Chapter 4

Balochi: A cross-dialect investigation of post-verbal elements

Maryam Nourzaei[a] & Geoffrey Haig[b]
[a]Uppsala University [b]University of Bamberg

This chapter investigates word order in three Balochi varieties: Coastal Balochi (Coastal), Koroshi Balochi (Koroshi) and Turkmenistan Balochi (Turkmen). Although all three are closely related, they are areally widely dispersed, making Balochi an interesting test case for investigating the effects of areality on closely related varieties. All three varieties are predominantly OV. However, pronominal direct objects show a stronger tendency to post-verbal placement, especially in Coastal, echoing similar findings for other Iranian languages of the region. All three varieties exhibit predominantly post-verbal Goals (VG), with the highest values found in Koroshi, confirming the expected correlation between higher frequency of VG and geographic proximity to the southwestern Mesopotamian region of the Western Asian Transition Zone.

1 Introduction: Language background and data sources

Balochi is a northwestern Iranian language, which belongs to the Indo-Iranian branch of Indo-European. Syntactically, Balochi is OV, but shows mixed adpositional typology (see Section 3.2) as well as dialectally differentiated alignment systems (see below). Word order in the NP is generally head-final: adjectives precede nouns, and take an attributive linker *-ēn/ẽ*. Possessors also precede the possessed, and the possessor takes a so-called genitive case, though this may vary in the westernmost dialects; see Section 3 for details.

Balochi provides an excellent window on the interplay of areal and genetic influence in shaping word order. The unity of 'Balochi' as the descendants of a

Maryam Nourzaei & Geoffrey Haig. 2024. Balochi: A cross-dialect investigation of post-verbal elements. In Geoffrey Haig, Mohammad Rasekh-Mahand, Donald Stilo, Laurentia Schreiber & Nils N. Schiborr (eds.), *Post-predicate elements in the Western Asian Transition Zone: A corpus-based approach to areal typology*, 121–154. Berlin: Language Science Press. DOI: 10.5281/zenodo.14266337

historically reconstructable sub-group of Iranian is justified in Korn (2005: 21), and there is no doubt that the three doculects under consideration here can be assigned to Balochi. Today, however, due to successive population movements, varieties of Balochi are spoken across a vast area, including Southeastern and Southwestern Iran, Southwestern Pakistan, and also in Afghanistan, India, Africa, Turkmenistan, Oman and the UAE. Almost all Balochi speakers are multilingual, with contact languages belonging to four different language families, and different genera within them: Indo-European (Iranian, Indo-Aryan and Slavic), Dravidian, Turkic and Semitic. Disentangling what is inherited Proto-Balochi from the multiple contact effects is methodologically challenging, for syntax just as it is for phonology and morphology. We address the implications of our study for broader questions of Iranian diachronic syntax in Section 5.

Research on Balochi recognizes three main dialects: southern, eastern and western Balochi. Each of these dialects demonstrates its own sub-divisions (see Jahani & Korn 2009: 636–637). In addition, a group of dialects to the southwest is distinguished, collectively referred to as Koroshi. The total number of Balochi speakers is uncertain, though, Jahani (2013) reports an estimate of at least 10 million speakers. For the comparisons undertaken in this chapter, we have selected data from three geographically dispersed locations, each of which lies within some larger dialect region of Balochi. We refer to the data from these locations as doculects (the variety documented in a specific data set), rather than dialects, because we cannot assume that each doculect is necessarily representative of its larger dialect region. The three doculects are labelled Turkmen (Turkmenistan Balochi), Coastal Balochi (Coastal Balochi), and Koroshi (Koroshi Balochi) respectively. The data for Coastal Balochi come from two villages, Korsar and Sedighzahi, Dashtiyari, Iran. For Koroshi, data comes from from Deh Piyaleh, Shiraz, and Marvdasht, Iran, while for Turkmen the data come from the Mari region, Turkmenistan. The location of the three doculects, and the main dialect divisions of Balochi, are provided in Fig. 1.

Turkmen belongs to the larger Western Balochi dialect group, Coastal Balochi dialect belongs to the Southern dialect group, and Koroshi is part of Koroshi, but would probably also be considered part of Southern Balochi in a broad dialect division (see also Nourzaei et al. 2015: 22).

It is important to bear in mind that our doculects are not necessarily representative for the entire dialect group to which they belong, and in fact our results suggest that there is considerable internal variation (at least on the word order parameters that we have investigated) within the larger dialect groups that have traditionally been recognized. For each doculect, a WOWA data set was compiled as the basis for quantitative comparison. Text types are traditional narrative texts for Coastal Balochi and Koroshi, while the texts for Turkmen include

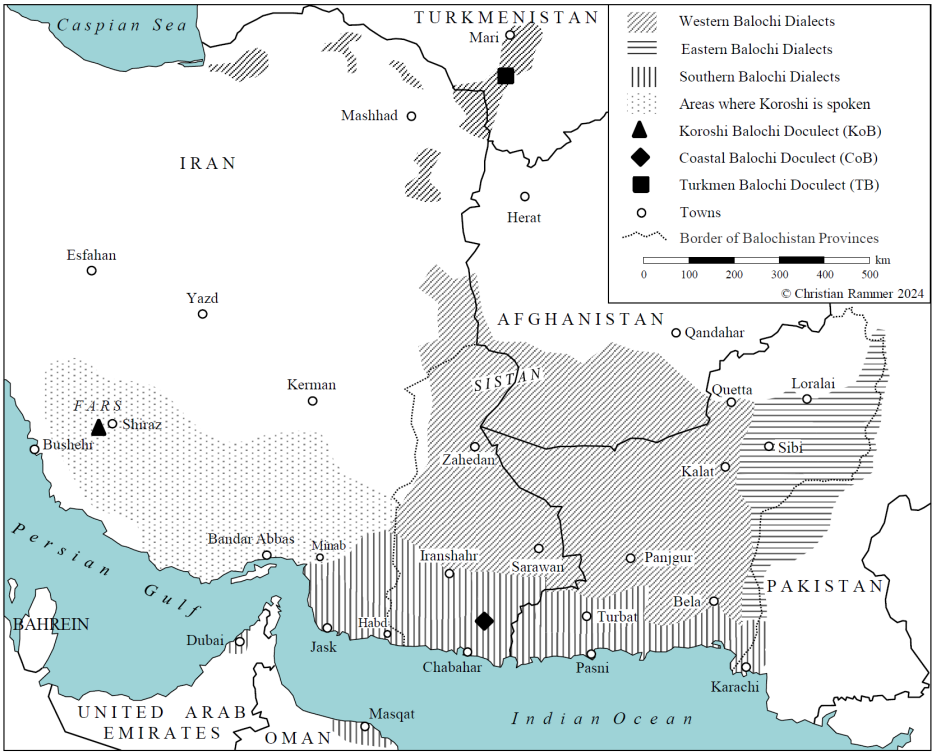

Figure 1: Location of the main Balochi dialects and doculects analysed

traditional narratives, a procedural text, and a life story; see Nourzaei (2021a,b) and Haig (2022a) for details and access to the data. In citing examples we follow the original transcriptions for Coastal Balochi and Koroshi but adapt the original transcriptions of Turkmen to bring it closer to the other two.

The doculects vary according to their alignment systems: Turkmen and Koroshi display the same alignment system in both past and present domain, while Coastal Balochi demonstrates ergative only for 3rd person in transitive clauses based on past-tense verb forms. Koroshi, due to reduction of its morphological case system, uses more person clitics than Turkmen and Coastal Balochi to index verb arguments. The most recent studies on Coastal Balochi are Nourzaei (2017, In press) and Korn & Nourzaei (2019). Earlier works on Koroshi include a grammatical sketch in Persian (Emādi 2005), Nourzaei et al. (2015), Nourzaei (2017, In press). It is assumed that speakers of Koroshi originally migrated from Balochistan to the Southwest of Iran. The most significant study on Turkmen dialect is Axenov (2006). Turkmen Baloch is the result of migration from either

Iran or Afghanistan and north-eastern parts of Iran. For the history of the Baloch migration to Turkmenistan, see Axenov (2000: 71), among others. In addition to the quantitative analysis based on the WOWA data sets, we also draw on additional published corpora and analysis (Barjasteh Delforooz 2010, Nourzaei et al. 2015, Nourzaei 2017, In press, Axenov 2006).

This chapter is organized as follows. Section 2 summarizes previous literature, while Section 3 provides a sketch of some relevant aspects of Balochi grammar, including NP-internal word order. In Section 4, we present the core of the quantitative analysis, based on the major clausal constituent types recognized in the WOWA framework, while Section 5 discusses the implications of the Balochi findings from a comparative (West) Iranian perspective.

2 Previous studies on word order in Balochi

With regard to clause-level constituents, empirical comparative research only commenced quite recently, in particular Jahani (2018) and Korn (2022). Jahani (2018) covers four dialects, viz., Koroshi, Sistani, Turkmen and Southern Balochi (Iran and Pakistan), and considers four argument types: goals, recipients, addressees, and final states of change-of-state predicates. She notes that across all varieties, goals of motion tend to be postverbal, while with regard to the other roles, there are differences: While Koroshi generally has all four argument types post-verbally, Sistani and Turkmenistan Balochi has goals, recipients, and addressees, but not final-states in post-verbal position. For Southern Balochi, also Jahani (2018) includes a more detailed study of spoken versus written genres, observing a very significant tendency towards strict pre-verbal placement of all argument types in written Southern Balochi, while spoken Southern Balochi places 90% of goals after the verb.[1] The differences between oral and written language are very much in line with the observations for Persian, where colloquial spoken Persian has a high frequency of post-verbal Goals (around 80%, see Rasekh-Mahand et al. 2024 [this volume]), while the frequency is much lower for formal written Persian. These pioneering observations suggest that the overall situation in Balochi shows some similarities with the situation that has been identified for languages further to the west (e.g. Kurdish, Haig 2022c), in particular the special status of goals (see Haig et al. 2024 [this volume] for summary), but there are also significant differences. We turn to explanations in Section 5 below.

[1]The data from Jahani (2018) are taken from the slides available from the conference talk. For Southern Balochi, the data for recipients on the overview slide do not match the findings reported from the oral vs. written case study, so we make no claim regarding recipients in Southern Balochi.

Korn (2022: 114) draws on data from Afro-Baloch varieties, taken from Nourzaei and Korn's unpublished Afro-Baloch data. Specifically, the data come from three locations: Shirgowaz (close to Bahukalat), Konark and Karewan. As such, they are only indirectly comparable to our Coastal Balochi doculect, although both Korn's data and our own Coastal Balochi doculect can be included under the broad umbrella term of Southern Balochi. Korn's data confirm the strong tendency to place goals post-verbally, though she suggests that goals of caused motion are overall less likely to be post-verbal, a finding which is confirmed in our own Coastal Balochi data in Section 4 below. Korn suggests that many of the caused goals in her data are human, making them overall more similar to recipients (Korn 2022: 104). Recipients are commonly post-verbal, especially when the recipient or benefactive is "mentioned for the first time "(Korn 2022: 105), but pre-verbal position is also well-attested, so that no clear conclusion can be reached with regard to a canonical or unmarked position; rather, the interaction of various factors such as animacy, verb semantics, overall presence of co-arguments, and information structure co-determine the placement. For addressees, only few overt examples were present in the data, but all occur preverbally (Korn 2022: 107). This finding confirms Jahani's (2018) observation that addressees are overall more likely to be preverbal than either goals or recipients. Again, this confirms observations from other languages in WATZ, according to which addressees are less likely to be post-verbal than recipients or goals (see Haig et al. 2024 [this volume], Section 4).

Korn also investigates direct objects, noting that post-verbal placement "is surprisingly frequent" (Korn 2022: 113), though no figures are provided; we return to this below. Korn also discusses the possible impact of flagging, and animacy, noting a tendency for [+human] arguments to be pre-verbal. She concludes with a suggestion for the pathway towards "generalization of the post-verbal position" in Balochi (Korn 2022: 118), according to which non-human goals would have been the first kinds of constituent in this position, extending then to include metaphorical uses of direction (purpose) and then other kinds of adverbial, while a second line of extension would proceed via [+human] goals to recipients and beneficiaries. Ultimately, an extension to direct objects is considered as a final possibility. Korn (2022) also discusses the impact of flagging, verb serialization, and interactions with subordination. In sum, Jahani (2018) and Korn (2022) provide a very informative overview of the relevant characteristics of Balochi, which already identifies some of the areas of cross-dialect variation. The present study builds on these observations but extends the the range of argument types considered, and is based on more accessible data.

3 Basic features of Balochi: Morphosyntax and NP-internal word order

3.1 Alignment, person marking, and nominal inflection

Among the three doculects considered here, only Coastal Balochi displays stem-sensitive ergativity, i.e. verbal forms with the present stem pattern accusatively (A in direct case and agreeing with the verb, P in oblique or object case) while verbal forms with the past stem pattern ergatively (P in direct case and agreeing with the verb (only for third person), A in oblique case) (see for details Nourzaei 2017). Both Koroshi and Turkmen Balochi exhibit only accusative alignment.

Doculects differ regarding the usage of person-marking clitics. Both Coastal Balochi and Turkmen only use the person marking clitics for the 3rd person, while Koroshi uses the person-marking clitics for all persons. The existence of person-marking clitics in Balochi has a strong correlation with the case system. In varieties such as Coastal Balochi and Turkmen (see above) which have a rich morphological case system, they are less commonly used, while in varieties such as Koroshi and Sarawani (see Baranzehi 2003: 86), which display a reduced case system, they are more common.

Balochi has a morphological case system containing at least a direct case, an oblique case, and a genitive case, which is the system found in Koroshi. In Sistani, Afghan and Turkmen Balochi, a locative case has developed from the genitive (cf. Buddruss 1988: 48, Axenov 2006: 80–82, Korn 2008, and the data in Barjasteh Delforooz 2010). In Coastal Balochi, it is sporadically attested as well, typically with human names (Nourzaei 2017: 61, Korn & Nourzaei 2019). In addition, another form appears in Sistani Balochi that is derived from the oblique case and is only used to mark direct objects, whence its name, "object case" (cf. Korn 2008: 61–63, Nourzaei 2017: 62). In Coastal Balochi and Koroshi doculects, object case is only available for pronouns (cf. Nourzaei 2017: 44).

With respect to plural marking, Koroshi differs from other varieties in that it has an innovated plural marker *-obār* in all cases (direct, genitive and oblique, Nourzaei et al. 2015: 29) while in Coastal Balochi and Sistani, the marker *-ān* does not mark the plural on nouns in the direct case, which are thus not directly marked for number. Instead, plural number can be indicated by plural agreement markers on the verb. Nouns in the oblique case, on the other hand, show a number opposition. No variety of Balochi has retained grammatical gender.

3.2 NP-internal word order

3.2.1 Adjective and noun

The most common pattern regarding adjective-noun ordering across the dialects is that adjectives precede nouns, and that attributive adjectives take a suffix *-ēn/**ẽ*** (adjective attribute suffix, ATTR).

(1) a. Coastal Balochi (Nourzaei 2021a, C, 1171)
zar'd-ẽ *negē'na*
yellow-ATTR stone
'the yellow stone'

b. Turkmen Balochi (Haig 2022a, A, 0049)
ǰwān-ēn *zāg=ē*
good-ATTR son=INDV
'a good son'

c. Koroshi Balochi (Nourzaei et al. 2015:42)
bōr-ēn *pašm-ā*
beige-ATTR wool-OBL
'the beige wool'

However, Koroshi exhibits borrowed ezafe constructions (N-ADJ order) from Persian shown in (2).

(2) a. Koroshi Balochi (Nourzaei 2021b, A, 0006)
ādam=e ***xūb***=*ī*
person=EZ good=INDV
'a good person'

b. Koroshi Balochi (UP)
ǰāhel=e nūrānī=ye
boy=EZ handsome=INDV
'a handsome boy'

3.2.2 Possessor and noun

Across the dialects, possessors normally precede the possessed noun (POSS-N order), and the possessor takes a so-called genitive case (for details regarding different forms of genitive case (see Nourzaei 2017, Nourzaei et al. 2015, Korn & Nourzaei 2019).

(3) a. Coastal Balochi (Nourzaei 2021a, A, 0315)
sī'morg-e *dap-ā*
fabulous_bird-GEN mouth-OBL
'(into) the fabulous bird's mouth'

b. Turkmen Balochi (Haig 2022a, A, 0083)
xānbādorr-ī *ǰenēn-ā*
Khanbadur-GEN wife-OBL
'the wife of Khanbadur'

c. Koroshi Balochi (Nourzaei 2021b A, 0007)
šāh-ay *awal-īn* *bač*
king-GEN first-ATTR son
'the king's first son'

In Koroshi, possessors may follow nouns (N-POSS) in a kind of ezafe construction, most likely borrowed from Persian. Most examples in Nourzaei et al. (2015) suggest not so much morphological borrowing as wholesale borrowing of phrases from Persian, as in (4). The actual productivity of N-EZ POSS constructions in Koroshi thus remains to be established.

(4) Koroshi Balochi (Nourzaei et al. 2015:212)
ya *banne=ye* *xodā=ī*
one servant=EZ god=INDV
'a fellow (lit. a servant) of God'

The ordering of pronominal possessors may differ from nominal possessors. In Coastal Balochi, pronominal possessors generally follow the possessed:

(5) a. Coastal Balochi (Nourzaei 2021a, B, 505)
gohār-ā ***otīg-a***
sister-OBL REFL-OBL
'your sister'

b. Coastal Balochi (UP)
māt *manīg*
mother mine
'my mother'

3.2.3 Demonstrative and noun

Across all dialects, demonstratives precede nouns, with some marginal exceptions which are ignored here (Korn & Nourzaei 2019).

(6) a. Coastal Balochi (Nourzaei 2021a, B, 0711)
ē *pet*
PROX father
'this father'

b. Turkmen Balochi (Haig 2022a, A, 0030)
ē ***bādešā***
PROX king
'this king'

c. Koroshi Balochi (Nourzaei 2021b, B, 0570)
ē *ǰo'ġlā*
PROX boy.OBL
'this boy'

3.2.4 Numeral and noun

Plural marking retains to a large extent the archaic pattern, also found in Kurdish, whereby only nouns in oblique cases are overtly plural marked with a suffix, but plural subject nouns are not overtly marked. All dialects share the commonality of ordering the numerals before head nouns.

3.2.5 Adpositions

The dialects have prepositions, postpositions, and circumpositions. Postpositions are generally relational nouns in the oblique case, with the NP complement in the genitive case. Similar to nouns in the oblique case, the adpositions can also be used alone as adverbs. Prepositions usually trigger the oblique case of the noun. Note that there is a tendency for losing the oblique case after prepositions in Koroshi. The respective frequencies of prepositional and postpositional use shows interesting cross-dialectal variation, which we sum up in Table 1 below. In general, Koroshi has the strongest tendency to use prepositions (see for more details Nourzaei et al. 2015: 43–46). Examples of prepositions and postpositions follow:

(7) a. Coastal Balochi (Nourzaei 2021a, B, 0678)
čāt-e ***tōkā***
well-GEN inside
'inside the well'

b. Turkmen Balochi (Haig 2022a, D, 0576)
gis-ay ***tā***
house-GEN inside
'inside the house'

c. Koroshi Balochi (Nourzaei et al. 2015:36)
dawr*=e* *ī* *mēdag-ā*
around=EZ PROX encampment-OBL
'around this encampment'

d. Koroshi Balochi (Nourzaei 2021b, B, 0553)
čāh-ay ***tōxā*** *por=e* *šamšīr=o* *nayza a=kan-t*
well-GEN in full=EZ sword=and spear VCL=do.PRS-3SG
'She fills the well with swords and spears.'

In contrast to Koroshi, Coastal Balochi and Turkmen possess circumpositions:

(8) Turkmen Balochi (Axenov 2006:150, glosses follow the source)
bi *diga* *gis-ē* *tā*
to other house-IND inside
'to another house'

For our quantitative analysis of adpositions in actual usage, based on the WOWA data, we will consider the respective frequencies of prepositional and postpositional flagging, across those functions that we would generally expect to favour adpositional over other types of flagging (e.g. case marking, or lack of any overt flagging).[2] Table 1 shows the respective frequencies of postpositional and prepositional flagging across the three dialects for NPs in these functions.

Table 1 suggests a two-way split across Balochi, between the predominantly prepositional Koroshi and Turkmen on the one hand, versus predominantly postpositional Coastal Balochi on the other. The high frequency of postpositional flagging in Coastal Balochi is intriguing. At this point we have no convincing explanation. It may be a retention of earlier Balochi structures, which has been lost in other varieties through greater contact with other west Iranian languages,

[2]The procedure was as follows: Taking the three Balochi WOWA data sets, we selected the total number of tokens in the following functions: ABL(ative); ADDR(essee); BEN(efactive); COM(itative); GOAL; GOAL-C(aused); INSTR(umental); LOC(ative); REC(ipient); REC-BEN. We then extracted those that were flagged with preposition or pre-nominal relational nouns (lumped together as "prepositional"), and those that were flagged with postpositions, or post-nominal relational nouns (lumped together as "postpositional"). Together this yielded 343 tokens.

Table 1: Balochi prepositional and postpositional flagging frequency

	Coastal		Koroshi		Turkmen		
Adpositional type	N	%	N	%	N	%	Totals
Postpositional	83	86	2	2	12	10	97
Prepositional	13	14	129	98	104	90	246
Totals	96		131		116		343

notably Persian. It may be connected to multilingualism with Urdu and other postpositional Indo-Aryan languages spoken in the region, but this is speculative. Interestingly, in our data Turkmen is dominant prepositional, although we might have expected higher rates of postpositions due to contact with Turkic. It is possible that Turkic has had less influence because the migrations of Turkmen speakers to Turkmenistan was relatively recent, and they retain contact with Sistani Balochi speakers in Iran. This would be in line with the findings of Haig et al. (In press), according to which adpositional type is a relatively conservative word order parameter that only shifts under intense and long contact.

3.3 Auxiliary and main verb, complement clause and matrix clause

TAM categories are expressed by the presence or absence of verbal prefixes (Jahani & Korn 2009, Axenov 2006, Nourzaei et al. 2015 among others). The perfect system uses the participle followed by the inflected copula, which cliticizes to the verb, while the progressive (e.g., Koroshi and Coastal Balochi) is built from infinitive plus clitic copula. Historically, then, we can assume at least some examples of V-Aux order, which have since univerbated through cliticization of the original auxiliary. In contemporary Balochi, however, prosodically independent (non-clitic) auxiliary verbs are infrequent, so establishing a regular order of auxiliary and verb is not straightforward. Examples with cliticized auxiliaries are the following:

(9) a. Turkmen Balochi (Haig 2022a, D, 0548)
ammā pa wat-ī māl-ān yakk yakk=ī nām
1PL for REFL-GEN animal-PL one one=INDV name
išt=at-an
put.PST=COP.PST-1PL
'We had given names to everyone of our sheep.'

b. Koroshi Balochi (Nourzaei 2021b, A, 0296)
hasan kačal faġat ***nay-āk-ag=en***
Hasan bald only NEG-come.PST-PP=COP.3SG
'Only Hasan the Bald has not come along.'

The Sistani variety of Western Balochi has an auxiliary verb *dāšten* 'have', a recent borrowed element from Persian, to build progressive construction in past and present domain (see for details Nourzaei 2024). As in Persian, this auxiliary precedes the main verb as in (10):

(10) a. Sistani Balochi (Nourzaei 2024)
dār-īn *wān-īn sāket be*
have.PRS-1SG read.PRS-1SG quiet IMP.become.PRS.2SG
'Be quiet, I am studying!'

b. Sistani Balochi (Nourzaei 2024)
dāšt-on *šot-on ke čākar āt*
have.PST-1SG go.PST-1SG CLM Chakar come.PST.3SG
'I was going when Chakar came.'

The subordinator *ke* may introduce various kinds of subordinate clause, i.e, relative, complement and adverbial, as well as quoted speech. Across the dialects, the complement clauses normally follow the main clause and are either linked to it by asyndetic subordination (juxtaposition) without any overt marker of subordination other than rising intonation, or with an overt complementizer. A number of compound conjunctions, composed of nouns or other elements plus ke, such as *mawġeī ke/waġteke* 'when', and *be šartī ke* 'on the condition that' are also used. Additional subordinating conjunctions include *tā/ta* 'until, so that' and *aga/aya* 'if'. In all dialects subordination closely follows the basic pattern of Persian and copies its compound conjunctions (see also Jahani & Korn 2009: 678). Examples of complement clauses with verbs of speech and perception, with and without complementizers follow:

(11) a. Coastal Balochi (Nourzaei 2021a, C, 1061)
pet-ā ***go*** ***ke*** *man-ī čo nī 'mã 'ta-rā 'sīr*
father-OBL say.PST CLM 1SG-GEN child now 1SG 2SG-OBJ wedding
da'y-ã̃
give.PRS-1SG
'The father said /that/, "My son, now I will marry you off." '

b. Turkmen Balochi (Axenov 2006)
gis-ay ***wāond gušt=ī*** ***ke*** *mnī* *piss* *iškārī=e*
house-GEN owner say.PST=PC.3SG CLM 1SG.GEN father hunter=INDV
at-ī
COP.PST-3SG
'The owner of the house said that my father was a hunter.'

c. Koroshi Balochi (Nourzaei et al. 2015:143)
a=genn-an ***bale*** *aždahā=am* *pīk-ay*
VCL=see.PRS-3PL yes dragon=ADD twist.PST.PP-COP.PRS.3SG
dawr=e *šāh-ay* *ǰanek-ay* *garden-ā*
around=EZ king-GEN daughter-GEN neck-OBL
'They see [that] indeed the dragon was wrapped around the neck of the king's daughter.'

4 Order of clause-level constituents: A quantitative analysis

In this section we present the results of the quantitative analysis, drawing on the set of constituent types defined in the WOWA framework; not all categories are considered, as some have too few tokens for meaningful quantitative analysis.

4.1 Direct object and verb

Across the dialects, nominal direct objects are overwhelmingly in the preverbal position (>90% OV in all three WOWA Balochi doculects). This confirms previous research on Balochi (Jahani & Korn 2009 and Korn 2022), and also reflects the overall tendency for Iranian languages in WOWA to be consistently OV in discourse, with the exception of Kumzari (see Haig 2022b, Haig et al. 2024 [this volume]). Our analysis does, however, identify some cross-dialectal differences. Table 2 provides the figures for direct objects, distinguishing between pronominal direct objects (12a) and nominal direct objects (12b). WH-words, clitic pronouns, and adverbials in object function have been excluded from the pronoun category, but we include reflexives. In Table 2, and in the following Tables, N refers to the absolute number of tokens included in each category, "Po" refers to the number of those tokens that were post-verbal, and "%" provides the percentage of post-verbal tokens in each category.

The absolute number of pronominal objects in the data is quite low, particularly in Koroshi, and those that are present are overwhelmingly human (82%, 142

Table 2: Nominal vs. pronominal post-verbal direct object frequencies

	Coastal			Koroshi			Turkmen			
	N	Po	%	N	Po	%	N	Po	%	Totals
Nominal	339	23	6.8	182	4	2.2	193	3	1.6	714
Pronominal	98	27	27.6	20	0	0	55	2	3.6	173
Totals	437			202			248			887

out of 173); we turn to the interplay of humanness and pronominality in Table 5 below. Examples of pronominal and nominal direct objects respectively are the following:

(12) a. Turkmen Balochi (Haig 2022a, A, 0007)
annūn b-raw-an ke ***šmā-rā*** *gis=a da-īn*
now SBJV-go.PRS-1PL CLM 2PL-OBJ house=VCL give.PRS-1SG
'Now let us go, I will marry you off.'

b. Koroshi Balochi (Nourzaei 2021b, A, 0022)
man 'wad=om ***as'p-ok-ā*** *'gott a=kan-ān*
1SG REFL=PC.1SG horse-DEF-OBL raise VCL=do.PRS-1SG
'I myself will raise the horse.'

With regard to nominal objects, the three dialects exhibit <10% levels of post-verbal placement. However, in Coastal Balochi, rates of post-verbal nominal objects are more than double the other two dialects. A pair-wise Fisher's Exact Test indicates that the difference between Coastal Balochi and the other two reach significance ($p<0.05$). For pronominal objects, the differences are much more pronounced, and again, it is Coastal Balochi that diverges from the other two.

Korn (2022: 112) had already noted that postverbal direct objects "are surprisingly frequent" in the geographically close variety of Southern Balochi that she investigates. Our data suggest that the spoken varieties of Southern Balochi, such as our "Coastal Balochi," may indeed differ from other varieties of Balochi, in particular with regard to pronominal objects. Other research (Stilo 2018, Haig et al. In press) has suggested that pronominal objects are the most mobile, in the sense that they are more likely to shift across the predicate from the canonical object position, and our findings provide further support for this assumption.

As Korn (2022: 113) notes, post-posing direct objects is probably related to information structure, but the exact nature of the triggering factors is "not always

obvious." The WOWA data base does provide a rough classification of direct objects into definite and indefinite, which we have analysed in Table 3.

Table 3: Definite vs. indefinite post-verbal nominal direct object frequencies

	Coastal			Koroshi			Turkmen			
	N	Po	%	N	Po	%	N	Po	%	Totals
Definite	110	9	8	126	3	2	89	2	2	325
Indefinite	229	14	6	56	1	2	104	1	1	389
Totals	339			182			193			714

Based on the admittedly blunt instrument of the definiteness classification in WOWA, the distinction between definite and indefinite does not contribute much to the explanation. Either the absolute figures are too low (Koroshi and Turkmen), or do not reach significance (Coastal Balochi).

A second factor that is often claimed to be relevant in placement of direct objects is weight. The WOWA data set distinguishes four levels of syntactic weight, based on orthographic words excluding clitics and adpositions: 1, 2, 3, and >3.[3] An example of a heavy direct object (three words) is 13, a light direct object is (12b) above.

(13) Coastal Balochi (Nourzaei 2021a, A, 0028)
'mã ***ō'tī*** ***'mačč-e*** ***'hōš*** *de'gar-ā*
1SG REFL.GEN date_palm-GEN cluster other-OBL
'na-dāt-ag=ã
NEG-give.PST-PP=COP.PRS.3PL
'I didn't give **my date palm clusters** to anyone.'

The figures comparing postverbal placement of the lightest (1 word) with the heaviest (>2 words) nominal direct objects are shown in Table 4 (all dialects combined).

Table 4 echoes findings from the WOWA spoken language corpora, which suggest that weight is not a significant factor in triggering object postposing (see Haig et al. 2024 [this volume]). In fact, the opposite tendency is suggested by our data.

[3] An additional measure of weight (number of characters) is also available, but was not applied here; see Haig et al. 2024 [this volume].

Table 4: Frequencies of post-verbal nominal direct objects, light vs. heavy

	All dialects		
	N	Po	%
Light (1 word)	329	16	4.9
Heavy (>2 words)	86	0	0
Totals	415		

Turning now to the factor of animacy, here reduced to human versus non-human, there is an interaction between humanness, pronominality, and post-verbal placement. Table 5 provides the relevant figures:

Table 5: Post-verbal direct objects (all doculects), according to humanness and pronominality

	+Hum			-Hum			
	N	Po	%Po	N	Po	%Po	Totals
Nominal	171	10	5.8	543	20	3.7	714
Pronominal	142	27	19	31	2	6.5	173
Totals	313			574			887

First, these figures suggest that there is a strong correlation between humanness and pronominality: more than 80% of all pronominal objects are human (142 out of 173). Thus non-human pronominal objects are a rarity in Balochi, confirming a cross-linguistic tendency to avoid non-human object pronouns in discourse (Haig et al. 2022). Second, pronominality generally increases the likelihood of post-verbal placement, irrespective of humanness. Third, humanness alone is only marginally relevant: a human, nominal direct object is not significantly more likely to be postverbal than a non-human, nominal object. The difference between human and non-human only becomes relevant when the direct object is pronominal.

Closer inspection of the data reveal that the general tendency to avoid non-human object pronouns is most pronounced in Coastal Balochi, where almost 90% of object pronouns are human. Furthermore, in Coastal Balochi the tendency

to place these [+human] object pronouns after the verb reaches around 30%, and it is the Coastal Balochi data which actually account for most of the effects shown in Table 5. In fact, Coastal Balochi exhibits a nascent tendency towards the system found in Kumzari (Haig 2022b, based on Anonby 2015), where the majority of pronominal objects are post-verbal, and nominal objects are still predominantly pre-verbal.

In sum, we have seen that all doculects exhibit dominant OV order, but closer inspection reveals a difference between Coastal Balochi and the other two: although all three dialects are dominant OV overall, Coastal Balochi shows a significantly higher number of postverbal objects than the other two, with the strongest effect found for human pronominal objects. This finding is difficult to reconcile with our earlier observation that Coastal Balochi is more postpositional than either of the other doculects (Section 3). If it is more strictly postpositional, we might have expected it to be more rigidly OV, but this is not the case. We see no obvious explanation for the relatively high degree of object mobility in Coastal Balochi. But we have seen elsewhere that dominant use of postpositions does not necessarily entail strict OV; both Georgian and Armenian are post-positional, yet object positioning is actually more flexible than it is for the mostly prepositional Iranian languages (Hodgson et al. 2024 [this volume]). Conversely, strict OV does not necessarily entail postpositions; witness some of the Iranian languages closer to the Mesopotamian core of WOWA, which are strictly OV but dominant prepositional, or the varieties of Neo-Aramaic which have shifted to OV, but remain prepositional (see Noorlander 2024 [this volume]). While we have no explanation for the divergent characteristics of Coastal Balochi, they further underscore the potential disconnect between adpositional order and verb/object ordering in discourse.

Our data further suggest that weight is not a significant factor in predicting post-verbal placement of objects. We have identified the difference between nominal and pronominal form as relevant, though only in conjunction with humanness. Basically, we find that neither humanness, nor pronominal form alone predicts post-verbal placement. However, with pronouns, a strong interaction with humanness becomes apparent; this tendency is strongest for Coastal Balochi, but we can only speculate on the reasons for this at present.

4.1.1 Copula complements

A similar picture emerges here as with direct objects: all dialects overwhelmingly prefer preverbal placement, as in the following examples:

(14) a. Coastal Balochi (Nourzaei 2021a, A, 0299)
te'yōk-ā *bī*
alone-OBL COP.PST.3SG
'He was alone.'

b. Turkmen Balochi (Haig 2022a, A, 0154)
xānbādorr*=o* *man=on*
Khanbadur=FOC 1SG=COP.PRS.1SG
'I am Khanbadur.'

c. Koroshi Balochi (Nourzaei 2021b, A, 0006)
ē *'šāh* *'ǰan=e* *da'wom=ī* ***'xeylī ā'dam=e*** ***'xūb=ī***
PROX king wife=EZ second=PC.3SG very person=EZ good=INDV
'na-bod-ag=en
NEG-become.PST-PP=COP.PRS.3SG
'This king's second wife was not a very good person.'

The relevant figures are provided in Table 6.

Table 6: Frequencies of post-verbal copula complements

	Coastal			Koroshi			Turkmen			
	N	Po	%	N	Po	%	N	Po	%	Totals
Copula complements	310	13	4.2	53	0	0	66	2	3	429

Examples of a post-verbal copula from Coastal Balochi and Turkmen are the following:

(15) Coastal Balochi (Nourzaei 2021a, C, 0997)
ke *ā-'ī* *'nām=a* *rahīm'baxš*
CLM DIST-GEN name=COP.PST.3SG Rahimbakhsh
'who is called Rahimbakhsh'

(16) Turkmen Balochi (Haig 2022a, A, 0022)
man=om *ast-om* ***pādišā=īē***
1SG=ADD exist.PRS-1SG king=INDV
'I am a king.'

These findings confirm a broader tendency observable across the entire WOWA data base, namely that in dominant OV languages, the position of copula

complements is generally conservative in the sense that they are even less prone to post-posing than direct objects are. In our sample, we find near-categorical pre-verbal placement of copula complements.

4.2 Verb and goal

Korn (2022) and Jahani (2018) suggest that among endpoint constituents, goals of verbs of movement (e.g., 'go', 'come') and goals of verbs of caused motion (e.g., 'put', 'throw', 'bring') have the highest rates of post-predicate position among all argument types, and across all data. Our data confirm this result. Examples of goals and goals of caused motion are the following:

(17) a. Goal
Coastal Balochi (Nourzaei 2021a, A, 0043)
'šo ***mē'tag-ā***
go.PST.3SG home-OBL
'He went home.'

b. Goal
Koroshi Balochi (Nourzaei 2021b, A, 0047)
ma'rō 'raft-ay ***lō'g-ā***
today go.PST-2SG home-OBL
'Today (when) you go home.'

c. Goal
Turkmen Balochi (Haig 2022a, A, 0160)
šot *bi* ***yak ǰā=ē***
go.PST.3SG to one place=INDV
'He went to a certain place.'

d. Caused Goal
Coastal Balochi (Nourzaei 2021a, A, 0066)
ǰat=e *'zahm=e* ***'dast=e***
beat.PST=PC.3SG sword=INDV hand=PC.3SG
'He struck the sword at its hand.'

e. Caused Goal
Koroshi Balochi (Nourzaei 2021b, B, 0510)
ma-ba'r-ā *bod-a* *ma-prē'n-ā*
IPFV-take.PRS-BACKG.3SG become.PST-PP IPFV-throw.PRS-BACKG.3SG
bod-a *mā* ***dar'yā-hā***
become.PST-PP into sea-OBL
'She used to take it [and] throw it into the sea.'

f. Caused Goal
Turkmen Balochi (Haig 2022a, D, 0609)
mn-ā ǰat ***be ḍigār-ā***
1SG-OBL beat.PST.3SG on ground-OBL
'(The donkey) threw me onto the ground.'

However, the dialects in our sample differ in the degree to which goals are postposed. Table 7 compares nominal goals and caused-motion goals, excluding adverbs:.

Table 7: Frequencies of post-verbal nominal goals

	Coastal			Koroshi			Turkmen			
	N	Po	%	N	Po	%	N	Po	%	Totals
Goal	85	60	70.6	68	63	92.6	23	8	34.8	176
Caused Goal	21	6	28.6	17	13	76.5	13	7	53.8	51
Totals	106			85			36			227

For Koroshi, the results are typical for an Iranian language close to the Mesopotamian core of the Western Asian Transition Zone, and also for colloquial spoken Persian (Haig et al. 2024 [this volume]): more than 80% of all goals are post-verbal. For the other two dialects, the results are somewhat puzzling. For Turkmen, significantly lower rates of post-verbal Goals can plausibly be related to effects of Central Asian varieties of Turkic, in line with the predictions of Haig et al. (In press) and Haig et al. 2024 [this volume], which suggest that the placement of Goals is highly sensitive to language contact. However, the absolute number of goals in the Turkmen data set is quite low, so this is a provisional conclusion. For Coastal Balochi, the most puzzling aspect is the overall low levels of postverbal caused goals, an observation that echoes Korn's (2022) finding for another Southern Balochi variety. For most data sets in the WOWA sample, the difference between caused and simple goals is not significant, and the two roles can be unified for most analyses. But for Coastal Balochi, the difference is striking; we have no explanation for this; it definitely merits further investigation (verb-specific effects, for example).

4.3 Recipients and addressees

A number of previous studies have shown that goals, recipients, and addressees do not necessarily pattern alike in the OV languages of Western Asia (Haig et al. 2024 [this volume], Korn 2022, Stilo 2018). Particularly addressees may exhibit quite different word order properties, even between closely related varieties (Jahani 2018 for Balochi, Haig 2022c for Kurdish). Small corpora of naturalistic spoken language are problematic for testing word order of addressees and recipients, because relevant tokens are not particularly frequent, and often pronominal. In varieties with clitic pronouns, the position of the pronoun is determined by language-specific clitic-placement principles, which may be quite distinct from the principles governing word order of prosodically independent constituents. Our findings here are correspondingly tentative. Examples of recipients and addressees are provided below, while Table 8 gives the frequencies per doculect.

(18) a. Recipient
Coastal Balochi (Nourzaei 2021a, A, 0 527)
k-ā'r-ã *de'-yã* ***'pet-a***
K.IPFV-bring.PRS-3PL give.PRS-3PL father-OBL
'They bring [and] give [them] to [their] father.'

b. Addressee
Coastal Balochi (Nourzaei 2021a, A, 0023)
o'š-ī ***čo'k-ān***
say.PRS-3SG child-OBL.PL
'He says to [his] children.'

c. Recipient
Koroshi Balochi (Nourzaei 2021b, B, 580)
pu'l-ā *a='dā* ***āle'm-ok-ā***
money-OBL VCL=give.PRS.3SG wise_man-DEF-OBL
'She gives the money to the doctor (lit. wise man).'

d. Addressee
Koroshi Balochi (Nourzaei 2021b, A, 0320)
'ya 'rō 'šāh ***ba dūmād-o'bār=ay*** *a='š-īt*
one day king to son_in_law-PL=PC.3SG VCL=say.PRS-3SG
'One day, the king says to his sons-in-law.'

The findings do not entirely match those of Jahani (2018) and Korn (2022: 109), who suggest regular pre-verbal addressees in "Southern Balochi", to which our Coastal Balochi would belong. The findings for Coastal Balochi are also puzzling

Table 8: Frequencies of post-verbal nominal addressees and recipients (includes recipient/benefactives)

	Coastal			Koroshi			Turkmen			
	N	Po	%	N	Po	%	N	Po	%	Totals
Addressees	40	16	40	12	4	33	5	0	0	57
Recipients	19	4	20	3	2	66	5	1	20	27
Totals	59			15			10			84

in view of a general tendency noted in Haig et al. 2024 [this volume], according to which addressees are generally less likely to be post-verbal, which would align with the findings of Jahani (2018) and Korn (2022). However, in our Coastal Balochi sample, this is not the case; we have no explanation for this mismatch; this requires further research.

Turning to pronominal addressees and recipients, Table 9 provides the relevant figures.

Table 9: Frequencies of post-verbal pronominal addressees and recipients (includes recipient/benefactives)

	Coastal			Koroshi			Turkmen			
	N	Po	%	N	Po	%	N	Po	%	Totals
Addressees	13	2	15	8	1	1	6	0	0	27
Recipients	40	9	25	15	8	50	6	2	33	48
Totals	43			23			12			75

With pronominal arguments, the expected trend for higher frequency of postposed recipients is confirmed in all doculects. Examples of nominal addressees are provided in (19a-19c) and a pronominal example is shown in (19d).

(19) a. Coastal Balochi (Nourzaei 2021a, B, 0504)
’nī lōṭā-’ēn-ī ***brā’t-ā̃***
now call-CAUS.PRS-3SG brother-OBL.PL
‘Then he called the brothers.’

b. Coastal Balochi (Nourzaei 2021a, A, 0188)
’nī ǰa’nek-ā go’š-ī ***’to*** *’dar ā*
now girl-OBL say.PRS-3SG 2SG PREV SBJV.come.PRS.2SG
‘Then he [the boy in the well] said to the girl, “You get out.”

c. Koroshi Balochi (Nourzaei 2021b, A, 0320)
’ya ’rō ’šāh ***ba dūmād-o’bār=ay*** *a=’š-īt*
one day king to son_in_law-PL=PC.3SG VCL=say.PRS-3SG
‘One day, the king says to his sons-in-law.’

d. Turkmen Balochi (Haig 2022a, D, 0540)
mnī piss ***pamman*** *sendbad-ī nakl-ā kort*
1SG.GEN father for.1SG Sindbad-GEN story-OBL do.PST.3SG
‘My father told me the story of Sindbad.’

In summary, the findings for addressees and recipients largely confirm a trend observed across other Iranian languages of the region, according to which recipients are generally more likely to be post-verbal than addressees. However, in Coastal Balochi for nominal arguments (Table 8) only, the trend is reversed. We currently lack an explanation for this, which definitely requires more research.

4.4 Complements of ‘become’, Place, Source, Instrument, Benefactive, Comitative

The absolute numbers of tokens for the remaining roles are quite low in several doculects, so quantitative analysis is not always meaningful. We have combined the results in an overview Table 10. Due to low absolute figures, we include all possible POS types, including pronouns, adverbs etc. Figure 2 provides an overview visualization of all roles considered.

With regard to complements of change-of-state verbs (‘become’), Coastal Balochi and Turkmen are overwhelmingly pre-verbal, while Koroshi shows around one third post-verbal placement. This would be expected given that the attested Iranian languages which have dominant post-verbal ‘become’-complements are all from the southwestern Mesopotamian periphery of WATZ (Haig et al. 2022, Haig et al. 2024 [this volume]). An example of post-verbal ‘become’-complement from Koroshi is provided in (20).

(20) Koroshi Balochi (Nourzaei 2021b, A, 0491)
ba ha’m=ī kasa-ō-’ēn ga’hār=eš a=’b-ant
to EMPH=PROX small-DIM-ATTR sister=PC.3PL VCL=become.PRS-3PL

Table 10: Other roles: frequencies of post-verbal placement

	Coastal			Koroshi			Turkmen			
	N	Po	%	N	Po	%	N	Po	%	Totals
'become'-compl.	14	1	7	12	4	33	10	0	0	36
Place	61	12	19.7	21	7	33.3	25	0	0	107
Source	12	1	8.3	7	3	43.9	21	2	9.5	40
Instrument	18	5	28	5	0	0	12	1	8	35
Comitative	18	3	17	13	4	30	18	2	11	49
Benefactive	13	2	15	12	1	8	2	0	0	27

ka'nīz=o ***naw'kar***
maidservant=and male_servant
'[You know, these six sons-in-law and their wives came and] became servants to this their youngest sister.'

Turning to local roles source and place, it has been suggested that oblique arguments generally tend to prefer post-predicate position (Jing et al. 2021, based predominantly on written-language corpora). Above, we have shown that this certainly applies to goals, but the data for source and place provide only weak support for assuming a general tendency applying to all obliques. Nevertheless, it is noteworthy that the other obliques in Table 10 do show notably higher rates of post-verbal placement than direct objects. Examples for place and source are the following:

(21) a. Coastal Balochi (Nourzaei 2021a, A, 0327)
ha'm=ē *'gōšt* *pa'dā* *kašt=ī* ***da'p-ā*** ***'če***
EMPH=PROX meat again pull.PST.3SG=PC.3SG mouth-OBL from
'It took out this meat from its mouth again.'

b. Coastal Balochi (Nourzaei 2021a, A, 0063)
'hanga *'nešt* *ha'm=ē* ***'mačč-e*** ***'čērā***
still sit.PST.3SG EMPH=PROX date_palm-GEN under
'Still he [the boy] was sitting under this date-palm.'

c. Koroshi Balochi (Nourzaei 2021b, A, 0217)
'kār *a=kan-ān* ***'mā*** ***'ī*** ***bā'ġ-ā***
work VCL=do.PRS-1SG in PROX garden-OBL
'I will work in this garden.'

d. Koroshi Balochi (Nourzaei 2021b, B, 0612)
ba mad'rasā=om *ġada'ġan=eš* *kod-a*
to school.OBL=ADD forbidden=PC.3PL do.PST-PP
'At school they have actually forbidden (your) coming.'

The data for instruments, comitatives, and benefactives are quite thin, but suggest a tendency for all three roles to be predominantly pre-verbal. This is particularly noteworthy for benefactives, which are >80% pre-verbal in all doculects. Benefactives are sometimes included under the umbrella term of "target" (Asadpour 2022a,b, Korn 2022) together with goals, but the Balochi data suggest that benefactives are far less likely to be postverbal than goals, and in fact no more likely to be postverbal than sources and locations. However, it should be noted that the majority of benefactives (about 90%) in our data are pronominal, so the bias towards preverbal position may be linked to the pronominal status, but given the paucity of nominal benefactives in the data, this remains to be clarified.

5 Summary: Post-verbal constituents in Balochi

Figure 2 summarizes the quantitative data from post-verbal positioning of various non-subject constituents considered in the preceding section. The values for the roles Recipient, Addressee, Benefactive, and Comitative are combined to "Oblique" in Figure 2; see the preceding section for individual roles.

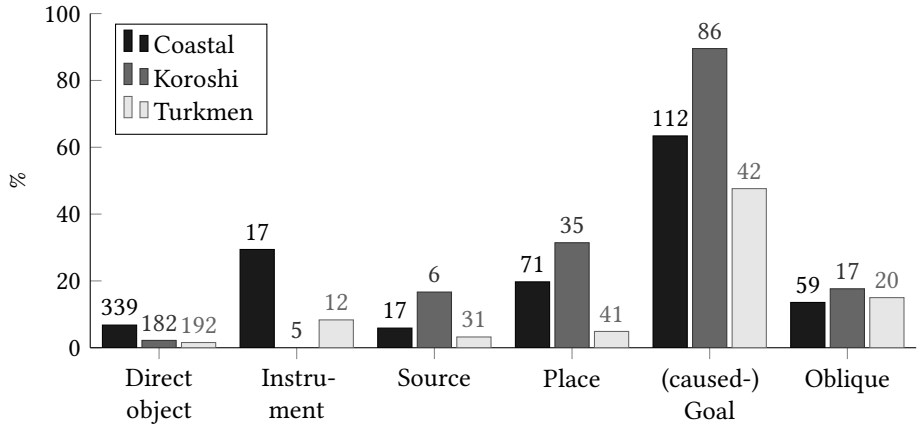

Figure 2: Frequency of post-verbal arguments in three Balochi doculects. *Note*: Ticks above bars indicate absolute count.

As expected, the highest rates of post-verbal placement for all doculects are found for goals, confirming the by-now familiar finding across the entirety of WOWA. Note that it is only in Koroshi that both recipients and goals are dominantly (>60%) post-verbal, a pattern that has parallels with varieties of Kurdish (Haig 2022c). In our data, addressees do not pattern with goals in any doculect, and overall do not differ from other obliques, showing less than 50% post-verbal placement. With regard to direct objects, all doculects are overwhelmingly OV, with post-verbal objects accounting for less than 10% of all objects. However, Coastal Balochi shows a significantly higher number of post-verbal objects than either of the other two doculects. In our sample, weight and definiteness do not seem to be relevant in accounting for post-verbal direct objects, in any doculect. We do, however, find an effect of humanness, particularly in combination with pronominal direct objects (Table 5 above). Overall, the most likely direct objects to be postposed are human, pronominal objects, with the difference most obvious for Coastal Balochi.

6 Conclusions and implications for comparative Iranian syntax

Table 11 summarizes the overall degree of head-initial syntax in various phrase types in the three doculects, based on the discussion in Section 3.

Table 11: Distribution of head-initial syntax in three Balochi doculects

Parameter	Koroshi	Turkmen	Coastal
N/Adjective	N-Adj	Adj-N	
N/Possessor	N-Poss	Poss-N	
Adp/N	Prep-N	Prep-N	N-Postp
Main clause/Compl. clause	Main-Compl. Clause		
Aux(Modal)/Main verb	Aux(Modal)-Main verb		
Complementizer/Clause	Compl-Clause		

Across these typological parameters, Koroshi is the only variety that is consistently head-initial, presumably linked to a greater degree of contact with Persian and other western Iranian languages exhibiting head-initial NPs. All doculects linearize main and complement clauses, complementizers, and auxiliaries alike, sharing these values with probably all other Iranian languages of WOWA. The most surprising aspect of Table 11 is the dominance of postpositions in Coastal

Balochi, for which we have no ready explanation; this is certainly an area for future research.

Finally, we consider the relevance of the Balochi data for addressing the larger question of how post-verbal syntax in Iranian may have emerged: What kind of historical scenario is compatible with the areal distribution that is now becoming evident? First, let us consider the Balochi data in the light of a feature that characterizes all other contemporary spoken western Iranian language in the WOWA sample: in connected discourse, nominal goals (including caused-motion goals) are 60–100% post-verbal. The sole exception turns out to be Turkmen Balochi (Turkmen), where the figure drops to around 40%. Turkmen Balochi is thus different from its relatives in this respect, and from any of the Iranian languages spoken in Mesopotamia, where post-verbal Goals are >85%. The most obvious explanation for the outlier status of Turkmenistan Balochi is its outlier geographic position, beyond the northeastern periphery of WATZ, and farthest from the Mesopotamian core that has been identified as harbouring the highest levels of post-verbal Goals. We take this as tentative support for the hypothesis that the placement of goals is particularly sensitive to contact influence.

Nevertheless, 40% post-verbal Goals is higher than many other OV languages in our sample (vernacular standard Turkish of Ankara, Iefremenko 2024 [this volume] for example), so geography cannot be the whole story. Generally, it appears that Iranian languages are just more prone to post-verbal Goals than Turkic languages, and this is a matter of shared inheritance, rather than geography alone. What we might therefore assume is that the shared ancestor of all western Iranian languages had some initial baseline level of post-verbal spatial goals, which may already have exceeded 50% in some of the (unattested) languages of that period. This feature was thus present prior to the dispersion of the various branches of what is traditionally termed 'West Iranian' (see also Korn 2022: 122). The result would be traces of post-verbal Goals (in the narrow sense) in all West Iranian languages, with actual levels dependent on their respective contact biographies over the last 2000 years. In those Iranian languages that shared territory with with Semitic languages (in particular Aramaic) for at least a millennium in the Mesopotamian and western Zagros regions, levels of post-verbal Goals converge to the near 100% that characterize the Semitic contact languages. Where Semitic influence is lacking, and contact with other OV languages (e.g. Indo-Aryan) was present, levels would have dropped, as in Balochi of Turkmenistan, or at least not risen further. Alternatively, the unrelated OV contact languages might converge to Iranian in this respect, as is the case for Qashqai Turkic, which exhibits 70% post-verbal nominal goals (Schreiber 2021), a figure close to its Iranian neighbours. The actual outcome would thus depend

on the local patterns of multilingualism and power relations among the speech communities, but a broad geographical tendency is nonetheless discernible.

This is essentially a refinement of the account outlined in Haig (2015), who noted that the wide distribution of post-verbal "Goals" (see below on terminology) across the Iranian languages of western Asia is suggestive of "an old trait of West Iranian that was inherited by the daughter languages," rather than independent parallel development (Haig 2015: 421). post-verbal Goals are attested in the Old Iranian period, as in the following Old Persian examples, cited from Haig (2015: 421); see also Jahani (2018) for further examples:

(22) a. Old Persian (DB V, 9–10, Kent 1953: 133)
pasāva Gaubaruva hadā kārā ašiyava ***ūvjam***
then Gobryas with army marched to_Elam
'Then Gobryas marched to Elam with an army.'

b. Old Persian (DB V, 21, Kent 1953: 133)
pasāva hadā kārā ašiyavam abiy ***sakām***
then with army went.1SG to Scythia
'Then I went with an army to Scythia.'

However, the formulation in Haig (2015) was still framed in terms of an assumed macro-role sense of "Goal", which included addressees, and recipients, among other things. More recent research (Iefremenko 2024 [this volume]), including this chapter, demonstrates that this approach is not tenable. Instead, a finer-grained approach is required that consistently distinguishes Goal in the narrower sense of motion and caused motion from other roles. As shown above in Figure 2, outside of spatial goals, no other role surpasses 50% rates of post-verbal placement in Balochi, with the exception of recipient, and that only in Koroshi, and this finding is replicated for several other Iranian languages.

Our revised suggestion would be that in proto-West Iranian, post-verbal Goals (in the narrow sense) were already a frequent option (perhaps majority), while postposing other arguments was possible, but determined by pragmatic and semantic principles, with no clear role-driven preferences; the multi-variate analysis of spoken Persian in Rasekh-Mahand et al. 2024 [this volume], demonstrates such a system empirically. In some Iranian languages, most notably Kurdish of the Mesopotamian region, post-verbal placement spread from goals to other argument types, ultimately encompassing a bundle of semantically-related roles. The initial focus on these languages led to the (misleading) assumption of a macro-role ("Goal" in Haig & Thiele 2014 and Haig 2015; "Target" in Asadpour 2022b and Jügel 2022; see Haig et al. 2024 [this volume], for discussion. The role-specific

analysis adhered to here suggests an areally-mediated shift in the frequency of post-verbal Goals in Balochi (lowest in Turkmen), and in the range of arguments that behave similarly to goals (the Koroshi clustering of recipient and Goal).

The developments just sketched are speculative to the extent that they reconstruct possible diachronic developments, based on synchronic feature distributions. The little historical data that is available is unfortunately difficult to interpret. There is a general problem with projecting from written texts to spoken language; an investigation of contemporary formal written Persian, for example, would fail to register the fact that spoken contemporary Persian has a fairly stable rate of around 80% post-verbal Goals (Haig et al. 2024 [this volume]). Early Classical Persian texts (11–13 centuries AD) have virtually no trace of post-verbal Goals (Parizadeh & Rasekh-Mahand 2024 [this volume]), but we doubt that these figures reflect the spoken language of the time, given the comparative evidence from other Iranian languages. Similarly, Jahani (2018) compares oral and written Southern Balochi, noting that in written Balochi, goals are never post-posed, while in spoken Southern Balochi the relevant figure is 90%. These differences raise a methodological problem with regard to reconstructing the history of post-verbal syntax; differences between yesterday's written and today's spoken language may just reflect differences between written and spoken registers that have always existed, rather than being evidence of any change across time.

This chapter provides a tentative cross-dialect survey of word order in Balochi, based on three doculects representing geographically dispersed varieties, with differing contact profiles. The picture that emerges confirms some of the previous observations in Jahani (2018) and Korn (2022), but extends the purview of these studies. We investigated a larger number of roles, and considered additional factors such as weight and humanness. On the other hand, the pan-dialectal approach means that we are unable to zoom in on finer semantic, pragmatic, or stylistic factors, such as the effects of individual verb classes, register, or information structure that figure in Korn (2022). Generally, our findings confirm the overall prediction that both overall frequencies of post-verbal Goals, and the number (and nature) of other role types that are prone to post-verbal placement, decrease with increasing distance from the Mesopotamian core of WATZ. We have also identified Coastal Balochi as exhibiting a significantly higher frequency of post-postions in discourse, and interestingly, a higher number of post-verbal direct objects, with a strong interaction with pronominal form and humanness. These findings raise interesting questions regarding the possible role of contact with Indo-Aryan languages in the region, which have yet to be explored.

Abbreviations

ADD	additive particle	NEG	negation
ATTR	attributive	OBJ	object case
BACKG	backgrounding	OBL	oblique case
CAUS	causative	PC	person-marking enclitic
CLM	complementizer	PL	plural
Coastal	Coastal Balochi doculect	PP	past participle
COP	copula	PROH	prohibitive prefix
DEF	definite	PROX	proximal
DIST	distal	PRS	present
EMPH	emphasis	PST	past stem
EZ	ezafe particle	REFL	reflexive pronoun
FOC	focus particle	SBJV	subjunctive
GEN	genitive	SG	singular
IND	indefinite	Turkmen	Turkmenistan Balochi doculect
INDV	individuation clitic	UP	Unpublished text
IPFV	imperfective	VCL	verbal clitic
K.IPFV	*k*-variant of imperfective	WATZ	Western Asian Transition Zone
Koroshi	Koroshi Balochi doculect		

References

Anonby, Christina van der Wal. 2015. *A grammar of Kumzari: A mixed Perso-Arabian language of Oman*. Leiden: Leiden University. (Doctoral dissertation).

Asadpour, Hiwa. 2022a. *Typologizing word order variation in northwestern Iran*. Frankfurt am Main: Goethe-Universität. (Doctoral dissertation).

Asadpour, Hiwa. 2022b. Word order in Mukri Kurdish - The case of incorporated targets. In Hiwa Asadpour & Thomas Jügel (eds.), *Word order variation: Semitic, Turkic and Indo-European languages in contact*, 63–87. Berlin: De Gruyter Mouton.

Axenov, Serge. 2000. Balochi orthography in Turkmenistan. In Carina Jahani (ed.), *Language in society – Eight sociolinguistic essays on Balochi* (Studia Iranica Upsaliensia 3), 71–78. Uppsala: Acta Universitatis Upsaliensis.

Axenov, Serge. 2006. *The Balochi language of Turkmenistan: A corpus-based grammatical description (StudiaIranica Upsaliensia, 10)*. Uppsala: Acta Universitatis Upsaliensis.

Baranzehi, Adam Nader. 2003. The Sarawani dialect of Balochi and Persian influence on it. In Carina Jahani & Agnes Korn (eds.), *The Baloch and their neighbours: Ethnic and linguistic contact in Balochistan in historical and modern times*, 75–111. Wiesbaden: Reichert.

Barjasteh Delforooz, Behrooz. 2010. *Discourse features in Balochi of Sistan (oral narratives)*. Uppsala: Uppsala Universitet. http://uu.divaportal.org/smash/record.jsf?pid=diva2:345413.

Buddruss, Georg. 1988. *Aus dem Leben eines jungen Balutschen, von ihm selbst erzählt*. Stuttgart: Steiner.

Emādi, Nezam. 2005. *Guyeš-e Korowš (The vernacular of the Korosh)*. Shiraz: Entešārāt-e Širāz Āvand Andišeh.

Haig, Geoffrey. 2015. Verb-goal (VG) word order in Kurdish and Neo-Aramaic: Typological and areal considerations. In Geoffrey Khan & Lidia Napiorkowska (eds.), *Neo-Aramaic and its linguistic context*, 407–425. Piscataway, NJ: Gorgias Press.

Haig, Geoffrey. 2022a. Balochi (Turkmenistan). In Geoffrey Haig, Donald Stilo, Nils N. Schiborr & Mahîr Dogan (eds.), *WOWA — Word Order in Western Asia: A spoken-language-based corpus for investigating areal effects in word order variation*. Bamberg: University of Bamberg. multicast.aspra.uni-bamberg.de/resources/wowa/.

Haig, Geoffrey. 2022b. Kumzari (Musandam). In Geoffrey Haig, Donald Stilo, Nils N. Schiborr & Mahîr Dogan (eds.), *WOWA — Word Order in Western Asia: A spoken-language-based corpus for investigating areal effects in word order variation*. Bamberg: University of Bamberg. multicast.aspra.uni-bamberg.de/resources/wowa/.

Haig, Geoffrey. 2022c. Kurdish (Northern, Muş). In Geoffrey Haig, Donald Stilo, Nils N. Schiborr & Mahîr Dogan (eds.), *WOWA — Word Order in Western Asia: A spoken-language-based corpus for investigating areal effects in word order variation*. Bamberg: University of Bamberg. multicast.aspra.uni-bamberg.de/resources/wowa/.

Haig, Geoffrey, Paul M. Noorlander & Nils N. Schiborr. In press. Which word order features are stable in a contact setting? Corpus-based evidence from the Western Asian transition zone. In Jeroen Darquennes, Joe Salmons & Wim Vandenbussche (eds.), *Language contact: An international handbook*, vol. 2. Berlin: De Gruyter Mouton.

Haig, Geoffrey, Mohammad Rasekh-Mahand, Donald Stilo, Laurentia Schreiber & Nils Schiborr. 2024. Post-predicate elements in the Western Asian Transition Zone: Data, theory, and methods. In Geoffrey Haig, Mohammad Rasekh-Mahand, Donald Stilo, Laurentia Schreiber & Nils N. Schiborr (eds.), *Post-predicate elements in the Western Asian Transition Zone: A corpus-based approach to areal typology*, 3–54. Berlin: Language Science Press. DOI: 10.5281/zenodo.14266331.

Haig, Geoffrey, Donald Stilo, Mahîr C. Doğan & Nils N. Schiborr. 2022. *WOWA — Word Order in Western Asia: A spoken-language-based corpus for investigating areal effects in word order variation.* Bamberg. multicast.aspra.uni-bamberg.de/resources/wowa/.

Haig, Geoffrey & Hannah Thiele. 2014. Post-predicate Goals in Northern Kurdish and neighbouring languages: A pilot study in quantitative areal linguistics languages. Paper presented at the *Second International Conference on Variation and Change in Kurdish, mardin artuklu university, turkey,* 8-9 October 2014.

Hodgson, Katherine, Victoria Khurshudyan & Pollet Samvelian. 2024. Post-predicate arguments in Modern Eastern Armenian. In Geoffrey Haig, Mohammad Rasekh-Mahand, Donald Stilo, Laurentia Schreiber & Nils N. Schiborr (eds.), *Post-predicate elements in the Western Asian Transition Zone: A corpus-based approach to areal typology*, 375–410. Berlin: Language Science Press. DOI: 10.5281/zenodo.14266355.

Iefremenko, Kateryna. 2024. Word order in the speech of Kurmanji-Turkish bilinguals. In Geoffrey Haig, Mohammad Rasekh-Mahand, Donald Stilo, Laurentia Schreiber & Nils N. Schiborr (eds.), *Post-predicate elements in the Western Asian Transition Zone: A corpus-based approach to areal typology*, 55–85. Berlin: Language Science Press. DOI: 10.5281/zenodo.14266333.

Jahani, Carina. 2013. The Balochi language and languages in Iranian Balochistan. *Journal of the Middle East and Africa* 4(2). 153–167.

Jahani, Carina. 2018. Post-verbal arguments in Balochi. Paper presented at *the International Conference on Anatolia-Caucasus-Iran: Ethnic and linguistic contacts (ACIC),* Yerevan University, 10–12 May.

Jahani, Carina & Agnes Korn. 2009. Balochi. In Gernot Windfuhr (ed.), *The Iranian languages*, 634–692. London & New York: Routledge.

Jing, Yingqi, Paul Widmer & Balthasar Bickel. 2021. Word order variation is partially constrained by syntactic complexity. *Cognitive Science* 45(11). 1–25. DOI: 10.1111/cogs.13056.

Jügel, Thomas. 2022. Word order variation in Middle Iranic: Persian, Parthian, Bactrian, and Sogdian. In Hiwa Asadpour & Thomas Jügel (eds.), *Word or-*

der variation – Semitic, Turkic and Indo-European languages in contact, 39–62. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110790368-003.

Kent, Roland. 1953. *Old Persian*. 2nd revised edition. New Haven, CT: American Oriental Society.

Korn, Agnes. 2005. *Towards a historical grammar of Balochi*. Wiesbaden: Reichert.

Korn, Agnes. 2008. A new locative case in Turkmenistan Balochi. *Iran and the Caucasus* 12(1). 83–99.

Korn, Agnes. 2022. Targets and other post-verbal arguments in Southern Balochi: A multidimensional cline. In Hiwa Asadpour & Thomas Jügel (eds.), *Word order variation – Semitic, Turkic and Indo-European languages in contact*, 89–128. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110790368-005.

Korn, Agnes & Maryam Nourzaei. 2019. Notes on the speech of the Afro-Baloch of the southern coast of Iran. *Journal of the Royal Asiatic Society* 29. 623–657. DOI: 10.1017/S1356186319000300.

Noorlander, Paul M. 2024. Neo-Aramaic in Iran and northeastern Iraq. In Geoffrey Haig, Mohammad Rasekh-Mahand, Donald Stilo, Laurentia Schreiber & Nils N. Schiborr (eds.), *Post-predicate elements in the Western Asian Transition Zone: A corpus-based approach to areal typology*, 431–469. Berlin: Language Science Press. DOI: 10.5281/zenodo.14266359.

Nourzaei, Maryam. 2017. *Participant reference in three Balochi dialects: Male and female narrations of folktales and biographical tales*. Uppsala: Acta Universitatis Upsaliensis. https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-314090.

Nourzaei, Maryam. 2021a. Balochi (Coastal). In Geoffrey Haig, Donald Stilo, Nils N. Schiborr & Mahîr Dogan (eds.), *WOWA — Word Order in Western Asia: A spoken-language-based corpus for investigating areal effects in word order variation*. Bamberg: University of Bamberg. multicast . aspra . uni - bamberg . de / resources/wowa/.

Nourzaei, Maryam. 2021b. Balochi (Koroshi). In Geoffrey Haig, Donald Stilo, Nils N. Schiborr & Mahîr Dogan (eds.), *WOWA — Word Order in Western Asia: A spoken-language-based corpus for investigating areal effects in word order variation*. Bamberg: University of Bamberg. multicast . aspra . uni - bamberg . de / resources/wowa/.

Nourzaei, Maryam. 2024. Progressive constructions in different Balochi dialects from a diachronic and an areal linguistic point of view. *Journal of the Bulletin of the School of Oriental and African Studies*. 1–26. DOI: 10 . 1017 / S0041977X24000144.

Nourzaei, Maryam. In press. *Documenting orality: Collecting folk narratives in Iran*. Uppsala: Acta Universitatis Upsaliensis.

Nourzaei, Maryam, Carina Jahani, Erik Anonby & Abbas Ali Ahangar. 2015. *Koroshi: A corpus-based grammatical description*. Uppsala: Acta Universitatis Upsaliensis.

Parizadeh, Mehdi & Mohammad Rasekh-Mahand. 2024. Post-predicate elements in Early New Persian (10–13th Century CE). In Geoffrey Haig, Mohammad Rasekh-Mahand, Donald Stilo, Laurentia Schreiber & Nils N. Schiborr (eds.), *Post-predicate elements in the Western Asian Transition Zone: A corpus-based approach to areal typology*, 231–244. Berlin: Language Science Press. DOI: 10.5281/zenodo.14266345.

Rasekh-Mahand, Mohammad, Elham Izadi, Mehdi Parizadeh, Geoffrey Haig & Nils Schiborr. 2024. Post-predicate elements in modern colloquial Persian: A multifactorial analysis. In Geoffrey Haig, Mohammad Rasekh-Mahand, Donald Stilo, Laurentia Schreiber & Nils N. Schiborr (eds.), *Post-predicate elements in the Western Asian Transition Zone: A corpus-based approach to areal typology*, 197–230. Berlin: Language Science Press. DOI: 10.5281/zenodo.14266343.

Schreiber, Laurentia. 2021. Oghuz (Qashqai). In Geoffrey Haig, Donald Stilo, Nils N. Schiborr & Mahîr Dogan (eds.), *WOWA — Word Order in Western Asia: A spoken-language-based corpus for investigating areal effects in word order variation*. Bamberg: University of Bamberg. multicast . aspra . uni - bamberg . de / resources/wowa/.

Stilo, Donald. 2018. Preverbal and postverbal peripheral arguments in the Araxes-Iran linguistic area. Paper presented at the *Conference Anatolia-Caucasus-Iran: Ethnic and Linguistic Contacts* 10-12 May 2018. Yerevan University, Yerevan, Armenia.