

Secondary Publication



Schreiber, Laurentia; Haig, Geoffrey L. J.; Taheri-Ardali, Mortaza; u. a.

Contact-induced change in the morphosyntax of Turkic in Boldaji, Chahar Mahal va Bakhtiari Province, Iran

Date of secondary publication: 17.09.2026

Accepted Manuscript (Postprint), Article

Persistent identifier: urn:nbn:de:bvb:473-irb-117212x

Primary publication

Schreiber, Laurentia; Haig, Geoffrey L. J.; Taheri-Ardali, Mortaza; u. a. (2021): Contact-induced change in the morphosyntax of Turkic in Boldaji, Chahar Mahal va Bakhtiari Province, Iran, in: Turkic Languages, Wiesbaden: Harrassowitz, Vol. 25, Nr. 2, pp. 210–242, <https://doi.org/10.13173/tl.25.2.210>

Legal Notice

This work is protected by copyright and/or the indication of a licence. You are free to use this work in any way permitted by the copyright and/or the licence that applies to your usage. For other uses, you must obtain permission from the rights-holders.

This document is made available with all rights reserved.

Contact-induced change in the morphosyntax of Turkic in Boldaji, Chahar Mahal va Bakhtiari Province, Iran

Laurentia Schreiber, Geoffrey Haig, Mortaza Taheri-Ardali,
Erik Anonby

Schreiber, Laurentia & Haig, Geoffrey & Taheri-Ardali, Mortaza & Anonby, Erik. 2021. Contact-induced change in the morphosyntax of Turkic in Boldaji, Chahar Mahal va Bakhtiari Province, Iran. *Turkic Languages* 25, 210–242.

This paper investigates effects of language contact on a previously undocumented variety of Turkic (Southern West Oghuz) spoken by sedentarized nomads in the city of Boldaji, Chahar Mahal va Bakhtiari Province, south-west Iran. Located in an area of high linguistic diversity including the Iranian language Bakhtiari and the Charmahali variety of Persian, Boldaji Turkic exhibits numerous contact-induced features. Here, drawing on data collected using the *Atlas of the Languages of Iran* (ALI) questionnaire, we focus on variation in five primarily morphosyntactic features, several of which have not figured prominently in the literature on contact-induced change in Turkic: (i) restructuring of vowel harmony; (ii) the postposed emphatic definiteness marker *jōgāz*; (iii) an additional oblique pronominal stem based on *bela-* + possessive suffix; (iv) ‘can/be able’ expressions; and (v) a reduplicated progressive verbal form. In line with well-established criteria from contact linguistics, we assess the probability of contact-induced change for these five features by considering differences with other West Oghuz languages outside the area, namely Turkish and Azeri, to rule out an internal explanation. Unlike previous studies, which use standard-type Persian as the contact language, we show the importance of considering model structures in geographically contiguous languages, with a focus on Bakhtiari and the local Persian variety, to account for areal convergence. Finally, we consider the plausibility of various contact-based explanations in the light of our knowledge of related Turkic varieties in the area. Our data bring into question the common view that contact-induced change is preferentially associated with simplification; rather, we find a mix of structural changes involving both loss and gain of morphological matter, and patterns.

Keywords: contact-induced change, morphosyntax, Turkic of Iran, Southern West Oghuz Turkic, Qashqai (Kashkay)

Laurentia Schreiber, Department of General Linguistics, Otto-Friedrich-Universität Bamberg, Schillerplatz 17, 96047 Bamberg, Germany. Email: laurentia.schreiber@uni-bamberg.de. Geoffrey Haig, Department of General Linguistics, Otto-Friedrich-Universität Bamberg, Schillerplatz 17, 96047 Bamberg, Germany. Email: geoffrey.haig@uni-bamberg.de. Mortaza Taheri-Ardali, Department of English/Bakhtiari Studies Research Institute, Shahrekord University, Rahbar Blvd, Shahrekord, Chahar Mahal va Bakhtiari Province, Iran. Email: taheri@sku.ac.ir. Erik Anonby, School of Linguistics and Lan-

guage Studies (SLaLS), Carleton University, 1125 Colonel By Drive, Ottawa, Ontario, K2C 2L4 Canada. Email: erik.anonby@carleton.ca.

1 Introduction

Over the past fifteen centuries, Turkic and Iranic¹ languages have been in contact across a large expanse of central and western Asia, and the varieties spoken in present-day Iran bear witness to successive waves of contact influence (Golden 2006).² Bulut (2014, 2018a) provides the most detailed available overview of Turkic varieties of western Iran, most of which belong to a Western dialect group of Oghuz (or “Southwestern”) Turkic. However, documentation of the many varieties is relatively incomplete; notable exceptions include South Azerbaijani, mainly Azerbaijani of Tabriz (Dehghani 2000; Kiral 2001; Lee 2008; Erfani 2012; Hashemi Zarabad 2015), Qashqai (Kashkay)³ of Fars Province (Soper 1996; Csató 2005; Dolatkahh 2016, 2019; Bulut 2016), Sonqor Turkic (Bulut 2005), and Khalaj (Doerfer 1971, 1988; Kiral 2000; and ongoing work by E. Ragagnin and colleagues).

As Johanson (2002) and Bulut (2016) have pointed out for Iran, the contact relationship between Turkic and Persian has been largely asymmetrical: Turkic varieties have undergone contact-induced changes due to pressure from the Iranic lingua franca Persian, whereas changes in the other direction (for examples, see Stilo 1994, 2006) are less evident. It is nevertheless noteworthy that most previous research has essentially adopted Persian itself as the baseline of comparison in identifying contact-induced change in Turkic (see, for example, Kiral’s (2001) important study on

- 1 Following Windfuhr (1979) and Perry (1998), among others, we use the linguistically specific label “Iranic” rather than “Iranian” to avoid the double confusion of ignoring related languages located outside of Iran, and including languages within Iran which are unrelated but commonly referred to as “Iranian” languages in the literature.
- 2 We would like to express our sincere appreciation to all of the consultants who made this study possible, in particular Ebrāhim and Esmā’il Rahmāni of Boldaji, who shared their language with us and worked with us patiently, over many hours and on many occasions, to enable us to complete this study. Earlier versions of this paper were presented by the authors at the conference on Language Contact and Language Change in Western Asia (LCLCWA), Goethe-Universität Frankfurt, 10 March 2017, and at the 24th LIPP (Graduate School Language and Literature) Symposium “Language Contact: Yesterday – Today – Tomorrow”, Ludwig-Maximilian-Universität Munich, 22 June 2017. Various elements of the research were supported by SSHRC (Social Sciences and Humanities Research Council of Canada), Carleton University, Leiden University, Otto-Friedrich-Universität Bamberg and the Alexander von Humboldt Foundation. Transcription of phonological data follows the Iranian linguistics conventions set out in <https://carleton.ca/iran/> transcription, with the following phonemic symbols differing from their phonetic counterparts in IPA: *č* [tʃ], *ǵ* [g]/[ɣ], *ǰ* [dʒ], *š* [ʃ], *v* [v]/[v]/[w], *y* [j], *ÿ* [j̥] and *x* [χ].
- 3 In the literature, this term is also spelled Qashqā’i, Qashqay, and Kashkai. As for most Turkic varieties in Iran, the auto-glossonym for this variety is *torki*.

the Turkic of Tabriz). While the pre-eminence of Persian as lingua franca and language of administration and literature is undisputed, the emphasis on Persian obscures the fact that in all regions where Turkic languages are spoken in Iran, they are also in contact with Iranic languages other than Persian (e.g., Tatic languages in the Qazvin region, Kurdish varieties around Sonqor, or Bakhtiari in Chahar Mahal va Bakhtiari Province; see below). Prior to the 20th century, these other languages were languages of wider communication across extensive regions and have retained varying degrees of vitality down to the present; we can reasonably expect them to have exerted considerable influence on the respective regional varieties of Turkic. Any survey of Iranic influence on Turkic languages of Iran, then, remains incomplete without appropriate attention to the non-Persian languages of each region.

In this paper, we focus on a previously undocumented variety of Southern West Oghuz Turkic from Boldaji (Turkic *boldajî*), a predominantly Turkic-speaking rural community in eastern Chahar Mahal va Bakhtiari (C&B) Province in Iran. Although language contact effects on Southern West Oghuz, and in particular its Kashkay varieties, have already been extensively documented in the literature (Bulut 2003, 2006, 2014, 2016; Soper 1996; Csató 2005; Dolatkah 2016, 2019), previous research is largely based on comparison with Persian, as pointed out above. Our work considers specifically the influence of Bakhtiari, a Southwestern Iranic language (Anonby & Asadi 2014; Anonby & Taheri-Ardali 2018; Taheri-Ardali & Anonby 2019), on neighbouring varieties of Turkic. Along with the well-established contact-induced features of Turkic varieties spoken in Iran (cf. Bulut 2018), we are able to identify a number of features that appear to be related specifically to language contact with Bakhtiari.

The paper is organized as follows. In Section 2 we outline the theoretical background and the methodology of data gathering within the framework of the *Atlas of the Languages of Iran* (ALI), and Section 3 summarizes the sociolinguistic background of Turkic in C&B Province. Section 4 presents a series of candidate features for contact-induced change in phonology and morphosyntax, including some features that have hitherto not figured prominently in the literature on Iranic-Turkic contact. Section 5 discusses the implications of these findings from a more explicitly theoretical perspective, considering criteria for establishing contact influence, the issue of simplification under contact conditions, and the relevance of borrowing versus shift or imposition (see van Coetsem 2000; and subsequently Lucas 2012). Section 6 summarizes, together with some brief methodological considerations, the main conclusions and research prospects opened up through this study.

2 Data sources and theoretical considerations

This study has been carried out in the context of the *Atlas of the Languages of Iran* (ALI) research programme (Anonby, Taheri-Ardali et al. 2015–2021; Anonby, Taheri-Ardali & Hayes 2019; <http://iranatlas.net>). A recent overview of the work

taking place specifically with Turkic languages in ALI appeared in Anonby et al. (2020).

Field research for the present study was conducted in C&B Province between May 2015 and September 2016. The first step was to conduct initial assessments of language and dialect distribution for each of the province's some 800 cities and villages (Taheri-Ardali 2020; Anonby, Taheri-Ardali & Stone, in press). This allowed us to gain a geographically detailed picture of the language situation and resulted in an overview of language distribution, as presented in Figure 1 (in Section 3) below. It also helped us to select research sites for language data collection from across the province.

In addition to learning more about the language situation, we collected notes regarding internal diversity among the province's dialects. After completing a language distribution map for the entire province (Taheri-Ardali et al. 2016), including notes on communities and districts with high levels of linguistic diversity, we were able to choose 30 language varieties from 26 locations across the nine *shahrestān* (provincial sub-districts). Sites were selected to highlight situations of three types: (i) homogeneous rural dialects in communities with a single dominant language; (ii) bi- or multilingual communities with a balanced proportion of speakers, in order to explore patterns of language contact and change; and (iii) district centres, which we projected would show the greatest impact of Persian on speaker proportions and linguistic structures.

Following pilot studies on several Turkic varieties of C&B Province (e.g., Anonby, Schreiber & Taheri-Ardali 2020), we have chosen the variety of Boldaji as the focus of this paper because the vitality of the language is high here, despite being spoken alongside Iranic languages (see Section 3), and because it exhibits striking areally-defined typological features. The present data were gathered by means of a pilot questionnaire provided in the *Atlas of the Languages of Iran*, developed specifically for the descriptive analysis of language variation in Iran, and which has been recently updated with a refined version (Anonby, Taheri-Ardali et al. 2015–2021). This questionnaire is based on translation tasks from Persian and was administered orally in recorded interviews. This tool is useful for efficiently gathering a uniform set of various types of data: sociolinguistic context, lexicon, phonology, and morphosyntax. Still, we are aware that the use of a translation task from the contact language Persian will impact the data. For example, speakers may unintentionally or even deliberately calque structures of the prestige variety Persian either to accommodate the (assumed) expectation of the researcher or to demonstrate the productivity of their language. (For a discussion of other possible limitations of methodology, see Section 6.) The questionnaire for Boldaji was administered and recorded by co-author Mortaza Taheri-Ardali and colleague Peiman Pishyardehkordi in Boldaji in August 2016. The data were provided by two 29-year-old male consultants, twins from Boldaji. These consultants have always lived in the community, but they work as teachers in Lordegān, a predominantly Bakhtiari speaking area. They are multilingual in Turkic (L1), Persian (L2), and Bakhtiari (adult L3 acquisition).

A number of authors have outlined criteria for identifying contact-induced change (Thomason 2001: 91–95; Winford 2005; Poplack & Levey 2010; Lucas 2012), but it is important to bear in mind that contact-induced changes generally interact with internally-motivated developments. We should not expect a simple linear chain of cause and effect. Rather, a contact situation yields a set of conditions which will nudge the structural developments within the recipient language in a certain direction, making a particular outcome more likely than it would have been in the absence of the contact language. This can be difficult to demonstrate in an absolute sense and must involve consideration of the overall structure of both recipient and donor languages, as well as the contact situation. We also need to consider the general plausibility of contact-based explanations in light of our knowledge of contact-induced change in Turkic (Johanson 2002) as well as known grammaticalisation pathways, both generally (Heine & Kuteva 2002; for borrowing hierarchies of grammatical structures in language contact, see also Matras 2007) and in particular for genealogically related varieties of linguistic areas such as Turkic of Iran (Doerfer 1988; Kırıl 2000; Bulut 2003, 2005, 2016).

Here we will focus on the following five features: (i) restructuring (but not complete loss) of vowel harmony in plurals, past suffixes, and infinitives; (ii) a postposed emphatic definiteness marker *joğaz*, possibly arising through reanalysis of an existing diminutive, which replicates a Bakhtiari model; (iii) an additional oblique pronominal stem based on *bela-* + possessive suffix (as noted in Bulut 2003), which has structural parallels in some Iranian languages; (iv) a reduplicated continuative verb construction, e.g., *galillar galalar* ‘they are coming’, again possibly modelled on similar constructions in Bakhtiari; and (v) ‘can/be-able’ expressions consisting of a preposed auxiliary *bilir* and a fully inflected lexical verb.

3 Sociolinguistic background

C&B Province is nestled in the heights of the Zagros Range in western Iran, with the mountains opening down onto the Iranian Plateau. The topography is reflected in the linguistic situation. Whereas the Southwestern Iranian language Bakhtiari dominates the mountainous areas that cover most of the province, two other main linguistic groups are intermingled in the lower areas of the north-east: Chārmahāli, a distinct Southwestern Iranian dialect group more closely related to Persian; and Turkic, spoken in over 40 cities and villages in the province (Taheri-Ardali et al. 2016; Anonby, Taheri-Ardali & Stone, in press). All Turkic varieties spoken here fall within the West (or Saljuqian) branch of Oghuz, with representatives of its Central sub-branch found in a few locations but the majority of varieties aligned with the Southern sub-branch (Taheri-Ardali and Anonby, field notes, 2019). The Turkic variety of Boldaji, alongside Kashkay, falls into this latter group. For the distribution of languages in C&B Province see Figure 1.

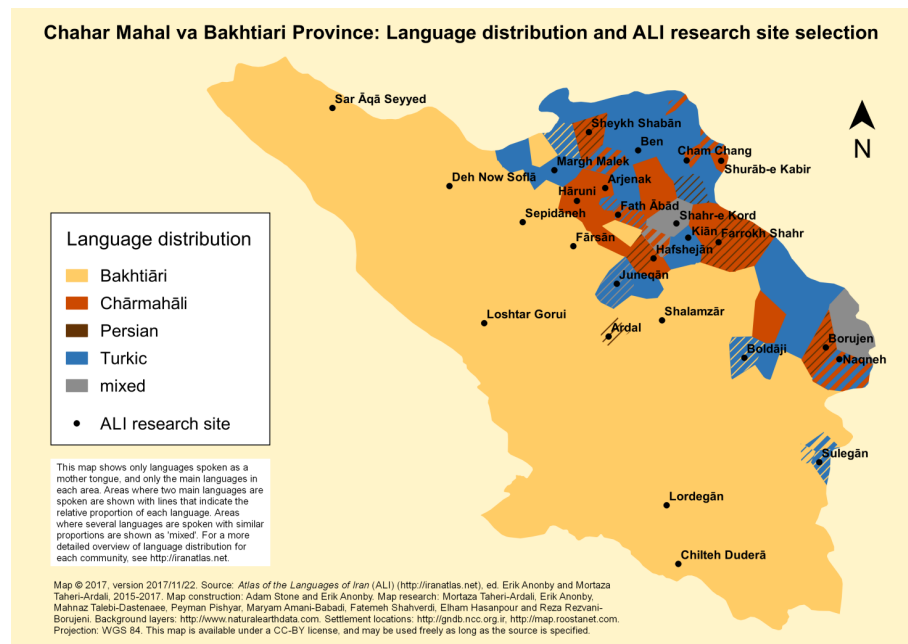


Figure 1: Language distribution in C&B Province (from Erik Anonby & Mortaza Taheri-Ardali, *Atlas of the Languages of Iran* (ALI), <http://iranatlas.net>, 2015–2021)

Like many Turkic peoples across south-western Iran, the Turkic community of Boldaji was traditionally nomadic, but has become sedentary over the course of the last century. According to the speakers consulted for this study, their community is ethnically Kashkay, but no longer maintains close relations with the larger body of the Kashkay tribe in Fars Province. The distance of this connection in space, time, and culture is evidenced by the lack of a precise position within the hierarchy of the Kashkay tribal confederation. And while the Turkic language of Boldaji is aligned with Kashkay in the Southern sub-group of West Oghuz Turkic, it shows some distinctive structural characteristics, as this study will make clear.

We can define Boldaji as a “Turkic community” in the sense of a Turkic-dominant multilingual community, based on the criterion of the most frequently acquired L1 being Turkic (see also the explanation below the legend in Figure 1). Boldaji is a rural town of almost 12,000 inhabitants (Iran Statistics Centre 2016) where an estimated 85% of people speak Turkic as their mother tongue.⁴ Approx. 5% of people speak Bakhtiari (local auto-glossonym: *lori*), which is the dominant contact language in and around the town of Boldaji, as L1 (Taheri-Ardali 2016). Another 10%

4 The percentages presented in this paragraph are based on self-assessed estimations of the consultants elicited as part of the ALI questionnaire (see Section 2).

of the population are native speakers of Chārmahāli or local types of Persian. Although the prevalence and vitality of Turkic varies greatly between communities in C&B Province, there is a general tendency whereby Tehran-type Persian, which was previously spoken only by immigrants to the province, is emerging as a mother tongue in all communities (see Taheri-Ardali 2015 for a detailed case study). In the case of Boldaji, our consultants estimated that about 10% of the population is now acquiring Persian as an L1. All children in Boldaji are bilingual at the point of school entry at the latest. Still, the language vitality of Turkic in Boldaji can be assessed as relatively high, based on the fact that the language is passed on to the children in most of the homes: consultants estimate that about 90% of the population have acquired their parent's L1 as their first language.

Regarding individual multilingualism, the consultants estimated that 60% of the population in Boldaji are fluent in Persian. About 20% are bilingual in both Bakhtiari and Turkic, with the learning of Turkic being proportionally more common among native speakers of Bakhtiari than the learning of Bakhtiari by native Turkic speakers.

4 Candidate features for contact-induced change

4.1 Restructuring of vowel harmony⁵

The Turkic variety of Boldaji (Boldaji Turkic, henceforth abbreviated BT) has fewer vowel-harmonic distinctions than Turkic varieties outside Iran.⁶ While vowel harmony is absent from the infinitive suffix, most other suffixes have some kind of “small” vowel harmony, sensitive to backness of the stem-final vowel but impervious to its specification for roundness. Back stem-final vowels trigger /o/, /ɑ/; and front vowels trigger /e/, /a/. /o/ versus /e/ are selected for possessive suffixes other

5 The editors of the journal have put forward that consonants also have an effect on the outcomes of the historical process of levelling and reconfiguration of vowel harmony and that therefore the term “sound harmony” would be more adequate. Although we acknowledge this phenomenon more generally, we maintain the narrow term “vowel harmony” since we have observed little of this effect in the data corpus, and consonants are not at the core of our analysis here.

6 As an anonymous reviewer points out, that does not necessarily imply that earlier vowel-harmonic distinctions have been “lost” in BT, because in older varieties of Turkic, a number of the relevant suffixes were invariable (so-called unique suffixes, see Bulut 1997: 52). BT may therefore simply have preserved an older state, rather than losing vowel harmony. On the other hand, it can be argued that contact influence from Iranic was instrumental in the maintenance of non-harmonic suffixes, which in other varieties of Turkic developed vowel harmony. While we cannot resolve this issue here, it seems uncontroversial that there is considerable phonological influence from neighbouring Iranic languages in the vowel system of BT, and this will undoubtedly have affected vowel harmony, regardless of its initial state in BT.

than 3SG and /ɑ/ versus /a/ for plural, ablative, locative, dative and negation suffixes (cf. “type A” in Dolatkhan 2016: 29).

The alternations in the BT system, as compared to Standard Turkish, pattern in detail as follows:

A. Where Standard Turkish⁷ has a “small” (front/back) vowel harmony, in BT this either completely lacks vowel-harmonic variation, as in infinitives (*vur-mağ* ‘to hit’ vs. *ged-mağ* ‘to go’), or exhibits vowel harmony, but with some variation in the output vowels, as in plural (*arvad-lar* ‘women’ vs. *geççi-lar* ‘goats’), ablative (*bazar-dan* ‘from the market’ vs. *ey-dan* ‘from the house’), locative (*madrassa-de* ‘at school’ vs. *şiraz-da* ‘in Shiraz’), (presumably) also dative (*bazar-a* ‘to the market’), or negation (*ger-ma* ‘don’t see’ vs. *vur-ma* ‘don’t hit’).

B. Where Standard Turkish has the more complex “big” vowel harmony, sensitive to backness and roundness: in BT this is either (a) absent, with the relevant suffixes showing a uniform closed unrounded /i/, as in the 3SG of the possessive (*kitab-i* ‘his/her book’ vs. *maşin-i* ‘his/her car’), the 3SG of the simple past (see Table 1),⁸ and presumably accusatives (*gapu-yi* ‘the door’), or (b) has been adapted to the small type used in A, such as the possessive suffixes other than 3SG (*kitab-om* ‘my book’ vs. *maşin-em* ‘my car’). It is possible that complete uniformity, i.e., option (a), may be preferred when the affix vowel is word-final.

1SG	<i>vur-dom</i>
2SG	<i>vur-doyñ</i>
3SG	<i>vur-di</i>
1PL	<i>vur-dog</i>
2PL	<i>vur-doyoz</i>
3PL	<i>vur-dolar</i>

Table 1: Past tense verbal paradigm of the verb ‘to hit’ in BT

Similar morphophonological features have been found in the genealogically related variety of Sonqor Turkic, a Southern West Oghuz variety of the Afshar group, spoken in Kermanshah Province of Iran. There, we find an invariable open back unrounded infinitive *-mağ*, ablative and locative (both *-de/dæ*, *-ne/næ*, *-re/ræ*) as well as the dative *-æ/e* displaying a middle front unrounded vowel, invariable high un-

7 In this article, we use the term “Standard Turkish” to refer to the standardized variety of the national language of Turkey, as codified in grammars such as Göksel & Kerslake (2005). Our use of the term “Turkish of Turkey” comprises different varieties of Turkish spoken in Turkey.

8 Following a comment by the editors, we would like to highlight that, in contrast to many other Turkic varieties in Iran, the vowel written as /o/ throughout the BT data presented here is indeed a back round mid vowel /o/ rather than its near-high lax counterpart /ʊ/ [ʊ] in all contexts, including suffixes and other non-initial syllables.

rounded *-mi(š)* past perfect, and an invariable plural *-le(r)* (Bulut 2005: 245). While palatal harmony (sensitive to backness) seems to be disregarded in Sonqor Turkic, labial harmony (sensitive to roundness) – a more recent development in Western Oghuz – probably never did develop in Sonqor Turkic (Bulut 2005: 245). Bulut (2005: 249) suggests that the differences in the vowel system of Sonqor Turkic (compared to Turkic varieties outside Iran) are due to Iranian influence by means of copying Iranian vowels into the Turkic system and, thus, disturbing it, especially with regard to roundness. As pointed out by Bulut (2005: 249), similar analyses of Iranian-Turkic contact have previously been put forward in relation to Uzbek-Tajik contact situations by Boeschoten (1983) and Johanson (1986).

We follow Bulut (2005) in arguing that the restructuring of vowel harmony in BT is related to contact-induced changes in the overall architecture of the vowel system. The (assumed) perfect symmetry of the 8-member Turkish system, exhibiting all possible combinations of three binary features ($\pm$ back, $\pm$ high, $\pm$ round) is disturbed by the introduction of an additional back vowel /ɑ/, because it leaves /ɑ/ in indeterminate position with regard to the feature of backness (cf. Bulut 2005: 251). The wholesale loss of front rounded vowels /ø/ and /y/ in some Turkic varieties in contact with Iranian leads to a collapse of original labial harmony, since these vowels generally merge with their non-rounded equivalents throughout (Bulut, p.c., 2014: 35–36, 2018: 411). Furthermore, the loss of contrast between /e/ and /i/ in Kashkay is possibly due to Iranian influence (Bulut 2016: 248). In general, the changes in vowel harmony in Turkish varieties in contact with Iranian, and Kashkay in particular, are among the most complex systemic changes in the phonology and require more detailed research.

4.2 The postposed emphatic definiteness marker *joğaz*

BT exhibits a postposed definiteness marker *joğaz*, attested with singular nouns, which has not been previously discussed in the literature (but see Anonby, Schreiber & Taheri-Ardali 2020; an anonymous reviewer has also noted the presence of this suffix with similar functions in texts from “Turkmen in Iran”, but we have been unable to substantiate this). Historically, Turkic lacks any dedicated inflectional marking of definiteness or specificity, except that these functions are involved in the use of the accusative suffix in Differential Object Marking (DOM). In BT, however, *joğaz* is used in emphatic and explicit definiteness contexts, and is generally attested in our data in response to the colloquial Persian definite suffix *-e* (e.g., P. *moalem-e* ‘the teacher’). Examples (1)–(3) below are illustrative. We provisionally assume that *joğaz* is a suffix rather than a clitic or particle, because it precedes the accusative suffix in (3); hence we gloss it as part of the host word. However, this analysis conflicts with the position of *joğaz* following the plural suffix in (4b) below.⁹ Further-

9 We thank an anonymous reviewer for pointing out to us that in Sonqor Turkic, the definiteness marker *-aka*, borrowed from Kurdish, also precedes the plural suffix (Bulut 2005: 254), just as in Central Kurdish.

more, it appears to be impervious to vowel harmony. The analysis as suffix is therefore provisional; this requires further research.

Example (1) illustrates *joğaz* with the head noun of a relative clause, and (2a/b) with comparatives and superlatives. When attached to a definite direct object, the accusative suffix follows *joğaz* (3), which, as mentioned, is suggestive of suffix rather than clitic status for *joğaz*. Unfortunately, we lack any occurrences of *joğaz* in BT with plural direct objects, even when the plural noun phrase can be assumed to be notionally definite, e.g., *siblar-i yedi* ‘s/he ate the apples’. A plural form of *joğaz* should be possible, however, since the analogous form *jağaz* is attested for plural in data from a Bakhtiari-Turkic bilingual speaker of the Turkic variety of Juneqan elsewhere in C&B Province (Anonby, Schreiber & Taheri-Ardali 2020), although it appears to be rare there as well and its variability indicates the instability of this pattern (4a/b). Singular direct objects with a definite interpretation are still found without *joğaz*, e.g., *sib-i yedi* ‘s/he ate the apple’. This suggests that *joğaz* is not (yet) an obligatory definiteness marker, but rather an optional definiteness marker, similar to recent analyses of definiteness marking in Central Kurdish (Haig 2018) and Balochi (Nourzaei 2020); this topic is briefly taken up below. However, *joğaz* seems to be obligatory in contact-induced Iranic type relative clauses with *ki* that obligatorily require a suffix *-ī* on the head noun.¹⁰ In other Turkic varieties in Iran, the head can be marked, for example, by a demonstrative (see Type B in Bulut 2006: 168) or by the definite suffix *-aka* (where Turkic is in contact with Kurdish; for a detailed discussion of relativization strategies of Turkic varieties in contact with Iranic see Kiral 2000; Bulut 2006; cf. ex. 1).

- (1) *moalem-joğaz* *ke* *mašhad-dan* *galdi*
 teacher-DEF PRT Mashhad-ABL come.PST.3SG
 ‘the teacher that came from Mashhad’
- (2) a. *beyig-tar* *ey-joğaz*
 big-COMP house-DEF
 ‘the bigger house’
- b. *bo* *ey* *beyig-tarin* *ey(-joğaz)* *de*
 DEM house big-SUP house(-DEF) COP
 ‘This is the biggest house.’

10 The precise nature of this suffix is controversial in Persian linguistics; however, cross-linguistically the association between definiteness and relativization is well established, though its precise nature remains somewhat controversial. This is not to discount the possibility of non-restrictive relative clauses, which undoubtedly exist. However, they often exhibit structural properties distinct from those found with restrictive relative clauses, and in some frameworks are actually considered a discrete phenomenon; cf. the summary discussion in Haig (1998: 10–13).

- (3) *sib-joğaz-i* *yedi*
 apple-DEF-ACC eat.PST.3SG
 ‘S/he ate the apple.’
- (4) Turkic of Juneqan (Anonby, Schreiber & Taheri-Ardali 2020: 22, ex. 3)
- a. *sib-jağaz-ley-i* *yiye*
 apple-DEF-PL-ACC eat.PRS.3SG
 ‘He is eating the apples.’
- b. *sib-la-jağaz-i* *yeyi*
 apple-PL-DEF-ACC eat.PST.3SG
 ‘He ate the apples.’

While many details of the function of *joğaz* remain to be elucidated, it is nevertheless appropriate to consider the question of its origins. The most plausible candidate for the etymological source is a reanalysis of the Turkic diminutive suffix *-cAğIz*, pronounced [ˈdʒɒɣɒz] (C. Bulut, p.c.). Elsewhere in Turkic, this suffix is not merely an indication of “small size”, but rather an expression of the speaker’s emotional engagement and evaluation, typically indicating endearment, compassion and empathy (or negatively connotated with disdain), for example, Standard Turkish *çocuk-cağız* ‘the poor/dear child’ (Göksel & Kerslake 2005: 60).¹¹ The use of diminutives to express evaluative content is well-documented cross-linguistically (Ponsonnet 2018). If this analysis is correct, then we assume that the evaluative component of the original marker has undergone a semantic shift along the following lines:

empathy/endearment > familiarity > definiteness (i.e., assumed discourse-recoverability)

Although a development like this has not previously figured in the extensive literature on the grammaticalization of definiteness (Heine & Kuteva 2002; De Mulder & Carlier 2015), more recently, two cases exemplifying a development along these lines have come to light. First, Pakendorf & Krivoschapkina (2014) describe the use of diminutives and augmentatives to express definiteness in Even, a North Tungusic language of northeastern Siberia, which reinforces the claim of a conceptual link between diminutive (in their terms, evaluative) morphology and definiteness marking. Second, and more pertinent in the present context, is the presence of definiteness marking in a number of West Iranian languages, some of which can be plausibly related to an earlier diminutive or evaluative suffix, e.g., Balochi *-ok* (Nourzaei et al. 2015: 32; Nourzaei, to appear) and Central Kurdish *-aka* (Haig 2018). A similar development is probably behind definiteness marking in Bakhtiari, which is achieved via a suffix *-(e)ke* (Anonby & Asadi 2014: 91–93; Anonby & Taheri-Ardali 2018: 452). We should also note that in both Balochi and Central Kurdish,

marking of definiteness through these suffixes is not fully obligatory (also in Bakhtiari), and is subject to both structural and functional constraints. For example, it does not occur with plural-marked nouns in the Koroshi dialect of Balochi (Nourzaei, 2020), which seems to echo a similar constraint in BT. This is a particularly interesting parallel, because a constraint against definiteness marking with plural nouns appears to be cross-linguistically rare (see Becker 2018).

Our preliminary analysis thus assumes that while the etymological source of the BT definiteness marker is the Turkic diminutive suffix, its development towards a marker of definiteness replicates a pattern that is attested in neighbouring West Iranian languages of the region, including Bakhtiari. If this analysis is correct, this would be a particularly interesting case of “pattern replication” (Matras & Sakel 2007; “selective copy” in Johanson 2002), based on a cross-linguistically rare pattern: i.e., the speakers of the replicating language identify the functionally parallel item in their language, *-(e)ke* in the case of Bakhtiari, and link it to a novel function in a manner that replicates the assumed grammaticalization path in Bakhtiari. This is rather different from the other instances of Turkic languages adopting definiteness markers through contact with Iranian, which generally involve “matter borrowing” (Matras & Sakel 2007). For example, Turkic varieties in contact with Kurdish in Iraq (Bulut 2018b) and western Iran (Bulut 2014) borrow the actual definiteness suffix of Central Kurdish *-aka* (Bulut 2014). Similarly, a definiteness marker *-(y)aki* is documented in some Kashkay varieties (as discussed in Dolatkhah 2016, 2019; Dolatkhah, Csató & Karakoç 2016; Taheri-Ardali, in preparation). This form is clearly likewise borrowed from neighbouring West Iranian languages, including Balochi, Gorani, and Central and Southern Kurdish, which exhibit a variety of cognate definiteness marking suffixes (e.g., *-aka*, *-aku*, *-ok*), most likely related to Old Iranian diminutive and evaluative morphology (see Haig & Mohammadirad 2019, and Nourzaei 2020 for discussion). Kuribayashi (2006: 64), who relates the Iranian source forms to the numeral ‘one’, e.g., Persian *yek* and its cognates, is clearly mistaken. In general, then, there is a well-documented internal development within West Iranian whereby definiteness marking develops from some form of diminutive or evaluative morphology, and this appears to have affected Turkic contact varieties in two ways: either through matter borrowing (borrowing of the actual Iranian definiteness suffixes into Turkic, e.g., the case of Kashkay *-(y)aki*, as discussed in Dolatkhah 2016), or through a more indirect process involving a parallel development of inherited Turkic diminutives towards definiteness marking. Note that both pathways yield a superficially similar output: suffixal marking of definiteness.

4.3 Double set of strong and weak object pronouns: Additional oblique pronominal stem based on *belə* + possessive suffix

In BT, we found two sets of object pronouns. Alongside the inherited Turkic set of pronouns, there is an additional emphatic oblique pronominal stem based on *belə* + possessive suffix (see Table 2) which has been attested elsewhere in Turkic under

Iranic influence (see Doerfer 1988 on *bilä*; Bulut 2003, 2007: 174 (on Iraqi Turkmen), 2016, 2018a on *bilä/bile*; Dolatkah 2016, 2019: 48 on *belä* in Kashkay; Ragagnin 2016; see also Anonby, Schreiber & Taheri-Ardali 2020 on Turkic of Juneqan).¹²

	Strong object pronoun	Default object pronoun
1SG	<i>bela-m</i>	<i>mani</i>
2SG	<i>bela-ỹ</i>	<i>sani</i>
3SG	<i>bela-seni</i>	<i>oni</i>
1PL	<i>bela-mez</i>	<i>bizi</i>
2PL	<i>bela-ỹez</i>	<i>sizi</i>
3PL	<i>bella-y(e)ni</i>	<i>ollari</i>

Table 2: Strong and weak object pronouns in BT

A possible contact model is the use of a dummy pronominal stem, to which possessive endings are added and – depending on the language – also case forms, which are used as non-subject pronouns in various West Iranian languages. For example, in one variety of Gorani, the invariable stem is *wan*= (ex. 5), and in Delvari, the invariable stem is *si*= (ex. 6).¹³

- (5) Gorani (Mahmoudveysi et al. 2012: 56, ex. 169)
m-wā-y *ča* *ma-t-ī* *wan = im*
 IND-say.PRS-3SG what IND-give-2SG wan=1SG
 ‘(He) says: “What will you give to me?”’
- (6) Delvari (Haig & Nemati 2013)
ersond = ešū *si = t*
 send.PST=3PL si=2SG
 ‘They sent you.’

However, both of these languages are outside the immediate region of Boldaji, so we can only cite this as a structural possibility that is well-attested in West Iranian, rather than as direct contact influence.

12 An anonymous reviewer suggests that the 1SG, 1PL and 2PL strong object forms in Table 2 would also carry a vocalic case marker, possibly schwa. Our data were insufficient to confirm this observation, though it is possible that the expected case marker was absent or not perceptible in the contexts elicited.

13 The base *si*= in Delvari is most likely related to a preposition with similar form, and fairly general benefactive semantics, as found in other West Iranian languages. However, this point is not at issue here, and we do not assume that the etymologies of these “dummy pronouns” are transparently available to speakers.

The *bile-/belä-* pronoun described in the literature (i.e., *beläsi* for 3SG and *beläläri* for 3PL) differs from the BT forms in Table 2. The BT data suggest that the third person form has generalized an accusative ending to become the default form for the third person pronouns, regardless of actual function. For first and second persons, the *bela-* stem remains unmarked for case, for example: *gerdi belamez-Ø* ‘he saw us’. However, in other Turkic varieties of C&B Province such as in that of Margh Malek, all *bela-* stems are inflected for accusative – which is also what Bulut (2014) reports for the form *bilä*. Furthermore, Dolatkah (2016: 75) reports a nasal form of the possessive suffix for the second person (i.e., *beläj* for 2SG and *beläjiz* for 2PL) which is retained in the current BT data only in the form of a nasalised glide.

Bulut (2014: 44) reveals that *bilä* occurred already in 17th century Turkic texts. It can be used in direct and indirect object position (ex. 7), where it inflects for accusative, dative, ablative, and instrumental, or combines with *için* ‘for’ in beneficial constructions, but never as a subject. In addition, the BT data exhibit its use as a resumptive pronoun following a relative clause (8) (cf. Dolatkah 2016: 75). The *bela-* stem is in general postposed after the predicate,¹⁴ though its position does not seem to be fixed, as preposing is also possible (9).

- (7) *bičāg-onan kasdem belasani*
 knife-with cut.PST.1SG PRON.ACC.3SG
 ‘I cut it with a knife.’
- (8) *moalem-joğaz ke mašhad-dan galdi*
 teacher- DEF PRT Mashhad-ABL come.PST.3SG
gerdem belaseni
 see.PST.1SG PRON.ACC.3SG
 ‘I saw the teacher that came from Mashhad.’
- (9) *istiram belaȳ / belaȳ istiram*
 want.PRS.1SG PRON.ACC.2SG
 ‘I love you.’

The stem appears to be etymologically related to Standard Turkish *böyle* ‘in this manner’ (*bela/bilä* in Azeri, Soper 1996), an adverb which can indeed carry possessive suffixes, although Bulut (2003: 325) also mentions a possible relation to an archaic New Persian term *bile*, which includes ‘side’ among its many meanings. A relation to the adverb *bile* ‘even’ in Standard Turkish has also been suggested (see Bulut 2003: 325). A factor in favour of the first proposed etymological source is that

14 The post-verbal position of the *bela-* stem is a feature also shared with the comparable pronominal stems in Gorani and Delvari mentioned in examples (5) and (6).

in the variety of Ben in C&B Province the phonological form of *bela-* alternates phonetically between [bɛlæ] and [bɔlæ].

Dolatkah (2016) suggests, for the emergence of the *bela-* construction, a contact scenario originating in a Persian structure: prepositional case marker + epenthetic h + possessive suffix, e.g., *be-h-eš* ‘to him’, cf. *belä-sin-a* ‘to him’ (Dolatkah 2016: 77). Note that the /h/ is phonologically conditioned in order to prevent the hiatus of /e/+e/. However, in this scenario the carrier of the Iranian person marker is the preposition *be*, whereas in the Turkic forms, there is a dummy stem followed first by person, then by case. What both constructions share is that the person value of the pronoun is expressed through some kind of possessive suffix or clitic. There is also a similarity between the use of *bela-* in (7) and the emphatic Persian construction *man-rā did* ‘s/he saw me’, which is more commonly used than the unmarked form *did=am* ‘s/he saw me’, since the latter is homophonous with ‘I saw’.

4.4 A reduplicated progressive verbal form

Bakhtiari of Ardal and nearby BT share an unusual grammatical construction, namely a reduplicated progressive verbal form, which is not reported in other dialects of Bakhtiari (cf. Anonby & Taheri-Ardali 2018: 461) and Turkic. The construction in BT consists of two inflected verbs: the first in present tense and the second in optative form (for the paradigm see Table 3).

	Imperfective/ present continuous	Past continuous
1SG	<i>ged-ir-am ged-em</i>	<i>ged-ir-d-em ged-em</i>
2SG	<i>ged-ir-aỹ ged-aỹ</i>	<i>ged-ir-d-eỹ ged-aỹ</i>
3SG	<i>ged-ir ged-a</i>	<i>ged-ir-d-i ged-a</i>
1PL	<i>ged-ir-ag ged-ag</i>	<i>ged-ir-d-eg ged-ag</i>
2PL	<i>ged-ir-aỹez ged-aỹez</i>	<i>ged-ir-d-eỹez ged-aỹez</i>
3PL	<i>ged-il-lar ged-ala</i>	<i>ged-ir-d-ela ged-ala</i>

Table 3: Reduplicated present and past continuous verbal paradigm of the verb ‘to go’ in BT

The reduplicated form appears in the BT data in both the present and past verbal paradigms (i.e., with the first verb in past tense and the second in the optative) for motion verbs (i.e., ‘go’, ‘come’, and ‘fall’, e.g., *diširam dišem* ‘I am falling’), but only in the past of the verbs ‘see’ and ‘hit’, and not for the verb ‘sleep’. Furthermore, this reduplicated form is not used in periphrastic verb constructions; rather, it

seems to alternate with the default present tense form (10a/b).¹⁵ We still lack testing in negative contexts. Note that these forms were elicited in response to colloquial Persian progressives, e.g., *dāram miram* ‘I am going’. The speaker associated (10a) with the Persian present progressive *dāram miram* ‘I am going/about to go’, and (10b) with the unmarked present *miram* ‘I go/am going’.

- (10) a. *ali-yenan gediram gedem bazar-a*
 Ali-with go.PRS.1SG go.OPT.1SG market-DAT
- b. *ali-yenan gediram bazar-a*
 Ali-with go.PRS.1SG market-DAT
 ‘I am going to the market with Ali.’

Considering the contact languages in the area, the most likely candidate model for this construction appears to be Bakhtiari (Ardal and Lordegān varieties, M. Taheri-Ardali, field notes), where a kind of reduplicated structure is used with ‘come’ and ‘go’ (but not ‘fall’) expressing continuous action, e.g., *irahdom berom* (Ardal, Lordegān) versus *iram berom* (Dastenā) ‘I was going’ (see Table 4 for present forms). An invariant particle *hey* is used in present forms, and dropping it yields a prospective: *erom berom* ‘I am about to leave’. The second verb form (*berom* etc.) is identical to the present subjunctive (and is also used in the past continuous).

	‘you (etc.) are going’	‘you (etc.) are coming’
1SG	<i>hey erom berom</i>	<i>hey iyam biyam</i>
2SG	<i>hey eri beri</i>	<i>hey iyey biyey</i>
3SG	<i>hey ere bere</i>	<i>hey iya biya</i>
1PL	<i>hey erim berim</i>	<i>hey iyeym biyeym</i>
2PL	<i>hey erin berin</i>	<i>hey iyeyn biyeyn</i>
3PL	<i>hey eren beren</i>	<i>hey iyan biyan</i>

Table 4: Reduplicated present continuous verbal paradigm ‘go’ and ‘come’ in Bakhtiari, Ardal variety (M. Taheri-Ardali, field notes)

Typologically, some languages use reduplication to express continuative meaning. Reduplication for aspectual purposes is not fully alien to Turkic, either (cf. manner adverbials in Standard Turkish: *ağlaya ağlaya* ‘crying, sobbing’).

15 An anonymous reviewer suggests that the forms in (10a/b) are actually not interchangeable but that rather (10a) has a progressive reading while (10b) has a habitual reading. This is in line with the potential semantics of the Persian forms used as the basis for elicitation of the Turkic construction.

The reduplicated construction found in BT has not been mentioned in the literature so far, but other Turkic varieties of C&B Province exhibit different strategies, some of which have previously figured in the literature, to express progressive aspect periphrastically. For instance: (i) Turkic of Kiān and Margh Malek use the existence noun *var* ‘existent’ inflected for person and tense in present and past;¹⁶ (ii) Kiān Turkic exhibits a periphrastic progressive/prospective construction for the verb ‘fall’ consisting of the verb *istemek* ‘want’, which is used to express future/prospective meaning and inflects for person and tense (both present and past), with a second verb in the optative/voluntative, *isir duş-sun* ‘he is falling’, *isir-di duş-sun* ‘he was falling’; (iii) a Kashkay variety of Shiraz uses periphrastic *innā* to express progressive meaning. It is striking that Turkic varieties of the area apply periphrastic modal constructions to express mood or aspect, such as innovated progressives, following Iranic patterns, e.g., the restructuring of the colloquial Persian model *dāštan* ‘have’ + present indicative (Csató 2005; Bulut 2016: 266); note also links to prospective aspect (cf. the contributions in Korn & Nevskaya 2017).

4.5 ‘Can/be able’ expressions

For BT, a type of ‘can/be able’ expression has been attested whose form differs from what has been previously described (cf. Doerfer 1988; Bulut 2000, 2018a: 425; Kırıl 2001, 2005: 288–291; Dolatkhah 2016 on modality in Turkic in Iran). In general, ‘can/be able’ expressions in Turkic are formed by the main verb stem and a grammaticalized modal suffix historically derived from the univerbation of lexical verbs with the verb *bil-* ‘to can/be able’ (ex. 11).

- (11) Standard Turkish
 <onu görebilir>
 PRON.ACC.3SG see.ABIL.PRS.3SG
 ‘S/he can see it.’

In BT, ‘can’ expressions are analytical constructions consisting of two separate verbs. The modal *bilir* precedes the main verb in the optative form on *-(y)a* (12) (cf. Kırıl 2001: 84). The question raised here is the historical sequence that led to the current construction in BT: whether the Standard Turkish construction, with the univerbated sequence of lexical verb + modal, was re-modelled to modal verb + lexical verb (involving a violation of the unidirectionality of grammaticalization (Heine & Kuteva 2002)), or whether the BT type represents a continuation of an earlier stage prior to univerbation. In our data, the only attested forms have the modal *bilir* in a third person singular aorist form. We are thus currently unable to ascer-

16 Note that Dolatkhah (2016, 2019: 121) also describes double-verb predicates in Kashkay where a finite form of *var* precedes the finite lexical verb; this construction does not seem to express progressive aspect in the first instance, though.

tain the forms of this auxiliary for other tenses and person values. Examples with third singular subjects are provided in (12)–(13) below.

- (12) *bilir ğali toxiya*
 can.3SG carpet weave.OPT
 ‘She can weave carpets.’
- (13) a. *bilir gera belaseni*
 can.3SG see.OPT PRON.ACC.3SG
- b. *bilir belaseni gera*
 can.3SG PRON.ACC.3SG see.OPT
 ‘She can see it.’

The position of the object shows some variation: at least the strong object pronominal stem *bela-* can occur both after or in front of the lexical verb (13a/b). It is not clear whether this flexibility in word order is due to greater word-order freedom for pronominal as opposed to full NP objects, or whether the post-verbal position is related to the presence of *bela-*, which is widely attested in post-verbal position (cf. ex. 7–9 above). For example, in the Turkic variety of Kiān in C&B Province, we find *belar oni gersen* ‘she can see it’, where a pronominal object without *bela-* precedes the verb. Note the variation in the lexical verb in Turkic of Kiān, which is in the voluntative form rather than the optative.¹⁷

The ‘can/be able’ structure in BT seems to be modelled after a Persian periphrastic ‘can’ construction consisting of a modal verb ‘can’ and the inflected main verb in the subjunctive (Bulut 2000: 165), e.g., *mitune māšīn bebare* ‘he can drive a car’, lit. ‘he can (that he) shall drive a car’ (example from Kırıl 2001: 84). This analysis is in line with Johanson (2014: 259) who claims that the preservation in some Oghuz varieties of the Turkic optative, which has otherwise often merged with the voluntative (see for example the form used in Kiān above), is due to contact with Persian. This case of “pattern borrowing” is plausible due to the similar semantic distribution of potentiality and desire which overlaps in optatives and subjunctives. Doerfer (1988) reports for Khalaj that the *bilir* potential form is mainly used for negative expressions. This restriction on negative contexts seems not to hold true for BT, though (see ex.12–13).

17 Bulut (2000: 166) also describes an analytic construction of possibility with the imperative third person singular form *-sIn* in Iraq-Turkic.

4.6 Other candidates for contact-induced morphosyntactic change

Adnominal possession

In BT, there is no genitive marking on the possessor of proper noun genitive NPs; we only find possessive marking on the possessee (14a) (see also Bulut 2018a: 418, and cf. Bulut 1999: 18, 2007: 173, 2018b: 364 for Iraq-Turkic). This differs from the Standard Turkish model in (14b), which requires marking on both the possessor and the possessee.

- (14) a. *maryam-Ø* *derd* *kuçik* *kaka-layi*
 Maryam four small brother-PL.3SG.POSS
- b. Standard Turkish
 ⟨*Maryam-ın* *dört* *küçük* *kardeş-i*⟩
 Maryam-GEN four small brother-3SG.POSS
- c. Bakhtiari of C&B Province (M. Taheri-Ardali, field notes)
çar *ta* *gavu* *koçir-e* *maryam-Ø*
 four CLASS brother small-EZ Maryam
 ‘Maryam’s four small brothers’

Note that the lack of genitive marking on the possessor is not only the case for proper noun NPs but for NPs in general (15a/b).

- (15) a. *rayat* *gaçi-si*
 farmer-Ø goat-3SG.POSS
- b. Standard Turkish
 ⟨*çiftçi-nin* *keçi-si*⟩
 farmer-GEN goat-3SG.POSS
- c. Bakhtiari of C&B Province (M. Taheri-Ardali, field notes)
boz-e *keşavarz*
 goat-EZ farmer
 ‘(the) farmer’s goat’

The fact that BT lacks genitive marking on the possessor and only marks the possessee with the possessive suffix coincides with the structures of neighbouring Iranic languages, namely Bakhtiari (14c and 15c) and Persian, which also solely mark the possessee by means of the *ezafe* suffix. The loss of genitive marking on the possessor in BT could be considered a case of “pattern borrowing” (Matras & Sakel 2007) from Bakhtiari, where only the possessee is marked: the 3SG POSS marker on the

possessee in BT mirrors the Bakhtiari *ezafe* (cf. Bulut 2018a: 428). Apart from the analysis of this structure as contact-induced, we could also assume a more general process of simplification of inflection in a contact situation, leading to the loss of the genitive case marker.

However, the order of possessor and possessee remains intact, with the possessor preceding the head in BT, but following the head in the Iranian languages, a distinction which Bulut (2005: 256) considers crucial for Turkic and Iranian languages. This is in line with Haig's (2001: 219) suggestion that NP-internal word order is a comparatively conservative feature of syntax, while clause-level and inter-clausal syntax are more susceptible to contact-induced change. It also runs counter to the observation of Matras (2007: 43) that it is common for word order in possessive constructions to adapt to a contact language. From this perspective, the maintenance of the Turkic word order in this construction, despite heavy contact influence, is surprising.

Predicative possession

As has been previously described by Bulut (2014: 41, 2018a: 427), 'have' constructions in Iran-Turkic are expressed by the clause-final 'existence adjective' (Bulut 2014: 41) *var* 'there is' + possessive suffix that indicates the possessor, while the possessee remains unmarked (ex. 16a). This contrasts with the common Turkic type shown in (16b), involving possessive marking on the possessee and lack of person marking on the non-verbal predicate *var*. Unlike other varieties of Turkic in Iran (Bulut 2018a), BT seems to lack the Turkic type entirely.

- (16) a. *bir gečči var-om*
 INDEF goat EXIST-1SG.POSS
- b. Standard Turkish
 ⟨keçi-m var⟩
 goat-1SG.POSS EXIST
 'I have a goat.'

We suggest a contact explanation based on restructuring of the common Turkic type, with an existential predicate coupled with a possessive-marked NP, which is restructured to approximate the Iranian model of lexical verb *dāštan* (present stem *dār-*) 'have' + verbal agreement suffixes. These parallels have been regularly noted in the literature, and we assume they are non-controversial. Bulut (2014: 42) also points to the similarity between the Turkic existential particle *var* and the Iranian derivational noun suffix *-var*, expressing possession of an attribute, e.g., *omīdvār* 'having hope' (example from Bulut 2014: 42). The phonological similarity may have contributed to the development of the new strategy for possessive constructions in contact-affected Turkic, but it may also of course be due to pure chance.

Relative clauses

Relative clauses in BT are formed by means of the borrowed Persian complementizer *ke*. This particle introduces the post-nominal relative clause and the finite verb (17a, repetition of ex. 8). The BT structure differs from the Standard Turkish strategy of using left-branching non-finite subjunctors (participles or verbal nouns), which lack an overt complementizer (17b), and mirrors the Iranic type of relative constructions consisting of post-nominal relative clauses introduced by the complementizer *ke* (ex. 17c) (cf. Kırıl 2000: 182; Bulut 2006, 2016: 272).

- (17) a. *moalem-joğaz ke mašhad-dan galdi*
 teacher-DEF CMPL mashhad-ABL come.PST.3SG
gerdem belaseni
 see.PST.1SG PRON.ACC.3SG
- b. Standard Turkish
 〈Mašhad-dan gelen öğretmen-i gördüm〉
 Mashhad-ABL come.PTCP teacher-ACC see.PST.1SG
- c. Persian
man moallem-i rā ke az mašhad
 1SG teacher-RELS ACC CMPL from Mashhad
āmad didam
 came.PST.3SG see.PST.1SG
 'I saw the teacher that came from Mashhad.'

Comparatives and superlatives

Comparatives in BT are marked on the adjective by the borrowed Iranic comparative suffix *-tar* (18a, repetition of ex. 2a). This contrasts with Standard Turkish, which has periphrastic comparative marking with the intensifier *daha*, plus marking of the standard NP with the ablative suffix *-dan* (19). Note also that the old Turkic comparative suffix *-raq* is used in some other varieties of Turkic in C&B Province, such as in Kiān (18b).

- (18) a. *beyig-tar ey-joğaz*
 big-COMP house-DEF
- b. Kiān
beheg-raq oša
 big-COMP house
 'the bigger house'

(19) Standard Turkish

⟨bu ev o ev-den daha büyük⟩
 DEM house DEM house-ABL more big
 ‘This house is bigger than that house.’

BT superlatives are either of the periphrastic form *hamisendan X-tar* (lit. ‘from all X-er’) (20), or borrow the Iranian superlative suffix *-tarin* (see 21, repetition of ex. 2b) while Standard Turkish superlatives are periphrastically formed with the adverb *en* ‘the most’ (22).

(20) *hamisen-dan beyig-tar ey-joğaz*
 everything-ABL big-COMP house-DEF
 ‘This is the biggest house.’ (lit. ‘From everything the bigger house is this.’)

(21) *bo ey beyig-tarin ey de*
 DEM house big-SUP house COP.3SG
 ‘This is the biggest house.’

(22) Standard Turkish
 ⟨en büyük ev bu (dur)⟩
 SUP big house DEM (COP.3SG)
 ‘This is the biggest house.’

The borrowability of the Iranian comparative suffix *-tar* and the Iranian superlative suffix *-tarin* is well attested in literature (Bulut 2005: 253, 2018a: 419; Doerfer 2006: 103); the use of a “source” flag (e.g., an ablative case marker, or a preposition such as Persian *az* ‘from’) for the standard of comparison (lit. ‘bigger from X’) is possibly universal in the region (Haig 2017).

Interrogatives

Polar questions in BT are marked by intonation only and lack the question particle *mI* (23a) which is otherwise common in Turkic (23b).

- (23) a. *ğapu-yi açdı*
 door-ACC open.PST.3SG
 ‘Did he open the door?’
- b. Standard Turkish
 ⟨kapı-yı açtı mI⟩
 door-ACC open.PST.3SG PRT
 ‘Did he open the door?’

The lack of the interrogative particle and the adoption of Iranian intonation patterns for questions is, as Bulut (2018a: 426) states, a common pattern in Turkic varieties under Iranian influence.

Constituent order

In BT, we found goal arguments in post-predicate position (24a, repetition of ex. 10b). The default word order in Standard Turkish (24b) is verb-final, although pragmatically determined variation is possible, and post-predicate goals are frequent in spoken Turkish of Turkey.

- (24) a. *Ali=yenan gediram bazar-a*
 Ali=with go.PRS.1SG market-DAT
- b. Standard Turkish
 <Ali=yale pazar-a gidiyorum>
 Ali=with market-DAT go.PRS.1SG
 'I am going to the market with Ali.'

The verb-goal order is typical of West Iranian, in particular in much of Kurdish (Haig 2014a, to appear), but also in colloquial Persian (Frommer 1981). As Bulut (2018a: 428) states, this type of word order is common across the whole area including Anatolia, Azerbaijan, Iran and Iraq; see also Haig & Khan (2018: 23–25) for data from other language families in the region. This is part of a striking case of widespread areal convergence evident in all languages spoken in western Iran, northern Iraq, and eastern Anatolia.

5 Theoretical implications

5.1 Assessing the likelihood of contact influence

Turning to the theoretical implications of our findings, our first step is to justify a contact explanation for these features, through (a) careful consideration of the relevant structures in potential donor languages, such as Bakhtiari and regional (Chārmahāli) Persian; (b) consideration of the relevant structures in genetically related languages outside the area, such as Turkish of Turkey and Turkic in Iraq (Bulut 1999, 2000, 2007, 2018b); and (c) considerations of general typological plausibility in assessing the possible causes and outcomes of different changes. In particular, (b) is a necessary step, though often ignored, because it provides us with a benchmark, albeit approximate, for evaluating how the state of the language might have been in the absence of Iranian contact influence.

The first five candidate features, examined in the preceding discussion, are as follows. With the exception of the first item, these features have not figured prominently in the literature on Turkic-Iranic contact:

- (i) Restructuring of vowel harmony
- (ii) The postposed emphatic definiteness marker *joğaz*
- (iii) An additional oblique pronominal stem based on *bela-* +possessive suffix
- (iv) ‘Can/be able’ expressions
- (v) A reduplicated progressive verbal form

Applying the above criteria (a)–(c) for assessing the plausibility of contact-induced change, we posit the following approximate hierarchy of likelihood for a contact-induced origin for features (i)–(v):

- (25) Highest > Lowest plausibility of a contact-induced origin
 (iv / v) > (ii) > (i) > (iii)

With regard to (iii), the dummy pronominal stem, the main objection to a contact explanation is the lack of an obvious model in the known, immediately geographically adjacent languages. For (i), vowel harmony, we are dealing with a general reduction of morpho-phonological complexity, which may or may not be contact induced, while (ii), definiteness marking, is somewhat difficult to assess due to the cross-linguistic paucity of comparable processes. However, given that one of the only known examples of the evaluative-to-definiteness pathway is in West Iranian languages (Nourzaei, to appear), it would seem an unlikely coincidence if the BT development had arisen independently of West Iranian. For (iv) and (v), contact influence seems beyond doubt.

Section 4.6 surveyed six further, well-attested candidates for Iranian contact influence. These included (vi) the lack of genitive marking in adnominal possession following Iranian models; (vii) “pattern borrowing” from Iranian in predicative possession; (viii) non-finite relative constructions introduced by *ke*, which are an established areal feature (Bulut 2006, 2016, 2018a); (ix) “matter borrowing” of Persian comparative and superlative suffixes alongside a tendency for periphrastic expressions of superlatives; (x) lack of the Turkic question particle in the whole area under Iranian influence; and (xi) post-predicate goals as a potential areal feature common to various Turkic varieties and West Iranian languages.

5.2 The role of simplification in language contact

There has been considerable debate on the role of simplification – however defined – in language contact (Meakins et al. 2019). In language shift scenarios (see Section 5.3), it is commonly predicted that some simplification will occur as adult second-language learners fail to fully acquire the full spectrum of morphological complexity in the recipient language, which is for them an L2. For other scenarios, predictions vary. Our data shed some light on these issues, though we emphasize the preliminary nature of our conclusions at this point.

Some of the changes noted above involve a loss of inherited Turkic morphological categories, or a reduction in available oppositions, which we will provisionally

categorize as “simplifications”. This type of change occurs in (i) simplification of vowel harmony; (vi) loss of genitive marking on adnominal possessors; and (x) omission of the interrogative particle. However, changes involving the introduction of novel morphological categories are in fact more frequent in our data: (ii) the post-posed emphatic definiteness marker *jōğaz*; (iii) an additional oblique pronominal stem based on *bela-* + possessive suffix; (v) the reduplicated progressive verbal form; and (ix) “matter borrowing” of Persian comparative and superlative suffixes. Other changes involve pattern shifts (word order changes) or replacement of one construction by another (relative clauses)¹⁸, and are thus more difficult to classify on an (admittedly overly simplistic) dichotomy between simplification and complexification.

In other words, inherited Turkic structures have not uniformly undergone major reconfiguration and simplification. Instead, they have experienced selective loss of functionally less integrated, and readily compensated morphology, in cases where contact languages lack the corresponding category. This is clearest in relation to the loss of the interrogative particle: interrogative structure is readily signalled through intonation, and it seems intuitively plausible that this is an attractive strategy (in the sense of Johanson 2002) to adopt. Interrogative particles in Turkic are syntactically and prosodically peripheral, and their loss has minimal ramifications for the core of morphosyntax. The adoption of additional morphology, for example comparative, seems to follow a similar pattern: morphology that is functionally transparent, fairly peripheral, with little interaction with the core areas of morphology, appears to be readily adopted in the recipient language. Other aspects of Iranic morphology are most notably absent, in particular non-predictable and non-concatenative morphology, such as irregular tense marking of highly frequent verbs (see Haig 2014b on the non-borrowability of non-transparent morphology). Other changes, such as (xi) convergence in word order of goal arguments, do not lend themselves to explanation in terms of simplification. They are of sufficient frequency to warrant further discussion and are briefly taken up in Section 5.3.

Overall, the BT example shows evidence of both simplification and complexification. Speakers who master both source and recipient languages are evidently capable of adopting morphosyntactic matter and pattern into their contact language, though not fully unconstrained. Our data do not, however, support an overall gain in simplification in a contact scenario.

5.3 Borrowing versus shift, or imposition

In contact linguistics, a fundamental distinction is drawn between two processes involved in a contact situation. These two processes are referred to as (i) borrowing

18 In the case of BT relative clauses, the new strategy is actually less complex than the original, involving a change from a more synthetic to a more analytic construction. Therefore, technically speaking, this could be classified as a simplification, if simplification is defined as a reduction in linguistic complexity.

(by most authors), and (ii) “interference through shift” (Thomason & Kaufman 1988: 37), or “imposition” (Van Coetsem 2000). Borrowing refers to a scenario in which native speakers of the recipient language adopt items from the donor language into the recipient language, and they are less proficient in the donor language. An example would be when English speakers with limited (or no) knowledge of French use a word or phrase of French origin in their English, as in *déjà vu*. It is unlikely that the English speaker will articulate this expression in a manner that fully matches a native French speaker’s pronunciation, and she will generally be unaware that the French item has a particular grammatical gender (masculine). Both in phonology, and morphosyntax, then, the expression will be adapted to the rules of English. Imposition, on the other hand, involves a situation in which native speakers of the donor language produce utterances in the recipient language, in which they are less proficient. In doing so, they will tend to “impose” the deeply entrenched (and generally subconscious) phonological and morphosyntactic system of their native language onto their recipient language speech. On a larger scale, this process may lead to persistent changes in the recipient language, when “a group of speakers shifting to a target language fails to learn the target language (TL) perfectly. The errors made by members of the shifting group in speaking the TL then spread to the TL as a whole when they are imitated by original speakers of that language” (Thomason & Kaufman 1988: 39). Although the details of the processes involved, and the terminology with which they are expressed, remain disputed (see Winford 2005), there is a widespread consensus regarding the necessity of a conceptual distinction along these lines with regard to what forms the driving force behind a contact-induced change, called “agentivity” in Van Coetsem’s (2000) framework (see also Lucas 2012). Its crucial significance is that it involves a prediction regarding the respective outcomes in the recipient language: under a borrowing scenario (or “recipient language agentivity”, in the terms of Van Coetsem 2000), we expect the adoption of lexical items from the donor to the recipient language to be the main outcome. Under imposition (or “source language agentivity”, Van Coetsem 2000), we expect phonology and syntax to be affected (the role of morphology is disputed), with correspondingly fewer lexical borrowings.

Returning now to the BT scenario, we know that there has been some form of coexistence among the various speech communities over many centuries. However, the details of the sociolinguistic situation(s) that prevailed over these centuries are inadequately documented; we can only project backwards from the historical, anthropological and other literature available to us today (e.g., Beck 1986), and indirectly, from the nature of the linguistic outcomes. With regard to the linguistic outcome, we have primarily documented contact influence on morphosyntax in this paper, and we do not have comprehensive data from the lexicon that would be pertinent here (but for initial results see Anonby, Taheri-Ardali & Stone, in press). The contact effects in morphosyntax would be compatible with an imposition scenario, and in fact this does not seem improbable, given what is known about the composition of the tribal groups speaking these Turkic languages, which have over the years

incorporated speakers of neighbouring Iranic languages.¹⁹ Along with this kind of imposition, we would also assume some measure of long-standing mutual bilingualism between the co-existing groups, yielding gradual convergence among the affected languages. However, given the paucity of historical records, it is ultimately impracticable to establish with certainty the nature of the prevailing conditions. Lucas & Manfredi (2020: 18) note in the context of languages in contact with Arabic that: “for some historical contact situations we simply do not have enough sociolinguistic information to be able to infer what kind of agentivity must underlie a given change. In such cases we must content ourselves with merely identifying the changes that are (likely) due to contact and, for the time being at least, give up on the goal of actually explaining how and why they were actuated.”

But it is not simply a matter of insufficient sociolinguistic data. There is in fact a fundamental methodological error inherent in the attempt to reconstruct sociolinguistic initial conditions based on the attested later outcome of language contact. Recall above that the borrowing vs. imposition distinction implies a prediction from social and cognitive states of speakers to the subsequent linguistic outcome. The prediction is unilateral: a particular state A leads to a linguistic outcome X. Our primary observable data is the linguistic outcome. Nothing in the framework actually permits us to reliably “predict” backwards what the initial state would have been, given a particular outcome; this is a classic error of logic. The framework only predicts that state A will yield outcome X. It does not rule out other causes of the outcome X. Of course, the outcomes that we document can be used as “good-enough” heuristics for formulating some kind approximation of likely initial states. But strictly speaking, this is methodologically questionable. Rather, what the framework permits is to work from an observable, given sociolinguistic configuration and predict the linguistic outcome. For obvious reasons, this is rarely a viable option.

A further claim of this paper is a shift in focus from “Persian influence on Turkic”, to a more nuanced approach that focusses on the role of local (and generally under-described) vernaculars as the primary donor languages. Thanks to the broad geographical and typological coverage of ALI (see Section 2), we are able to more directly identify candidate sources for contact phenomena in the language under consideration. This approach also reflects much more closely the linguistic realities for the period of contact history: for centuries, up to the comparatively recent advent of compulsory schooling and mass media, direct influence of Persian, especially in rural communities, would not have been particularly high. However, the growth and increasing interconnectedness of larger urban centres across Iran have introduced

19 We owe these observations to an anonymous reviewer who points out that, historically, speakers of other languages have “integrated into the host sub-tribes and learned Turkic as their second language”.

new dynamics into language contact, and these will require careful empirical monitoring.²⁰

6 Conclusion

This paper investigates the Turkic variety of Boldaji (BT), with a focus on identifying a set of morphosyntactic features and assessing the probability of contact-induced change against well-established criteria from contact linguistics.

The data compilation is embedded in the ALI project, and includes sociolinguistic metadata and a corpus of naturalistic recordings, although these have not been tapped in the current paper.

With regard to the methodology applied, we readily acknowledge that the use of a translation task from the contact language Persian, while conducive to the collection of a robust and comparable data set, is less than ideal for a contact study and will impact the data that are elicited. Similarly, the availability of a single dataset, as analysed in this study, allows only for a preliminary window into the Turkic variety of Boldaji. An in-depth further investigation of the described features in BT would need to include spontaneous speech data such as narratives or conversations and collect data from different speakers of the community. Notwithstanding the limitations of this study, we have identified two hitherto undocumented types of morphosyntactic change in Turkic, namely the development of definiteness marking in evaluatives with genuine Turkic material, and reduplication-based aspectual marking, both of which we tentatively ascribe to Iranian influence. We have found little support for the view that contact-induced change is preferentially associated with simplification; rather, we find a mix of structural changes involving both loss and gain of morphological matter, and patterns. We have also considered our findings in the light of the dichotomy between borrowing and imposition, but conclude that this is not a particularly insightful framework for approaching scenarios of long-term coexistence, characterized by maintenance of the respective in-group language coupled with wide-spread multilingualism in out-group languages. Our focus is thus on identifying the linguistic output, rather than attempting to project backwards to an assumed sociolinguistic or cognitive initial state. We have also noted, in passing, that such a procedure is in fact based on a logical fallacy.

Finally, our research also underscores the crucial role of Iran's mostly undocumented "other languages" in shaping contact effects on Turkic. Only when a more comprehensive approach to the complexities of Iran's multilingual landscape is adopted can we meaningfully address the issue of how an Iranicized version of Turkic has emerged. There are obviously convenience-related factors in favour of adopting Persian as the benchmark for "Iranic", but with the growing availability of data within the ALI framework and other documentation efforts, there is no sound

20 See Anonby, Schreiber & Taheri-Ardali (2020) for a first attempt at a bilingual case study of Turkic and Bakhtiari spoken in the town of Juneqan, further to the north-west in C&B Province.

justification for continuing to overlook the respective contributions of the diverse minority languages that constitute the linguistic tapestry of Iran.

Abbreviations

ABIL	abilitive	IND	indicative
ABL	ablative	INDEF	indefinite
ACC	accusative	OPT	optative
CLASS	classifier	PL	plural
COMP	comparative	POSS	possessive
CMPL	complementizer	PRON	pronoun
COP	copula	PRS	present tense
DAT	dative	PRT	particle
DEF	definite	PST	past tense
DEM	demonstrative	PTCP	participle
EXIST	existential marker	RELS	relativizing suffix
EZ	ezafe	SG	singular
GEN	genitive	SUP	superlative

References

- Anonby, Erik & Asadi, Ashraf 2014. Bakhtiari studies: Phonology, text, lexicon. *Studia Iranica Upsaliensia* 24. Uppsala: Acta Universitatis Upsaliensis. <http://uu.diva-portal.org/smash/record.jsf?pid=diva2:758171> (accessed 20 November 2020).
- Anonby, Erik & Schreiber, Laurentia & Taheri-Ardali, Mortaza 2020. Balanced bilingualism: Patterns of contact influence in L1 and L2 Turkic and Bakhtiari Speech in Juneqan, Iran. *Iranian Studies* 53: 3–4, 589–622. DOI: 10.1080/00210862.2020.1755957.
- Anonby, Erik & Taheri-Ardali, Mortaza 2018. Bakhtiari. In: Khan, Geoffrey & Haig, Geoffrey (eds.) *The languages and linguistics of western Asia. An areal perspective*. Berlin: De Gruyter Mouton. 445–480.
- Anonby, Erik & Taheri-Ardali, Mortaza et al. 2015–2021. *Atlas of the Languages of Iran (ALI)*. Ottawa: Geomatics and Cartographic Research Centre, Carleton University. <http://iranatlas.net> (accessed 23 September 2020).
- Anonby, Erik, Taheri-Ardali, Mortaza & Hayes, Amos 2019. The *Atlas of the Languages of Iran (ALI)*: A research overview. *Iranian studies* 52: 1–2, 199–230. DOI: 10.1080/00210862.2019.1573135.
- Anonby, Erik & Taheri-Ardali, Mortaza & Schreiber, Laurentia et al. 2020. Turkic languages in the *Atlas of the Languages of Iran (ALI)*. *Turkic Languages* 24: 2, 290–308.
- Anonby, Erik & Taheri-Ardali, Mortaza & Stone, Adam In press. Toward a picture of Chahar Mahal va Bakhtiari Province, Iran, as a linguistic area. *Journal of Linguistic Geography*.
- Beck, Lois 1991. *Nomad: A year in the life of a Qashqa'i tribesman in Iran*. Berkeley: University of California Press.
- Becker, Laura 2018. *Articles in the world's languages*. [PhD dissertation, University of Leipzig].
- Boeschoten, Hendrik E. 1983. *Özbekisches aus Aibak (Samangan)* [Uzbek from Aibak (Samangan)]. (Utrechtse turkologische reeks 1). Utrecht: Utrecht State University.

- Bulut, Christiane 2000. Optative constructions in Iraqi-Turkman. In: Göksel, Aslı & Kerslake, Celia (eds.) *Studies on Turkish and Turkic languages. Proceedings of the IXth International Conference on Turkish Linguistics*, Lincoln College, Oxford, 12–14 August, 1998. (Turcologica 46.) Wiesbaden: Harrassowitz. 161–169.
- Bulut, Christiane 2003. Pronominal systems in the transitional varieties of the Turkic dialects in East Anatolia, Iraq and Western Iran. In: Özsoy, Sumru (ed.) *Studies in Turkish linguistics: Proceedings of the Tenth International Conference in Turkish Linguistics*, Boğaziçi University Istanbul, 16–18 August, 2000. Istanbul: Boğaziçi University Press. 321–335.
- Bulut, Christiane 2005. Iranian influences in Sonqor Turkic. In: Csató, Éva Á. & Isaksson, Bo & Jahani, Carina (eds.) *Linguistic convergence and areal diffusion. Case studies from Iranian, Semitic and Turkic*. London: Routledge Curzon. 241–269.
- Bulut, Christiane 2006. Syntactic traces of Turkic-Iranian contiguity: An areal survey of language contact-induced shift in patterns of relativization. In: Johanson, Lars & Bulut, Christiane (eds.) *Turkic-Iranian contact areas. Historical and linguistic aspects*. (Turcologica 62.) Wiesbaden: Harrassowitz. 165–208.
- Bulut, Christiane 2007. Iraqi Turkman. In: Postgate, John N. (ed.) *Languages of Iraq. Ancient and modern*. Cambridge: Cambridge University Press. 159–187.
- Bulut, Christiane 2014. Turkic varieties in West Iran and Iraq. Representatives of a South Oghuz dialect group? In: Stein, Heidi (ed.) *Turkic in Iran. Past and present*. Wiesbaden: Harrassowitz. 15–99.
- Bulut, Christiane 2016. Convergence and variation in the Turkic varieties of Iran. Examples from Qashqā'ī. In: Csató, Éva Á. & Johanson, Lars & Róna-Tas, András & Utas, Bo (eds.) *Turks and Iranias. Interactions in language and history*. (Turcologica 105.) Wiesbaden: Harrassowitz. 235–282.
- Bulut, Christiane 2018a. The Turkic varieties of Iran. In: Haig, Geoffrey & Khan, Geoffrey (eds.) *The languages and linguistics of Western Asia. An areal perspective*. Berlin: De Gruyter Mouton. 398–444.
- Bulut, Christiane 2018b. Iraq-Turkic. In: Haig, Geoffrey & Khan, Geoffrey (eds.) *The languages and linguistics of Western Asia. An areal perspective*. Berlin: De Gruyter Mouton. 354–384.
- Coetsem, Frans Van 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.
- Csató, Éva Á. 2005. On copying in Kashkay. In: Csató, Éva Á. & Isaksson, Bo & Jahani, Carina (eds.) *Linguistic convergence and areal diffusion. Case studies from Iranian, Semitic and Turkic*. London: Routledge Curzon. 271–283.
- Dehghani, Yavar 2000. *A grammar of Iranian Azari*. Munich: LINCOM.
- De Mulder, Walter & Carlier, Anne 2012. The grammaticalization of definite articles. In: Heine, Bernd & Narrog, Heiko (eds.) *The Oxford handbook of grammaticalization*. Oxford: Oxford University Press. Online Publication Date: Sept 2012. DOI: 10.1093/oxfordhb/9780199586783.013.0042.
- Doerfer, Gerhard 1971. *Khalaj materials*. Bloomington, Indiana: Indiana University Publications.
- Doerfer, Gerhard 1988. *Grammatik des Chaladsch*. Wiesbaden: Harrassowitz.
- Dolatkhah, Sohrab 2016. *Le qashqay. Langue turcique de l'Iran*. North Charleston, South Carolina: CreateSpace.
- Dolatkhah, Sohrab 2019. *Kashkay Turkic. A comprehensive corpus-based grammar*. Munich: LINCOM.

- Dolatkhah, Sohrab & Csátó, Éva Á. & Karakoç, Birsal 2016. On the marker *-(y)akî* in Kashkay. In: Csátó, Éva Á. & Johanson, Lars & Róna-Tas, András & Utas, Bo (eds.) *Turks and Iranians. Interactions in language and history*. (Turcologica 105.) Wiesbaden: Harrassowitz. 283–295.
- Eberhard, David & Simons, Gary & Fennig, Charles 2020²³. *Ethnologue. Languages of the world*. Dallas, Texas: SIL International. <http://www.ethnologue.com> (accessed 20 November 2020).
- Erfani, Parisa 2012. Azeri Morphosyntax. The influence of Persian on a Turkic language. [MA Thesis, Simon Fraser University].
- Frommer, Paul 1981. Post-verbal phenomena in colloquial Persian syntax. [PhD dissertation, University of Southern California].
- Golden, Peter 2006. Turks and Iranians: An historical sketch. In: Johanson, Lars & Bulut, Christiane (eds.) *Turkic/Iranian contact areas. Historical and linguistic aspects*. Wiesbaden: Harrassowitz. 17–38.
- Göksel, Aslı & Kerslake, Celia 2005. *Turkish. A comprehensive grammar*. London: Routledge.
- Haig, Geoffrey 1998. *Relative constructions in Turkish*. Wiesbaden: Harrassowitz.
- Haig, Geoffrey 2001. Linguistic diffusion in present-day East Anatolia: From top to bottom. In: Aikhenvald, Alexandra & Dixon, R.M.W. (eds.) *Areal diffusion and genetic inheritance. Problems in comparative linguistics*. Oxford: Oxford University Press. 195–224.
- Haig, Geoffrey & Nemati, Fatemeh 2013. *Clitics at the syntax-pragmatics interface: The case of Delvari pronominal enclitics*. Paper presented at the 5th International Conference on Iranian Linguistics (ICIL5), Bamberg, 24–26 August, 2013. https://www.academia.edu/4484033/Clitics_at_the_syntax_pragmatics_interface_the_case_of_Delvari_pronominal_clitics (accessed 25 May 2021).
- Haig, Geoffrey 2014a. Verb-Goal (VG) word order in Kurdish and Neo-Aramaic: Typological and areal considerations. In: Khan, Geoffrey & Napiorkowska, Lidia (eds.) *Neo-Aramaic and its linguistic context*. New York: Gorgias Press. 407–425.
- Haig, Geoffrey 2014b. East Anatolia as a linguistic area? Empirical and conceptual issues. In: Behzadi, Lale & Franke, Patrick & Haig, Geoffrey et al. (eds.) *Bamberger Orient Studien*, 1. Bamberg: Bamberg University Press. 13–36.
- Haig, Geoffrey 2017. Western Asia: East Anatolia as a transition zone. In: Hickey, Raymond (ed.) *The Cambridge handbook of areal linguistics*. Cambridge: Cambridge University Press. 396–423.
- Haig, Geoffrey & Masoud Mohammadirad 2019. Definiteness in Central Kurdish: Sources and outcomes. Paper presented at the 4th International Conference on Kurdish Linguistics (ICKL4), University of Rouene, September 2–3, 2019.
- Haig, Geoffrey To appear. Post-predicate constituents in Kurdish. In: Matras, Yaron & Öpengin, Ergin & Haig, Geoffrey (eds.) *Structural and typological variation in the dialects of Kurdish*. London: Palgrave MacMillan.
- Hashemi Zarajabad, Hossein 2015. Copied passives in Southern Azerbaijani, Northern Azerbaijani and Turkish. In: Zeyrek, Deniz & Sağın Şimşek, Çiğdem & Ufuk, Atas & Rehbein, Jochen (eds.) *Ankara papers in Turkish and Turkic linguistics*. Wiesbaden: Harrassowitz. 573–582.
- Heine, Bernd & Kuteva, Tania 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

- Iran Statistics Centre. 2016. *Sālnāmeḥ-ye Āmār-e Irān* [Iran statistical annals]. Tehran: Ministry of the Interior, Iran Statistics Centre. <http://www.amar.org.ir> (accessed 20 November 2020).
- Johanson, Lars 1986. Reproduktion, Widerstand und Anpassung: Zur lautlichen Iranisierung im Türkischen. In: Schmitt, Rüdiger & Skjaervø, Prods O. (eds.) *Studia Grammatica Iranica. Festschrift für Helmut Humbach*. München: Kitzinger. 185–201.
- Johanson, Lars 2002. *Structural factors in Turkic language contacts*. Richmond: Curzon Press.
- Johanson, Lars 2014. The decline of the Ottoman optative. In: Demir, Nurettin & Karakoç, Birsal & Menz, Astrid (eds.) *Turcology and linguistics. Éva Ágnes Csató festschrift*. Ankara: Hacettepe Üniversitesi Yayınları. 253–260.
- Kıral, Filiz 2000. Reflections on *-mis* in Kalaj. In: Johanson, Lars & Utaş, Bo (eds.) *Evidentials. Turkic, Iranian and neighbouring languages*. Berlin: De Gruyter Mouton. 89–102.
- Kıral, Filiz 2001. *Das gesprochene Aserbaidŝanisch von Iran. Eine Studie zu den syntaktischen Einflüssen des Persischen*. Wiesbaden: Harrassowitz.
- Korn, Agnes & Nevskaya, Irina 2017. *Prospective and proximative in Turkic, Iranian and beyond*. Wiesbaden: Reichelt.
- Kuribayashi, Yuu 2006. Syntactic Borrowing of Azerbaijani and Qashqay in Iran. 岡山大学文学部プロジェクト研究報告書 [Research Report, Faculty of Letters, Okayama University] 2006: 2, 63–72. <http://ousar.lib.okayama-u.ac.jp/en/18899> (accessed 23 September 2020).
- Kuribayashi, Yuu 2012. Grammaticalized topics in Kashkay: The implication for the relativization of Turkic languages. In: Kincses-Nagy, Éva & Biacsi, Mónika (eds.) *The Szeged Conference: Proceedings of the 15th International Conference on Turkish Linguistics held on August 20-22, 2010 in Szeged*. Szeged: University of Szeged. 311–318.
- Lee, Noah S. 2008. *A grammar of Iranian Azerbaijani*. Seoul: Korean Altaic Society.
- Lucas, Christopher 2012. Contact-induced grammatical change: Towards an explicit account. *Diachronica* 29: 3, 275–300.
- Lucas, Christopher & Manfredi, Stefano 2020. Introduction. In: Lucas, Christopher & Manfredi, Stefano (eds.) *Arabic and contact-induced change*. Berlin: Language Science Press. 1–33.
- Meakins, Felicitas & Hua, Xia & Algy, Cassandra & Bromham, Lindell 2019. Birth of a contact language did not favor simplification. *Language* 95: 2, 294–332.
- Mahmoudveysi, Parvin & Bailey, Denise & Paul, Ludwig & Haig, Geoffrey 2012. *The Gōrāni language of Gawraju, a village of West Iran. Grammar, texts and lexicon*. Wiesbaden: Reichert.
- Matras, Yaron 2007. The borrowability of structural categories. In: Matras, Yaron & Sakel, Jeanette (eds.) *Grammatical borrowing in cross-linguistic perspective*. Berlin: De Gruyter Mouton. 31–73.
- Matras, Yaron & Sakel, Jeanette 2007. Investigating the mechanisms of pattern replication in language convergence. *Studies in Language* 31: 4, 829–865.
- Maryam Nourzaei 2020. Definiteness marking from evaluative morphology in Balochi: Internal variation and diachronic pathway. *Iranian Studies* DOI:10.1080/00210862.2020.1813555.
- Nourzaei, Maryam & Jahani, Carina & Anonby, Erik & Ahangar, Abbas A. 2015. *Koroshi. A corpus-based grammatical description*. Uppsala: Uppsala University.

- Pakendorf, Brigitte & Krivoschapkina, Ija V. 2014. Even nominal evaluatives and the marking of definiteness. *Linguistic Typology* 18: 2, 289–331.
- Perry, John R. 1998. A review of Encyclopaedia Iranica: Languages and dialects, Islamic Period. *Iranian Studies* 31: 3–4, 517–525.
- Ponsonnet, Maïa 2018. A preliminary typology of emotional connotations in morphological diminutives and augmentatives. *Studies in Language* 42: 1, 17–50.
- Poplack, Shana & Levey, Stephen 2010. Contact-induced grammatical change: A cautionary tale. In: Auer, Peter & Schmidt, Jürgen E. (eds.) *Language and space. An international handbook of linguistic variation*, 1. Berlin: De Gruyter Mouton. 391–419.
- Ragagnin, Elisabetta 2016. Azeri Turkic. Paper held at the workshop *Language Contact and Language Change in Cross-Border Kurdistan (CLiCK)* 9–10 December, 2016, Wolfson College, Cambridge.
- Soper, John 1996. *Loan syntax in Turkic and Iranian. The verb systems of Tajik, Uzbek, and Qashqay*. Bloomington, Indiana: Eurolingua.
- Stilo, Donald L. 1994. Phonological systems in contact in Iran and Transcaucasia. In: Marashi, Mehdi (ed.) *Persian studies in North America. Studies in honor of Mohammad Ali Jazayeri*. Bethesda, Maryland: Iranbooks. 75–94.
- Stilo, Donald L. 2006. Circumpositions as an areal response: The case of the Iranian zone. In: Johanson, Lars & Bulut, Christiane (eds.) *Turkic-Iranian contact areas. Historical and linguistic aspects*. Wiesbaden: Harrassowitz. 310–333.
- Taheri-Ardali, Mortaza et al. 2016. Chahar Mahal va Bakhtiari Province. In: Anonby, Erik & Taheri-Ardali, Mortaza et al. *Atlas of the Languages of Iran (ALI)*. Ottawa: Geomatics and Cartographic Research Centre, Carleton University. [http://iranatlas.net/module/ language-distribution.chahar_mahal_va_bakhtiari](http://iranatlas.net/module/language-distribution.chahar_mahal_va_bakhtiari) (accessed 23 September 2020).
- Taheri-Ardali, M. 1399/2020. Ostān-e Chāhār Mahāl va Bakhtiāri dar atlas-e zabānā-ye irān: raveshshenāsi-ye pazhuhesh va parākandegi-ye zabāni [Chahar Mahal va Bakhtiari Province in the Atlas of the Languages of Iran: Research methodology and language distribution]. *Iranian Journal of Comparative Linguistic Research* 10: 19, 47–69. DOI: 10.22084/rjhl.2020.20969.2004.
- Taheri-Ardali, Mortaza In preparation. Definiteness marking in the languages of Chahar Mahal va Bakhtiari Province, Iran. Paper to be held at the *Ninth International Conference on Iranian Linguistics (ICIL9)*, Vienna.
- Taheri-Ardali, Mortaza and Erik Anonby 2019. Some salient features of the Bakhtiari language in Chahar Mahal va Bakhtiari Province. *Proceedings of the National Conference on Bakhtiari Language, Culture and Literature*, Payam-e Noor University of Shahr-e Kord, 9 May, 2018. Tehran, Iran: Payam-e Noor University. 993–1001. <http://conference.pnu.ac.ir/ChaharMahal-bakhtiari/post.aspx?id=1312> (accessed 20 November 2020).
- Thomason, Sarah G. & Kaufman, Terrence 1988. *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.
- Windfuhr, Gernot L. 1979. *Persian grammar: History and state of its study*. Berlin: De Gruyter Mouton.
- Winford, Donald 2005. Contact-induced changes: Classification and processes. *Diachronica* 22: 2, 373–427.