Post-predicate elements in Iranian and neigbouring languages: Inheritance, contact, and information structure

Position paper, for the project on
Post-predicate constituents in Iranian and neighbouring languages.

Geoffrey Haig   Mohammad Rasekhmahand

February 2019

0   Introduction

In terms of conventional word-order typology, the Iranian languages are collectively characterized as 'SOV', rendering them superficially on a par with rigidly verb-final languages such as Japanese. However, a significant number of Iranian languages regularly places certain constituents after the verb. While this is well documented for complement clauses (e.g. in Persian, where complement clauses are consistently post-verbal), it also applies to certain nominal arguments, yielding various kinds of verb-medial structures. These pass under the radar of conventional word-order typology, and if mentioned at all, are generally dismissed as pragmatically-motivated scrambling, rather than a significant syntactic trait of the languages. However, our research to date indicates that a purely performance-based, pragmatic explanation fails to account for the internal structural consistency, the commonalities across different Iranian languages, and the token frequency of the relevant constructions in the available corpora.

The main goal of this project is to compile a rich and structured data set from a sample of Iranian languages, combining corpus-based and questionnaire methodologies. Parallel to this we will investigate selected Middle Iranian corpora, and unrelated contact languages, in order to reconstruct the genesis and historical development of this aspect of Iranian syntax. The outcome will deepen our understanding of syntactic typology by going beyond the coarse-grained parameters of conventional word-order typology, but will also impact on our understanding of the nature of syntactic change, and the respective roles of inheritance and

contact in shaping syntax. This paper provides a state-of-the art overview of the topic, and outlines the intended structure and outcomes of the project.

1   Theoretical background

Within word-order typology, the relative ordering of direct object and verb (OV vs. VO) has retained a central role as the most predictive parameter of cross-linguistic syntactic variation. The ordering of other verbal arguments has attracted much less attention, with the exception of Dryer and Gensler (2011), and Hawkins (2008). Hawkins, working in a performance-based explanatory framework, focusses on the cross-linguistically attested ordering of additional obliques (e.g. Recipients, Goals, Locatives etc., abbreviated here X) in transitive clauses. He points to an asymmetry between OV and VO languages in this respect: while OV languages frequently permit disharmonic orderings of other arguments (i.e. OVX), in VO languages, disharmonic orders (i.e. XVO) are almost unattested. Explanations for this asymmetry can be formulated in terms of minimizing processing domains, i.e. the length of strings that are required to be parsed before syntactic resolution is achieved, but can also be related to the relative heaviness of constituents, and the tendency to prefer shorter before longer constituents. This line of research has a long tradition in corpus linguistics and psycho-linguistics, where processing effects are reflected in statistically significant correlations between constituent length and ordering preferences (Arnold et al 2000, Wasow 2002), or psycholinguistically measurable variables such as reaction times. Within Minimalism, attempts have been made to account for these asymmetries within the Final-over-Final Condition (see Sheehan 2013, Sheehan et al 2017 for recent discussion and references). This approach likewise predicts asymmetries in branching directionality, when structures with opposing head directionality are combined, but considers purely processing-based accounts to be inadequate. Rather, the Final-over-Final-Condition is taken to be part of universal syntactic constraints, which are also operative in e.g. auxiliary/verb ordering, or constituent ordering within the NP/DP. A third approach to accounting for asymmetries in the VP is in terms of the outcome of contingent historical processes. For example, Nikitina (2011) accounts for OVX order in Mande languages in these terms. In Mande, direct objects obligatorily precede the verb, while other verbal arguments must follow (OVX order). Nikitina relates this to the origin of the current finite verbs from nominalized forms, which previously headed NPs and inherited the word-order properties of NPs when they were re-analyzed as finite verbs.

2 Post-predicate constituents in Iranian: research overview

2.1 Persian

Word-order variation in Iranian has been most intensely researched for Persian. Early work by Birner and Mahootian (1996) considered the pragmatics of fronting in Persian (XSV) order, while Karimi (2003) investigates scrambling in Persian within a Minimalist framework, focussing on the relative ordering of direct and indirect objects (DO-IO/PP) (see also Ghomeshi 1997 and Ganjavi 2007). According to Karimi (2003), a DO is marked with *rā* when 'specific', and appears before the IO/PP; but, if it is non-specific (without *rā*), it appears after the IO/PP, adjacent to the verb. The attested variation is thus motivated by informational-discoursal factors of the DO, such as being in focus or topicalized. More recently, Faghiri and Samvelian (2015) and Faghiri et al (2014) have applied a corpus-based, quantitative methodology to the DO-IO/PP alternation, adopting a multi-factorial approach. They find that a simple specific/non-specific distinction does not adequately explain the attested variability. They also examine the effect of heaviness, following Hawkins (1994, 2004, 2014), assuming the "short-before-long" tendency that has been confirmed in other SOV languages like Japanese and Korean. The effect of syntactic length is detectable with indefinite and bare-modified DOs, confirming the short-before-long tendency observed for Japanese (Yamashita and Chang 2001, 2006). However, they consider cognitive accessibility to be the prime determinant, rather than syntactic length per se. Note however that both the generative approach of Karimi (2003) and the corpus-based work of Faghiri and Samvelian (2015) and Faghiri et al (2014) only consider the ordering of **pre**-predicate elements; like most contemporary research on Persian syntax, the basic assumption is that Persian is verb-final (e.g. Taleghani 2008), and post-predicate elements are relegated to the pragmatic margins, outside the purview of regular syntactic processes.

Post-predicate placement of relative clauses is investigated by Rasekh-Mahand et al (2016). They show that information status, grammatical weight and verb class are three main motivations for relative clause extraposition. Grammatical weight of the RC in relation to VP plays the main role in relative clause extraposition; verb class and information status are ranked lower. These findings support Hawkins' (2004) domain minimization principle and provide additional evidence for the hypothesis that Persian, nevertheless regularly employs the post-predicate position for pragmatic alternations.

Turning to nominal post-predicate constituents, it has regularly been noted (e.g. Lazard 1957 [2002]) that in spoken and informal Persian, certain constituents may optionally be placed post-verbally. The following examples illustrate goal arguments in Persian, with the assumed canonical pre-predicate placement in (1a) contrasting with the post-predicate placement in (1b):

(1a) ø   *mixād*   *(be)*   *madrese*   *bere*
      PRO   want.PRS.3SG (to)   school   SUBJ.go.3SG
      'He/she wants to go to school.'

(1b) ø   *mixād*   *bere*   *(be)*   *madrese*
      PRO   want.PRS.3SG SUBJ.go.3SG (to)   school
      'He/she wants to go to school.'

Structures such as (1b) were first systematically investigated for spoken Persian by Frommer (1981), based on a corpus of spontaneous colloquial speech, and non-spontaneous, but informal, language (Frommer 1981: 68-79), with a total of 5,784 clauses. Frommer documents high rates of post-predicate positions, but with the variation sensitive to both degree of formality, and to the nature of the constituents and the verb type. His results are summed up in the following hierarchy of post-posability for casual speech:

(2)   POSTPOSABILITY HIERARCHY FOR CASUAL STYLE (Frommer 1981: 172)
       Goal (without preposition) > Goal (with preposition) > PP (non-Goal, including
       IO) > DO (with *râ*) and ADV (without preposition) > SU > DO (without *râ*)

Frommer shows that the goals with the most frequent verbs of motion *raftan* (to go) and *âmadan* (to come) are overwhelmingly post-predicate (87.1% and 91.4% respectively), while the goal of the caused-motion verb *gozâštan* 'to put, place' is almost categorically post-verbal, a finding that is replicated for contemporary spoken Persian in Haig (2017), based on Adibifar (2016). Frommer explores the effect of other factors, noting that clause type, main or subordinate, or the verb type, simple or compound, have no effect on the frequency of postposing (Frommer 1981: 167-169), and syntactic weight is only weakly correlated with post-verbal position in the more formal registers investigated by Frommer. In other styles, Frommer finds no evidence for the effect of constituent heaviness (Frommer 1981: 170).

In sum, post-predicate position of goals in the spoken language is both highly frequent, and relatively impervious to pragmatic factors; it is difficult to reconcile these findings with the prevalent view, according to which post-verbal constituents are the result of occasional scrambling in informal registers and pragmatically marked contexts. To our knowledge, no convincing alternative explanations have been formulated, even in studies purporting to analyse Persian discourse structure (Roberts 2009). To date, with the exception of Frommer

(1981), all serious research on Persian word order is based on the written language,[1] a fact which we believe distorts the overall picture.

2.2 Other Iranian languages

The documentary record of other Iranian languages, most of which lack an officially codified written norm, reveals very frequent instances of post-predicate constituents, most typically goals. The following are typical (goals in bold type):

(3) *bævæsd*   *man*   *awe*
     jump.PST.3SG   into   water.OBL
     'He jumped **into the water**.' (Vafsi; Stilo 2005: 231)

(4) *sova*   *vamigardom*   *delvar*
     tomorrow   return.PRS.1SG   Delvar
     'Tomorrow I will return **to Delvar**.' (Delvari, Bušehr; Haig and Nemati 2013)

(5) *wat-ī*   *mardumānā*   *dēm*   *day*   *mnī*   *baggay*   *sarā*
     REFL-GEN   people.PL.OBL   face   give.IMP.SG   my   camel_herd   to
     '... send your people **to my herd of camels**.' (Sistan Balochi; Delforooz 2010: 331)

Haig (2014, to appear) focusses on the position of 'Goal' arguments in Kurdish, whereby Goal is a cover term for goals of verbs of (caused) motion, recipients of verbs of transfer, and addressees of verbs of speech. In Kurdish, goals of verbs of motion, and of the verb 'to give' are obligatorily in post-predicate position. Unlike Persian, then, this cannot be interpreted as pragmatically-driven scrambling, but is part of the hard facts of syntax. This is illustrated in (6)-(9) for the Behdinî dialect of Northern Kurdish (cf. Haig, to appear, for details). Example (6) illustrates the only possible position for the goal *Dohuk*, which takes the Oblique case, and immediately follows the predicate:

---

[1] The recent dissertation of Vafaeian (2018) draws on 'spoken Persian', but the material is actually from Persian film screenplays, i.e. is pre-scripted spoken language, rather than spontaneous speech.

(6)   *Dihî*   *ez*   *ču-bû-m = e*   **Dohuk-ê**
        Yesterday   1SG   go.PST-PLPRF-1SG = DRCT   Duhok-F.OBL
        'Yesterday I went to Dohuk.'

The Goal argument cannot be separated from the verb by the adverb (7):

(7)   **Ez*   *ču-bû-m = e*   *dihî*   **Duhok-ê**
        1SG   go.PST-PLPRF-1SG = DRCT   yesterday   Duhok-F.OBL
        Intended: 'I went to the market yesterday'

The placement of other arguments is relatively flexible, for example following the Goal in (8), but they cannot intervene between Goal and verb:

(8)   *Ez*   *ču-bû-m = e*   **Duhok-ê**   *[digel*   *deyk-a*   *xwe]*
        I   go.PST-PLPRF-1SG = DRCT   Duhok-F.OBL   with   mother-F.EZ   REFL
        'I went to Duhok [with my mother].'

Nor can the goal argument be moved in front of the verb (9):

(9)   **Dihî*   *ez*   **Duhok-ê**   *ču-bû-m = e*
        Yesterday   I   Duhok-F.OBL   go.PST-PLPRF-1SG = DRCT

Furthermore, the position of the Goal argument remains unchanged if the clause is embedded under a matrix verb like 'say' or 'believe', or is in a relative clause. There is no doubt that post-predicate position of these arguments is syntactically determined, and that Northern Kurdish (and probably most other varieties of Kurdish) could be classified as OVX basic word order.

    Notably, the syntactically fixed post-predicate arguments in Kurdish are of the same type as those that are most frequently post-posed in colloquial Persian (e.g. goals of verbs of motion, see Frommer's (1981) hierarchy in (2) above). Thus the 'soft constraints' of Persian discourse (statistically significant, but not fully obligatory) which determine positional variation are mirrored by the hard constraints of Kurdish syntax. Similar patterns are observed for a number of other Iranian languages, including Vafsi, Delvari, Gorani, and certain dialects of Balochi (Haig 2014, 2017, Jahani 2018, Stilo 2018), though it is generally not been established whether post-predicate position in these languages is grammatically fixed, or only discourse-frequent. Stilo (2018) considers data from 29 languages from the Araxes-Iran Linguistic Area, including

both Iranian and non-Iranian (e.g. Armenian, Turkic, Semitic). He suggests the following hierarchy of likelihood for post-predicate positioning:

(10)   Goal   >   Recipient   >   Benefactive ~Addressee

Haig, to appear, independently formulates a similar hierarchy for Kurdish:

(11)	goal of verbs of motion and caused motion / recipient of *dan* 'give'	>	addressee of *gotin* 'say, tell'	>	other obliques, light verb complements	>	direct objects, copular complements

It seems unlikely that this kind of similarity would arise across an area including most of West Iran, North Iraq, Eastern Turkey and Armenia by chance. Nor is it very plausible to assign the patterns to universally operative pressures of, for example, information packaging. There is no obvious link between the semantics and syntax of goal and recipient arguments and, for example, pragmatically relevant factors such as 'afterthought', 'focus', or a particularly heavy internal structure, which might favour a statistical preponderance of post-predicate placement. In fact, investigations into the information status of locatives and goals in Schnell et al (2018) indicate that these argument types are actually very likely to code given information. In particular, recipients are overwhelmingly pronominal in discourse, thus would not be expected to occur later in the clause. In other words, there is no plausible reason in terms of information packaging why post-predicate position should be so consistently associated with goals and recipients in the Iranian OV languages. Nor is there any known typological parallel for this phenomenon (Haig, to appear, §4). At present then, we consider it unlikely that the Iranian pattern can be meaningfully explained in terms of information structure, though of course different information structural configurations may serve to reinforce or weaken the observed tendencies.

2.3   The areal dimension

The available data offer suggest that patterns of post-predicate elements are significantly sensitive to geographic location. More generally, Iranian languages of the Mesopotamian region (modern North Iraq, West Iran, and southeastern Turkey), tend to be more likely to exhibit post-predicate constituents than those further north and east. Figures reflecting the general tendency for post-predicate position of recipients in Iranian languages are provided in Table 1 (see also Haig & Khan 2018, for additional support).

Table 1: Frequencies post-posed recipients in selected Iranian languages (Stilo 2018)

POST-PREDICATE POSITION	VARIETY, APPROXIMATE LOCATION
>90%	Zakho (Northern Kurdish), North Iraq
	Mukri (Central Kurdish), West Iran
>70%	Vafsi (Southern Tatoid), West Iran
	Zazaki (?), Central Anatolia
>40%	Lerik (Caspian),
	Caucasian Tat, Azerbaijan

The initial indications are that geographic location, and the nature of the respective contact languages, is a better predictor of degree of post-positioning than genetic inheritance. Contrasting for example Turkic of Central Anatolia (Erzincan) with that of Tabriz, under heavy Iranian contact influence, also revealed a very significant increase in the extent to which various kinds of non-direct-object occur post-verbally (Reetz 2015) in the Tabriz variety as opposed to the Erzincan dialect of Turkish. Recent work by Schroeder (2018) has pointed to contact influence among Turkish speakers with an L1 Kurdish, who display significantly higher rates of post-verbal clausal complements than comparable groups of monolingual Turkish speakers.

2.4 Historical aspects

Given the current broad distribution of post-predicate elements in Iranian, the question arises as to the possible diachronic developments that led to the present situation. Post-predicate position of goals was certainly an available option in Old and Middle Iranian, but we lack a systematic survey. The following Middle Persian examples shows pre-predicate placement of an addressee, and post-predicate placement of a goal of motion:

(12) **Ō** **man** **guft**
To me say.PST.3SG

čē **rāy** **ne** **hē** **šud** **ō** **xwēš** **šahr**
what for NEG AUX.2SG go to self country
'He said **to me**: Why have you not gone **to your country**?'
(Middle Persian, Turfan, Durkin-Meisterernst 2014: 409)

We must assume, pending a more systematic investigation of Old and Middle Iranian, that both options were available. The question is: what were the relative frequencies, and what may have triggered the shift to regular post-predicate placement in languages like Kurdish? Haig (2014)

tentatively suggests early Aramaic influence, pointing to striking parallels in the word-order patterns of geographically contiguous Neo-Aramaic and Kurdish varieties, but this needs more detailed investigation. It is nevertheless undoubtedly the case that post-predicate placement of goals is a diffusible feature across language boundaries, found for example in Turkic varieties under Iranian influence (Haig & Khan 2018).

3   Research questions

Our project will address the following four research questions:

a. What is the range of cross-language variation, and what are the recurrent parallels across contemporary Iranian languages, and their historically attested forerunners, with regard to (i) the syntactic status, (ii) the semantic role, and (iii) the information status of post-predicate constituents?

b. Are the existing hierarchies of post-posability (cf. (2), (10) and (11) above) viable for a more representative sample of varieties?

c. To what extent do non-Iranian contact languages share a similar profile with regard to post-predicate elements? Which aspects are most likely to diffuse in a contact situation, and which appear relatively stable?

d. Can the findings of Frommer (1981) for contemporary Persian, based on diaspora Persian speakers in California from the 1970's, be confirmed on a more recent corpus of contemporary spoken Persian? In particular, what is the impact of discourse factors such as information status (given, bridging, new etc., cf. Arnold et al 2013), and syntactic weight, in determining post-posability?

4   Methodology

To tackle the research questions (a), (b) and (c) we will design a questionnaire that is applicable to Iranian, and to contact languages, aimed at identifying the main kinds of post-predicate constructions. The questionnaire will be administered by a team of specialists for a sample of approximately eight contemporary Iranian languages, four contact languages, and historically attested stages of Iranian. The team will meet at a workshop in 2019 to refine questionnaire and methodology, and again in 2020 to report the findings. This data will form the basis for an edited volume, with single chapters authored by the respective experts, and an extensive overview chapter synthesizing the main results.

Question (d) will be approached using the methods of corpus-based typology (Haig et al 2011, Haig et al 2017, Haig 2018). We are currently compiling a corpus of contemporary spoken Persian, the HamBam corpus (Hamedan-Bamberg) corpus, in order to counteract the paucity of empirical research into the syntax of spoken Persian. The corpus is designed and

annotated in line with the methodology already implemented in the Multi-CAST archive (Schiborr 2016). The corpus will be made freely available and will employ a modified version of our existing annotation for information status and syntactic structure.

5 International collaboration and workshops

In designing and implementing the questionnaire we rely on the support of an international team of researchers, who have already expressed their willingness to participate in the two workshops and the questionnaire development and implementation:

Dr. Cormac Anderson (MPI for the Science of Human History, Jena)	Phylogenetic methods in comparative Indo-European
Prof. Mohammad Dabirmoghaddam (Allameh Tabataba'i University, Tehran)	Iranian languages within Iran, Early New Persian
Prof. Carina Jahani (Uppsala University)	Balochi, colloquial Persian
Prof. Agnes Korn (CNRS, Paris)	Old and Middle Iranian, Balochi
Prof. Geoffrey Khan (Cambridge University)	Northeastern Neo-Aramaic
Bettina Leitner, PhD candidate (Vienna)	Vernacular Arabic in Khuzestan province, Iran
Dr. Robin Meyer (Oxford University)	Classical Armenian and Middle Iranian contact-induced syntactic change
Dr. Ergin Öpengin (University of Uppsala)	Northern, Central and Southern Kurdish
Dr. Zakharia Pourtskhvanidze (Universität Frankfurt)	Georgian language islands in Iran, quantitative methods in language variation
Prof. Christoph Schroeder (Universität Potsdam, DFG-Forschergruppe *Emerging Grammars in Language Contact Situations: A Comparative Approach*)	Turkish/Kurdish language contact and multilingualism

The workshop is open to other interested persons such as PhD candidates (e.g. Hiwa Asadpour (Frankfurt), Shirin Adibifar (Bamberg), Kataryna Iefremenko (Potsdam)).

6 Deliverable research outcomes of the project

1. A questionnaire on post-predicate constituents, freely available for other researchers;
2. An edited volume based on a sample of Iranian languages, relevant contact languages, and historically attested Iranian languages (intended for an open access publisher; we would submit a proposal to Language Science Press);

3. The HamBam corpus of spoken Persian. We have already approximately two hours or recorded spontaneous speech, compiled by a team of graduate students at the Dept. of Linguistics in Hamedan under the supervision of Rasekhmahand. In accordance with the principles of Open Science and the Multi-CAST archive, this will likewise be made freely available using existing web infrastructure;

4. Two international workshops with approximately 15 invited participants, combining knowledge of individual languages with a variety of theoretical perspectives (2019, 2020);

References

Adibifar, Shirin. 2016. Persian. In: Haig, Geoffrey & Schnell, Stefan (eds.), Multi-CAST (Multilingual Corpus of Annotated Spoken Texts), https://lac.uni-koeln.de/multicast-persian/, date accessed 25.11.2018.

Arnold, J. E., Wasow, T., Losongco, A. & Ginstrom, R. (2000). Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language* 76, 28-55.

Birner, B. & Mahootian, S. (1996). Functional constructions in inversion in English and Farsi. *Language Science* 18, 127-136.

Delforooz, Barjasteh (2010). *Discourse Features in Balochi of Sistan (Oral Narratives)*. Uppsala: Acta Universitatis Upsaliensis..

Dryer, M. S. & Gensler, O. D. (2011). Order of object, oblique, and verb. In Dryer, M. S. & Haspelmath, M. (eds.), *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, Chapter 84. Available online at: http://wals.info/chapter/84. (Accessed on 17/09/2011).

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

Faghiri, P. & Samvelian, P. (2015). How much is determined by the syntax? An empirical approach to the position of the direct object in Persian. In Dabir-Moghaddam, M. (ed.), *Proceedings of the 9th Iranian Conference on Linguistics*. Tehran: AllameTabataba'i University Press. 1419-1435.

Faghiri, P., Samvelian, P. & Hemforth, B. (2014). Accessability and word order: The case of ditransitive constructions in Persian. In Müller, S. (ed.), *Proceedings of the 21st International Conference on Head-Driven Phrase Structure Grammar*. Stanford, CA: CSLI Publications. 217-237.

Frommer, P. (1981). *Post-verbal phenomena in colloquial Persian syntax*. PhD thesis, University of Southern California.

Ganjavi, S. (2007). *Direct objects in Persian*. PhD thesis, University of Southern California.

Ghomeshi, J. (1997). Topics in Persian VPs. *Lingua* 102(2-3), 133–167.

Haig, G. (2014). Verb-Goal(VG) word order in Kurdish and Neo-Aramaic: Typological and areal considerations. In Khan, G. & Napiorkowska, L. (eds.), *Neo-Aramaic and its linguistic context*. New York: Gorgias Press. 407-425.

Haig, G. (2017a). Post-predicate arguments in Iranian languages. Keynote lecture, *Seventh International Conference on Iranian Languages (ICIL7)*, Lomonosov Moscow State University, 28-30th August 2017.

Haig, G. (2017b). Western Asia: East Anatolia as a transition zone. In Hickey, R. (ed.), *The Cambridge Handbook of Areal Linguistics*. Cambridge: Cambridge University Press. 396-423.

Haig, G. & Khan, G. (2018). Introduction. In Haig, G. & Khan, G. (eds.), *The languages and linguistics of Western Asia: An areal perspective*. Berlin, New York: Mouton de Gruyter. 1-29.

Haig, Geoffrey and Fatemeh Nemati. (2013). *Clitics at the syntax-pragmatics interface: The case of Delvari pronominal enclitics.* Revised handout, paper held at the ICIL 5, Bamberg, 24-26th August, 2013.

Haig, G., Schnell, S. & Schiborr, N. (2017). The limits of accessibility: A corpus-based typological approach. Paper at the *workshop Advances in corpus-based typology: Exploring corpora of semi-parallel and indigenous text.* 12th Conference of the Association of Linguistic Typology. Australian National University, Canberra, 12-14th December 2017.

Haig, G., Schnell, S. & Wegener, C. (2011). Comparing corpora from endangered language projects: Explorations in typology based on original texts. In Haig, G., Nau, N., Schnell, S. & Wegener, C. (eds.), *Documenting endangered languages: Achievements and perspectives.* Berlin, New York: Mouton de Gruyter. 55-86.

Haig, Geoffrey (to appear). Post-predicate constituents in Kurdish. In: Matras, Yaron & Ergin Öpengin & Geoffrey Haig (eds.), *Structural and Typological Variation in the Dialects of Kurdish.* London: Palgrave MacMillan.

Hawkins, J. A. (1994). *A performance theory of order and constituency.* Cambridge: Cambridge University Press.

Hawkins, J. A. (2004). *Efficiency and complexity in grammars.* Oxford: Oxford University Press.

Hawkins, J. A. (2008) An asymmetry between VO and OV languages: The ordering of obliques. In Corbett, G. & Noonan, M. (eds.), *Case and grammatical relations: Essays in honor of Bernard Comrie.* Amsterdam: Benjamins. 167-190.

Hawkins, J. A. (2014). *Cross-linguistic variation and efficiency.* Oxford: Oxford University Press.

Karimi, S. (2003). On object positions, specificity, and scrambling in Persian. In Karimi, S. (ed.), *Word order and scrambling.* Hoboken, NJ: Blackwell. 91–124.

Lazard, Gilbert. ([1957]2006). *Grammaire du persan contemporain.* (Nouvelle édition, avec la collaboration de Yann Richard, Rokhsareh Hechmati et Pollet Samvelian). Tehran: Institut Français de recherche en Iran. [orig. 1957].

Matras, Yaron et al. 2016. The Dialects of Kurdish. Web resource, University of Manchester. http://kurdish.humanities.manchester.ac.uk/

Nikitina, T. (2011). Categorial reanalysis and the origin of S-O-V-X order in Mande. *Journal of African Languages and Linguistics* 32, 251–273.

Rasekh-Mahand, M., Alizadeh-Sahraei, M. & Izadifar, R. (2016). A corpus-based analysis of extraposition of relative clauses in Persian. *Ampersand* 3, 21-31.

Reetz, S. (2015). Post-predicate Goals in spoken East-Anatolian Turkish of Erzincan: A corpus-based analysis. MA-Thesis, Universität Bamberg.

Roberts, John. (2009). *A study of Persian discourse structure.* Uppsala: Acta Universitatis Upsaliensis..

Schiborr, N. (2016). *Multi-CAST Multilingual Corpus of Annotated Spoken Texts: Corpus overview and description.*https://lac2.uni-koeln.de/corpora/Multi-CAST/_multicast_background/Annotations/Multi-CAST_corpus-overview.pdf. (Accessed on 30/11/2018).

Schnell, S., Schiborr, N. & Haig, G. (2018). Is intransitive subject the preferred role for introducing new referents? Evidence from corpus based typology. Paper held at the workshop Comparative corpus linguistics: New perspectives and applications, 51st AGM of the SLE, Tallinn, 29th August-1st September 2018.

Schroeder, C. (2018). The post-verbal position in heritage Turkish in Germany: Focus on register. Paper held at *EuroSLA 2018, Colloquium Turkish Heritage Speakers in Germany.* Münster, 5-8th September 2018.

Sheehan, M., Biberauer, T., Roberts, I. & Holmberg, A. (2017). *The Final-over-Final Condition: A syntactic universal.* Cambridge, MA: MIT Press.

Sheehan, M. (2013). Explaining the Final-over-Final Constraint: Formal and functional approaches. In Biberauer, T. & Sheehan, M. (eds.), *Theoretical approaches to disharmonic word orders.* Oxford: Oxford University Press. 407–468.

Stilo, D. L. (2005). *Vafsi folk tales.* Wiesbaden: Reichert.

Stilo, D. L. (2018). Preverbal and Postverbal Peripheral Arguments in the Araxes-Iran Linguistic Area. Invited lecture at the conference *Anatolia-Caucasus-Iran: Ethnic and Linguistic Contacts,* Yerevan University, 10-12th May 2018..

Taleghani, Azita. (2008). Modality, Aspect and Negation in Persian. Amsterdam: John Benjamins.

Vafaeian, G. (2018). *Progressives in use and contact: A descriptive, areal and typological study with special focus on selected Iranian languages*. PhD thesis, Stockholm University.

Wasow, T. (2002). *Postverbal behavior*. Stanford, CA: CSLI Publications.

Yamashita, H. & Chang, F. (2001). 'Long before short' preference in the production of a head-final language. *Cognition* 81, B45–B55.

Yamashita, H. & Chang, F. (2006). Sentence production in Japanese. In Nakayama, M., Mazuka, R. & Shirai, Y. (eds.), *Handbook of East Asian Psycholinguistics, vol. 2*. Cambridge: Cambridge University Press. 291–297.