *beit* as an example of Rumi's rhyming *ē* and *ī* in the same words:

آن یکی شیرست کادم میخورد
و این یکی شیرست کادم میخورد

That one is a *shīr* (lion) which eats man,
while this one is a *shīr* (milk) which man eats

The CC also sheds light on an earlier, non-absolute, stage in the mergers, revealing the linguistic contexts in which they first took hold. For example, while Bodrogligeti (1971: 46–47) notes that *ē* has generally been retained in the CC, he indicates that sometimes it surfaces as */ī/*, especially in the verbal prefix *mē* (می) before stem-initial *ā*, as in */mīāmīzam/* 'I mix' (میامیزم). He also notes that *ō* has generally been retained, but sometimes surfaces as *ū*: e.g. */pūst/* 'skin' (پوست), */frāmūf/* 'forget' (فراموش). Finally, while noting that *ā* has generally been retained, he provides a few examples where it has raised to */ō/* before nasals: e.g. */xōm/* 'unripe' (خام), */paʃaxōna/* 'bed curtain' (پشه خانه). Bodrogligeti's (1971) presentation of the CC's snapshot of sound change in progress is notable by its reference to various phonetic and morphological environments which later became irrelevant once the mergers reached completion, such as in the case of the merger of *majhul* and *ma'ruf* vowels in CIP.

A Latin transcription of a Persian Koran (LPK) by an unnamed Spaniard in the early part of the 17th century provides a later snapshot of the state of the changes described here (Bodrogligeti 1961). Table 3 below illustrates some of the data in this work. The column labelled LPK employs the symbols used in the manuscript. The *j* is equivalent to IPA *x*, and there is usually no indication of vowel length. As can be seen, there is variation in the text between forms more similar to the those found in ENP and those found in CIP. For example, Bodrogligeti (1961) provides examples of both the preservation of *majhul ē* and its merger with *ī*, even for the same word. He also provides some interesting data regarding the raising of *ā* before nasals.

⁴ Text: Nicholson 1925–1940, Vol. 1, p. 18, line 263; translation: Nicholson 1925–1940, Vol. 2, p. 18, line 263. The author uses an acute accent to indicate length.

Some words appear to have the intermediate target *o*, whereas others reach the modern pre-nasal raising target, *u*. The words in which this raising is exemplified show some variation with respect to modern practice. While the morpheme نامه ‘letter, book’ does undergo raising in some modern words, e.g. روزنامه ‘newspaper’ /ruznume/ (Peisikov 1960: 83), the animate plural morpheme نا ān has been claimed to rarely be subject to raising, due to its replacement by the general plural هā in the colloquial language (Kahn and Bernstein 1981: 136).

Phenomenon	Word	LPK	CIP
<i>majhul</i> ē preserved	ريختن ‘pour’	rejten	rixtaen
<i>majhul</i> ē variation	کنيد ‘do’ (2 nd person plural)	kuned, kunid	konid
<i>majhul</i> ē > i	می (present/durative prefix)	mi	mi
3. ā > o, u / ____ [+nasal]	نامه ‘book’	nome, nume	nome, ruznume ‘newspaper’ (Peisikov 1960)
ā > o / ____ [+nasal]	دانا ‘knowing’	dona	donon
ā > u / ____ [+nasal]	نيکان ‘good men’	nikun	nikon

Table 3. Variation in LPK

The direction of changes between ENP short vowels and CIP is shown in Figure 4. Two of these changes, *i* > *e* and *u* > *o*, result in lowering.⁵ This is in accordance with Labov’s (1994: 116) Principle II: in chain shifts, short vowels fall. The chain aspect here is that the /e/ and /o/ positions are vacant due to the CIP long vowel shifts described above where ē > i and ō > u. However, two changes involving ENP *a* do not adhere to Principle II. In general, *a* has raised to /æ/ in CIP, while word-finally it has raised even further to /e/ (Perry 1996: 272-273). Labov (1994: 116) notes that while this principle of short vowels falling applies to most available examples, there are exceptions.

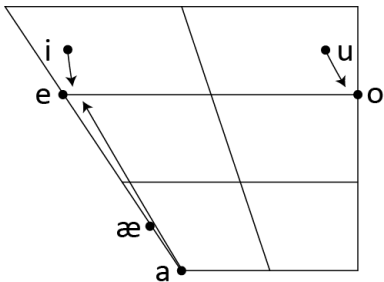


Figure 4. Short vowel changes between ENP and CIP

These changes are also foreshadowed in earlier documents. Horn (1901: 19) cites a Judeo-Persian text from Ahvaz, Khuzestan, from 1021 AD where *a* is written with the Hebrew letter *yod* to indicate a raised initial vowel in /kārd/ (כירד ‘did’), /häst/ (היסת ‘is’) and /fārmān/ (פירמאן ‘order’).⁶ While noting that ENP *a* was retained in

⁵ Thanks to Bruce Hayes for suggesting an investigation of this phenomenon.

the majority of cases, Bodrogligeti (1971: 43) cites several examples from CC of **a** having raised to what he also transcribes with *ä*, which is presumably /æ/ (Peisikov 1960 uses the same symbol to transcribe the CIP short *a*): e.g. /kärädäm/ (کردم ‘I did’), /zän/ (زن ‘woman’). Regarding word-final **a**, Bodrogligeti (1971: 43) notes that /a/ is predominant, but there was raising in /fambä/ (شنبه ‘Saturday’) in CC. In the later LPK (Bodrogligeti 1961), /e/ is very common for ENP **a** in all positions, except in the environment of uvulars, as noted by Pisowicz (1985: 80); e.g. /kerdenha/ (گردن ها ‘necks’) vs. /katere/ (قطره ‘drop’).⁷

While observing that the majority of ENP **i** vowels in CC are retained, e.g. /gil/ (گل ‘earth’), /dil/ (دل ‘heart’), Bodrogligeti (1971: 44) notes some examples of lowering to /e/: e.g. /xedmat/ (خدمت ‘service’), /āfeq/ (عاشق ‘in love’). Similarly, while reporting that generally ENP **u** is retained, e.g. /gul/ (گل ‘flower’), /gurg/ (گرگ ‘wolf’), he provides examples of lowering to /o/: e.g. /honarmand/ (هنرمند ‘skillful’), /moft/ (مشت ‘fist’). Similarly in LPK, **i** and **u** are generally retained, but here are lowered examples of both: /ke/ (که ‘that’), /jodauenda/ (خداوندا ‘O God’).⁸

The direction of changes between ENP diphthongs (or VC sequences) and CIP is shown in Figure 5. The chain aspect here is that the ENP diphthongs are now occupying the space previously occupied by ENP *ē* and *ō*, which have shifted to /i/ and /u/ respectively. These changes do not adhere to Labov’s (1994: 116) Principle IIA: in chain shifts, the nuclei of upgliding diphthongs fall, as exemplified in the English Great Vowel Shift where Middle English *ī* > Shakespearean /ej/ > modern /aj/ as in *bite* and Middle English *ū* > Shakespearean /ow/ > modern /aw/ as in *foul* (Jespersen 1949: 232). Labov (1994: 116–117) finds that this principle applies to a larger number of cases than Principle II (in chain shifts, short vowels fall), and ultimately restates the input to this principle as “short nuclei of upgliding diphthongs”. We have no reason to believe the nuclei of these diphthongs in ENP were not short, so these facts remain in opposition to Labov’s Principle IIA. However, it does not seem to be a coincidence that ENP **a** and **aj** are raised in tandem in CIP.

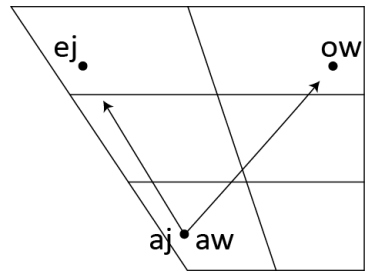


Figure 5. Diphthongs in CIP

⁶ Horn’s transcriptions are retained. Since Horn does not provide Judeo-Persian script, similar examples were taken from Margoliouth (1899). Thanks to Rachel Strong and Melissa Fox for help finding relevant examples in this document.

⁷ Original LPK transcription retained; vowel length generally not distinguished. As Bodrogligeti (1961: 265–266) notes, the transcriber exhibits variation in the transcription of uvulars and velars, in this case we see that the transcriber has used the same symbol, *k*, for the velar represented by گ /g/, and the uvular represented by ق /q,ɣ/.

⁸ Original LPK transcription retained; LPK *j* = IPA *x*; vowel length generally not distinguished.

Jahangiri (2000: 69–79) discusses the monophthongization of both /ej/ → [e] and /ow/ → [o] in contemporary Tehrani Persian. He finds that monophthongization is more common with /ow/ than /ej/, and that the more formal the style, and the higher the social group, the less the monophthongization encountered. Jahangiri mentions a further process among working class Tehranis whereby /ow/ is monophthongized to [o] and then raised to [u], e.g. [ʃuhar] for standard CIP /ʃowhar/ (شوهر ‘husband’). He stresses that this process only occurs to [o]’s resulting from monophthongization and not to [o]’s in general. Given that it would be reasonable to assume that [o]’s deriving from the diphthong /ow/ would be fairly long to start with (while Jahangiri discusses the lengthening resulting from monophthongization of /ej/ he does not mention it with respect to /ow/), this raising to [u] fits with Principle I, whereby long vowels rise. Its failure to apply to “short” [o], or those that have not resulted from monophthongization, fits with Principle II, which expects short vowels to fall rather than rise.

3. Afghan Persian

The standard Kaboli Afghan Persian, or Dari, vowel system is generally the same as that of ENP, though it is not clear whether the short vowels have retained their original quality. In Table 4, we compare the qualities assigned by several studies, adding phonemic equivalents using IPA symbols when the authors provide foreign example words. We see that the majority of transcription variation is in the mid vowels.

ENP	Farhadi (1955) French example	Neghat (1993) English example	Rees (2008) acoustic study	Kiseleva and Mikolaichik (1978)	Henderson (1972)
a	<i>salle</i> a	<i>up</i> ʌ, ə ⁹	a	a	æ, a
i	<i>geste</i> ε	<i>sit</i> ɪ	ε, ɪ	e	ε
u	<i>poste</i> ɔ	<i>put</i> ʊ	ʊ	o	ʊ
ā	Swedish <i>dag</i> ɒ ¹⁰	<i>fall</i> ɒ ¹¹	ɑ, ɔ	â	ɔ, ɑ
ē	<i>maire</i> ε: ¹²	<i>pay</i> e	e	ê	e
ī	<i>dit</i> i	<i>see</i> i	i	î	i
ō	<i>taux</i> o	<i>note</i> o	o	ô	o
ū	u	<i>moon</i> u	u	u	u

Table 4. Transcription of Dari vowels

Farhadi (1955: 8) observes that short /a/ retains its quality word-finally, and does not raise to /e/ as in CIP. However, he notes several words where ENP **a** does cor-

⁹ Cf. Ladefoged (1982: 28–30).

¹⁰ Farhadi’s choice of a Swedish example is interesting. French traditionally has a distinction between *pâte* /pat/ and *patte* /pat/, however this distinction is becoming less reliable in metropolitan France. In addition to providing a surer exemplar of a back /a/, Farhadi probably sought to indicate rounding. According to Haugen (2009), “Swedish *a* has moved closer to *â*, being backed and rounded.”

¹¹ American English *fall* has a range of pronunciations, including /a/, /ɒ/ and /ɔ/.

¹² Following /r/ serves as a *consonne allongante*, lengthening the preceding vowel (Walker 1984: 26).

respond to Dari /ɛ/: آتش ‘fire’ /āteʃ/, چشم ‘eye’ /tʃeʃm/, دست ‘hand’ /dest/. Also, in Dari, *ā* does not raise to *u* before nasals. Finally, in contrast to CIP, majhul *ē* and *ō* are preserved, as shown in Table 5. However, Farhadi (1975: 15) notes the word /firdān/ ‘faucet’, literally ‘lion mouth’, exhibiting the CIP merger, which he ascribes to Iranian influence.

Word	Afghan pronunciation	CIP pronunciation
نیست ‘is not’	/nest/	/nist/
گیتی ‘universe’	/geti/	/giti/
ابرو ‘eyebrow’	/abro/	/æbru/
روز ‘day’	/roz/	/ruz/

Table 5. Preservation of *ē* and *ō* in Dari

The diphthongs *aj* and *aw* are preserved in Dari, e.g. پیدا ‘find’ /pajdā/, چلو ‘rice’ /tʃalaw/, though Farhadi (1955: 8–9) notes a set of words where the spoken language uses /ē/ and the “traditional educated pronunciation” uses both /ē/ and /aj/: e.g. امید ‘hope’ /omed, omajd/ and پریشان ‘distressed’ /perefɒn, perajfɒn/. Similarly, he notes variation between /o/ and /aw/: e.g. روشن ‘clear’ /roʃan, rawʃan/, روغن ‘oil’ /roqan, rawqan/. This variability, combined with examples like /firdān/, indicates that the Dari vowel system may be in motion along the lines established earlier for CIP.

4. Tajik Persian

The Tajik vowel system, depicted in Figure 6, represents divergences from the ENP system in ways that are different from CIP (Perry 2005). The majhul *ē* and *ō* are preserved as /e/ and /u/ (for this symbol, see below) respectively, but ENP short *i* and long *ī* have merged as /i/, and ENP short *u* and long *ū* have merged as /u/. Tajik /o/ represents a raised and rounded ENP *ā* in all positions, in contrast to the socially and stylistically gradient *ā* and *u* of CIP (Modaressi-Tehrani 1978: 74–109, Perry 1996: 274, Miller 2011). Just as a qualitative system replaced a quantitative one in CIP, a different set of mergers has resulted in a qualitative system in Tajik.

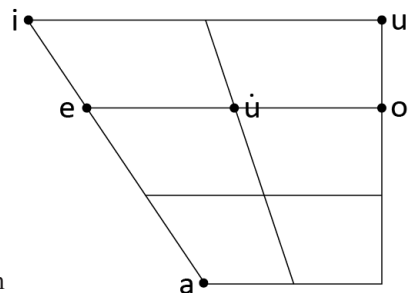


Figure 6. Tajik vowel system

Perry notes that Tajik /ü/ is also used for borrowings from Uzbek (in turn derived from common Turkic *ö* and *ü*). The vowel chart placement and symbol for this vowel are taken from Perry, who reprises a symbol used by Lazard (1956). Accord-

ing to Perry (2005: 18), /ü/ “lies phonetically between [u] and [y] ... higher than French *peu* (/ø/) ... less rounded and more lax than /u/.” Lazard (1956: 126) simply says it is lower and fronter than /u/. In any case, it is clear that it is fronted with respect to **ō**, which is our first example of Labov’s (1994: 116) Principle III: in chain shifts, back vowels move to the front.

In other Persian dialects we see fronting as well. For example Okati et al. (2009) report on **ū** fronting to [u] and [y] in Iranian Sistani. Miller and Moats (2011) report on data from a Herati (Afghanistan) speaker with a merger between **ū** and **ō** (cf. Farhadi 1955: 10 and Ioannesyan 2007: 268) realized as [y] as shown in Table 6. LeCoq (1989: 250) mentions other dialects where **ū** fronting is attested, including unrounding to /i/ in Semnani, e.g. /pil/ ‘money’, and Mazandarani, e.g. /dir/ ‘far’. Haig and Opengin (2012: 13, 42) cite Kurdish dialects where the cognate sound of **ū** is fronted to [ø:] or [y:] (Shemzinani) and others where it additionally unrounds to [i:] (Badinani in northern Iraq and south-eastern dialects of Kurmanji in Turkey). In apparent contrast to a view whereby fronted variants derived from an original /u/, Bodrogligeti (1961: 267) ascribes to Németh the supposition that ENP **u** had the quality [y], based on the presence of **ü** in Persian words borrowed into Turkish.

ENP vowel	Herati
ū	بودیم /bydim/ ‘we were’
ō	دوست /dyst/ ‘friend’
ō	فراموش /farāmyʃ/ ‘forget’

Table 6. Herati fronting

5. Synthesis

The vowel transformations between ENP and the various contemporary dialects we have discussed so far could be encoded in a pandialectal pronunciation key as shown in Table 7. Due to the preliminary nature of this aspect of the study, we have employed a mixture of IPA and traditional symbols, and in some cases, we have included a set of possible symbols encountered in the literature. Each word’s pronunciation could be provided underlyingly in ENP, and the various contemporary dialects’ pronunciations could be generated automatically by a rule-based system effecting the transformations indicated in the key, and then presented to users in one or more of the contemporary dialects in an electronic or online dictionary. We have not discussed consonantal variation, but a similar approach could be taken to account for the presence for example of /v/ in CIP and /w/ in Dari.

Of course, such a presentation ignores the fact that there are more complicated divergences among the dialects. For example, there are idiosyncratic differences between Dari and CIP, such as خنک ‘cool’, pronounced /xonak/ in CIP and /xonok/ in Dari. While the first vowel could be represented with ENP **u**, the variation in the second vowel cannot properly be derived using the method described above. Another example is the word عوض ‘exchange’, which is pronounced /avaz/ in CIP and

ENP	CIP	Kaboli	Herati ¹³	Tajik
ū	u	ū	y, u	u
ō	u	ō	y, ō, u	u
ā	ɒ, u (before nasals under certain conditions)	ɒ	ɒ, u (before and after nasals), a (unstressed)	o
ē	i	ē	ē, i	e
ī	i	ī	i	i
a	æ, e (word-finally)	a	æ, ɛ (word-finally)	a
i	e, ɛ	ɛ, i	i, e, ɛ	i
u	o	ū, o	u	u

Table 7. Rudiments of a pandialectal pronunciation key for Persian

/ēwaz/ in Dari. In this case, Farhadi (1955: 20) indicates that there is compensatory lengthening in Dari resulting from the non-pronunciation of ɛ.

6. Conclusion

We have attempted to show the utility of the ENP vowel system in understanding the variation among the vowel systems of several diachronically and synchronically separated varieties of Persian which are derived from it. The vowel changes have been considered with respect to Labov's (1994) principles of vowel shifting, and they appear both to corroborate his findings and to provide some alternative directions of change in particular cases. Foreign-language descriptions of Persian in different periods have been adduced to capture sound change in progress in cases where the contemporary situation exhibits sound changes that have reached completion. Finally, a practical way of using this knowledge to provide learners and scholars with a dynamic picture of Persian vowels in a number of distinct varieties has been suggested in the form of a pandialectal pronunciation dictionary. In the future, we hope by means of a pandialectal acoustic study to shed light on the exact quality of vowels, as well as the appropriate symbols to use to describe them when viewed from a holistic Persian perspective.

Abbreviations

CC: *Codex Cumanicus*
 CIP: Contemporary Iranian Persian
 ENP: Early New Persian
 IPA: International Phonetic Alphabet
 LPK: Latin transcription of a Persian Koran

References

- Bodrogligeti, A. (1961). "The Persian Translation of the Koran in Latin Letters." *Acta Orientalia Academiae Scientiarum Hungaricae* 13, pp. 261–276.
 Bodrogligeti, A. (1971). *The Persian Vocabulary of the Codex Cumanicus*. Budapest: Akadémiai Kiadó.

- Browne, Edward G. (1895). "Review of *A Grammar of the Persian Language*, by John T. Platts." *Journal of the Royal Asiatic Society* 27(1), pp. 235–245.
- Farhadi, Abd-ul-Ghafur (1955). *Le Persan parlé en Afghanistan*. Paris: Klincksieck.
- Farhadi, A. G. Rawan (1975). *The Spoken Dari of Afghanistan*. Kabul: Peace Corps.
- Haig, Geoffrey, and Öpengin, Ergin (2012). "Kurmanji Kurdish in Turkey: Structure, Varieties and Status" (unpublished manuscript).
- Haugen, Einar (2009). "Danish, Norwegian and Swedish." In: Comrie, Bernard (ed.), *The World's Major Languages* (2nd edition). London: Routledge, pp. 124–144.
- Henderson, Michael M. T. (1972). *Dari (Kabul Persian) Phonology*. PhD dissertation, University of Wisconsin.
- Horn, Paul (1901). "Neupersische Schriftsprache." In: Geiger, Wilhelm and Kuhn, Ernst (eds), *Grundriss der iranischen Philologie* 1:2. Strassburg: Trübner, pp. 1–200.
- Ioannesyan, Youli (2007). "Situating the Khorasani Dialects within the Persian-Dari-Tajiki Dialect Continuum." In: Lewis, Franklin, Sharma, Sunil, and Moayyad, Heshmat (eds), *The Necklace of the Pleiades*. Amsterdam: Rozenberg, pp. 267–278.
- Ioannesyan, Youli (2009). *Afghan Folktales from Herat*. Amherst, New York: Cambria Press.
- Jahani, Carina, and Korn, Agnes (2009). "Balochi." In: Windfuhr, Gernot L. (ed.), *The Iranian Languages*. London: Routledge, pp. 634–692.
- Jahangiri, Nader (2000). *A Sociolinguistic Study of Persian in Tehran*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- Jespersen, Otto (1949). *A Modern English Grammar on Historical Principles*. London: George Allen and Unwin.
- Kahn, Margaret, and Bernstein, Jared (1981). "Progress on the uN Change in Persian: Dæbestan vs. Tabestun." *Acta Orientalia Academiae Scientiarum Hungaricae*, 35(1), pp. 133–140.
- Kiseleva, L., and Mikolaichik, V. (1978). *Dari-Russkii Slovar*. Moscow.
- Labov, William (1994). *Principles of Linguistic Change: Internal Factors*. Oxford: Blackwell.
- Ladefoged, Peter (1982). *A Course in Phonetics*, Second Edition. New York: Harcourt Brace Jovanovich.
- Lazard, Gilbert (1956). "Caractères distinctifs de la langue Tadjik." *Bulletin de la Société de Linguistique de Paris* 52(1), pp. 117–186.
- Lazard, Gilbert (1957). *Grammaire du Persan Contemporain*. Paris: Klincksieck.
- LeCoq, Pierre (1989). "Le classement des langues irano-aryennes occidentales." In: de Fouchécour, Charles-Henri and Gignoux, Philippe (eds), *Études irano-aryennes offertes à Gilbert Lazard* [Studia Iranica Cahier 7]. Paris: Association pour l'avancement des études iraniennes, pp. 247–264.
- Majidi, Mohammad-Reza, and Ternes, Elmar (1999). "Persian." In: *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press, pp. 124–125.
- Margoliouth, D. S. (1899). "A Jewish-Persian Law Report." *The Jewish Quarterly Review* 11(4), pp. 671–675.
- Meier, Fritz (1981). "Aussprachefragen des älteren Neupersisch." *Oriens*, 27–28, pp. 70–176.
- Miller, Corey (2011). "A Holistic Treatment of /ān/ to [un] in Persian." *Proceedings of ICPHS XVII*, Hong Kong, pp. 1386–1389.
- Miller, Corey, and Moats, Najiba (2011). "Convergence on [u] across Persian Dialects." Poster presented at International Conference on Iranian Linguistics 4, Uppsala, Sweden.
- Modaressi-Tehrani, Yahya (1978). *A Sociolinguistic Analysis of Modern Persian*. Ph. D. dissertation, University of Kansas.
- Monchi-Zadeh, Davoud (1969). *Das Persische im Codex Cumanicus* [Studia Indoeuropaea Upsaliensia, 1]. Uppsala: Acta Universitatis Upsaliensis.
- Neghat, Muhammad Nassim (1993). *Dari-English Dictionary*. Omaha: University of Nebraska.
- Nicholson, Reynold A. (1926). "On Editing the *Mathnawi*." *Bulletin of the School of Oriental and African Studies*, 4(2), pp. 421–430.
- Nicholson, Reynold A. (1925–1940). *The Mathnawi of Jalálu'ddīn Rūmī*. London: Brill.
- Okati, Farideh, Ahangar, Ali Abbas, and Jahani, Carina (2009). "Fronting of /u/ in Iranian Sistani." *Orientalia Suecana*, 58, pp. 120–131.

- Okati, Farideh (2012). *The Vowel Systems of Five Iranian Balochi Dialects* [Studia Linguistica Upsalien-
sia, 13, Studia Iranica Upsaliensia, 20]. Uppsala: Acta Universitatis Upsaliensis.
- Paul, Ludwig (2002). "A Linguist's Fresh View on Classical Persian." *Cahiers de Studia Iranica*, 26, pp.
21–34.
- Peisikov, L. S. (1960). *Tegeranskii Dialekt*. Moscow.
- Perry, John R. (1996). "Persian in the Safavid Period: Sketch for an *Etat de Langue*." In: Melville, Charles
(ed.), *Safavid Persia: The History and Politics of an Islamic Society*. London: I.B. Tauris & Company,
pp. 269–283.
- Perry, John R. (2005). *A Tajik Persian Reference Grammar*. Leiden: Brill.
- Phillott, D. C. (1919). *Higher Persian Grammar*. Calcutta: Baptist Mission Press.
- Pisowicz, Andrzej (1985). *Origins of the New and Middle Persian Phonological Systems*. Krakow: Jagiel-
lonian University.
- Rees, Daniel A. (2008). *Towards Proto-Persian: An Optimality Theoretic Historical Reconstruction*. PhD
dissertation, Georgetown University.
- Stein, Heidi (2006). "Palatal-Velar Vocalism of Arabic-Persian Loanwords in 16th-Century Ottoman Turk-
ish." In: Johanson, Lars and Bulut, Christiane (eds), *Turkic-Iranian Contact Areas: Historical and Lin-
guistic Aspects*. Wiesbaden: Harrassowitz, pp. 143–157.
- Thiesen, Finn (1982). *A Manual of Classical Persian Prosody*. Wiesbaden: Harrassowitz.
- Toosarvandani, Maziar Doustdar (2004). "Vowel Length in Modern Farsi." *Journal of the Royal Asiatic
Society*, Third Series 14(3), pp. 241–251.
- Walker, Douglas C. (1984). *The Pronunciation of Canadian French*. Ottawa: University of Ottawa Press.
- Wells, J. C. (1982). *Accents of English 1*. Cambridge: Cambridge University Press.
- Windfuhr, Gernot (1979). *Persian Grammar: History and State of its Study*. The Hague, Paris and New
York: Mouton.
- Windfuhr, Gernot (2009). "Persian." In: Comrie, Bernard (ed.), *The World's Major Languages* (2nd edi-
tion). London: Routledge, pp. 445–459.
- Windfuhr, Gernot, and Perry, John (2009). "Persian and Tajik." In: Windfuhr, Gernot (ed.), *The Iranian
Languages*. London: Routledge, pp. 416–544.

The Distribution and Role of the Verb Clitic =a/a= in Different Balochi Dialects

Maryam Nourzaei, Uppsala University

maryam.nourzaei@lingfil.uu.se

Carina Jahani, Uppsala University

carina.jahani@lingfil.uu.se

Abstract

This study investigates the distribution and role of the verb clitic =a/a= in a variety of Balochi dialects. We summarize earlier findings, argue for the interpretation of this morpheme as a clitic, and present new data for four Balochi dialects spoken in Iran. In these four dialects, the verb clitic is variously consistently proclitic (one dialect), restricted proclitic (two dialects), and restricted enclitic (one dialect). We also find that there is a stronger tendency for enclitic attachment of the verb clitic =a/a= the further eastwards a certain Balochi dialect is spoken.

The basic semantic component of the verb clitic =a/a= is one of imperfectivity, and it is used with the non-past indicative and the past imperfective indicative verb forms. It can be linked to the homophonous verb clitic found in several other Western Iranian languages.

Keywords: Balochi, Iranian Balochi dialects, Iranian languages, verb clitic, imperfectivity, enclitic, proclitic

1. Introduction¹

One important feature of different Balochi dialects is the distribution and function of the element =a/a=, which until now has been described in various terms, e.g. as a prefix (Pierce 1875: 9; Elfenbein 1982: 86–89), suffix (Buddruss 1977: 11), affix (Buddruss 1988: 62), verbal element (Baranzehi 2003: 92; Bashir 2008: 56; Jahani and Korn 2009: 661), enclitic particle (Axenov 2006: 166) and clitic (Barjasteh-Delforooz 2010: 79). The aim of the present study is to discuss the status and semantic function of this element in a number of Balochi dialects, including four dialects spoken in Iran for which we present hitherto unpublished data. These four dialects are spoken in Fars Province around Shiraz (known as Koroshi) [BKIr], in Hormozgan Province close to Minab [BMIr], in Jashk [BJIr], and in Habd [BHIr]. A possible diachronic development of the element more globally in Balochi will also be discussed in view of its occurrence in other Iranian languages as well.

The descriptions of BKIr, BMIr, BJIr, and BHIr are based on data gathered by Maryam Nourzaei and transcribed by Maryam Nourzaei and Carina Jahani. The corpus consists of recorded folktales, life stories, and procedural texts told by male and female linguistic consultants between approximately 40 and 60 years of age with different social backgrounds.

The fieldwork for BKIr was carried out in 2009 and for BMIr, BJIr, and BHIr in

¹ Sincere thanks to the anonymous referees and to Geoffrey Haig, Bamberg, for valuable comments on earlier versions of this article. In fact, Geoffrey Haig's comments contributed to a considerable restructuring of our argumentation. We are also very grateful to Agnes Korn and Christian Rammer for producing the map.

2011. For BKIr we have six texts: four folktales, one procedural text, and one life story – amounting to a total of about 70 minutes of recordings. There are three linguistic consultants for BKIr, two men and one woman. For BMIr we have four texts: two procedural texts, one text about local traditions, and one life story – amounting altogether to about 10 minutes of recordings. Here we have two female linguistic consultants. The data for BJIr consist of three texts: one folktale and two life stories – amounting to about 25 minutes. Three linguistic consultants, two men and one woman, contributed with these texts. For BHIr there are five texts: one life story, one procedural text, one text about old traditions, and two folktales – amounting to about 25 minutes, and told by three linguistic consultants, one man and two women.

First, we summarize findings about the verbal element =a/a= in previous studies with a concluding discussion of the cliticization versus affixation of this verbal element in different Balochi dialects and a comparison with the Persian verb prefix *mi-*. Then we present the distribution and function of =a/a= in BKIr, BMIr, BJIr, and BHIr. The article concludes with a discussion of possible areal distribution of proclitic and enclitic cliticization of the verbal element in different Balochi dialects.



Map 1. Distribution of Balochi dialects

2. Previous descriptions of the verbal element =*a/a*=

In one of the earliest descriptions of Southern Balochi (based on the dialect spoken east of Gwadar, Pakistan), Pierce (1875: 9–10) testifies to the presence of what he calls an *a*-prefix, noting that “[w]hen preceded by a word ending in a short vowel, the aorist usually loses its *a*”. Elfenbein (1982: 86–89) discusses this so called *a*-prefix at length in connection with the fact that it is also found in Brahui. He concludes that “[t]here is no doubt that the *a*-prefix in Balochi is Iranian: it occurs in many other languages of the family as well (in most Central languages as well as in Kurdish, where its presence is very plain)” (Elfenbein 1982: 89),² something which also Buddruss (1988: 62) agrees with in his description of Afghani Balochi [BAf]. Buddruss, however, finds that it should not be regarded as a prefix, but as added to the word preceding the verb. He calls it a suffix (Buddruss 1977: 11) or an affix (Buddruss 1988: 62) but symbolizes it as =*a*, which is the way clitics are normally represented. Elfenbein (1982: 87), however, “cannot agree with Buddruss that ‘prefix *a*-’ should be re-interpreted as ‘suffix *a*-’ on the previous word”.

In their description of the Balochi dialect spoken in Noshke, Pakistan, Barker and Mengal (1969: 1: 149) state that often when the substantive element of a complex predicate ends with a consonant (and especially when it ends with two consonants), a brief *ə*-sound is heard between it and the verb. They hold that this *ə*-sound “is optional and has no discernible meaning.” Elfenbein (1990: 2: IX), on the other hand, is of the opinion that in what he calls Rakhshani, a conglomerate of dialects of Western Balochi including the dialect of Noshke, the *a*-prefix, as he calls it, “has definite durative value” but in all other dialects has no semantic function (Elfenbein 1990: 2: X–XVIII).

In a number of recent studies of various Balochi dialects, following Buddruss (1977, 1988) the verbal element =*a/a*= has been treated as attached to the word preceding the verb rather than to the verb itself. Baranzehi (2003) describes this verbal element in Balochi of Sarawan, Iran [BSaIr], Axenov (2006) in Turkmenistan Balochi [BT], and Bashir (2008) in Eastern Balochi (spoken in Pakistan) [EBP], and Barjasteh Delforooz (2010) also mentions its presence in Sistani Balochi [BS].³

Axenov (2006: 166) uses the term “enclitic particle” and Barjasteh Delforooz (2010: 80) the term “clitic” for this morpheme. Baranzehi (2003: 89) argues for it being prosodically attached to the word preceding the verb rather than to the verb itself, and Bashir (2008: 56–57) also attaches it to the word preceding the verb; so it is clear that they also regard it as a clitic, or, more specifically an enclitic, even if they do not employ this term. It attaches to nouns, adjectives, pronouns, adverbs, numerals, prepositions (functioning as preverbs), and past participles directly preceding the finite verb.⁴ They also note that the clitic never occurs in clause initial position and that there are also other restrictions on where it can occur.

In BT, the clitic can only appear after a noun, pronoun, adjective or adverb (see

² McCarus (2009: 608) describes *a*- in Kurdish as an imperfective aspect prefix rather than as a clitic. Lecoq (1989: 315, 316, 319) also mentions a verb prefix *a*- in several other Iranian dialects, such as Vafsi, Khonsāri, Farizandi, and Bādrudi, to mention a few. Gershevitch (1985: 225) likewise reports the presence of a present-future “particle *a*-” in Bashkardi.

³ Since Barjasteh Delforooz’s work is on discourse features rather than on morphosyntax, he does not describe the =*a* clitic in detail.

Axenov 2006: 272–292). Also in BS, there are restrictions on when it can occur (see Barjasteh Delforooz 2010: 286–391). It is not found in connection with the verb *dāštin* ‘to have’, after the clitic =*ē* indicating indefiniteness, after the suffix *-rā* added to the indefiniteness marker =*ē* to mark a direct object, after the word *am/um* ‘also, even, and’, and, due to its enclitic status, at the beginning of a sentence or a clause. Neither Ahangar (2007) nor Baranzehi (2003) mentions any restrictions regarding its use, but since it is described as an enclitic, it is likely that similar restrictions are found in BGlr and BSaIr as in BT and BS. In Karachi Balochi, Pakistan [KBP] described by Farrell (1990 and 2003) there is no mention of the verbal element.

The semantics of the clitic are unanimously described as marking imperfective aspect. In the past system, the clitic expresses imperfective aspect, whereas in the non-past⁵ system, it has no obvious aspectual meaning but simply marks indicative mood, regardless of aspectual distinctions.

In the dialects described above in which the verb element =a/a= exists, it thus appears as an enclitic and is absent if there is no preceding word to attach to, or if the preceding word cannot host it (because, for example, it is prosodically too weak, such as the complementizer *ke*). It denotes indicative mood in the non-past tense, whereas it marks imperfective aspect in the past tense (see also Jahani and Korn 2009: 661–662). It is thus parallel in use to the Persian verb prefix *mi-*. The imperfective aspectual marking of the clitic =a/a= must be regarded as its primary semantic component.

The origin of the Modern Persian verb prefix *mi-* is easy to trace in view of the long literary tradition of New Persian. It was an adverb, *hamē(v)* ‘also’ in Middle Persian, which gradually became more and more closely attached to the verb to denote imperfective aspect both in the non-past and past tense. In Early New Persian it is found as *hamē-* or *mē-*, which normally attaches before the verb stem, i.e. *(ha)mēkonam* ‘I keep doing’, but in poetry also often as a free adverb *hamē* after the verb, i.e. *(ha)mēkonam* or *konam hamē*, and only gradually did it develop into an obligatory prefix in the present indicative and past imperfective in Persian.

The origin of the Balochi =a/a= may likewise be an adverb that gradually became more and more closely attached to the verb, but it seems that prosodically =a/a= has remained a clitic rather than developing into a prefix, at least in the dialects of Western Balochi for which we have a considerable corpus (e.g. BAf, BT, BS). As an enclitic, there are also constraints to where the element can occur. The most natural constraint is, of course, clause initial position, but there are also other restrictions on where the clitic occurs, as we have seen above. This creates a certain instability, which may lead to even further reductions and finally a total dropping of the verb element, if it was ever present, as we see in, e.g., Karachi Balochi.

⁴ An affix is a bound morpheme that functions on the word level both syntactically and phonologically and attaches to one specific category of words, e.g. only to nouns or only to verbs. A clitic, on the other hand, functions above the word level syntactically and can attach to words belonging to various syntactic categories, as is the case with the morpheme under discussion. For characteristics of clitics and a contrastive comparison between clitics and affixes, see also: <http://www.sil.org/linguistics/GlossaryOfLinguisticTerms/WhatIsACliticGrammar.htm>

⁵ The term “non-past tense” is here used for the present-future tense, even if the descriptions referred to use the latter term, or the term “present tense”.

However, our new data will reveal that there are also Balochi dialects further to the west in which $=a/a=$ is more firmly attached to the verb and thus neither sensitive to clause initial position nor to the presence of another preceding word that could act as a host for the element. In these dialects, it would indeed be possible to describe the element as a prefix. However, in conformity with other Balochi dialects where it is an enclitic, even in dialects where it attaches to the verb we continue to describe it as a proclitic rather than a prefix. We will also see that it is only in one of the dialects described below that the clitic is never omitted in the non-past indicative and the past imperfective. The $=a/a=$ is thus also somewhat unstable in two of the dialects where the verbal element occurs as a proclitic.

Furthermore, the clitic $=a/a=$ could in the dialects described below be interpreted as placed on the pre-verbal element and act like an enclitic whenever possible and only as a proclitic when there is no pre-verbal element to attach it to, like in other Balochi dialects described previously.⁶ However, since it is prosodically very difficult to determine whether it attaches to the verb or the element preceding the verb, in dialects containing cases where the only possible interpretation is proclitic, this interpretation is applied in all instances. This applies to three of the dialects described below. If, on the other hand, the clitic does not appear in contexts where it must be described as a proclitic, i.e. does not occur in clause initial position or after a clear pause, it is here described as an enclitic. This happens in one dialect. In all dialects the clitic is always unstressed.

3. The clitic $=a/a=$ in Koroshi Balochi, and Balochi of Minab, Jashk, and Habd (Iran)

As in the dialects described above, the clitic indicates indicative mood in the non-past tense and imperfective aspect in the past tense in the four dialects to be described below. It is also interesting to note that in BJIr, probably under the influence of Persian, the clitic has also acquired irrealis modality in the past tense.⁷

3.1. *Koroshi Balochi, Iran [BKIr]*

In BKIr, the clitic must at least in some instances be described as a proclitic $a=$, because it occurs in positions where it cannot be described as an enclitic (since there is no host to which it could attach). For this reason we classify the clitic as proclitic in all its occurrences. There are no clear instances where it cannot be proclitic and there are no restrictions on its use with indicative verbs in the non-past tense and on verb forms with imperfective aspect in the past tense indicative. It has the allomorph $ar=$, which occurs in almost all cases when the verb stem begins with r -.⁸ It is also worth noting that there is a free variant $\bar{a}=$ as well, which is

⁶ This analysis was adopted in Jahani and Nourzaei (2011).

⁷ Irrealis modality expressed by the verb clitic is also observed in Nourzaei's unpublished data from Southern Balochi (Iran).

⁸ In fact, the presence of an allomorph indicates that the clitic is becoming more closely attached to the verb and that it therefore may be on the verge of acquiring the status of a prefix. See also fn. 10.

frequently used with a number of verbs beginning in *k-* (both original and as a marker of imperfectivity, see Axenov 2006: 166–167) and with a few other verbs as well.

In example (1), the clitic occurs twice in a clause initial position (*ar=rant*⁹ ‘they go’, *a=natwānt* ‘it cannot’), where it can only be a proclitic, and twice in ambiguous cases which can be either enclitic or proclitic (*a=žanant* ‘they arrive’, *sost a=kant* ‘it gets tired’). In the first of the ambiguous cases, it occurs after a conjunction, where it would not be found in dialects where =a is only an enclitic. We also see the allomorph *ar=* in *ar=rant* ‘they go’.

- (1) *ar=r-ant* *tā* *a=žan-ant* *ye* *sarābālāī-yā*
VCL=go.PR-3PL until VCL=arrive.PR-3PL one hill-OBL
kar-ok *sost* *a=kan-t*
donkey-DIM weak VCL=do.PR-3SG
*a=na-twān-t*¹⁰ *bār-ā* *be-bā*
VCL=NEG-can.PR-3SG load-OBJ SUBJ-carry.PR.3SG

They go until they arrive at a hill. The donkey gets tired; it cannot carry the load.

Example (2) provides three tokens of clause initial *a=* (*a=šī* ‘it says’, *a=kafey* ‘you will fall’, *a=merey* ‘you will die’), and one ambiguous instance where dialects with enclitic =a would also have had the clitic (*a=keyt* ‘it comes’).

- (2) *dobāra* *korrag* *a=keyt* *ǰolo=ē*
again foal V.CL=come.PR.3SG in.front.of=PC.3SG
a=š-ī *rū=ye* *ǰālī-yā* *na-nen-ey* *ā*
V.CL=say.PR-3SG on=EZ carpet-OBL NEG-SUBJ.sit.PR-2SG EMPH
čēr=ey *čāh=en*
under=PC.3SG well=COP.PR.3SG
a=kaf-ey *mā* *čāh-ā* *a=mer-ey*
V.CL=fall.PR-2SG in well-OBL V.CL=die.PR-2SG

Again the foal comes before him [and] says: Don’t sit on the carpet; beneath it is a well; you will fall into the well [and] die.

Examples (3) and (4) show the free variation between *a=* and *ā=* on the verb form *a=kāyant/ā=kāyant* ‘they come’.

- (3) *ā=kāy-ant* *had=e* *īš-ī*
V.CL=come.PR-3.PL to=IZ DEM.PROX-OBL

They come to him...

⁹ In the verb forms quoted above the examples, only the clitic is indicated as a separate morpheme. For a detailed morpheme analysis of the verb forms, please refer to the glossing in the examples.

¹⁰ Note that, unlike the Persian prefix *mi-*, the =a/a= clitic in Balochi precedes the negation. This suggests a rather late grammaticalization and is an argument for regarding the element as a clitic. Note, however, the counterargument presented in fn. 8.

- (4) *ya rō* *šāh=o* *īš-ān* *a=kāy-ant* *tafrī-yā*
 one day king=and DEM.PROX-PL V.CL=come.PR-3.PL leisure-OBL
mā ē *bāḡ-ā*
 in DEM.PROX garden-OBL

One day the king and his men come to this garden for leisure.

Example (5) demonstrates $a=/\bar{a}=$ as a marker of imperfective aspect in the past tense indicative. The sentence tells about something that habitually took place in the past. In none of the verbs here (*ēreš ā=koda* ‘they put down’ (twice), $a=zortadeš$ ‘they took’) is there any indication whether the clitic is enclitic or proclitic. In view of (6), though, it is interpreted as proclitic also in the past tense. Though outside the scope of the present study, it may be interesting to note the ergative construction with the agent clitic in this and other examples of transitive verbs in the past tense. Furthermore, the COP.PT.3SG frequently loses its final *-d* when there is no enclitic pronoun or personal ending attached to it.

- (5) *ya nāh=o nagan=ē ēr=eš ā=kod=a*
 one date=and bread=IND down=PC.3PL VCL=do.PT=COP.PT.3SG
ya māst=o nagan=ē ēr=eš ā=kod=a
 one yoghurt=and bread=IND down=PC.3PL VCL=do.PT=COP.PT.3SG
ya māst=o xormā=e a=zort=ad=eš
 one yoghurt=and date=IND VCL=take.PT=COP.PT.3SG=PC.3PL

They used to put some bread and dates down [in their bag] they used to put some yoghurt and bread down [in their bag]. They used to take some yoghurt and dates.

In example (6), there are two clear examples of a proclitic, namely the second occurrence of $a=čedadeš$ ‘they sheared’, which is a clause initial verb, and $a=restaden$ ‘we spun’, where there is a clear pause after the previous word (*towsānān* ‘in the summer’). In view of this, the clitic is regarded as a proclitic not only in the non-past, but also in the past tense. In the first occurrence of $a=čedadeš$ both enclitic and proclitic interpretations are possible, but in view of the two instances of unambiguous proclitic use, a proclitic interpretation is chosen in all instances.

- (6) *bozz-ey mūd-ā a=čed=ad=eš*
 goat-GEN hair-OBJ VCL=shear.PT=COP.PT.3SG=PC.3PL
a=čed=ad=eš towsān-ān
 VCL=shear.PT=COP.PT.3SG=PC.3PL summer-PL
a=rest=ad=en
 VCL=shear.PT=COP.PT.3SG=PC.3PL

They sheared the goat’s hair; they sheared [it] in the summer; we spun [it].

3.2. *Balochi of Minab, Iran [BMIr]*

In BMIr the verb clitic *a=* (also here with the allomorph *ar=*) is not as stable as in BKIr. It is occasionally absent when the verb is in clause initial position or when it is only preceded by a conjunction in the clause. There are, however, instances where it is clearly proclitic, for which reason it is here analysed as a proclitic in all instances.

Example (7) provides two examples of proclitic use of *a=*, where the verb occurs in clause initial position (*a=rēč-en* ‘we pour’, *a=dranjene* ‘we hang it up’) and two examples of ambiguous (proclitic or enclitic) use (*a=dōč-en* ‘we sew’ (twice)). All the non-past indicative verbs in this example are marked with the clitic.

- (7) *ya čī=ye pārče-i a=dōč-en kisa*
 one thing=IZ cloth-ADJZ VCL=sew.PR-1PL bag
pārče-i a=dōč-en
 cloth-ADJZ VCL=sew.PR-1PL
a=rēč-en be āhī lāfā
 VCL=pour.PR-1PL to DEM.DIST.GEN inside
a=dranj-en=e da āf=eš b-reč-ī
 VCL=hang.PR-1PL=PC.3SG until water=PC.3PL SUBJ-pour.PR-3SG

We sew something of cloth, a bag; we sew [it] of cloth [and] pour [it, i.e. boiled whey which has turned into lumps] into it. We hang it up [i.e. the bag] until the (lit. its) water [of the boiled whey] dries up (lit. pours out).

Example (8) contains three verbs in the non-past indicative, two of which are not marked with the clitic. One of them, *mānī* ‘she stays’, occurs immediately after a conjunction, and the other one, *kārenī* ‘we bring her’, occurs in clause initial position. In neither of these cases would there be any clitic in, e.g., BT (Axenov 2006: 169–170). The third verb, *dāwatī a=kanen* ‘we invite her’, takes the clitic. Here, of course, enclitic interpretation is possible, but in view of (7), also here the clitic is interpreted as a proclitic.

- (8) *se rōz ke mān-ī ōdānā*
 three day CLM stay.PR-3SG there
bād deya¹¹ dāwat=ī a=kan-en
 then again invitation=PC.3SG VCL=do.PR-1PL
k-ār-en=ī ba lūg-ā
 IMPk-bring.PR-1PL=PC.3SG to house-OBL

When she (i.e. the new bride) has stayed there (i.e. at her parents’ place) for three days, then we invite her, you know, and bring her home (i.e. to the parents-in-law’s place where she will remain).

In example (9), the clitic is found as a marker of imperfectivity in the past tense indicative on the verbs *ar=raptaden* ‘we went’, *a=čedaden* ‘we picked’, and *hošken a=koda* ‘we dried’. In the two first examples, only proclitic interpretation is possible, whereas with the third verb enclitic interpretation would be possible. In view of

¹¹ This word functions here as a discourse marker rather than an adverb meaning ‘again’. It corresponds to ‘you know’ in the free translation.

the two unambiguous proclitic examples, the third clitic is also regarded as a proclitic. The sentence tells about repeated (habitual) actions in the past. Here we also find the allomorph *ar=* before a verb stem beginning in *r-*. It may also be interesting to note that the person ending in the 1PL is homophonous with the 1PL enclitic pronoun (agent clitic) in this dialect.

- (9) *ar=rapt=ad-en* *a=čed=ad=en*
 VCL=go.PT=COP.PT-1PL VCL=pick.PT=COP.PT.3SG=PC.1PL
hošk=en *a=kod=a*
 dry=PC.1PL VCL=do.PT=COP.PT.3SG
 We went; we picked [herbs] and dried [them].

3.3. *Balochi of Jashk, Iran [BJIr]*

In BJIr, the verb clitic is found in instances where it can only be interpreted as a proclitic, and is therefore treated as a proclitic in this dialect, but it is even less stable here than in BMIr. It is sometimes found after a pause and after the conjunction *ke*, where it is normally absent in, e.g., BT (Axenov 2006: 169–170),¹² but it is also frequently missing after a pause and *ke*. After the personal pronoun *man* ‘I’ and in connection with the verb *bayag* ‘to be, to become’, it is sometimes missing, as well as after the OBL marker *-ā*.¹³ No allomorphs are found in BJIr.

In example (10), there are four verbs in the non-past indicative, three of which are marked with the clitic (*baranda a=bīd* ‘he wins’, *a=gī* ‘he gets, he buys’, *a=zīrī* ‘he takes’). In all these instances, enclitic interpretation is possible, but in view of example (12), where the clitic is unambiguously proclitic, the latter interpretation is selected for all occurrences of the clitic in BJIr. The verb *borrīd* ‘he cuts’, which follows directly after an OBL suffix *-ā*, is not marked with the clitic, as is also the case in BT (see Axenov 2006: 168).

- (10) *dāxe* *kōštī-yā* *malek.mahmad* *baranda*
 in wrestling-OBL PN winner
a=b-īd=o *ya* *mēš=e* *a=g-ī*
 VCL=become.PR-3SG=and one sheep=IND VCL=get.PR-3SG
ā *sarī* *mēš-ay* *sar-ā* *borr-īd=o*
 DEM.DIST quickly sheep-GEN head-OBL cut.PR-3SG=and
a *tamam=e* *āhī-ay* *jön-a* *fağa*
 from all=IZ DEM.DIST.OBL-GEN corpse-OBL only
lāf=ay *a=zīr-ī*
 stomach=PC.3SG VCL=take.PR-3SG

In the wrestling, Malek Mahmad wins. He takes a sheep. He quickly cuts off the head of the sheep [and] from the whole slaughtered sheep (lit. corpse) [he] takes only its stomach.

¹² Axenov’s (2006: 169) remark that “the aspectual particle *-a* is never used before the conjunction of general subordination *ki*” is a *lapsus calami*. It should read: the aspectual particle *-a* is never used after the conjunction of general subordination *ki*.

¹³ In BT, the clitic is only missing after the OBL (in BT classified as OBJ) marker *-ā*, not in connection with the verb ‘to be, to become’ or after the personal pronoun ‘I’.

In example (11), the clitic is absent on the sentence initial verb *gušt* ‘he says’, as in BT (Axenov 2006: 170), whereas in (12), the clitic is present on the verb *a=ra* ‘he goes’, where it is impossible to consider it enclitic, since there is a considerable pause after the word *hamedān* ‘right here’. This shows a certain instability in the use of the clitic, but at the same time it is clear that it functions as a proclitic in BJIr.

- (11) *gušt-ī* *yā* *hazrat*¹⁴=*e* *musā*
say.PR-3SG VOC presence=IZ PN
He says: “Your Holiness Moses...”

- (12) *ya...* *do* *nafar* *ye* *nafar*
one... two person one person

a *mošomay* *molk-ā* *ham=edān*
from PRON.1PL.INCL¹⁵ land-OBL EMPH=here

a=ra-Ø *be* *emārāt*
VCL=go.PR-3SG to PN

One... two persons, one person from our region (i.e. from the local area), right here, goes to the Emirates.

In example (13), the clitic occurs on the verb *a=gozar-ī* ‘it passes’ after the conjunction *ke*, whereas in (14) it is missing on the verb *lōt-ēn* ‘we want’ after the same conjunction, like in, e.g., BT (Axenov 2006: 169) and BS (see Barjasteh Delforooz 2010: 287, 299 etc.). This, again, shows that there is a certain instability in the use of the clitic in BJIr.

- (13) *čān* *rōč* *ke* *a=gozar-ī*
some day CLM VCL=pass.PR-3SG
After some days... (lit. when some days pass)

- (14) *čīz=ē* *ke* *lōt-ēn* *be* *mardom=e*
thing=IND CLM want.PR-1PL to person=IND

yā *brād* *yā* *gohār=ē-ā* ???¹⁶
or brother or sister=IND-OBL ???

Something that we want (possibly: to be given) to someone, a brother or a sister...

Example (15) shows that the clitic is absent on the verb *nazānon* ‘I don’t know’ after *man* ‘I’, at least in this example, which is the only example of this environment in the data from BJIr. This is contrary to BT (see Axenov 2006: 49, 127 etc.) and BS

¹⁴ The word *hazrat* is an honorific title used in Islam to refer to, e.g., prophets, among whom Moses is counted.

¹⁵ The reason why the narrator uses a 1PL INCL pronoun is that he is telling the story to his uncle, who comes from the same region as himself, rather than to the researcher, who also was present during the recording.

¹⁶ It is impossible to discern this word, but that has no bearing on the discussion.

(see Barjasteh Delforooz 2010: 287, 292 etc.), where the clitic is actually present after *man* ‘I’. The negation may, of course, be a crucial factor for the absence of the clitic.¹⁷

- (15) *man* *gušt-a=Ø* *man* *na-zān-on*
 PRON.1SG say.PT-PP=COP.PR.3SG PRON.1SG NEG-know.PR-1SG
 I said: ‘I don’t know.’

Example (16) provides an example where the clitic is absent in connection with the verb *bayag* ‘to be, to become’. This is, again, an environment where it occurs in, e.g. BS (see Barjasteh Delforooz 2010: 377 etc.).

- (16) *malek.mahmad* *ham* *xošhāl* *b-īd*
 PN also happy become.PR-3SG
 As for Malek Mahmad, he becomes happy.

In example (17), two instances are found of the clitic as a marker of imperfective aspect in the past tense (*a=šoda* ‘he went’, *darse a=wānta* ‘he studied’). It would, of course, be possible to interpret the clitic as an enclitic in both these instances, but in view of the unambiguous proclitic in (12), the clitic is interpreted as a proclitic here as well.

- (17) *malek.mahmad* *hameša* *ba* *maktab* *a=šod=a=o*
 PN always to school VCL=go.PT=COP.PT.3SG=and
dars=e *a=wānt=a*
 lesson=PC.3SG VCL=read.PT=COP.PT.3SG
 Malek Mahmad always went to school and studied.

Example (18) shows that the clitic is also used to denote irrealis modality in the past tense (*a=bo*, *a=dā*). This phenomenon occurs in Persian as well, where the *mi*-marked past form (e.g. *mikard*) is used both for imperfective aspect and irrealis and for counterfactual modality, and it is therefore not unlikely that this semantic extension is copied from Persian.¹⁸

- (18) *aga* *čūpān* *a=bo* *ā* *fad-ī*
 if shepherd VCL=win.PT.3SG DEM.DIST REFL-GEN
xanjar-ā *ba* *āhī* *a=dā*
 dagger-OBL to DEM.DIST.OBL VCL=give.PT.3SG
 If the shepherd won, then he (i.e. Malek Mahmad) would give his dagger to him (i.e. to the shepherd).

¹⁷ Note, however the presence of both the clitic and the negation in example (1). See also fn. 23, where the verb form ‘I don’t know’ is discussed.

¹⁸ The Persian sentence *agar bačče zende mimānd hame xošhāl mišodand* can be interpreted in two ways depending on the context, either as irrealis modality ‘if the child stayed alive, everybody would be happy (sometimes the child stayed alive, sometimes he/she died)’, or as counterfactual modality ‘if the child had stayed alive, everybody would have been happy (but he/she died)’.

3.4. *Balochi of Habd, Iran [BHlr]*

In BHlr the verb clitic =a can be interpreted as an enclitic in all instances found in the corpus. This interpretation is adopted here.¹⁹ It is not found after a pause, nor does it occur after the clitic =o ‘and’, the OBL suffix -ā, or the CLM *ke*. The data for BHlr displays no allomorphs for the clitic.

In example (19), there are two verbs with the clitic, *jedā=a kanēn* ‘we separate’ and *āf=a dahene* ‘we give them (lit. it) water’, and one verb in clause initial position, where the clitic is missing, *rahēn* ‘we go’.²⁰ As already stated, the clitic is interpreted as an enclitic in BHlr, since there are no occurrences that are clearly proclitic.

- (19) *šanek-ān-a* *jedā=a* *kan-ēn*
goat-PL.OBL-OBJ separate=VCL do.PR-1PL
rah-ēn *āf=a* *dah-en=e*
go.PR-1PL water=VCL give.PR-1PL=PC.3SG

We separate the goats. We go and give them (lit. it)²¹ water.

Just like in (19), in (20) we find the verb *rahēn* ‘we go’ in clause initial position (twice) without the clitic. In the same way, there is no clitic in connection with the verb *bandene* ‘we tie them (lit. it)’, which occurs after the enclitic conjunction =o ‘and’. The same is true in BS (see Barjasteh Delforooz 2010: 314), and probably in BT as well, although Axenov does not mention this case.

- (20) *rah-ēn* *banjāh=o* *band-en=e*
go.PR-1PL stable=and tie.PR-1PL=PC.3SG
rah-ēn *kāsag-ā* *zīr-ēn* *šōd-en=ī*
go.PR-1PL bowl-OBL take.PR-1PL wash.PR-1PL=PC.3SG

We go to the stable and tie them (lit. it). We go [and] take a bowl [and] wash them.

Example (21) provides an example of absence of the clitic on the verb *bī* ‘it becomes’ after the conjunction *ke*, like in BT (Axenov 2006: 169) and BS (see Barjasteh Delforooz 2010: 287, 299 etc.) and sometimes also in BJlr (see (14)). The other two verbs, *rahen* ‘we go’ and *kanen* ‘we do’ occur in clause initial position and also lack the clitic.

- (21) *sard* *ke* *b-ī* *rah-en*
cold CLM become-PR.3SG go.PR-1PL
kan-en *ma* *hīzak-ā*
do.PR-1PL in goat.skin-OBL

When it gets cold, we go and put it (i.e. the yoghurt) in the goat skin.

¹⁹ There is, however, nothing (e.g. a pause after the clitic) that prohibits interpreting it as a proclitic. However, in view of the enclitic interpretation in previous studies (see, e.g., Buddruss 1988, Axenov 2006, Barjasteh Delforooz 2010), this interpretation is also adopted for BHlr.

²⁰ In the examples from this dialect, the vowel in the 1PL personal ending is sometimes *e*, sometimes *ē*.

²¹ The coreferent of the plural noun *šanekāna* ‘the animals’ is here a 3SG form of the PC. Animals and things often treated as grammatically singular when they are viewed collectively.

In examples (22) and (23), we see that the clitic is missing after the OBL/OBJ marker *-ā/-rā* in connection with the verbs *pačan* ‘they bake’ and *naylan* ‘they don’t allow’, as is the case in BT (see Axenov 2006:168) and sometimes also in BJIr (see Ex. 10).

- (22) *nagan-ā* *kapsul-ān-ī* *sarā* *pač-an*
 bread-OBL gas.bottle-PL.OBL-GEN on²² cook.PR-3PL
 They bake the bread on a tin-plate [with a flame underneath] burning by means of gas.

- (23) *mā-rā* *nay-l-an* *ke* *mā*
 PRON.1PL.EXCL-OBJ NEG-allow.PR-3PL CLM PRON.1PL.EXCL
wat-ī *boz-ān* *āp* *day-en*
 self-GEN goat-PL.OBL water SUBJ.give.PR-1PL
 They don’t allow us to water our goats.

In example (24), the use of the verb clitic to denote imperfective aspect in the past tense indicative is demonstrated with the verbs *=a pakkān* ‘they baked’ and *=a wārtān* ‘they ate’. The text from which (24) is taken is about old traditions and habits rather than single events in the past.

- (24) *komāč=a* *pakk-an* *komāč=a* *wārt-an*
 a kind of bread=VCL bake.PT-3PL a kind of bread=VCL eat.PT-3PL
 They baked *komāč* (a kind of bread baked on heated sand), they ate *komāč*.

4. Summary

Dialect	Proclitic	Allomorphs	Status
BAf	–	–	not consistently present
BGIr	–	–	not consistently present
BHIr	–	–	not consistently present
BJIr	+	–	not consistently present
BKIr	+	+	consistently present ²³
BMIr	+	+	not consistently present
BSaIr	–	–	not consistently present
BS	–	–	not consistently present
BT	–	–	not consistently present
EBP	–	–	not consistently present
KBP	Ø	Ø	absent

Table 1. Features of the *=a/a=* clitic in eleven Balochi dialects

²² The postposition *sarā* was originally the noun *sar* ‘head’ in the OBL case.
²³ During the finalization of this article, we actually came across one verb in the present indicative which occasionally has no clitic. The verb form we have found in two recently transcribed texts is *nazānān* ‘I don’t know’ (after a pause). In both texts there are also negative present indicative forms after a pause with the verbal clitic, and in one of the texts the very form *a=nazānān* is encountered. It thus seems that the use of the verb clitic with the verb ‘to know’ is somewhat unstable, at least when this verb is negated.

Table 1 shows that there is a considerable strengthening of the clitic the further westwards a certain dialect is spoken. BKIr, spoken in Western Iran, exhibits such a strong status for the clitic that it could in fact be re-analysed as a prefix. It is consistently present as a proclitic in the verb forms where it is expected and it also exhibits allomorphs. A slightly weaker status is found in BMIr, also spoken west of Balochistan proper (in Hormozgan Province), and the same status, although without displaying any allomorphs, is found in BJIr (spoken further to the east in Hormozgan Province). In the mainstream dialects, the clitic is not consistently present, and only in instances where it is possible to regard it as an enclitic (i.e. never if there is no word that can host it preceding the verb) and there are no reports of allomorphs. In all the dialects where the clitic is found, it always precedes the negation. The south-easternmost dialect, KBP, exhibits no clitic.

5. Conclusions

The verb clitic in Balochi has in recent studies been described as an enclitic =a. In this study, the verb clitic in Balochi is found to be a proclitic occurring in clause initial position. This holds for three of the four Iranian Balochi dialects for which new data are presented here, namely BKIr, BMIr, and BJIr. In these dialects the clitic is thus symbolized as *a=*, whereas in the dialect where it only occurs as an enclitic, BHlr, it is symbolized as *=a*.

There are several factors that lead us to believe that the verb clitic =a/a= is not a recent innovation in Balochi. The strongest argument is the fact that the same morpheme, *a-*, is found, and described as a verb prefix, in several other Western Iranian languages in addition to Balochi, both North-Western (e.g. Kurdish) and South-Western (e.g. Bashkardi). Another argument is that it is widespread throughout the Balochi speaking region and found in all the three main dialect groups, although more strongly in Western Balochi than in Southern and Eastern Balochi.²⁴ A third argument is that there is no strong candidate for where it should have been copied from. None of the languages of wider communication and/or national languages that Balochi is under the influence of has a verb clitic =a/a= denoting imperfective aspect.

The origin of the clitic is, however, not clear. It is not unlikely that it, like the Persian *mi-*, originates from an adverb which gradually became more and more closely attached to the verb. It is interesting to note that the westernmost dialects BKIr, BMIr and BJIr, spoken in Fars and Hormozgan Provinces, can attach the clitic as a proclitic, whereas in dialects spoken farther to the east, it only occurs as an enclitic. In KBP, the dialect spoken in the very south-east of the Balochi speaking area, the clitic is totally absent. There thus seems to be a stronger preference for proclitic at-

²⁴ For the main dialect division of Balochi into Western, Eastern and Southern dialects, see Jahani and Korn (2009: 636–638). The clitic has been attested in descriptions of several Western Balochi dialects as well as in Eastern Balochi. For Southern Balochi, it has not been attested in KBP, but it is present in several Southern Balochi variants spoken in Iran. The dialects described here must, e.g., be classified as variants of Southern Balochi. It is also found as an enclitic in the Lashari dialect described by Yusefian (1383) (see sample text, p. 186).

tachment the further westwards the dialects are spoken. Enclitic attachment is the preference further eastwards. Enclitic attachment also imposes limitations on when the clitic can be present, resulting in weaker grammaticalization and potential loss (if it ever was present). In BKIr, spoken farthest to the west, there is full grammaticalization of the clitic as a proclitic, possibly reinforced by contact with Persian.²⁵

The semantic function of the verb clitic =*a/a*= in Balochi to denote durativity and imperfective aspect is found in a much wider variety of Balochi dialects than what could be ascertained when Elfenbein (1990: 2: IX–XVIII) wrote that this clitic only carries a semantic meaning in some Western Balochi dialects. In the past tense, this aspectual meaning is realized as an imperfective indicative verb form, which stands in contrast to the past perfective indicative form, while in the non-past tense, where durativity is an inherent aspect of verb forms including the present moment, the form with the clitic has instead come to denote indicative mood. There are also examples in the corpus where the form marked with the clitic in the past tense has an irrealis modal meaning. This must be seen as an extension of its semantic field, probably copied from Persian.

List of abbreviations and symbols

-	indicates a morpheme break
=	separates off a clitic
()	explanations in the English translation
[]	additions in the English translation
1	first person
2	second person
3	third person
ADJZ	adjectivizer
BAf	Afghani Balochi
BHlr	Balochi of Habd, Iran
BGlr	Balochi of Granchin, Iran
BJlr	Balochi of Jashk, Iran
BKlr	Koroshi Balochi, Iran
BMlr	Balochi of Minab, Iran
BSaIr	Balochi of Sarawan, Iran
BS	Sistani Balochi
BT	Turkmenistan Balochi
CLM	clause linkage marker
COP	copula
DEM	demonstrative
DIST	distal
EBP	Eastern Balochi, spoken in Pakistan
EXCL	exclusive
GEN	genitive
IMP _k	imperfective prefix <i>k</i>
INCL	inclusive

²⁵ Special thanks to Geoffrey Haig whose suggestions resulted in this paragraph. Haig (private communication 17 Dec. 2012) also finds that the Kurdish corresponding morpheme “is clearly attached to the verb itself, but nevertheless retains certain clitic tendencies.”

IND	indefinite (specific) form of a noun
IZ	izāfa
KBP	Karachi Balochi, spoken in Pakistan
NEG	negation
OBL	oblique
OBJ	object
PC	pronominal clitic (enclitic pronoun)
PL	plural
PN	proper noun
PP	past participle
PR	present
PRON	personal pronoun
PROX	proximal
PT	past
REFL	reflexive pronoun
SG	singular
SUBJ	subjunctive
VCL	verb clitic
VOC	vocative particle

References

- Ahangar, Abbas Ali (2007). "A Description of the Verbal System in Sarhaddi Balochi of Granchin." *Orientalia Suecana* 56, pp. 5–24.
- Axenov, Serge (2006). *The Balochi Language of Turkmenistan. A Corpus-based Grammatical Description* [Studia Iranica Upsaliensia, 10]. Uppsala: Acta Universitatis Upsaliensis.
- Baranzehi, Adam Nader (2003). "The Sarawani Dialect of Balochi and Persian Influence on It." In: Jahani, Carina, and Korn, Agnes (eds), *The Baloch and Their Neighbours: Ethnic and Linguistic Contact in Balochistan in Historical and Modern Times*. Wiesbaden: Reichert, pp. 75–111.
- Barjasteh Delforooz, Behrooz (2010). *Discourse Features in Balochi of Sistan. Oral Narratives* [Studia Iranica Upsaliensia, 15]. Revised version. Uppsala: Acta Universitatis Upsaliensis. Online at: <http://uu.diva-portal.org/smash/record.jsf?pid=diva2:345413>
- Barker, Muhammad Abd-al-Rahman, and Mengal, Aqil Khan (1969). *A Course in Baluchi*. Montreal: McGill University Press. 2 vols.
- Bashir, Elena (2008). "Some Transitional Features of Eastern Balochi: An Areal and Diachronic Perspective." In: Jahani, Carina, Korn, Agnes, and Titus, Paul (eds), *The Baloch and Others: Linguistic, Historical and Socio-political Perspectives on Pluralism in Balochistan*. Wiesbaden: Reichert, pp. 45–79.
- Buddruss, Georg (1977). "Buttern in Baluchistan." In: Benzing, Brigitta, Böcher, Otto, and Mayer, Günter (eds), *Wort und Wirklichkeit. Studien zur Afrikanistik und Orientalistik* [Eugen Ludwig Rapp zum 70. Geburtstag]. Teil II, Linguistik und Kulturwissenschaft. Meisenheim am Glan: Hain, pp. 1–16.
- Buddruss, Georg (1988). *Aus dem Leben eines jungen Balutschen, von ihm selbst erzählt*. Stuttgart: Steiner.
- Elfenbein, Josef (1982). "Notes on the Balochi-Brahui Linguistic Commensality." *Transactions of the Philological Society* 80, pp. 77–98.
- Elfenbein, Josef (1990). *An Anthology of Classical and Modern Balochi Literature*. Harrassowitz, Wiesbaden. 2 vols.
- Farrell, Tim (1990). *Basic Balochi. An Introductory Course* [Baluchistan Monograph Series, 1]. Naples: Istituto Universitario Orientale.
- Farrell, Tim (2003). "Linguistic Influences on the Balochi Spoken in Karachi." In: Jahani, Carina, and Korn, Agnes (eds), *The Baloch and Their Neighbours. Ethnic and Linguistic Contact in Balochistan in Historical and Modern Times*. Wiesbaden: Reichert, pp. 169–210.

- Gershevitch, Ilya (1985). "The Crushing of the Third Singular Present." In: Gershevitch, Ilya, *Philologica Iranica*, selected and edited by Nicholas Sims-Williams. Wiesbaden: Reichert, pp. 223–236.
- Jahani, Carina, and Korn, Agnes (2009). "Balochi." In: Windfuhr, Gernot (ed.), *The Iranian Languages*. London and New York: Routledge, pp. 634–692.
- Jahani, Carina, and Nourzaei, Maryam (2011). "A Folktale in Koroshi Dedicated to Joy Edelman." In: Dodyxudoeva, L. R., Vinogradova, S. R., and Bayandur, A. S. (eds), *Leksika, étimologiya, yazykovye kontakty. K yubileyu doktora filologičeskix nauk professora Džoy Yosifovny Édel'man*. Moscow: Akademiya Nauk, pp. 62–70.
- Lecoq, Pierre (1989). "Les dialects du centre." In: Schmitt, Rüdiger (ed.), *Compendium Linguarum Iranicarum*. Wiesbaden: Reichert, pp. 313–326.
- McCarus, Ernest N. (2009). "Kurdish." In: Windfuhr, Gernot (ed.), *The Iranian Languages*. London and New York: Routledge, pp. 587–633.
- Yusefian, Pakzad (Yusefiān, Pākzād (1383 [2004])). *Towṣīf=e radešenāxti=ye zabān=e baluči*. Ešfahān (unpublished PhD thesis).
- <http://www.sil.org/linguistics/GlossaryOfLinguisticTerms/WhatIsACliticGrammar.htm> [Retrieved Sept. 25, 2012].

Adpositions and Argument Indexing in the Mukri Variety of Central Kurdish: Focus on Ditransitive Constructions

Ergin Öpengin, Lacito CNRS, Université Paris III, and Otto-Friedrich-University, Bamberg

ergin.opengin@etud.sorbonne-nouvelle.fr

Abstract

A set of adpositions in Central Kurdish interact with the argument indexing system in a peculiar manner. Previous research (Edmonds 1955; MacKenzie 1961; Samvelian 2007; Haig 2008; Jügel 2009) has described argument indexing in Central Kurdish, but ditransitive constructions are largely omitted. Here I study the adpositions and argument indexing in Central Kurdish in relation to semantic roles, valency and modality. I argue that the adpositions in question have obtained the semantically specified syntactic function of introducing a participant, usually a third-participant, into the speech event. Then I describe and analyse two previously unattested constructions, namely a periphrastic causative construction, where an absolute adposition introduces a causee in an indirect causation construction, and a periphrastic potentiality construction with passive inflection, where the adpositional complement is a “weakened actor” of the given event.

Keywords: Central Kurdish, absolute adpositions, argument indexing, ditransitive constructions, periphrastic constructions, semantic roles

1. Introduction

A set of adpositions in Central Kurdish (CK), often called absolute prepositions, interact with the argument indexing system, and also operate in the valency and modality systems of the language.¹ When the complements of these adpositions are pronominal, rather than full NPs, the rules governing the realization of the complements are complex, depending among other things on the transitivity value and tense of the verb in the clause concerned. The pronominal complements of absolute adpositions may be expressed through clitic person markers attached to their governing adposition, they may be non-locally attached to other constituents, or they may be expressed through verbal affixes on the verb. While previous research (Edmonds 1955; MacKenzie 1961; Haig 2008; Jügel 2009) has described argument indexing and the alignment with transitive verbs in CK, little attention has been paid to ditransitive verbs. Samvelian (2007), on the other hand, dealt specifically with absolute prepositions focusing on the syntactic status of their complements.

With a view to complementing the previous scholarship, I will study the adpositions and argument indexing in the Mukri variety of CK in relation to semantic roles, valency and modality.² I will argue that the adpositions in question have obtained the semantically specified syntactic function of introducing a participant, usually a third participant, into the speech event. The introduced participant can be an

¹ I would like to express my gratitude to Geoffrey Haig for his rich feedback at different stages of the writing of this paper. Many thanks also to Carina Jahani and Diana Forker as well as two anonymous reviewers for their careful reading and valuable suggestions for improvements. Naturally, all remaining errors are my responsibility.

argument, subcategorized for by the verb, or not, in which case it is a meaning adjustment strategy. Finally, two further previously unattested constructions will be described in which adpositions are deployed in valency and modality: (i) a periphrastic causative construction where an absolute adposition introduces a causee; (ii) a periphrastic potentiality construction with passive inflection where a specific type of actor is added to the event.

2. Argument indexing in Mukri Kurdish

I follow Siewierska (2004) and subsequent work in treating free pronouns, clitic pronouns, and traditional agreement affixes under the cover term of “person markers” (PM).³ Four types of person markers are distinguished in Mukri. The first comprises free forms, while the other three are bound (all are shown in Table 1. Forms and paradigms of person markers in Mukri Central Kurdish). Here we are interested only in bound forms.

Clitic PMs are restricted to occurring at the right edge of certain other constituents. Their status as clitics derives primarily from their mobility and from the fact that they are relatively free in host selection (e.g. their hosts may be nouns, verbs, preverbal particles, or inflectional affixes).⁴

Verbal affixes are inflectional suffixes that occur solely after the verb stem. They are divided into two sets: Set₁ used in the present tenses, and Set₂ used in the past tenses. Finally, copular endings are clitics, yet they behave like verbal affixes in terms of argument indexing. They will be discussed along with verbal affixes.⁵

		INDEPENDENT PMs		DEPENDENT PMs			
		Free pronouns		Clitic PMs	Verbal Affix PMs		Copular PMs
		Strong	Weak		Set ₁	Set ₂	
SG	1	emin	min	= <i>(i)m</i>	- <i>(i)m</i>	- <i>(i)m</i>	- <i>(i)m</i>
	2	eto	to	= <i>(i)t</i>	- <i>ī/-y</i>	- <i>ī/-y</i>	- <i>ī/-y</i>
	3	ew	(e)wī (obl.)	= <i>ī / =y</i>	- <i>ē(t)/-ā(t)</i>	- <i>ø</i>	- <i>(y)ē</i>
PL	1	eme	me	= <i>mān</i>	- <i>īn/-yn/-n</i>	- <i>īn/-n</i>	- <i>īn/-n</i>
	2	engo	=ngo	= <i>tān</i>	- <i>(i)n</i>	- <i>(i)n</i>	- <i>(i)n</i>
	3	ewān	wān	= <i>ū / =w</i> = <i>yān</i>	- <i>(i)n</i>	- <i>(i)n</i>	- <i>(i)n</i>

Table 1. Forms and paradigms of person markers in Mukri Central Kurdish

² The variety analysed here is Mukri Central Kurdish. It is spoken around the city of Mehabad in the north of Iranian Kurdistan. The data come from my three-hour-long Mukri corpus of annotated connected speech (including folktales, anecdotes, autobiographical narratives and procedural texts), that I collected in the field during two field trips in March–April 2011 and October 2011. Examples from my field notes are also included and native speakers’ judgements are consulted for grammaticality. The abbreviations in the square brackets following the English translation of the examples stand for individual texts and sentence numbers in the database. The source texts are listed at the end of this article.

³ See Haspelmath (2012) for recent discussion and references.

⁴ See Korn (2009) for the diachrony and extent of pronominal clitics in West Iranian languages and Öpengin (in preparation) for an extensive description of the morphosyntax of person markers in Mukri Kurdish.

⁵ Despite their clitic status, the copular PMs will not be shown in the glosses by a preceding equal sign (=), the sign being reserved for clitic PMs, which are of crucial importance for the current study.

Person markers indicate arguments bearing different syntactic functions. I will follow the widespread practice of referring to S, A, P (cf. Comrie 1989) for the core arguments of intransitive and monotransitive clauses. More recently, alignment typology has extended the inventory of basic syntactic functions to cover three-participant constructions (cf. Margetts and Austin 2007; Malchukov *et al.* 2010). Accordingly, T refers to the theme argument of a three-participant event, while R is any (direct or oblique) argument covering recipient-like semantic roles.

Having introduced the terminology, some examples showing the way person markers in Mukri are deployed will be provided. First, it must be noted that the S in all tenses (1) and the A in the present tense constructions (1) are obligatorily indexed by verbal affixes.⁶

- (1) *de-č-m* *ew* *kābrā-e-ī* *de-hēn-m*
 IND-go.PRS-1SG:S DEM fellow-DEM2-OBL IND-bring.PRS-1SG:A
 I shall go and fetch that fellow. [NZ.063]

In past tense constructions, however, an A is obligatorily indexed by a clitic PM (2). A free pronoun or NP in A function may also occur in the clause, depending on the information structure (e.g. contrastive focus), but this does not block the occurrence of the clitic PM. In this, the construction is clearly clitic-doubling.

- (2) (*eme*) (...) *gōdirēž-eke-ān=mān* *best-ewe*
 (1PL) (...) donkey-DEF-PL=1PL:A tie.PST-PTCL
 (We) tied the donkeys. [ŽB.047]

Unlike the PMs indexing an A-past, clitic PMs indexing a P in the present tense do not allow clitic doubling. That is, the P in present tense can be indexed by clitic PMs, but only if there is no coreferent-free expression of the P in the same clause, as illustrated in (3a). The P in the past tense is indexed by verbal affixes (3b), and again a coreferent-free pronoun or NP cannot occur in the same clause.⁷ Thus, in (3c), the plural NP in P function is not indexed on the verb (the 3SG copular form being the default person index with participial verb forms).⁸

- (3) a. *de-č-m* [**ew kābrā-e-ī*]_i *de=ī-hēn-m*
 IND-go.PRS-1SG [DEM fellow-DEM2-OBL] IND=3SG:P-bring.PRS-1SG
 I shall go (and) bring him. [NZ.066]
- b. *hīč* *kes* (**emin*)_i *řā=ī-ne-de-girt-m_i*
 no person (1SG) PVB=3SG:A-NEG-IPFV-keep.PST-1SG:P
 Nobody would accept me in. [ŽB.024]
- c. *ser=ī* *bēčo-eke-ān=yān* *helpisānd-ū-e*
 head=EZ chick-DEF-PL=3PL:A cut.off.PST-PTCP-COP.3SG
 They have cut off the chicks' heads. [FN.25]

⁶ Notice that a rather morphemic transcription is adopted here; allomorphic variation and epenthesis are usually not shown.

⁷ There are instances where a non-obligatory P index occurs with a coreferent NP (but not with a free pronoun) in the same clause. This obviously contradicts the neat presentation of P indexing above, but the issue is complex and not necessarily pertinent to the purposes of this paper.

⁸ Copular endings are completely parallel to verbal affixes in terms of person marking in the past tenses. The difference is distributional: copular endings are used on past participle verb forms whereas verbal affixes are employed elsewhere.

Finally, the R participants, which for the purposes of this study include traditional indirect object and oblique arguments introduced by adpositions and other strategies, may also be expressed through clitic PMs in the present tense. With R, this kind of indexing is not obligatory, but only occurs in the absence of a coreferent-free expression (see examples (6a) and (6c) below).⁹ In the past tense, similar to P indexing, instead of clitic PMs we find verbal affixes in this function.¹⁰ Thus, in (4b) and (8b), a verbal affix alternatively marks the R and a coreferent-free form is precluded.

Now that argument indexing has been described, I will continue by describing the mechanisms by which the third participants are introduced.

3. Three-participant events in Mukri Central Kurdish

The concept of “ditransitive constructions” presumes that the R participant is a syntactic argument. However, distinguishing an argument from an adjunct is acknowledged to be a task with peculiar difficulties (cf. Comrie 1993: 906–907). The concept may also exclude derived ditransitives like causative and applicative constructions (e.g. Malchukov *et al.* 2010: 2). To avoid the complexities of these issues and also to account for the derived ditransitives in CK, I will employ, following Margetts and Austin (2007: 401) and Margetts (2007: 71), the somewhat semantic concept of “three-participant event” and analyse events which are dynamic states of affairs that crucially involve three entities in their conceptualization. The concept thus accounts not only for the obligatory constituents but also for those whose expression is merely possible.

Of the cross-linguistic strategies of encoding three-participant events diagnosed in Margetts and Austin (2007: 402–403) three are relevant to Kurdish: (i) the oblique and adjunct strategy, where a third participant is expressed as an oblique argument or an adjunct; (ii) the causative strategy, where a third participant is added to the monotransitive verb root through causativization; (iii) the possessive strategy, where the third participant is an adnominal dependent of one of the arguments. The first two deploy adpositions, whereas the possessive strategy does not, hence it will not be considered here.

3.1. *Oblique and adjunct strategy*

3.1.1. Absolute adpositions introducing the R

As mentioned in the introduction, a set of adpositions, called absolute prepositions following MacKenzie (1961), hold a key role in indexical expression of event participants in Mukri and throughout CK.¹¹ These absolute prepositions are related to a sub-class of simple prepositions. Both types are presented in Table 2.

⁹ Note that non-clitic PM expression of the R requires a complete restructuring of the sentence. Furthermore, it will be discussed below that sometimes clitic PM expression of the R is the only possible strategy for introducing an R, hence clitic expression can be obligatory. Nevertheless, the generalization above captures the principal tendency of alignment in the language.

¹⁰ See MacKenzie (1964), Sims-Williams (2012: 34–35) and Korn and Nourzai (2011) for examples of this peculiarity in other Iranian languages.

¹¹ However, in addition to a further Mukri-specific *wē* ‘to’ and other formal peculiarities, the adpositional behaviour in Mukri diverges from Standard CK or other CK varieties in some respects.

SIMPLE ADP	ABSOLUTE ADP	
<i>le</i>	<i>lê</i>	‘from’ ‘at’
<i>be</i>	<i>pê</i>	‘to’
-	<i>wê</i>	‘to’
<i>bo</i>	<i>bo</i>	‘to’ ‘for’
<i>de</i>	<i>tê</i>	‘in’ ‘into’
<i>degel</i>	<i>degel</i>	‘with’

Table 2. Two sub-classes of adpositions in Mukri Kurdish

The simple adpositions have free pronoun or NP complements (4a), whereas absolute adpositions can have only bound complements: clitic PMs in the present (4b) and non-locally realized verbal affixes in the past tenses (4c).

- (4) a. *ewrô bo to=m hênā-w-e*
today for you=1SG:A bring.PST-PTCP-COP.3SG
Today (I) have brought (them) to you. [KF.043]
- b. *bo=mān de-bîn-î-ewe ĉāk-e*
to=1PL IND-see.PRS-2SG-PTCL nice-COP.3SG
(You) had better find (it) for us. [MK.168]
- c. *ew beserhāt-e=î bo gērā-m-ewe*
DEM story-DEM2=3SG:A to narrate.PST-1SG:R-PTCL
(He) narrated this story to me. [ŽB.005]

However, the difference between the two sets of adpositions is not only a matter of the distribution and lexical class of their complements. For instance, in (5a), where the absolute adposition *lê* enables a malefactive reading of its complement, an alternative free form expression of the R on the corresponding simple preposition is not possible, as shown in (5b). Even in cases where the free form complement is allowed, which are quite frequent, there is often a semantic distinction between a construction with a bound complement and one with a free form complement. Thus, the absolute adpositions will be discussed independently of simple prepositions, although they are clearly related to the latter.

In past tense transitive constructions, in line with the argument indexing principles sketched in Section 2, the complement of an absolute adposition expressing the R is not a clitic PM, but a verbal affix realized on the verb stem. Thus, the sentence (5c), where, in addition to the A, the R is also a clitic PM, is clearly ungrammatical in Mukri.¹²

- (5) a. *pišîle-ke=yān [lê kušt-im]*
cat-DEF=3PL:A from kill.PST-1SG:R
They killed my cat on me (i.e. they killed my cat, and I perceived this as indirectly affecting me in a negative manner). [FN.10]

¹² See Edmonds (1955: 501) for an example from Suleimani Kurdish, where in addition to the A, the R is also expressed by a clitic PM.

- b. **pišīle-ke=yān* [le min kušt]
 cat-DEF=3PL:A from 1SG:R kill.PST
 They killed my cat on me.
- c. **pišīle-ke=yān* *lē=m* *kušt*
 cat-DEF=3PL:A from=1SG:R kill.PST
 They killed my cat on me.

As for the syntactic status of the complements of absolute adpositions, in some constructions the verb clearly subcategorizes for the third participant, whereas in some others the verb lends itself both to two-place and three-place predication. For instance, the verbs *hestāndin* ‘to take’ (6a), *hestāndinewe* ‘to take back’, *čākī kirdin* ‘do (someone a) favour’ (6c) are always used with an R flagged by an adposition. Similarly, some adpositions have been lexicalized with the verb stem to form a meaning distinct from that of the base verb, such as *lē-dān* (ADP-give) ‘to hit’. With those verbs, the adpositional complement is obligatory. However, some other verbs, such as *kutin* ‘to tell’, *gērānewe* ‘narrate’ and *froštīn* ‘to sell’, can be used without an overt R; usually the discourse provides the knowledge of participants. The third participant is added by the incorporation of adpositions to adjust the meaning of the construction.

3.1.2. Semantic roles of oblique arguments and polysemy

The adpositions usually add a certain semantic role interpretation to the adpositional complement, sometimes non-identifiable from the habitual meaning of the adposition. Thus, in (4c) the 1SG verbal affix marks the adpositional complement in an addressee or recipient role. In (6a) the adpositional complement has a source role that easily lends itself to a malefactive reading. The same adposition *lē* can introduce a participant that is obligatorily interpreted as malefactive (5a). A recipient or benefactive reading of the adpositional complement in (6b) is possible. Finally in (6c), where the comitative adposition *degel* ‘with’ is used, only a benefactive reading of the adpositional object is possible. Now the benefactive reading is not encoded in the meaning of the adposition, but rather obtained through the specific construction it is used in.

- (6) a. *āxā zewī-ye-ke=y lē hestānd-im*
 agha field-DEM2-DEF=3SG:A from take.by.force.PST-1SG:R
 The seigneur took the/my field from me by force. [FN.09]
- b. *ew kārāsk-e=m āwā pē frošt-ūw-in*
 DEM gazelle-DEM2=1SG:A thus with sell.PST-PTCP-2/3PL:R
 That’s how I have sold this gazelle to them. [EP.667]
- c. *ege ew čākī-ye=y degel kird-bū-n*
 when DEM favour-DEF=3SG:A with do.PST-COP.PST-2/3PL:R
 As he had done this favour for them. [EP.324]

A survey of similar constructions in the corpus showed that there are some recurrent adposition-semantic role associations. These are presented in Table 3.

FORMAL DEVICE		SEMANTIC ROLE
ADPOSITION	<i>lē</i> ‘from’	SOURCE/MALEFACTIVE/POSSESSOR
	<i>pē</i> ‘to’	ADDRESSEE/RECIPIENT/CAUSEE (cf. Section 3.2)
	<i>wē</i> ‘to’	RECIPIENT
	<i>bo</i> ‘for’	BENEFACTIVE
	<i>degeṭ</i> ‘with’	BENEFACTIVE

Table 3. Semantic roles associated with the complements of absolute adpositions

There is an apparent polysemy between possessor on the one hand and benefactive and malefactive interpretations of adpositional complements on the other. In (5a) the adpositional complement in a malefactive role is necessarily the possessor of the direct object. Yet, the polysemy is especially evident when used in intransitive constructions as in (7). The adpositional complement in (7b) is both the possessor and the negatively influenced participant. The malefactive reading is clearly the semantic contribution of the adposition. In (7a), without the adposition, the NP is S and the 1SG PM is an adnominal possessor; in (7b), with the incorporation of the adposition into the argument structure of the verb *sūtān* ‘to burn’, the erstwhile possessor expresses also the malefactive role. Note that (7b) is not a three-participant event, however the construction neatly illustrates the semantic role assigning feature of absolute adpositions.

- (7) a. *māl-eke=m* *sūtā-ø*
house-DEF=1SG:POSS burn.PST-3SG
My house burnt down.
- b. *māl-eke=m* *lē* *sūtā-ø*
house-DEF=1SG:POSS/R from burn.PST-3SG
My house burnt on me. [FN.17]

3.2. Periphrastic causative construction

There are two processes of causativization in Mukri that serve as valency increasing mechanisms. The morphological causative deploys a causative/transitivizing suffix *-ēn* in the present tense and *-ānd* in the past tense. Thus, the S of the intransitive construction in (8a) is the P of the monotransitive verb in (8b).

- (8) a. *hetā* *žin-e* *bi-tirs-ē*
until woman-DEM2 IRR-fear.PRS-3SG:S
so that the woman would be scared. [GD.91]
- b. *bi-č-in* *lewē* *b=ī-tirs-ēn-in*
IRR-go.PRS-2PL there IRR=3SG:P-fear.PRS-CAUS-2PL
You (pl.) go (and) scare her there. [GD.81]

In Mukri, and in the rest of Kurdish, the morphological causative operates only on intransitive bases, rendering them transitive and introducing an A into the clause. However, there is a further periphrastic causative construction in Mukri that oper-

ates on transitive bases. It deploys the absolute adposition *pē* ‘to, with’. Here, the adposition introduces a causee into the event, while the verb itself remains morphologically unchanged. This is illustrated in (9).¹³

- (9) a. *pē=yān* *de-be-m-e* *memleket=ī* *xo=mān*
 to=3PL:R IND-take.PRS-1SG-DRCT country=EZ REFL=1PL
 I will make them take (the treasure) to our country. [MK.326]
- b. *xānū-eke=y* *pē* *sāz* *kird-im*
 house-DEF=3SG:A to NVP.build do.PST-1SG:R
 He made me build the/his house. [FN.12]

In accordance with the cross-linguistic tendency, the morphological causative (8b) illustrates a direct causation, whereas the periphrastic causative (9) expresses indirect causation, i.e. one requiring a special effort on the part of a causer and, possibly, involving unwillingness on the part of the causee (cf. Aikhenvald 2000: 165–166).

The expression of the causee adheres to the same structural principles that were outlined above for the R of a ditransitive construction. In the present tense, the causee remains as a clitic attaching to the preposition (9a) or to a preceding element, while in the past transitive constructions (9b), the marker of the causee shifts from the preposition and is expressed through a verbal affix. It is of course cross-linguistically common for causees to be expressed through dative-type markers (cf. Shibatani and Yeo Chung 2002), so the formal identity with the R-marker is not surprising.

4. Periphrastic potentiality expression

In Mukri, and in the rest of CK, there is a lexical modal expression of potentiality through the verb *twānīn* ‘to be able to’ (10). This is also the pattern typically found for West Iranian languages (cf. Mahmoudveysi and Haig 2009).

- (10) *ne=t-de-twānī* *dest=ī* *lē* *de-y*
 NEG=2SG:A-IPFV-be.able.PST hand=3SG:R at give.PRS-2SG
 You would not be able to touch her. [ŽB.223]

In Mukri, however, a further potentiality expression is seen in which the adpositions *pē* and *bo* are used. The verb of the construction can be a transitive or an intransitive one. However, if it is transitive, the verb is inflected for passive, carrying the suffixes *-rē* with present and *-rā* with past stems. The construction thus requires an intransitive base.

In (11) a usual passive sentence is given, without the *pē* adposition. The syntactic subject is the only argument and the verb is marked for passive. In (12)–(14), however, the adposition *pē* is incorporated into the structure and a further partici-

¹³ Note that this is radically different from the corresponding Kurmanji periphrastic causative, which also incorporates a similar adposition, albeit with different properties, and which utilizes an auxiliary verb coupled with an infinitive (cf. Haig and Öpengin forthcoming).

pant is introduced as a clitic PM adpositional complement. The construction is thus again valency-increasing and it clearly denotes potentiality.¹⁴ The adpositional complement, the underlying subject, lands on the leftmost constituent of the VP. In (12) and (14), the initial NPs must be interpreted as clause-external focussed elements.

- (11) *ew-e* *be* *šīr=ī* *gāmēš-ī* *bexēw*
 DEM-DEM2 with milk=EZ buffalo-OBL NVP.feed
k-rā-w-e
 do.PRS-PASS-PTCP-COP.3SG
 This (mare) has been fed with water-buffalo milk. [NZ.117]

- (12) *ew* *dū-wan-e=š* *čī=yān* *pē* *de-k-rē*
 DEM TWO-PL-DEM2=ADD what=3PL:R to IND-do.PRS-PASS
 As for these two, what can they do! (lit. What can be done to them?) [EP.630]

- (13) *pē=m* *berhelā* *nā-k-rē-n*
 to=1SG:R idle NEG-do.PRS-PASS-3PL
 I cannot give freedom to them. [FN.15]

- (14) *selte=y* *qočāx* *wekū* *māšīn=ī* *elfān=ī* *pē* *de-kir-rā*
 single=EZ potent like car=EZ now=3SG:R to IND-do.PST-PASS

A potent single (man), he was able to do what a car does today. [QA.35]

Now the argument introduced by the adpositions in this construction is not a “third” participant, but rather a non-canonically marked subject,¹⁵ though it is introduced in a manner parallel to the third participants, i.e. through the oblique strategy.¹⁶

The adposition *bo* is usually employed with an intransitive verb and on an active verb form. In the main clause in (15), the underlying subject is the complement of the adposition *bo*, the sentence is in the past tense, and the verb is inflected in past imperfective. There are also rare cases where the adposition *bo*, like *pē*, is used with a transitive verb carrying passive inflection (16).

- (15) *herčend* *tē=y* *de-nūsā-n*
 no.matter.how.much to=3SG:R IPFV-stick.PST-3PL
bo=yān *ne-de-hāt-ewe* *der-ē*
 to=3PL:R NEG-IPFV-come.PST-PTCL out-OBL
 No matter how long/hard they stuck to it (sword), it would not come out.
 (lit. (...), to them it (the sword) would not come out.) [EP.644]

¹⁴ Note that this is remarkably different from periphrastic passive agents introduced within a PP such as *le lāyen ... =ewe* ‘from the side of ...’ (similar to English by-phrase agents). The latter is by no means obligatory; it does not denote any sense of potentiality and it is almost exclusively found in the written language.

¹⁵ See Öpengin (In prep.) for comprehensive treatment of similar constructions in Mukri Kurdish. See Haig (2008: 257–263) and Jahani *et al.* (2012) for non-canonical subjects in Badinani Kurdish and Balochi, respectively, where case marking is responsible for the particular treatment of subjects in a subset of related constructions.

¹⁶ That is why the same gloss (R) has been used for both of the argument types.

- (16) *ew suħāl-e=m bo juwāb nā-de-rē-t-ewe*
 DEM question-DEM2=1SG:R to answer NEG-give.PRS-PASS-3SG-PTCL
 I cannot answer this question. (lit. This question can't be answered to me.)
 [KF.103]

The argument introduced through *pē* and *bo* adpositions in periphrastic potentiality and causative constructions have certain similarities. In both constructions the adposition introduces an additional agentive participant into the clause, but the new argument is by no means a prototypical agent; it is a “weakened actor” in the potentiality construction in the sense that the action is merely potential (not actual), whereas in the causative construction it is a “causee” destined to carry out an action under duress. Hence, the argument introduced by the adposition *pē* cannot readily be identified through any of the conventional semantic roles, but fits well with Haig’s (2008: 58) notion of an “Indirect Participant” which can be understood as a discourse participant who lacks direct control over the event but whose interests are affected by the given event.

5. Conclusion

In this paper I have summarized the principles of argument indexing, focusing on three-participant events, so as to capture the alignment properties of a variety of constructions that deploy so-called absolute adpositions in Mukri Central Kurdish. I have then analysed the ways in which absolute adpositions are employed to express three-participant events. The semantico-syntactic contribution of the adpositions to the argument structure and semantics of the language was examined. Finally, two periphrastic constructions (causative and potentiality) were analysed in which two of the absolute adpositions introduce types of indirect participants. This paper has thus tried to contextualize and elaborate upon the extent of the adpositional functions in interaction with the argument-indexing system in a variety of Central Kurdish.

Source texts used in the examples

(Full details of the corpus are available in Öpengin (in prep.))

NN	Text type	Male/female	Age
NZ	folktale	M	73
EP	folktale	M	73
QA	autobiographical narrative	M	73
	regional anecdote		
ŽB	regional anecdote	M	32
GD	folktale	M	32
KF	folktale	M	32
MK	folktale	M	60

Abbreviations

A	agent-like argument of a transitive verb
ADD	additive clitic = <i>(i)ş</i>
ADP	adposition
CAUS	causative
COP	copular ending
DEF	definite
DEM	demonstrative <i>ew</i>
DEM2	demonstrative particle = <i>e</i>
DRCT	directional particle = <i>e</i>
EZ	ezafe particle / linker
FN	field notes
IND	indicative
INDF	indefinite
IPFV	imperfective
IRR	irrealis
NEG	negative prefix
NVP	non-verbal element in a complex predicate
OBL	oblique case
P	patient-like argument of a transitive verb
PASS	passive
PL	plural
PM	person marker
PRS	present
PST	past
PTCL	verbal particle - <i>ewe</i>
PTCP	past participle
PVB	preverb
R	recipient-like argument in a three-participant event
REFL	reflexive pronoun
S	the only argument of an intransitive verb
SG	singular
T	theme argument in a three-participant event
TAM	tense, aspect, modality marker

References

- Aikhenvald, Alexandra Y. (2000). "Transitivity in Tariana." In: Dixon, R. M. W. and Aikhenvald, A. Y. (eds), *Changing Valency: Case Studies in Transitivity*. Cambridge: Cambridge University Press, pp. 145–172.
- Comrie, Bernard (1989). *Language Universals and Linguistic Typology*. Chicago: University of Chicago Press.
- Comrie, Bernard (1993). "Argument Structure." In: Jacobs, J., von Stechow, A., Sternefeld, W., and Vennemann, T. (eds), *Syntax: An International Handbook of Contemporary Research*. Berlin: Walter de Gruyter, pp. 905–914.
- Edmonds, C. J. (1955). "Prepositions and Personal Affixes in Southern Kurdish." *Bulletin of the School of Oriental and African Studies* 17(3), pp. 490–502.
- Haig, Geoffrey (2008). *Alignment Change in Iranian Languages: A Construction Grammar Approach*. Berlin: Mouton de Gruyter.
- Haig, Geoffrey, and Öpengin, Ergin (forthcoming). "Kurdish in Turkey: An Overview of Grammar, Dialectal Variation and Status." In: Bulut, Christiane (ed.), *Minority Languages in Turkey* (special edition of *Turkologica*). Wiesbaden: Harrassowitz.

- Haspelmath, Martin (2012). "Argument Indexing: A Conceptual Framework for the Syntax of Bound Person Forms." Ms., Max Planck Institute for Evolutionary Anthropology. Online at: http://www.academia.edu/1208796/Argument_indexing_A_conceptual_framework_for_the_syntactic_status_of_bound_person_forms [Retrieved Nov. 15, 2012].
- Jahani, Carina, Barjasteh Delforooz, Behrooz, and Nourzaei, Maryam (2012). "Non-canonical Subjects in Balochi." In: Aghaei, B., and Ghanoonparvar, M.R. (eds), *Iranian Languages and Culture. Essays in Honor of Gernot Ludwig Windfuhr*. Costa Mesa: Mazda Publishers, pp. 196–218.
- Jügel, Thomas (2009). "Ergative Remnants in Sorani Kurdish?" *Orientalia Suecana* 58, pp. 142–158.
- Korn, Agnes (2009). "Western Iranian Pronominal Clitics." *Orientalia Suecana* 58, pp. 159–171.
- Korn, Agnes, and Nourzaei, Maryam (2011). "Those Were the Hungry Years: A First Glimpse at Coastal Balochi as Spoken by Blacks in Iran." Paper presented at ICIL 4, Uppsala University, 17–19 June.
- MacKenzie, David N. (1961). *Kurdish Dialect Studies, Vol. I*. London: Oxford University Press.
- MacKenzie, David N. (1964). "The 'Indirect Affectee' in Pahlavi." In: Jamaspasa, Kaikhusroo (ed.), *Dr. J.M. Unvala Memorial Volume*, pp. 45–48. Bombay: Kanga.
- Mahmoudveysi, Parwin, and Haig, Geoffrey (2009). "The Typology of Modality in Some West Iranian languages." In Allison, C., Joisten-Pruschke, A., and Wendtland, A. (eds), *From Daena to Din. Religion, Kultur und Sprache in der iranischen Welt*. Wiesbaden: Harrassowitz, pp. 41–52.
- Margetts, Anna, and Austin, Peter (2007). "Three-participant Events in the Languages of the World: Towards a Crosslinguistic Typology." *Linguistics* 45(3). 391–451.
- Margetts, Anna (2007). "Three-participant Events in Oceanic Languages." *Oceanic Linguistics* 46, pp. 71–127.
- Malchukov, Andrej, Haspelmath, Martin, and Comrie, Bernard (2010). "Ditransitive Constructions: A Typological Overview." In: Malchukov, A., Haspelmath, M., and Comrie, B. (eds), *Handbook of Ditransitive Constructions*. Berlin: Mouton de Gruyter, pp. 1–64.
- Öpengin, Ergin (in preparation). *Clitic/affix Interactions: A Corpus-based Study of Person Marking in Mukri Central Kurdish*. PhD Thesis: Sorbonne Nouvelle and University of Bamberg.
- Samvelian, Pollet (2007). "What Sorani Kurdish Absolute Prepositions Tell us about Cliticization." In: Hoyt, F., Seifert, N., Teodorescu, A., and White, J. (eds), *Texas Linguistics Society IX: The Morphosyntax of Understudied Languages*. Stanford: CSLI publications, pp. 263–283. Online at: http://csli-publications.stanford.edu/TLS/TLS9-2005/TLS9_Samvelian_Pollet.pdf [Retrieved Nov. 20, 2012].
- Shibatani, Masayoshi, and Chung, Sung Yeo (2002). "Japanese and Korean Causatives Revisited." *Japanese/Korean Linguistics* 10, pp. 32–49.
- Siewierska, Anna (2004). *Person*. Cambridge: Cambridge University Press.
- Sims-Williams, Nicholas (2011). "Differential Object Marking in Bactrian." In: Korn, A., Haig, G., Kari-mi, S., and Samvelian, P. (eds), *Topics in Iranian Linguistics*. Wiesbaden: Reichert, pp. 23–38.

Past and Non-past Structures in the Mazandarani Dialect Spoken by the Galesh of Ziarat

Guiti Shokri, Uppsala University
g_shokri@hotmail.com

Abstract

This article deals with the Mazandarani dialect spoken in Ziarat, a village near Gorgan in Golestan Province, Iran. This dialect is spoken by the Galesh cow herders in this area and is to be regarded as a variant of Mazandarani, which is spoken in Mazandaran Province and neighbouring regions south of the Caspian Sea in Iran.

I give a brief description of the past and non-past² verb forms in the Ziarat dialect and show the formation of the present/future, past, and past imperfect indicative verbs in this dialect. I conclude that the formation of the non-past indicative is influenced by Persian, whereas the past indicative and past imperfect indicative follow the Mazandarani structure. However, when the past imperfect has a derivational preverb, it uses a Persian structure with a Mazandarani stem. All personal endings, except the 3SG past ending, are copied from Persian.

On the whole, this hybrid dialect can basically be described as a dialect of Mazandarani, but it is affected by Standard Persian, Gorgani Persian³ and Khorasani Persian depending on the geographic situation.

Keywords: Mazandarani, Galesh, Ziarat dialect, Sari dialect, Persian, language contact, verb forms

1. Introduction¹

Ziarat is a village situated about 20 kilometres south of the town of Gorgan in Golestan Province, Iran. Golestan Province was called Estarabad and Gorgan in the past, and covers an area of 2,000,000 hectares east of Mazandaran and west of Khorasan. It has a population of approximately 1,650,000 (*Sāzmān-e modiriat va barnāmerizi* 1386). The presence of forests and good pastures enables the people living in this area to have an economy based on cattle breeding and agriculture.

The combination of ethnic groups that make up the population in Golestan is very interesting. The population includes the following groups: Fars, Turkmen, Sistani, Baluch, Kurd, Kazak (*Sāzmān-e modiriat va barnāmerizi* 1386). This situation creates a multi-ethnic population and thus provides an excellent opportunity for studying language contact and contact-induced linguistic phenomena.

The text corpus from Ziarat comes from a joint project on the Mazandarani spoken by the Galesh of Ziarat carried out by Carina Jahani, Uppsala University, Hossein Barani, Gorgan University of Agriculture, and myself. The study was funded by the Hans Rausing Endangered Languages project.⁴ After an initial journey to estab-

¹ Sincere thanks to Geoffrey Haig, Carina Jahani, and the referees for useful comments on earlier versions of this article.

² Following Comrie (1985: 48–49), the term non-past is here used to denote present/future. The tense system in Mazandarani, as well as in most other Iranian languages, a binary past versus non-past system.

³ It is interesting to note that Gorgani Persian, spoken in the city of Gorgan, is surrounded by Mazandarani dialect in several directions.

lish local contacts, carried out by Carina Jahani and Guiti Shokri in November 2006, we documented the language, ceremonies, lifestyle, and customs of the Galesh in Ziarat during four subsequent field journeys; in October–November 2008, April–May 2009, January–February 2010, and May–June 2011. The field journeys were carried out by Guiti Shokri and Hosein Barani, and the transcription and translation of the interviews was carried out by Guiti Shokri and Carina Jahani. Previous ethnological and linguistic research on the Galesh has been carried out by Purkarim (1357) Borjian and Borjian (2008) among others.

In 1385 A.H. (2006/07 A.D.), just before our first field journey, about 2,000 people lived in Ziarat village.⁵ The people who live in Ziarat and work as cattle breeders are called Galesh. In Mazandaran, Gorgan, and Ziarat, Galesh is the term for someone who breeds cows. This word is used in Gilan as well, but there it refers to people who breed both cows and sheep. There is thus a difference in the usage of the term Galesh between Gilan and Mazandaran (Ranjbar and Rādmard 1382: 14).

Because of its dry and cool weather, Ziarat is also a popular place for villas and summer houses. Nowadays, due to its proximity to Gorgan, Ziarat is growing and the village is going through a rapid modernization, which in the near future could cause the loss of all its rural and cultural characteristics, including its dialect. We chose this area for our work for the above reasons and also because of the people and their specialized work, which is on the verge of vanishing due to modernization and changing lifestyle and methods of production.

In this article, I focus on the structure of the verb in the dialect of Ziarat. First, I will present some examples of past and the non-past verb forms in the Mazandarani dialect of Sari, then I will give examples from the dialect of Ziarat.

2. Presentation of the data⁶

2.1. *Non-past indicative*

In the Mazandarani of Sari, the non-past indicative is formed by the non-past stem⁷ + non-past personal endings (see also Shokri 1374: 107–110; Yoshie 1996: 33–34).

⁴ The outcome of this study is deposited at: <http://elar.soas.ac.uk/deposit/jahani2010galesh>

⁵ <http://ziaratvillage.blogfa.com>

⁶ The personal endings shown in the tables below are the most common endings in each dialect. There are variants, but since the focus of this article is not on variants but on the structure of the verb forms, only the most common ending for each person has been included in each table. Other variants are found in the examples.

⁷ In the discussions of Mazandarani grammar, there has so far been no discussion of whether there is an augment element of the present stem with *-en-* (assimilated to *-em-* before *m* with secondary dissimilation of *mm > mb*) before the personal endings. Yoshie (1996: 33) describes the longer endings as “variation of the general personal ending” with stress on the first syllable of the personal ending. In view of the fact that *-en/-em-* takes stress, they may be re-analysed as belonging to the stem, in the same way that the *-en-* is analysed in Zazaki (see Paul 1998: 75–80). This discussion is, however, beyond the scope of the present article, and I follow the analysis of Mazandarani adopted so far for stems and endings (Shokri 1369: 217–231; Yoshie 1996: 27–34; Borjian and Borjian 2007: 205).

	Verb form	Non-past stem	Non-past personal endings
1SG	<i>rušembe</i>	<i>ruš</i>	<i>-embe</i>
2SG	<i>rušeni</i>	<i>ruš</i>	<i>-eni</i>
3SG	<i>rušene</i>	<i>ruš</i>	<i>-ene</i>
1PL	<i>rušembi</i>	<i>ruš</i>	<i>-embi</i>
2PL	<i>rušenni</i>	<i>ruš</i>	<i>-enni</i>
3PL	<i>rušenne</i>	<i>ruš</i>	<i>-enne</i>

Table 1. Non-past indicative in Mazandarani, dialect of Sari, for *baruten* ‘to sell’

In Ex. (1), the verb *xorne* ‘he eats’ follows the pattern outlined in Table 1. Note that the ending here is *-ne* rather than *-ene* (see also footnote 6).

- (1) *me piyer še yezā=re še xor-ne*
 PRON.1SG.GEN father REFL food=ACC REFL eat.PRS-3SG
 My father eats his food himself. [Sari]

In the dialect of Ziarat, on the other hand, the non-past indicative is mainly formed by the verb prefix *me-* + non-past stem + personal endings,⁸ or, occasionally, like the form found in Sari, without the verb prefix. Note also that the personal endings are copied from spoken Persian.⁹

	Verb form	Prefix	Non-past stem	Non-past personal endings
1SG	<i>merušem</i>	<i>me-</i>	<i>ruš</i>	<i>-em</i>
2SG	<i>meruši</i>	<i>me-</i>	<i>ruš</i>	<i>-i</i>
3SG	<i>meruše</i>	<i>me-</i>	<i>ruš</i>	<i>-e</i>
1PL	<i>merušim</i>	<i>me-</i>	<i>ruš</i>	<i>-im</i>
2PL	<i>merušin</i>	<i>me-</i>	<i>ruš</i>	<i>-in</i>
3PL	<i>merušen</i>	<i>me-</i>	<i>ruš</i>	<i>-en</i>

Table 2. Non-past indicative in Mazandarani, dialect of Ziarat, for *baruten* ‘to sell’

In Ex. (2) the form *mepejen* ‘they cook’ comes with the non-past stem of Mazandarani but with a Persian non-past indicative prefix and personal endings copied from Persian (cf. Persian *mipazan*).

- (2) *alān berenj čejur me-pej-en*
 now rice how PREF-cook.PRS-3PL
 How do they cook rice now? [Ziarat]

In Ex. (3) we have a non-past indicative verb, *nembue* ‘it doesn’t become’ with a Persian structure and a Mazandarani stem. The forms *davui* ‘you should be’ and *danibuin* ‘(if) they are not’ are subjunctive forms and are not treated in this paper. The last verb *nemekārim* ‘we don’t plant’ is a non-past indicative form with the Persian structure and stem described below. Note also the two pronunciation variants *šuppā* and *šupā* ‘night watchman’.

⁸ Note that there is no difference between the non-past and past endings in the dialect of Ziarat. The dialect of Sari makes such a distinction.

⁹ The personal endings are *-am*, *-i*, *-e*, *-im*, *-in*, *-an* in spoken Persian.

- | | | | | |
|-----|-------------------------|---------------------|----------------|-----------------|
| (3) | <i>ne-m-bu-e</i> | <i>bāyd</i> | <i>šuppā</i> | <i>da-vu-i</i> |
| | NEG-PREF-become.PRS-3SG | should | night.watchman | PREV-be.PRS-2SG |
| | <i>šuppā</i> | <i>da-ni-bu-in</i> | | <i>ke</i> |
| | night.watchman | PREV-NEG-be.PRS-2PL | | EMPH |
| | <i>ne-me-kār-im</i> | | | |
| | NEG-PREF-sow.PRS-1PL | | | |
- It doesn't work (lit. it doesn't become). There should be night watchmen, if there are no watchmen, we don't plant, you know. [Ziarat]

Also in (4), the verbs *megiren* ‘they get’ and *mezenen* ‘they hit’ are modelled on the Persian structure.

- | | | | | |
|-----|---|-------------------------|--|-------------------------|
| (4) | <i>sad-tā</i>
hundred-CL | <i>šomāre</i>
number | <i>me-gir-en</i>
PREFIX-get.PRS-3PL | <i>kāyaz</i>
paper |
| | <i>me-gir-en</i>
PREFIX-get.PRS-3PL | <i>un</i>
DEM.DIST | <i>sad-tā</i>
hundred-CL | <i>angušt</i>
finger |
| | <i>me-zen-en</i>
PREFIX-hit.PRS-3PL | | | |
| | They take a hundred numbers, they get documents, they take a hundred fingerprints. [Ziarat] | | | |

The only occurrence of a non-past indicative of the Sari Mazandarani type, i.e. lacking the prefix *m(e)-*, in the whole of the data (comprising about three hours of speech) is the verb *guim* ‘we say’ in Ex. (5). Even here, the verb ending is copied from Persian.

- (5) *emā* *gu-im* *nāmzed*
we say.PRS-1PL engaged
We call it *nāmzed* (i.e. engaged). [Ziarat]

2.2. Simple past indicative

The structure of the simple past indicative in the Sari dialect of Mazandarani comprises a prefix¹⁰ + past stem + past personal endings (see also Shokri 1995: 112–113; Yoshie 1996: 34–35).¹¹

	Verb form	Prefix	Past stem	Past personal ending
1SG	<i>baruteme</i>	<i>ba-</i>	<i>rut</i>	<i>-eme</i>
2SG	<i>baruti</i>	<i>ba-</i>	<i>rut</i>	<i>-i</i>
3SG	<i>barute</i>	<i>ba-</i>	<i>rut</i>	<i>-e</i>
1PL	<i>barutemi</i>	<i>ba-</i>	<i>rut</i>	<i>-emi</i>
2PL	<i>baruteni</i>	<i>ba-</i>	<i>rut</i>	<i>-eni</i>
3PL	<i>barutene</i>	<i>ba-</i>	<i>rut</i>	<i>-ene</i>

Table 3. Simple past indicative in Mazandarani, dialect of Sari, for *baruten* ‘to sell’

¹⁰ This prefix is only found in the positive form, in the negative form it is replaced by the negative prefix *na-*. If the verb has a derivational preverb, there is no prefix in the simple past.

¹¹ The marking in the past system of Mazandarani works in quite the opposite way as in Persian, and several other Western Iranian languages (for Balochi, see Nourzaei and Jahani, this volume). The simple past form and past participle are marked, i.e. take an additional prefix, while the imperfective is unmarked.

The verb *baxrime* ‘I bought’ in Ex. (6) follows the structure outlined in Table 3. Note that the ending is *-me* rather than *-eme* in this example (see also footnote 6).

- (6) *men* *in* *ketāb=re* *diruz* *ba-xri-me*
 PRON.1SG DEM.PROX book =ACC yesterday PREF-buy.PST-1SG
 I bought this book yesterday. [Sari]

In the dialect of Ziarat, the same pattern as in the dialect of Sari is found in the simple past indicative, but, again, the endings are copied from spoken Persian, except the ending of 3SG, which is *-Ø* in Persian. Here the dialect of Ziarat retains the Mazandarani ending *-e*.¹²

	Verb form	Prefix	Past stem	Past personal endings
1SG	<i>barutem</i>	<i>ba-</i>	<i>rut</i>	<i>-em</i>
2SG	<i>baruti</i>	<i>ba-</i>	<i>rut</i>	<i>-i</i>
3SG	<i>barute</i>	<i>ba-</i>	<i>rut</i>	<i>-e</i>
1PL	<i>barutim</i>	<i>ba-</i>	<i>rut</i>	<i>-im</i>
2PL	<i>barutin</i>	<i>ba-</i>	<i>rut</i>	<i>-in</i>
3PL	<i>baruten</i>	<i>ba-</i>	<i>rut</i>	<i>-en</i>

Table 4. Simple past indicative in Mazandarani, dialect of Ziarat, for *baruten* ‘to sell’

In Ex. (7) we have three simple past verbs with a Mazandarani structure, but with the personal endings copied from Persian (*baruten* ‘they sold’, *narutim* ‘we didn’t sell, *baxerien* ‘they bought’).

- (7) *ande* *mardem* *zamin* *ba-rut-en*
 so.many people land PREF-sell.PST-3PL
 emā *na-rut-im* *hečči* *hanu*
 PRON.1PL NEG-sell.PST-1PL nothing even
 bazi *az* *pesar-ā=m* *ba-xeri-en*
 some from son-PL=PC.1SG PREF-buy.PST-3PL
 So many people sold land, but we did not sell. Anyway! Some of my sons actually bought (the share of those who wanted to sell). [Ziarat]

In Ex. (8), there is one simple past verb, with a Mazandarani structure (*bāgutem* ‘I said’). Note the variant *bā-* of the prefix found in this example.

- (8) *ham=un* *avval* *vače=re* *bā-gut-em* *dige*
 EMPH=DEM.DIST first child=ACC PREF-tell.PST-1SG other
 I told (you) about the first child, you know. [Ziarat]

2.3. Past imperfect indicative

The structure of the past imperfect indicative in the Sari dialect of Mazandarani comprises a past stem + past personal endings (see also Shokri 1995: 113–115; Yoshie 1996: 35–36).

¹² All the other persons have the same ending in the non-past and past in spoken Persian, see footnote 9.

	Verb form	Past stem	Past personal endings
1SG	<i>ruteme</i>	<i>rut</i>	<i>-eme</i>
2SG	<i>ruti</i>	<i>rut</i>	<i>-i</i>
3SG	<i>rute</i>	<i>rut</i>	<i>-e</i>
1PL	<i>rutemi</i>	<i>rut</i>	<i>-emi</i>
2PL	<i>ruteni</i>	<i>rut</i>	<i>-eni</i>
3PL	<i>rutene</i>	<i>rut</i>	<i>-ene</i>

Table 5. Past imperfect indicative in Mazandarani, dialect of Sari, for *baruten* ‘to sell’

In Ex. (9), the verb form *bāzi kārde* ‘we played’ follows the structure outlined in Table 5.

- (9) *emā vače bi-mi xale bāzi kārde-mi*
 we child be.PST-1PL very playing do.PST-1PL
 (When) we were children, we played very much. [Sari]

The past imperfect indicative in the dialect of Ziarat is formed on the same pattern, but, again, with endings modelled on Persian except for the 3SG ending (see above).

	Verb form	Past stem	Past personal endings
1SG	<i>rutem</i>	<i>rut</i>	<i>-em</i>
2SG	<i>ruti</i>	<i>rut</i>	<i>-i</i>
3SG	<i>rute</i>	<i>rut</i>	<i>-e</i>
1PL	<i>rutim</i>	<i>rut</i>	<i>-im</i>
2PL	<i>rutin</i>	<i>rut</i>	<i>-in</i>
3PL	<i>ruten</i>	<i>rut</i>	<i>-en</i>

Table 6. Past imperfect indicative in Mazandarani, dialect of Ziarat, for *baruten* ‘to sell’

Ex. (10) is taken from a text about wedding customs, and the past imperfect verb *zien* ‘they played’ denotes a habit in the past. It follows the Sari pattern, but takes an ending modelled on Persian.

- (10) *are unvax sāz zi-en*
 yes then instrument play.PST-3PL
 Yes, then they played instruments. [Ziarat]

I found three derivational preverbs in the data that are retained in the past imperfect formation in the dialect of Ziarat, namely *de-*, *(h)e-* and *va-*. Here, the past imperfect indicative is formed by the addition of the prefix *-m(e)-* (which here turns into an infix) following the derivational preverb (cf. the Persian prefix *mi-*). This formation is not found in the dialect of Sari.¹³

¹³ According to my analysis of the data, *ba-* with its variants (*bo-*, *bā-*, *be-*, *bey-*, *bi-*) is an inflectional prefix; and *da-* with its variants (*de-*, *dā-*, *day-*), *ha-* with its variants (*hā-*, *he-*, *e-*, *ā-*), *va-* with its variant (*var-*) are derivational preverbs. The prefix does not change the meaning of the verb, whereas the preverb does. However I also believe that *ba-* used to be a derivational preverb, but that it later was re-analysed as a prefix. In the Sari dialect, *ha-* has also lost its role as a derivational preverb and can be regarded as a prefix, but in the Ziarat dialect, there are places where *ha-* adds the meaning ‘on top of’. The preverb *da-* adds the meaning ‘inside, in’, and *va-* ‘up, upwards’ (Shokri 1380).

	Verb form	Derivational preverb	Prefix	Past stem	Personal ending
1SG	<i>demeštem</i>	<i>de-</i>	<i>-m-</i>	<i>ešt</i>	<i>-em</i>
2SG	<i>demešti</i>	<i>de-</i>	<i>-m-</i>	<i>ešt</i>	<i>-i</i>
3SG	<i>demešte</i>	<i>de-</i>	<i>-m-</i>	<i>ešt</i>	<i>-e</i>
1PL	<i>demeštim</i>	<i>de-</i>	<i>-m-</i>	<i>ešt</i>	<i>-im</i>
2PL	<i>demeštin</i>	<i>de-</i>	<i>-m-</i>	<i>ešt</i>	<i>-in</i>
3PL	<i>demešten</i>	<i>de-</i>	<i>-m-</i>	<i>ešt</i>	<i>-en</i>

Table 7. Past imperfect indicative in Mazandarani, dialect of Ziarat, for *ešten* ‘to put’ with the preverb *de-* (*dešten* ‘to put something into’)

I use Ex. (11) to compare the two verbs *demešten* ‘they put’ and *yārden* ‘they brought’, both of which are in the past imperfect but have different structures. The first one has a derivational preverb + the prefix (here infix) *-m-* (the Persian imperfect marker) a Mazandarani stem and the Persian personal endings, while the second one is a past imperfect verb with a Sari structure but with Persian personal endings.

- (11) *ham amšu de-m-ešt-en dele majme yard-en*
 just tonight PREV-PREF-put.PST-3PL inside dish bring.PST-3PL
 On that very night, they put it (i.e. the gift for the bride’s parents) in a dish and brought it. [Ziarat]

In Ex. (12) the verb *hemeštim* ‘we put’ follows the same pattern as *demešten* ‘they put’. This verb can take either preverb, *he-* or *de-*, but with a change in meaning; *he-* denotes putting something onto something, while *de-* denotes putting something into something.

- (12) *zirbayal am ke zirbayali he-m-ešt-im*
 armpit also EMPH armpit.patch PREV-PREF-put.PST-1PL
 Under the arm, well, we put a patch there. [Ziarat]

There is also an additional formation of the past imperfect indicative in the dialect of Ziarat, namely the addition of the prefix *m(e)-* to the form shown in Table 6. This form must be regarded as a copy of the Persian past imperfect indicative (e.g. *mikardam*). This formation is found alongside the form without the prefix *m(e)-*.

	Verb form	Prefix	Past stem	Personal ending
1SG	<i>memeštem</i>	<i>me-</i>	<i>mešt</i>	<i>-em</i>
2SG	<i>memešti</i>	<i>me-</i>	<i>mešt</i>	<i>-i</i>
3SG	<i>memešte</i>	<i>me-</i>	<i>mešt</i>	<i>-e</i>
1PL	<i>memeštim</i>	<i>me-</i>	<i>mešt</i>	<i>-im</i>
2PL	<i>memeštin</i>	<i>me-</i>	<i>mešt</i>	<i>-in</i>
3PL	<i>memešten</i>	<i>me-</i>	<i>mešt</i>	<i>-en</i>

Table 8. Past imperfect indicative in Mazandarani, dialect of Ziarat, for *mešten* ‘to smear’

In Ex. (13) we have several verbs. The form *kand-im* ‘we dug’, which is a past imperfect form with a Mazandarani structure but a Persian personal ending, is found twice. This verb follows the pattern outlined in Table 6. Then we have *meštim* ‘we put’ with the prefix *m-* copied from Persian, a Mazandarani stem, and the Persian personal endings.¹⁴ The third verb is *memeštim* ‘we smeared’, again with the

¹⁴ In other texts, the form *eštim* is also found.

Persian prefix and endings, and a Mazandarani stem. These two verbs follow the pattern outlined in Table 8. The verb *naxize* ‘it should not fall’ is in the subjunctive mood, and is thus outside the scope of this article.

- (13) *are* *kele* *kele* *zamin=de*¹⁵
 yes stove stove earth =ACC
kand-im *zamin=dre*¹⁶ *kand-im* *čāle*
 dig.PST-1PL earth =ACC dig.PST-1PL hole
dey=re *m-ešt-im* *inje* *sang-ā=e*
 pot=ACC PREF-put.PST-1PL here stone-PL=EZ
xurd=re *me-mešt-im* *ke* *na-xiz-e*
 small=ACC PREF-smear.PST-1PL CLM NEG-subj.slide.PRS -3SG
 Yes, we made the stove by digging the earth, we dug a hole in the earth, we
 put the pot there, we attached (lit. smeared) small stones to it¹⁷ so that it (i.e.
 the pot) would not slide down. [Ziarat]

3. Conclusion

Table 9 below is a summary of the structures for the non-past indicative, simple past indicative, and past imperfect indicative in the two Mazandarani dialects of Sari and Ziarat. One of the most important characteristics of the non-past indicative verb forms in the dialect of Ziarat is that it is a hybrid construction which combines Persian indicative prefixes and personal endings with a Mazandarani stem. In the simple past, the dialect of Ziarat follows a Mazandarani structure, apart from the personal endings, which are Persian. The exception is the 3SG ending, where the dialect of Ziarat has the Mazandarani ending. In the past imperfect, a Mazandarani structure is normally used, but when the verb takes a derivational preverb examples have shown that the Persian prefix is added between the preverb and the stem. There are also examples of a structure with the Persian *me-* prefix without a derivational preverb. In all the verb forms found in the Ziarat dialect, the stem is Mazandarani and the personal endings (except in the past tense, 3SG) are copied from Persian. I therefore divide the past imperfect in Mazandarani of Ziarat into three structures in Table 9.

Verb form (all indicative mood)	Maz. preverb/ prefix	Prefix <i>m(e)-</i>	Pe. personal endings	Maz. personal endings
Non-past, Sari				+
Non-past, Ziarat		+	+	
Simple past, Sari	+			+
Simple past, Ziarat	+		+	
Past imperfect, Sari				+
Past imperfect, Ziarat (1)			+	
Past imperfect, Ziarat (2)	+	+	+	
Past imperfect, Ziarat (3)		+	+	

Table 9. Summary of verb prefixes and suffixes in Mazandarani of Sari and Ziarat

¹⁵ The form *=de* is a variant of *=re*, which is here used after *n*, where also the variant *=dre* is found.

¹⁶ See footnote 15.

¹⁷ Here “it” refers to a construction with a couple of big stones, on which the pot was placed. It was important that these stones were even, and therefore they were covered with small stones mixed with clay.

The general conclusion is that the dialect of Ziarat belongs to the Mazandarani language, but is affected by spoken Standard Persian, Gorgani Persian and/or Khorasani Persian depending on migration patterns, contact with the town of Gorgan, etc. A comparison with the dialect of Galesh herdsmen in Kherimesar, in Kordkheyl district, near Sari in Mazandaran, as documented by Borjani and Borjani (2008), shows that the verb system in the dialect of the Galesh of Ziarat is more influenced by Persian than that of the Galesh of Kherimesar. In the dialect described by Borjani and Borjani (2008), the verb forms are mostly identical to those found in the Sari dialect.

It is clear that the use of a Persian-modelled present indicative verb form is not a recent phenomenon which should be attributed to schooling and modern media, since it is found in the speech even of the elderly in the village. The same is true also of the Persian-modelled past imperfect of the 3rd type in Table 9. These forms, which are identical to the form in Persian, are most likely attributable to long and intense contact between Mazandarani and Persian all the way from the Caspian Sea region to Khorasan. This is also in accordance with what Persian sources report about the old lost Gorgani language described by Kiā (1330), which shares many features of structure and vocabulary with Mazandarani.

On the basis of previous research, Bisang (2006: 91–92) finds that the first markers copied from one language into another are discourse markers, and that markers operating on clause-level syntax are also more liable to copying than markers operating on phrase level. The marker copied here (the aspectual/modal marker *mi-*), as well as the personal endings, are operating on verb phrase level and are not the most likely markers to be copied according to these typological constraints. This supports the assumption that there has been intense contact between Persian and Mazandarani in and around Gorgan.¹⁸

We also note that acquisition of the Persian verb prefix has occurred in the places where the Sari dialect of Mazandarani has no prefix. The prefix is fully grammaticalized in the non-past, but it is also found in the past imperfect. In both these instances, the Sari dialect has no prefix. Where the Sari dialect already has a prefix, i.e. in the simple past, this prefix has been retained. In fact, Persian has no alternative prefix to offer, since the Persian simple past is unmarked. It is also interesting to note that the Persian endings have been acquired wherever Persian has an ending, but that the Mazandarani ending has been retained where there is no ending in Persian, namely in the 3SG past tense.

¹⁸ Further studies are needed to determine to what degree markers on higher levels are copied from Persian into the Ziarat dialects.

Abbreviations

-	separates a morpheme
=	separates a clitic
1	first person
2	second person
3	third person
ACC	accusative
CL	classifier
CLM	clause linkage marker
DEM	demonstrative
DIST	distal
EMPH	empathic particle
EZ	ezāfe
Maz.	Mazandarani
NEG	negative
PC	pronominal clitic (enclitic pronoun)
Pe.	Persian
PL	plural
PREF	prefix
PREV	derivational preverb
PRON	pronoun
PROX	proximal
PRS	present/future (i.e. non-past)
PST	past
REFL	reflexive pronoun
SG	singular

References

- Bisang, Walter (2006). "Contact-Induced Convergence: Typology and Areality." In: Brown, Keith (chief ed.), *Encyclopedia of Language and Linguistics*. Amsterdam and Oxford: Elsevier, pp. 88–101.
- Borjian, Habib, and Borjian, Maryam (2008). "The Last *Gālesh* Herdsman: Ethno-Linguistic Materials from South Caspian Rainforests." *Iranian Studies* 41(3), pp. 365–401.
- Comrie, Bernard (1985). *Tense*. Cambridge etc.: Cambridge University Press.
- Kiā, Moḥammad Šādeq (1330 [1951/52]). *Vāženāme-ye Gorgāni* (A Vocabulary of Gorgani). Tehrān: Dānešgāh-e Tehrān.
- Nourzaei, Maryam, and Jahani, Carina (2012). "The Distribution and Role of the Verb Clitic =a/a= in Different Balochi Dialects." *Orientalia Suecana*, this volume.
- Paul, Ludwig (1998). *Zazaki. Grammatik und Versuch einer Dialektologie*. Wiesbaden: Reichert.
- Purkarim, Hušang (1357 [1978/79]). "Gālešhā va taqsim-e kār (The Galesh and Division of Labour)." In: *Soxanrānihā-ye moḥaqqaqān-e markaz-e mardomšenāsi-ye irān dar haftomin kongre-ye taḥqiqāt-e irāni*. Tehrān: Vezārat-e farhang va honar, pp. 28–44.
- Ranjbar, Maḥmud, and Rādmard, Roqiye (1382 [2003/04]). *Barresi va towṣīf-e gūyeš-e gāleši* (A Survey and Description of the Galeshi Dialect). Rašt: Našr-e Gilakān.
- Sāzmān-e modiriāt va barnāmerizi* (1386 [2007/08]). Golestān.
- Shahidi, Minoo (2008). *A Sociolinguistic Study of Language Shift in Mazandarani* [Studia Iranica Upsaliensia, 11]. Uppsala: Acta Universitatis Upsaliensis.
- Shokri, Guiti (Šokri, Giti) (1369 [1990/91]). "Sāxt-e fe'1 dar gūyeš-e māzandarāni-ye sāri (The verb structure in the Sari dialect of Mazandarani)." *Farhang* 6. Tehran: Institute for Humanities and Cultural Studies, pp. 217–231.
- Shokri, Guiti (Šokri, Giti) (1374 [1995]). *Gūyeš-e sāri* (The Sari Dialect). Tehran: Institute for Humanities and Cultural studies.

Shokri, Guiti (Šokri, Giti) (1380 [2001]). “Barresi-ye barxi tağyirāt dar afʿāl-e morakkab-e esnādi-ye guyeš-e gilaki (Some Changes in Compound and Predicative Verbs in Gilaki).” *Farhang* 14 (37-38). Tehran: Institute for Humanities and Cultural studies, pp. 159–169.

Thomason, Sarah G. (2001). *Language Contact: An Introduction*. Edinburgh: Edinburgh University Press.

Yoshie, Satoko (1996). *Sāri Dialect* [Iranian Studies 10]. Tokyo: Institute for the Languages and Cultures of Asia and Africa.

<http://elar.soas.ac.uk/deposit/jahani2010galesh> [Retrieved April 27, 2012].

<http://ziaratvillage.blogfa.com/> [Retrieved April 27, 2012].

Book Reviews

Replik. Zu Annette L Heitmanns Besprechung des Buches Martin Delhey. *Samāhitā Bhūmiḥ*. Das Kapitel über die meditative Versenkung im Grundteil der Yogācārabhūmi. Teil 1–2. Wien: Arbeitskreis für Tibetische und Buddhistische Studien, 2009.

Kritische Rezensionen sind unbestritten ein unverzichtbares Instrument unserer Wissenskultur. Sie sind aber denselben Prinzipien verpflichtet wie die wissenschaftlichen Arbeiten, die der Kritik unterzogen werden, d.h. die in ihnen ausgesprochene Kritik muß begründet und belegt werden. Geschieht dies nicht, ist die Kritik unwissenschaftlich und wird zur (wie auch immer motivierten) Verunglimpfung.

Ein eklatantes Beispiel einer solchen unwissenschaftlichen Verunglimpfung ist die o.g. Besprechung der Arbeit von Martin Delhey durch Annette L. Heitmann, erschienen in *Orientalia Suecana* LX (2011): 198f.

Auf der ersten Seite der Besprechung geht es ja noch recht sachlich zu, wenngleich einige Bemerkungen inhaltlich nicht unproblematisch sind. Die Feststellung, daß in der *Yogācārabhūmi* grundlegende Konzepte der [Yogācāra-]Tradition enthalten seien, ist ja nicht unkorrekt, aber in concreto wäre doch festzustellen, daß von den hierbei aufgeführten Begriffen *viññāptimātra* außerhalb des in der *Viniścayasamgrahaṇī*, dem wohl spätesten Teil des Textes, in extenso zitierten *Samdhinirmocanasūtra* überhaupt nicht vorkommt, und *ālayavijñāna* sowie die drei *svabhāvas* in großen Teilen des Textes gänzlich fehlen, in anderen nur mehr oder weniger sporadisch vorkommen und nur in bestimmten Stücken der *Viniścayasamgrahaṇī* ausführlich behandelt werden. Problematisch erscheint es mir auch, im Falle der von Delhey edierten *Samāhitā Bhūmiḥ* von einer „umfangreichen Quellenlage“ zu sprechen, obwohl der Originaltext nur in einer einzigen Handschrift erhalten ist, von der wir zudem nur (häufig recht unscharfe) alte Fotos haben. Es muß wohl die Tatsache gemeint sein, daß außerdem nicht nur eine tibetische, sondern auch eine chinesische Übersetzung vorliegt.

Nun würde man der Rezensentin solche nicht ganz präzisen Formulierungen gewiß nachsehen, würde ihre Kritik nicht auf der zweiten Seite der Besprechung in teils kleinliche, teils ausgesprochen mißgünstige Beanstandungen ausarten, und dies überwiegend ohne Belege und ohne (oder jedenfalls ohne schlüssige) Begründung.

Daß der Edition „einige wenige Zeilen einer Teilübersetzung“ beigegeben seien, wird kaum ein Leser so verstehen, daß in Wirklichkeit ein Textstück von 244 Zeilen (wie die Rezensentin zugibt: mit umfangreichen Anmerkungen) übersetzt worden ist, d.h. 20% des Gesamttextes. Daß dies in Anbetracht der detaillierten Analyse des Gesamttextes (S. 451–468) und zusätzlicher Übersetzung zahlreicher Passagen in der Einleitung nicht ausreichen soll, um auf den Gesamttext schließen zu können, vermag ich nicht nachzuvollziehen. Es hat keinen Sinn, die angesichts der extrem schwierigen Manuskriptsituation zeitraubende Erstpublikation dieser Texte dadurch noch weiter zu verzögern, daß man gleichzeitig vollständige Übersetzungen fordert.

Wenn die Rezensentin Titelzusätze wie „einige Bemerkungen zu ...“ für nicht informativ

hält, so zeigt sie damit lediglich, daß ihr der Unterschied zwischen einem Anspruch auf Vollständigkeit und dem Bewußtsein, nur einen vorläufigen Beitrag zur Klärung eines Problems leisten zu können, nicht geläufig ist. Und der Zusatz „weitere Bemerkungen“ ist ja wohl notwendig, wenn bestimmte Aspekte eines Thema schon in einem früheren Abschnitt behandelt worden waren.

Daß ein Hinweis auf „Schmierflecken“ für die Beurteilung der Sicherheit von Lesungen keineswegs irrelevant ist (eben weil sie Akṣaras unkenntlich oder extrem undeutlich machen), kann nur bestreiten, wer die Fotos der Handschrift nicht gesehen hat.

Wenn im drittletzten Absatz Delhey vorgeworfen wird, er changiere in seinen Erörterungen zwischen Alltagssprache, Wissenschaftssprache und Erzählsprache, so wäre dies anhand konkreter Beispiele zu exemplifizieren. Dies gilt umso mehr, als mit diesem Vorwurf ja doch insinuiert wird, daß Delheys Darlegungen weithin unwissenschaftlich seien. Es wäre aber erst einmal aufzuzeigen, inwiefern die Verwendung des einen oder des anderen Sprachtyps (wodurch auch immer sich diese Typen voneinander unterscheiden mögen) oder der Wechsel zwischen ihnen überhaupt die Wissenschaftlichkeit der Darlegungen beeinträchtigen muß; denn Wissenschaftlichkeit hängt doch wohl weniger am Sprachstil als vielmehr daran, daß Feststellungen *belegt* und Thesen *begründet* werden; und von daher ist nicht Delhey, sondern der Rezensentin Unwissenschaftlichkeit vorzuwerfen. Und schon gar nicht hat Delheys im nächsten Satz desselben Absatzes kritisiertes Festhalten an der alten deutschen Rechtschreibung (der ich meinerseits ebenfalls den Vorzug gebe) etwas mit der Wissenschaftlichkeit seiner Arbeit zu tun.

Unklar bleibt mir, ob die anschließende Beanstandung der — auch mir selbst (ohne jeden Beleg!) unterstellten — Verwendung einer

Begrifflichkeit wie „Gemütsversenkung“, „Nicht-Vorstellung“, „Nicht-Nahrung“, „Nicht-Verunreinigt-werden“, „Nicht-Hervorbringen“, „Lebewesentöter“, „Einschrumpfung des Geistes“ oder „Nichtbetrachtung“ usw.

etwas mit dem Wechsel zwischen den genannten drei Sprachtypen zu tun hat. Wohl eher nicht, da die hier angeführten Begriffe anschließend als unbeholfene, für ein Verständnis des Gemeinten nicht förderliche „Wortschöpfungen“ bezeichnet werden. Außerdem sollen sie eine „Unfähigkeit zur philosophischen Reflexion aufweisen“. Eine Begründung ihrer Vorwürfe hält die Rezensentin nicht für nötig. Sie wäre aber einzufordern, zumal bei den beanstandeten indischen Ausdrücken recht unterschiedliche Situationen vorliegen. Selbst wenn man zugesteht, daß in einigen Fällen eine glücklichere bzw. gefälligere Wiedergabe denkbar wäre, schießt die Kritik insgesamt weit über das Ziel hinaus. Es mag genügen, dies an einigen Beispielen zu illustrieren.

Warum etwa soll *anāhāra* nicht mit „Nicht-Nahrung“ wiedergegeben werden, oder *cittābhisamkṣepa* mit „Einschrumpfung des Geistes“, wenn es dem Übersetzer in erster Linie darauf ankommt, den indischen Begriffen, die doch genau dies bedeuten, so nahe wie möglich zu bleiben? Soll der Übersetzer etwa die Metaphorik des indischen Ausdrucks durch eine abstrakte Formulierung zerstören? Im Falle von *cittābhisamkṣepa* macht Delhey übrigens die bewußte Beibehaltung der indischen Metapher dadurch deutlich, daß er seine Wiedergabe in Anführungsstriche setzt. Wie die Rezensentin aus diesem Vorgehen ohne irgendeine Begründung „Unfähigkeit zu philosophischer Reflexion“ abzuleiten vermag, entzieht sich meinem Verständnis und scheint mir eher die Frage nach ihrer eigenen Fähigkeit zu solcher Reflexion aufzuwerfen.

In beiden Fällen ist übrigens ein enger Anschluß an die indische Ausdrucksweise auch deshalb ratsam, weil damit deren Mehrdeutigkeit gewahrt bleibt. Wie nämlich Delhey auf S. 426 (Anm. 94) aufzeigt, wird *cittābhisamkṣepa* nicht nur (wie in der *Samāhitā Bhūmih*) im Sinne von Schläfrigkeit, sondern, im Kontext der Geistesberuhigung (*śamatha*), auch im positiven Sinn eines Sich-nach-innen-Wendens des Geistes gebraucht. Und zu der Feststellung, daß als „Nicht-Nahrung“ für Verlangen nach den Objekten sinnlicher Begierde das Widerwärtige

bzw. dessen Betrachtung und wiederholte Betrachtung zu verstehen sei, bemerkt der Text, die (einfache) Betrachtung führe dazu, daß noch nicht entstandenes Verlangen gar nicht erst entsteht, während die wiederholte Betrachtung bewirke, daß schon entstandenes Verlangen aufhört. Hier wird deutlich, daß die Negation in dem Begriff *anāhāra* in einer doppelten Funktion gemeint ist: zunächst als reine Negation (*prasajyapraṭiśedha*), insofern die (einfache) Betrachtung kein neues Verlangen nährt, dann aber als ‚ausgrenzende Negation‘ (*paryudāsa-praṭiśedha*), insofern die wiederholte Betrachtung, indem sie das schon vorhandene Verlangen zerstört, sozusagen als Anti-Nahrung wirkt.

Unverständlich ist mir auch, was die Wahl der Wiedergabe „Lebewesentöter“ (für Skt. *prāṇātīpātika*; cf. Pāli *pāṇātīpātīn*) mit „Unfähigkeit zu philosophischer Reflexion“ zu tun haben soll. Das Kompositum mag im Deutschen nicht der Alltagssprache angehören, aber es läßt sich doch (dank der noch kreativen Kompositionsbildung im Deutschen) ohne weiteres bilden und drückt genau das aus, was der indische Begriff besagt: eine Person, die habituell Lebewesen tötet. Daß wir dafür kein gängiges Wort haben, liegt daran, daß in unserem Kulturkreis das Töten nichtmenschlicher Lebewesen, das in dem indischen Begriff eingeschlossen ist, einfach einen anderen Stellenwert hat als dort.

Komplexer ist die Situation im Falle des Ausdrucks „Nicht-Vorstellung“, mit dem Delhey Skt. *a-saṃjñā* wiedergibt. Es ist nicht klar, ob die Rezensentin die Wiedergabe von *saṃjñā* mit „Vorstellung“ beanstandet oder die nominale Fügung mit „Nicht-“ (oder beides). Im ersten Fall handelt es sich in der Tat um ein echtes Problem; denn *saṃjñā* ist ein komplexer Begriff, bei dessen Verwendung unterschiedliche Nuancen im Vordergrund stehen können (vgl. etwa Tilmann Vetter. *The 'Khandha' Passages' in the Vinaya-piṭaka and the four main Nikāyas*. Wien: Österreichische Akademie der Wissenschaften 2000: 24–27). Wenn die Rezensentin mit Delheys Wiedergabe nicht einverstanden ist, so sollte sie dies begründen und angeben, mit welchem deutschen Begriff sie den von *saṃjñā* umspannten semantischen Bereich denn treffender fassen zu können glaubt. Es gibt nun einmal leider indische Begriffe, die sich mit unseren nicht decken, so daß man sich als Übersetzer mit Notlösungen behelfen und die involvierte Problematik entweder in einer Anmerkung diskutieren oder auf eine entsprechende Diskussion in der Literatur verweisen muß, was Delhey (S. 412, Anm. 17) ja auch tut. Geht es hingegen in erster Linie um die nominale Fügung mit „Nicht-“, so kann es sich nur um eine stilistische Beanstandung handeln, da die inhaltlichen Probleme, die in der Tat zu einer mehrere Deutungen offenlassenden Wiedergabe der Negation zwingen, von Delhey auf S. 60f gebührend dargelegt werden.

Wenn die Rezensentin, im Anschluß an die im vorigen besprochene Beanstandung bestimmter Begriffe, behauptet, in Delheys Arbeit sei „so manche Übersetzung ... verfehlt“, so darf der Leser doch auch bei dieser schwerwiegenden Behauptung erwarten, daß sie mit einer Anzahl griffiger Beispiele substantiiert wird. Davon kann aber keine Rede sein. Der einzige angeführte Fall betrifft nur die Wiedergabe eines Begriffs:

Z.B. *Übelwollen* (*vyāpāda*) mit *Ärger* (*āghāta*) auf eine Stufe zu stellen, ist fragwürdig; besser wäre, auch im Sinne der Erklärungen der SamBh, *malicious feeling* oder *bösartige Gesinnung* für *āghāta* anzusetzen.

Die Rezensentin hat offenbar nicht zur Kenntnis genommen, daß Delhey in einer längeren Anmerkung (S. 423f, Anm. 81) die Begriffe *vyāpāda* und *āghāta* keineswegs „auf eine Stufe“ stellt (was genau soll damit eigentlich gesagt sein?), sondern sehr differenziert behandelt und ihre jeweilige, allerdings keineswegs einfache Einordnung in einem Wortfeld, „das 1. Ärger, Zorn, 2. Abneigung und 3. den Wunsch, daß andere Wesen Schaden erleiden mögen, oder sogar eine konkrete Schädigungsabsicht beinhaltet“, diskutiert. Dabei ist *āghāta* oft der weitere Begriff, kann aber durchaus auch eine aktuelle Zornesaufwallung bezeichnen, und an der *Samāhitā-Bhūmi*-Stelle, um die es geht, wird *vyāpāda* (im Sinne der 3. Bedeutungs-facette) nicht einfach als bösartige Gesinnung, sondern als ein durch einen Tadel oder die Erinnerung an eine erlittene Schädigung oder Beleidigung (*apakāra*) ausgelöster *cetasa āghātaḥ* erklärt,

so daß das Moment des Ärgers oder (präziser vielleicht) Grolls („resentment“) durchaus mitschwingt.

Wenn die Rezensentin des weiteren (im vorletzten Absatz) Delhey pauschal (ohne auch nur ein einziges Beispiel anzuführen) Fehler der Rechtschreibung und Zitierweise sowie „ungenauere Angaben usw.“(!) vorwirft, so sollte man zumindest erwarten, daß ihr selbst solche Vorwürfe nicht gemacht werden können. Dem ist aber nicht so. Oder ist „rekuriert“ (S. 198,22, mit einem r) etwa neue Rechtschreibung? Und was soll, in der Aufzählung der grundlegenden Konzepte der Yogācāra-Tradition, das Stichwort *bījā* (sic)? Seit wann ist *bīja* (n.) femininum? Und was soll man davon halten, daß Delhey dafür getadelt wird, daß er in der Bibliographie auch „Anfängerwerke“ aufführt, deren Nutzen für die Arbeit nicht ersichtlich sei (was in concreto zu belegen und zu begründen die Rezensentin natürlich nicht für nötig hält)? Welche Art von Arbeiten soll denn damit gemeint sein? Etwa auch Dissertationen? Auch ist es ja nicht a priori so, daß „Anfängerarbeiten“ keine wesentlichen Einsichten enthalten könnten. Wenn der Verfasser Magisterarbeiten mit Gewinn benutzt hätte, würde es doch zweifellos zum Ethos eines Wissenschaftlers gehören, dies entsprechend zu dokumentieren. Man muß sich wirklich fragen, welcher Art von Umgang mit den Leistungen junger Wissenschaftler die Rezensentin hier das Wort redet.

Lambert Schmithausen, Asien-Afrika-Institut, Universität Hamburg

Korn, Agnes, Haig, Geoffrey, Karimi, Simin, and Samvelian, Pollet (2011). *Topics in Iranian Linguistics*. Beiträge zur Iranistik, 34. Wiesbaden: Reichert. Pp. 214. ISBN: 978-3-89500-826-9.

The volume *Topics in Iranian Linguistics* consists of 13 articles presented at the Third International Conference on Iranian Linguistics (ICIL3) held in Paris in September 2009. It is divided into three parts: Historical and Comparative Iranian Syntax, The Morpho-Syntax of Lesser-Known Iranian Languages, and Linguistics of Modern Persian. Containing interesting articles on a number of Iranian languages from different geographical areas and covering a time span of at least 3000 years, it is a highly valuable contribution to Iranian linguistics.

There are five contributions in the section on Historical and Comparative Iranian Syntax, three on specific languages, namely Bactrian and Sogdian, and two of a comparative nature discussing several Iranian languages.

The first article, by Saloumeh Gholami, is entitled “Definite Articles in Bactrian”. Gholami discusses to what extent Bactrian demonstratives also can function as definite articles. She bases her definition of definite articles on a theory that demonstrative adjectives occur in pragmatically definite contexts, whereas definite articles occur in semantically definite contexts. After investigating over 100 examples of two Bactrian demonstratives, she concludes that they occur both in semantically and pragmatically definite contexts, and that they therefore can be described both as definite articles and demonstrative adjectives.

In the second article, Nicholas Sims-Williams studies “Differential Object Marking in Bactrian” within the framework used by the “Leipzig ‘ditransitivity’ project”. Studying mono- and ditransitive constructions mainly in post-Kushan Bactrian, a language with ergative alignment in the past tense, he concludes that differential object marking (DOM) is found both for “patients” (direct objects of monotransitive verbs) and “recipients” (indirect objects of ditransitive verbs). He further notes that the same preposition is used to mark a personal patient and a non-personal recipient, but that a different preposition is used for a personal recipient. He concludes with a very interesting diachronic discussion about the handling of the ambiguity that arises when the “theme” (direct object of a ditransitive verb) and the recipient receive the same encoding.

In her article “The Emergence and Development of the Sogdian Perfect”, the third one in this section, Antje Wendtland discusses how a perfect with “to have”, first attested in the Ancient Letters, is grammaticalized as the auxiliary for transitive verbs, and even spreads to some intransitive verbs, though not to all. She finds that the best model for describing the intransitive verbs that take “to have” is one in which intransitive verbs are divided into unergative verbs (which have an underlying object) and unaccusative verbs (which have an underlying subject) and it is the unergative verbs that acquire a perfect with the auxiliary “to have”. This distribution is parallel to that found in, e.g., German and French.

The fourth article, by Agnes Korn, carries the title “Pronouns as Verbs, Verbs as Pronouns: Demonstratives and the Copula in Iranian”. Demonstrating an impressive command of a wide range of Old, Middle, and New Iranian languages, Korn discusses how pronouns have been incorporated into the verb morphology of several Eastern Iranian languages, where they are used as copulas. She also demonstrates parallel cases of this phenomenon in many other languages, such as Arabic and Modern Hebrew. In the second part of the article, she discusses the opposite development, namely copulas and personal endings of verbs making their way into the pronominal system of a language. Among the candidates that she presents for this latter development is Balochi. One way this may have happened is very nicely sketched in (19) on p. 66.

In the last article of the first section, “Counterfactual Mood in Iranian”, Arseniy Vydrin discusses whether there is a dedicated marker or mood for counterfactuality in six Iranian languages, namely Sogdian, Ossetic (Iron dialect), Pashto, Tat, Talysh, and Parači. After concluding that too few examples are available for Sogdian, Vydrin finds that Tat, Talysh, and Parači have a dedicated counterfactive, something which he believes is probably due to influence from Turkic/Indo-Aryan languages.

The four articles in the second section, *Morpho-Syntax of Lesser-Known Iranian Languages*, treat Iranian Taleshi, the Pamir languages, and Ossetic. Concerning the title of the section, one might, of course, ask by whom and to what extent are these languages “lesser-known”?

Daniel Paul is the author of the first article, “A Glance at the Deixis of Nominal Demonstratives in Iranian Taleshi”. Based on data from his own fieldwork, he explores the nominal demonstrative system of the Asalem and Masal dialects of Iranian Taleshi and demonstrates that the demonstratives in both these dialects have a number of endophoric and exophoric uses for proximate (I would prefer “proximal”) and distal senses.

In the second article, “Valence Sensitivity in Pamir Past-tense Inflection: A Realizational Analysis”, Gregory Stump and Andrew Hippisley propose a formal synchronic analysis of the past-tense auxiliary in the Pamir languages. The article, which is highly theoretical, argues that “the Pamir languages [...] are alike in possessing a past-tense auxiliary clitic expressing subject agreement” but “[t]he precise properties of this clitic [...] vary from language to language” (p. 114).

The two final articles in this section deal with Ossetic. Oleg Belyaev and Arseniy Vydrin study “Participle-Converbs in Iron Ossetic: Syntactic and Semantic Properties” and David Erschler and Vitaly Volk examine “On Negation, Negative Concord, and Negative Imperatives in Digor Ossetic”. Both are based on the authors’ fieldwork and provide solid descriptions of the research subjects.

The third section, *Linguistics of Modern Persian*, contains four articles, one on phonology, two on morphosyntax, and one on a sociolinguistic topic.

Navid Naderi and Marc van Oostendorp are the authors of “Reducing the Number of Farsi Epenthetic Consonants”. They describe consonant epenthesis as a hiatus-resolving feature in Iranian Persian (or Farsi). Discussing the status of the glottal stop in Iranian Persian, they follow the analysis of Kambuziya (p. 155, see bibl. for reference) that “words starting with glottal stops in Farsi should be divided into two groups of which one has phonemic /ʔ/ and the other epenthetic [ʔ].” I strongly question this analysis, because in my opinion there is no difference

between how the (potential) glottal stop is realized in, e.g., [ham(?)asr] “contemporary” and [ham(?)andiʃi] “conference”. There is normally no phonetic realization of the glottal stop, regardless of whether it occurs in an Arabic loanword or a Persian word. In the article, most words with a vowel onset have a ʔ as the first symbol in their phonetic transcription, but occasionally they are represented without this symbol (see pp. 162, 165). I see no reason for this distinction, and it may just be an unintentional omission. This does, however, indicate that it is unnatural (and in my opinion a feature of a Semitic language structure) to include the glottal plosive as a phone, or phoneme, in word-initial position in Persian. It is slightly unclear in the rest of the article if Naderi and van Oostendorp are describing spoken or written Persian, and from where they have taken their phonetic representations (e.g. under (15) on p. 157). The phrase “our home” is, in my opinion, never naturally pronounced [xɒneʔemɒn] in Persian, not even when a written text is read aloud.

The second contribution is “On Direct Objects in Modern Persian: The Case of the Non-*rā*-Marked DOs”. The issue of direct object marking in Modern Persian has been the subject of numerous investigations, but in this article Shadi Ganjavi studies the non-*rā*-marked direct objects rather than the *rā*-marked ones. She finds that “contrary to the widely held belief in the literature” (p. 180) they do not form a natural class, which is an interesting and noteworthy result. What is not entirely clear to me is where her examples come from, whether she has provided them by herself only, or if she also has involved other L1 speakers of Persian in judging the grammaticality of the sentences discussed in the article. In any case, her conclusions appear to be valid.

In the article entitled “Finite Control in Persian”, Mohammadreza Pirooz introduces “a Minimalist account of the Obligatory Control constructions in Persian” (p. 193). As in the previous article, the origin of the data remains unclear and, although I am not well versed in the Minimalist theory, it seems that there are things in this article that can be questioned, such as the statement in connection with (16) (p. 191). A sentence such as *una Mohsen-o majbur=kaerd-aen (ke) PRO xod-es-ro eslah=kon-e* “they forced Mohsen to correct himself (change his attitude)” (note that my transcription follows that in the article) seems to be a fully acceptable construction in Persian. Therefore PRO does not have to check the nominative case in the lower clause. Whether a direct object can be taken depends on the verb, and *haerf=zædaen* “to speak” in (16) does not take one. I also hold that the direct object marker =*rā* is a clitic, not a suffix.

Finally, Farzaneh Deravi and Jean-Yves Dommergues study “Bilingual Speech of Highly Proficient Persian-French speakers”. Basing their study on interviews with 30 “highly proficient Persian-French bilinguals” (p. 197), the authors find that, in basically Persian speech, nouns are by far the most common grammatical category for which code switching occurs (mainly to French but occasionally to English). In their results, they also list verbs and adjectives as categories where code switching commonly occurs (p. 206), although Table B (p. 204) indicates that adverbs are more common than verbs. It also seems that the percentages given in this table are somewhat wrong, which, of course, is unimportant for the general discussion.

A few remarks apply to several of the articles in the volume. It would have been useful to provide an abstract and keywords for each article (only one article has an abstract), and three of the articles lack a list of abbreviations. A further minor observation is that the distinction between suffixes and clitics is not upheld in the glossing in some of the articles. The reader occasionally gets the feeling that there has not been a proper check of the English accuracy in one or two of the articles written by non-native English speakers.

The volume does, however, represent a most valuable contribution to Iranian linguistics. The range of Iranian languages studied is impressive and most of the articles demonstrate high scholarly competence. The careful editorial work has produced an attractive and cohesive volume. Maps, diagrams, and tables make it easy to get an overview of the articles, which for the most part are innovative and treat a great variety of interesting topics. In addition to answering the research questions posed in the studies, in many cases they invite further research. One very

interesting topic for further studies is presented by Sims-Williams, namely the ditransitive construction and the encoding of addressees, recipients, and benefactives. Other topics treated in this volume that would be interesting to study in a variety of Iranian languages include the development of perfect forms, deixis, consonant epenthesis, and finite control.

Carina Jahani, Department of Linguistics and Philology, Uppsala University

Kreyenbroek, Philip G. (2009). *Yezidism in Europe. Different Generations Speak about their Religion*. In collaboration with Z. Kartal, Kh. Omarkhali, and Kh. Jindy Rashow. Göttinger Orientforschungen Iranica Neue Folge 5. Wiesbaden: Harrassowitz. Pp. 246.

ISBN: 978-3-447-06060-8.

When religious and ethnic minority groups live in a secular or foreign religious culture they face difficulties transferring their religious and cultural worldview to the next generation. When a large part of the group also has become migrants and the different generations have experienced diverse childhood conditions, an additional aspect is added to the question of religious and ethnic socialization. Philip G. Kreyenbroek, Professor of Iranian Studies, University of Göttingen, addresses in his book *Yezidism in Europe: Different Generations Speak about their Religion* the relationship between transferring religious beliefs and practices, and the experiences of being a minority in one's homeland and being a minority in the Diaspora.

It is important that studies of Yezidism not only focus on the religious belief system and everyday practices, but also treat the religious dimension as part of a broader social context. Moreover, in Kreyenbroek's study, we encounter both the religiously active Yezidis and those whose religious tradition has declined, the later mostly among the second generation in the Diaspora. Kreyenbroek makes a valuable contribution to the field with his nuanced description of the Yezidis' religious life. However, he does not discuss the reasons behind the rapid secularization processes among migrants and minority groups in the homeland. Nevertheless, the author's rich empirical examples compensate for the lack of theoretical analysis in that they accurately describe the differences between the generations and the differences between Yezidis in their homeland and in the Diaspora.

The book consists of three parts and a short conclusion. In the Introduction Kreyenbroek describes central aspects of the Yezidi religion: prayer, acts of worship, pilgrimage, sacred places, how to recite religious texts, and the religious year. The author also describes how the disciple should practise religion in everyday life in relation to clothing, taboos, prohibitions, and healing rituals. Furthermore, the author describes how the Yezidis should practise the rites of passage, for example initiatory rites, weddings, childbirth, the first forty days, teething, and funerals. In the Introduction, Kreyenbroek also gives a short overview of Yezidi societies in their homelands (Northern Iraq, Armenia, Georgia, Syria, and Turkey) and in the Diaspora (Germany, the Netherlands, Belgium, France, Russia, the Ukraine, and Scandinavia).

The second and third chapters, which constitute the main part of the book, present the empirical material based on qualitative interviews among Yezidis in Germany, Armenia/Russia and their homeland. The excerpts from the interviews have been selected according to two principles. Some excerpts reflect the most common view, and others reflect topics that are relevant for the predominant aim of the book, namely to examine shifts in beliefs and practices between generations and people socialized under different circumstances.

The second part focuses on Yezidis born in their homeland. The author describes general attitudes that are common in traditional societies, where community membership comes first and the social control is strong. We meet attitudes towards authority, respect and social control,

shame and honour, feuds and quarrels, conflict-resolution, eloping and abduction. Kreyenbroek further discusses Yezidi tribe and caste identity, how they view religious socialization, and how they act to transmit the religious worldview to the next generation.

In the third part, we meet Yezidis socialized in the Diaspora. The majority of the excerpts come from Germany, and the remainder from Armenia/Russia. The chapter begins with a description of what the Yezidis in the Diaspora know about their homeland, what the communication is like between the generations, what they know about their religion, and how the education in religious practices is organized. After that Kreyenbroek includes a subchapter about Yezidi attitudes toward their culture and religion. Here, we can follow what they think about purity, if they practise obedience to the older generation or prefer personal freedom of choice, but also how they feel about honour, blood feuds, and marrying out,¹ the relationship between the genders, the bride price, as well as how they perceive the caste system. In the next section, the author discusses life in the Diaspora. He describes how the Yezidis interpret the differences between life in the homelands and in the Diaspora, how the contacts between different groups in the Diaspora are organized, and in what sense the Yezidis in the Diaspora maintain religious rites and practices such as circumcision, weddings, funerals, fasting, and pilgrimage, but also the use of holy objects and taboos. The chapter ends with a section about how the members of the group view the future.

The Conclusions begin with a theoretical declaration. Kreyenbroek writes that the study is an exploratory work and that the design does not allow him to draw statistical sociological conclusions. However, the author also claims that the material enables us to see some tendencies that may help us to better understand the conditions of Yezidis in the Diaspora. Kreyenbroek chooses to present these tendencies as ideal types in the Weberian sense. This means that the tendencies presented cannot be reduced to single persons, nor will we encounter the variations of attitudes, values, and beliefs shown in the material in the discussion. The empirical material is presented on an aggregated level, with the most common value in each topic highlighted.

The author begins the Conclusions with a discussion about how the social environment influences the socialization of a religious worldview. What happens, for example, to the honour and shame norm system when it is not legitimated with a traditional authoritarian patriarchal system? The younger generation does not understand or accept the blood feuds, the bride price, or the custom of marrying within their caste. Then follows a discussion about the difficulties for a religion to survive in a modern society when the religion lacks not only a holy book, but also exegetical and dogmatic literature and a knowledgeable priesthood. A lack of communication between generations is another factor when it comes to transferring a religious worldview between generations. The generations do not have a common language for talking about religious things. The older generation is fluent in Kurdish and the younger generation in German. Furthermore the younger generation are often more educated and the older generation may feel ashamed of their old-fashioned beliefs. The chapter ends by pointing out three processes by which the Yezidis are trying to reconcile their religion with the modern world. First is the process of re-interpretation of the religion, whereby believers try to distinguish “real” religion from cultural artefacts.² Second is the process of establishing a new caste for those who want to convert, marry between castes, or “marry-out” and still belong to the Yezidi religion and culture. The third process is the establishment of religious institutions such as the Yezidi Association in Germany and the Yezidi Religious Council in Iraqi Kurdistan.

A strength of Kreyenbroek’s book is that it is based both on empirical studies, in which Yezidis have their say, and social facts about different aspects of the Yezidi religion. On the other hand the author could have provided more follow-up questions. For example, the men who have kidnapped their wives (as could happen when the bride price is considered too high),

¹ According to the tradition, it is forbidden to marry outside the Yezidi group. It is also forbidden to marry outside a person’s caste. It is also forbidden to convert to the religion. You can only be a Yezidi “by blood”.

² We see a similar process among many believers today, who try to find a “pure” religion which is not affected by culture.

say it was no problem for them, but what was it like for the women to be kidnapped? What is their experience of being kidnapped? What opportunities do they have to take active part in choosing a husband?

Another strength of the book is that the author identifies many interesting aspects of religious socialization. He both points out the gradual erosion of traditional religious observances and behaviours and discusses factors that are of significance in this process. It would be interesting to see a more theoretical discussion about what factors are significant when a religious tradition becomes more secular. I think this question is very important when religion is studied as a social construction. Not least because we also see the opposite tendency today: a process of radicalization of beliefs and practices when people with a traditional faith meet a western and/or secularized modern society. The question is also interesting on a societal level. How can a democratic society take into account different religious groups that demand political influence over the whole of society, as for example Islamism, Jihadism, and Christian fundamentalism do?

In conclusion, this book is an important contribution to the academic understanding of different ethnic and religious groups in contemporary society. It provides insight into the Yezidis' religious worldview and their experiences of being migrants. What is more, the author also gives us a broader understanding of the challenges a religious group faces in transferring a religious worldview to the next generation in a modern society.

Madeleine Sultán Sjöqvist, The Hugo Valentin Centre, Uppsala University

Wikander, Ola (2012). *Drought, Death and the Sun in Ugarit and Ancient Israel. A Philological and Comparative Study*. (Doctor of Theology dissertation, Lund University). Lund. Pp. 277.
ISBN 978-91-7473-317-4.

In his thesis, *Drought, Death and the Sun in Ugarit and Ancient Israel*, Ola Wikander treats the complex nature of the Ugaritic solar deity as the progenitor of both fertility and drought, and discusses similar roles of Yahweh who both “makes weal and creates woe” (Isa 45:7). The Ugaritic samples are picked from the Ba'al Cycle, from funerary texts, and from the epics of Aqhat and Kirta; the Bible is represented by a sample of “case studies”, the Carmel incident in 1 Kings 18, illustrative examples from the prophets, some instances from Job and Psalms, and a passage from Ben Sira. Subsequent to the discussion of these passages, Wikander contemplates the circumstance that the biblical texts represent faith in a God who both roars from Sion in a thunderstorm and inflicts drought on the fields of the shepherds (Am 1:2).

The introductory chapter (pp. 1–30) confidently holds out the prospect of demonstrating a common mythological legacy discernible both in Ugarit texts from the 13th century and in the multi-layered biblical texts from Ancient Israel. This goal is primarily meant to be achieved through a study of etymologically related Semitic verbs that relate to heat and drought. Chapter two (pp. 31–122) sets out to illustrate various roles of the divine sun in Ugarit, who is the dispenser of judicial and royal power, and at the same time (due to a partly chthonic nature), a menacing goddess. The point of departure is a phrase called the “Refrain of the Burning Sun”. It occurs three times in the Ba'al Cycle: CAT 1.3 V 17–18; 1.4 VIII 21–24 and 1.6 II 24–25, of which the last runs: *nrt ilm špš šhrt // la šmm byd bn ilm mt* “Shapshu the luminary of the gods did glow hot; the heavens were wearied by the hands of divine Mot”. This phrase is put forward as pivotal for the structure of the whole Cycle, whose basic contents are as follows: the victorious god Ba'al is in need of a palace; a palace is built on Mount Šaphon from where he reigns; Ba'al must surrender to his rival Mot (an entity personifying drought-and-death); the goddess 'Anat brings Ba'al back to life.

The reader is, however, not allowed to know exactly in what way the “refrain” is pivotal for the story, only that the first occurrence is poorly preserved, the second is found in relation to the building of Ba‘al’s palace, while the third relates to his descent to the netherworld. The discussion rather revolves around the “etymological poetics” of the word *šhrrt* “did glow hot”, which is not otherwise associated with the sun. Exception is taken to Johannes de Moor and Juliane Kutter who think that the verb refers to the sun as “dust-coloured”. Instead, by reference to the Arabic root *ṣ-ḥ-r*, manifested in *ṣahrā* “desert”, and (questionably) Akkadian *ṣarāḥu* “to heat, scorch”, it is argued that the radicals *ṣ-ḥ-r* signifies a burning, feverish heat, an idea that is said to find support in the related common biliteral Semitic root *ḥ-r* “be hot”. Wikander is well versed in the secondary literature, but for some reason he refers very sparsely to *A Dictionary of the Ugaritic Language* by G. del Olmo Lete and J. Sanmartín. For the passage under discussion, this lexicon suggests a transitive sense for *ṣ-ḥ-r-r*, viz., “to roast, burn”, which means that *la šmm* in the passage is analysed as the object: “the sun roasts the strength of the skies”. Among alternative renderings (in which *la* is assigned a verbal force) one finds the reasonable proposal “der Himmel wurde schwach (inf. abs.)”, which in the English form “were wearied” is presented as the correct interpretation. However, Wikander also maintains (pp. 50–51) that *nrt ilm špš šhrrt* may be interpreted as the apodosis of a preposed expression of wish: “May he (Mot) not serve you up like a lamb in his mouth [...] (for) the lamp of the gods, *will then* glow hot, the heavens will be powerless/dried up”. For this proposal, Tropper, *Ugaritische Grammatik* (Ugarit-Verlag, Münster, 2000, pp. 716f.) is called in, notwithstanding that the paragraph cited explicitly says that a future sense of the suffix conjugation presupposes that the verb in question takes a preposed particle, either *w* “and”, *k* “then”, or *l* “for sure”. Moreover, the Ba‘al epic refers to the sun in a graphic description of the devastating effect of Ba‘al’s absence in CAT 1.6 IV 1–3: *pl ‘nt šdm y špš // pl ‘nt šdm il* “Parched are the furrows of the fields, O Shapshu; parched are the furrows of the fields of El.” After a lengthy discussion of the meaning of the verb *p-l-l*, Wikander arrives at the conclusion that the sense is nothing but “be parched, cracked” – an opinion that is shared by del Olmo Lete and Sanmartín.

The next major section deals with the role of the solar deity in connection with drought-and-death in other Ugaritic texts (pp. 84–122). The discussion of the enigmatic CAT 1.12 focuses on a passage in col. II 31–45 that contains a description of a mythical fight in which certain monsters kill Ba‘al. Without mentioning the solar goddess, this passage intriguingly comprises a cluster of verbal roots, viz., *ḥ-r-r*, *ṣ-ḥ-n*, and *ṣ-ḥ-r-r*, all of which indicate the idea of “being hot, burning and scorching”, thus indicating the connection between these notions and death. Shapshu is also encountered in some ritual texts related to the funerary cult. The function of these texts was obviously to promote contact with the deceased king. In the middle of one of these, CAT 1.161, the solar deity is invoked by the imperative *išhn*. This form was earlier derived from *ṣ-w-ḥ* “to sink” or *ṣ-ḥ-ḥ* “to bow down”, but recent scholarship is inclined to derive it from *ṣ-ḥ-n* “to be hot”, which Wikander thinks is philologically more plausible. In fact, del Olmo Lete and Sanmartín quote the passage under discussion, viz., *išhn špš w išhn*, and propose: “Heat, O DN, yes heat”. Moreover, Juliane Kutter, in her book *nūr ilī: Die Sonnengötter in den nordwestsemitischen Religionen* (Ugarit-Verlag, Münster, 2008), thinks that Shapshu – being responsible for the cosmic order – keeps a check on the chaotic netherworld whose supernatural beings appear in the ritual texts. But in the opinion of Wikander, Shapshu is actually distinguished by a *liminal* character, which means that she represents both fertility and drought, a characteristic that is manifested by her perpetual crossing of the border between the realms of life and death.

As is duly pointed out, there are several linguistic parallels between descriptions of drought and ensuing death in the texts from Ugarit and Ancient Israel. Drought is conceptually associated with death in both societies. Lamenting the death of his son Aqhat, Danel says: *bl ṭl bl rbb bl šr’ thmtm* “without dew, without downpour, without surging of the double-depth” (CAT 1.19 I 44–45); a biblical counterpart is found in David’s bewailing the death of Jonathan:

'al-tal w' al-mātār 'alēkæm ūšēdē r'rumōt (2 Sam 1:21) – for *wšdy trwmwt* the Ugaritic phraseology suggests the widely accepted conjecture: *wšr' thmt* yielding “let there be no dew or rain upon you, *nor upsurging of the deep*”.

Chapter three (pp. 123–196) has a biblical focus, the aim of which is to trace underlying myths connected with drought-and-death in Ancient Israel that are reflected in the narrative, prophetic, and poetic portions of the Bible. The Carmel incident (1 Kgs 18:19–46) is discussed under the assumption that its present form represents an “unusually clear example” of an Israelite reinterpretation of the old myth, an idea that would explain the complete lack of a solar deity as progenitor of the drought. Likewise, the examples selected from Isaiah describe Yahweh as the supreme God, who punishes with drought, e.g., in Isa 5:13–14: “its multitude is parched with thirst”, and in Isa 42:15 “I will lay waste mountains and hills, and dry up all their herbage”, but still protects from the harms of the sun, as evident in Isa 49:10: “heat and sun (*šārāb wā-šamæš*) shall not strike them”. Jeremiah chs 14–15 focus on drought-and-war, since God has abandoned his people. Accordingly, the wording in v. 18: “behold, those killed by the sword”, links the hardships of drought with the ravages of war, and in vv. 8–9a God is depicted as “a stranger in the land”, and “a traveller, who has turned aside for the night”. Hosea 13:14 opens: “O Death, where are your plagues? O Sheol, where is your destruction?” The evils of *daēbær* “plague” and *qætæb* “destruction, ravage” are here attributed to Death and Sheol, a circumstance that brings Deut 32:24 to mind, since this latter passage combines *rā āb* “hunger”, *ræšæf* “flame”, and *qætæb* “destruction”, but this observation is discussed only in ch. 4. Instead, the subsequent discussion treats the concept of death as a semi-autonomous being, as well as the role of Yahweh as sender of the sirocco, and the derivation of the word Sheol. Much attention is given to Hos 13:5: *'anī r' 'lūkā bammīdḅār* “I shepherded you in the desert” (so the Septuagint), continued by, *b' æræš tal' ūbōt*, “in an uninhabited, or ravaged land”, as the ancient versions have it – but according to HALOT rather to be interpreted as “in a land infected by illness” (or the like), because the hapax *tal' ūbōt* is likely related to Akkadian *la'ābu* (which in the G-stem is “to harass”, and in the D-stem is “to infect s.o. with illness”). If correct – Wikander attentively says – this derivation provides yet another example of how the concepts of aridity and disease interrelate.

The first chapter of Joel depicts a completely ravaged land, afflicted by swarms of locusts with ensuing drought and death. Biblical scholars regard Joel 1:8: “Lament like a virgin dressed in sackcloth for the husband (*ba'al*) of her youth”, as a reminiscence of the ancient mythical scene in which virgin Anat laments for the dead Baal, (see CAT 1.6 I 2–10); and actually Arvid Kapelrud identified the monsters that killed Ba'al with locusts (cf. above). Having accounted for this clear connection between the Bible and the texts from Ugarit, Wikander continues with a discussion of the verbs *'ābal* “dry up, mourn” and *'umlal* “languish, wither” (both used in Joel ch. 1), and most reasonably infers that these verbs represent a common inherited stock of poetic tools used to describe the mythical motifs of disaster and feverish hotness.

Doomsday is depicted in Malachi 3:19–21 as a burning oven; but, interestingly, the sun brings blessing – not destruction – that is to say, the “sun of righteousness” (*šamæš šēdāqā*) that will rise with healing in its wings. The explanation given is that this sun that clears the believers of punishment represents a fusion between the Mesopotamian judging sun and the Ugaritic burning sun. The same synthesis, Wikander maintains, is present in Psalm 84, where v. 12 is quoted as follows: *kī šamæš ūmāgēn Yhwh // 'lōhīm hēn w'kābōd* “for YHWH is sun and shield, God is grace and glory”. Referring to the dangerous character of God's solar *kābōd* (evident in 1 Kings 8:11), he furthermore holds that this unique passage expresses an antithetic parallelism, in which *šamæš* opposes *māgēn*, and *hēn* opposes *kābōd*, which means that the menacing entities *šamæš* and *kābōd* are actually combined in the Yahwistic faith. Moreover, in Isaiah 60:19 Yahweh's *kābōd* is promised to be an *'ōr 'ōlām* “everlasting sun” (v. 19) – an idea that may relate to the Phoenician/Ugaritic phrase *šmš/špš 'lm*, and thus perhaps intimates a chthonic aspect of the phrase *'ōr 'ōlām*. Certainly, the poem of Job recurrently refers to dry-

ness and death, and to the imagery of the devouring jaws of Sheol. See *inter alia*, 24:19, 30:3, 30. (Also, it might be correct that *mibbēlī* in 18:15 is to be read *mabbēl* “burning heat”, and that the miners, who in 28:4 are depicted as *niškāhīm minnī-rāḡæl* “forgotten by feet”, are perhaps rather *niškāhīm min-Nergal*, i.e., stricken by fever inflicted by Nergal; cf. škh¹¹ “be parched” in HALOT.) Nevertheless, by and large, all the quoted biblical passages accord with Ben Sira 43:2–5, where the burning sun is described as part of God’s glorious creation, just like other natural phenomena. Save for a few passages, such as Isa 49:19, the sun is not presented as a destructive force either. In relation to drought-and-death one might actually say that *le soleil brille par son absence*.

This state of affairs is commonly held to be the result of an ideological development, which means that the drought-and-death theme gradually weakened and finally became a moral lesson taught by the monotheistic deity, an example of which is found in Hag 1:1–8 where drought is regarded as a mere consequence of the people’s unwillingness to build the Temple. Due to Egyptian or Mesopotamian influence the very image of Yahweh is also thought to have undergone a so-called solarization, Wikander informs, referring to Prov 15:11, where the solar image of Yahweh is depicted as gazing into the dark lands of death; the verse runs: “Sheol and Abaddon lie open before the Lord”.

The section devoted to conclusions is divided into several topics: The relationship between heat, drought, and death; the liminal role of the Ugaritic solar goddess Shapshu and her role as connector between living and dead; drought in historical tradition, as well as how punishment by drought is integrated into the Yahwistic faith; and the function of drought as a theological metaphor. Not discussed is the question for what purpose did the biblical writers employ culturally conditioned mythical phrases in their description of Yahweh and his relationship to his people.

In the case studies, there is obviously no trace of the sun in connection with the netherworld, and there are very few traces of a harmful sun. Nevertheless, as is clearly laid down at the beginning of the concluding ch. 4 (pp. 202f.), a small section in the archaic Song of Moses (Deut 32:22–24), in a metonymy using *’ēš* “fire” for *šæmæš* “sun”, actually exhibits all the signs of a common ancient cultural tool-box for describing evils inflicted by the burning sun. The passage runs: “For a fire (*’ēš*) is kindled (*qāḡ^ehā*) by my anger, and burns to the depths of Sheol; it devours the earth and its produce, and sets on fire (*ʿlaheṭ*) the foundations of the mountains. I will heap disasters upon them, spend my arrows (*hiṣṣay*) against them: [they will be] wasted with hunger (*m^ezē rā’āb*), devoured with plague (*ʿhūmē ræšæf*), and bitter destruction (*qæṭæf m^erīrī*)”. In the words of the author, this passage definitely refers to Yahweh’s flaming anger traversing the boundary of life and death and reaching into Sheol itself, for his fire – albeit not called sun – makes the same journey to the netherworld as Shapshu does, and it burns all in its path. It is not merely a question of a solarized Yahweh gazing into the land of the dead, but of the full burning capacity of divine anger. If genuinely archaic – an idea that is not *a priori* rejected – this passage actually weakens the whole concept of a gradual development by which Yahweh and his consuming fire replace a menacing solar deity. As observed, the word *qæṭæf* “destruction” links Deut 32:24 to Hos 13:14. In addition, however, it should be noted that the verb *qāḡah* “kindle” is met with in Jer 15:14: “in my anger, a fire is kindled (*’ēš qāḡ^ehā*)”; the verb *liheṭ* “devour, burn up” is found in Joel 1:19: “a flame has devoured (*læhābā lih^etā*) all the trees of the field”, and in Mal 3:19: “that day will burn them up (*w^elihaṭ ʿōlām*)”.

Regrettably, Wikander does not present specific criteria for his choice of passages, a circumstance that makes the actual scope of the whole investigation a bit unclear. A certain flaw is his neglect of some biblical passages that collocate destruction and the hot noonday; e.g., Jer 15:8 speaks about a destroyer at noonday (*šōḡēḡ baššohōrāyim*) and Ps 91:6 mentions a destruction that destroys at noon (*qæṭæb yāšūd šohōrāyim*). Also neglected are two rare poetic words for the sun indicating its harmful aspect, viz., *ḥammā* “sun” (in MH also “heat”, “fever”) derived form *ḥ-m-m* “be hot”, cf. *ḥōm* “heat”, as well as *ḥæræs* “sun” which – whether an

etymological cognate or not – also means “itch” (Deut 28:27). This said, the thesis has the merit of calling attention to lexical peculiarities indicating that shared hydrological conditions form the basis for more or less overt mythical notions connected with agrarian culture in the Levant.

Mats Eskhult, Department of Linguistics and Philology, Uppsala University

List of Contributors

Alexander Andrason, Ancient Studies, University of Stellenbosch, Private Bag X1, Matieland, 7602 South Africa
andrason@sun.ac.za

Erik J. Anonby, Department of French and School of Linguistics and Language Studies, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada, and Department of Linguistics and Philology, Uppsala University, Box 635, SE-751 26 Uppsala, Sweden
erik_anonby@carleton.ca

Ute Bohnacker, Department of Linguistics and Philology, Uppsala University, Box 635, SE-751 26 Uppsala, Sweden
ute.bohnacker@lingfil.uu.se

Mats Eskhult, Department of Linguistics and Philology, Uppsala University, Box 635, SE-751 26 Uppsala, Sweden
mats.eskhult@lingfil.uu.se

Geoffrey Haig, Lehrstuhl Allgemeine Sprachwissenschaft, Otto-Friedrich-Universität Bamberg, DE-96045 Bamberg, Germany
geoffrey.haig@uni-bamberg.de

Forogh Hashabeiky, Department of Linguistics and Philology, Uppsala University, Box 635, SE-751 26 Uppsala, Sweden
forogh.hashabeiky@lingfil.uu.se

Mahmoud Hassanabadi, University of Sistan and Baluchestan, Daneshgah Street, Zahedan, Iran, and Department of Linguistics and Philology, Uppsala University, Box 635, SE-751 26 Uppsala, Sweden
mahmoud.hassanabadi@yahoo.com

Pétur Helgason, Department of Linguistics and Philology, Uppsala University, Box 635, SE-751 26 Uppsala, Sweden
petur.helgason@lingfil.uu.se

Ayat Hosseini, Department of Language and Information Sciences, Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1, Komaba, Meguroku, Tokyo 153-8902, Japan
hosseini@phiz.c.u-tokyo.ac.jp

Carina Jahani, Department of Linguistics and Philology, Uppsala University, Box 635, SE-751 26 Uppsala, Sweden
carina.jahani@lingfil.uu.se

Martin Joachim Kümmel, Philosophische Fakultät, Friedrich-Schiller-Universität Jena, Zwätzengasse 12, DE-07743 Jena, Germany
martin-joachim.kuemmel@uni-jena.de

Stephen H. Levinsohn, SIL International, Stephen H. Levinsohn, SIL International, 60 Falcon Avenue, Bedford MK41 7DX, United Kingdom
stephen_levinsohn@sil.org

Corey Miller, Center for Advanced Study of Language, University of Maryland, 7005 52nd Avenue, College Park, MD 20742, USA
cmiller6@umd.edu

Somayeh Mohammadi, Azad University of Bam, Houzeh va Daneshgah Boulevard, Bam, Iran
somayemohammadi77@yahoo.com

Maryam Nourzaei, Department of Linguistics and Philology, Uppsala University, Box 635, SE-751 26 Uppsala, Sweden
maryam.nourzaei@lingfil.uu.se

Farideh Okati, University of Zabol, Emam Khomeini Boulevard, Zabol, Iran, and Department of Linguistics and Philology, Uppsala University, Box 635, SE-751 26 Uppsala, Sweden
farideh.okati@uoz.ac.ir

Ergin Öpengin, Lacito CNRS, Université Paris III, France, and Otto-Friedrich-Universität Bamberg, DE-96045 Bamberg, Germany
ergin.opengin@etud.sorbonne-nouvelle.fr

Lambert Schmithausen, Asien-Afrika-Institut, Universität Hamburg; priv.: Dobbertinweg 24, DE-21266 Jesteburg, Germany

Guiti Shokri, Department of Linguistics and Philology, Uppsala University, Box 635, SE-751 26 Uppsala, Sweden
g_shokri@hotmail.com

Madeleine Sultán Sjöqvist, The Hugo Valentin Centre, Uppsala University, Box 521, SE-751 20 Uppsala, Sweden
madeleine.sjoqvist@valentin.uu.se