Chapter 1

Post-predicate arguments in Modern Eastern Armenian

Katherine Hodgson[a], Victoria Khurshudyan[b] & Pollet Samvelian[c]

[a]University of Cambridge [b]Institut National des Langues et Civilisations Orientales (INALCO) [c]Université Paris III - Sorbonne Nouvelle

This study uses a corpus of oral narratives to investigate the positioning of DOs, as well as other arguments, including goals, in Modern Eastern Armenian (MEA). We find a preference for OV, very strong for indefinite DO (96.8% OV), weaker for definite DO (67% OV). Animacy and weight appear to have a slight effect favouring VO, but the numbers are not statistically significant. Definiteness, animacy, and weight also appear to have a slight effect favouring post-predicate position for other roles. Other roles articulated as pronouns occur less frequently in post-predicate position than lexical NPs. There is no clear evidence that any of the other factors investigated (givenness, topicality, crowding effect, verb type, clause type) affects the position of arguments. The (generally informal spoken) MEA data from this study exhibit similarities to the "OVX" pattern that characterizes comparable registers in languages of the Mesopotamia region, with goals showing a preference for post-predicate position (68.9%), in contrast to other arguments.

1 Introduction

1.1 Overview

Modern Eastern Armenian (MEA) is generally grouped with SOV languages based on the branching direction of its various constituents (quite consistently left-branching), but also some syntactic properties of its VP (e.g. the preverbal position of the focus and bare/indefinite objects), which are assumed to be characteristic of OV languages. The quantitative study of Stilo (2018) also shows a

sectioning has to be exhaustive. I added two sections "overview" Feel free to change the title

clear preference for OV order. However, according to some studies (Dum-Tragut 2009, Badikyan 1976, among others), definite DOs display a clear preference for the postverbal position, and grammars of literary MEA generally consider VO to be the unmarked word order. In a recent study, Samvelian et al. (2023) conducted a quantitative investigation of word order in MEA. Based on a small task-oriented corpus and two experiments, they conclude that (S)VO order cannot be qualified as a marked option in MEA, since definite DOs display a strong preference for the VO order in "out-of-the-blue" sentences. Their results confirm, on the other hand, the preference of non-definite DOs for the OV order, indicating furthermore that indefinite (with the indefinite article) and bare DOs are equally likely to favor the OV order. Apart from definiteness, length is also mentioned as a relevant factor, since long indefinite and bare DOs are more likely to behave like definite DOs by appearing in postverbal position.

Apart from DOs, other arguments, especially goals of verbs of motion and caused motion, may also occur in post-predicate position in MEA. This phenomenon has been shown to have an areal dimension, being characteristic of many languages of Western Asia; see contributions to this volume. The purpose of the present study is to investigate factors affecting the position (pre- vs. post-predicate) of DOs, as well as certain other types of arguments that may occur variably in pre- or post-predicate position. Specifically, we aim to check whether the same word order preferences for the placement of DOs reported in Samvelian et al. (2023) also hold in a corpus of unscripted spoken language, as well as to investigate the word order preferences of goals and other types of arguments that show variable position. Besides definiteness and length, we also investigate the relevance of some other factors in the placement of DOs and other arguments:

- Topicality: It has been claimed that word order in MEA is closely linked to information structure (Hodgson 2019). Given that our corpus is annotated for topic persistence and referential distance, as well as other factors connected to topicality, such as givenness and animacy, we can check whether there is a correlation between topicality and postverbal placement.
- Type of referential element (lexical NP vs. pronoun).
- Crowding effect: Hayrapetyan (1981) mentions the presence (or not) of the subject as a relevant feature in the pre- or postverbal placement of the DO in Classical Armenian. This factor may also be relevant in MEA.
- We also investigate the possibility that clause type (main vs. subordinate) and verb type (simple vs. periphrastic) could have an effect on the position of objects.

1.2 Typological preliminaries and background

The position of direct objects in Eastern Armenian is somewhat controversial. Although Armenian is often assumed to be an OV language, Eastern Armenian in particular shows considerable flexibility in the position of objects, especially definite objects, and in WALS is described as having no dominant order (Dryer 2013). The placement of DOs has been the focus of a few studies, including corpus-based and experimental ones. Samvelian et al. (2023) provide a thorough survey of the literature on the issue. They note that despite the fact that typological and theoretical syntactic studies have generally grouped Armenian with SOV languages (Der-Houssikian 1978: 227–228, Dryer 1998: 286, 310, Dum-Tragut 2002, Giorgi & Haroutyunian 2016: 190, Hawkins 1979: 625, 1983: 286, Hodgson 2013: 6, Kahnemuyipour & Megerdoomian 2011, 2017: 81, Kozintseva 1995: 8, Minassian 1980: 263, Tamrazian 1991: 101, 1994: 7, among others), Armenian grammars and descriptive studies, on the other hand, generally refer to the postverbal position as the "natural" position of the object. This is either overtly claimed based on small scale quantitative studies (Badikyan 1976), or is induced by the fact that in most examples illustrating transitive sentences and the DO, the latter is placed in the postverbal position.

Given the fact that OV-VO variation is not trivial in MEA, Samvelian et al. (2023) build on quantitative methods in order to provide a reliable picture of word order distribution in MEA, as well as different factors that may be involved in the choice of OV versus VO order. The main factor investigated by Samvelian et al. (2023) is definiteness. Indeed, several studies have mentioned the relevance of this factor in the placement of the DO (Badikyan 1976, Dum-Tragut 2009, Stilo 2018), with definite DOs occurring more frequently in postverbal position. In a preliminary corpus investigation based on a sample of 570 transitive sentences from the EANC (Eastern Armenian National Corpus, see below), Faghiri & Samvelian (2020) report that definite DOs are overwhelmingly postverbal (79.1% VO vs. 20.9% OV).

The results of Samvelian et al. (2023) are based on a small-scale corpus study and 2 experiments. Their task-oriented spoken corpus includes recordings of picture-based storytelling from 10 native speakers of MEA (see Khurshudyan 2006 and Samvelian et al. 2023 for a detailed description of the corpus and its annotation). Their corpus contains 231 finite declarative transitive sentences with SOV or SVO orders. The results confirm that definite DOs tend to appear in the postverbal position. Samvelian et al. (2023: 476) also show the relevance of length in the placement of the DO: "long and/or heavy bare and indefinite DOs appear frequently in the SVO order, while simple (that is short or minimal) bare and indefinite DOs are more likely to appear in the SOV order."

Samvelian et al. (2023) also provide two sentence production experiments. The first experiment, which compares definite and indefinite DOs, uses a cued sentence recall protocol, and the second one, which compares indefinite and bare DOs, uses constrained sentence production protocol. Furthermore, Samvelian et al. (2023) study the effect of animacy and length.

In the first experiment, 82.7% of definite DOs occur in the postverbal position, while the percentage of postverbal indefinite DOs is 38.3%. As for animacy, although its main effect is not significant, it nevertheless has an effect on the order with definite DOs: animate definite DOs favor SVO more than inanimate definite DOs do. The second experiment shows that both indefinite and bare objects are overwhelmingly preverbal: 74.4% vs. 75.7% respectively. The experiment also confirms the effect of length: non-simple (or long) DOs are more likely to appear in the SVO order.

To sum up, both Faghiri & Samvelian (2020) and Samvelian et al. (2023) report a clear-cut divide between definite DOs on the one hand and indefinite and bare DOs on the other hand. In both corpus-based and experimental studies, definite DOs are overwhelmingly postverbal. Therefore, not only does definiteness seem to be the best predictor for the pre- or postverbal placement of the DO, but also postverbal position seems to be the default placement for definite DOs.

The results of Samvelian et al. (2023) are in sharp contrast with the findings of Stilo (2018). In his data from Colloquial Yerevan Armenian, Stilo (2018) finds that both definite and indefinite DOs show a strong preference for preverbal position: indefinite DOs are 95% preverbal, while definite DOs show a slightly weaker, but still strong preference for preverbal position (86.1%). The other dialects of Armenian included in the study of Stilo (2018) also show a general strong preference for preverbal position in the case of both indefinite and definite DOs: indefinite DOs are preverbal in 100% of cases in the dialects of Erzurum and Stepanakert, and 81% preverbal in Lori, while definite DOs are 98% preverbal in Stepanakert, 91.9% preverbal in Erzurum, and 83.2% preverbal in Lori.

The contrast between the results of Stilo (2018) and those of Samvelian et al. (2023) can be accounted for by several facts: dialectal variation, the language register, and the "genre" of the corpus (discourse, oral, written, etc.), as well as the fact that the data in the experimental studies of Samvelian et al. (2023) include only "out-of-the-blue" declarative non-embedded sentences, which may be considered unmarked, while the present study includes all types of sentences. The data from colloquial Yerevan used by Stilo (2018), which have not been available to us, might also reveal other reasons for the discrepancies. A similar pattern has been observed in Romeyka, where those studies based on elicitation of out-of-the-blue sentences, with overt subjects and objects, yield predominantly VO

structures, while the data from connected spontaneous spoken discourse show a higher rate of OV (Schreiber & Janse 2024 [this volume]). As Samvelian et al. (2023) note in their conclusion, it is very likely that the rate of SOV is higher in spontaneous oral discourse. It is also likely that geographical variation, perhaps associated with different language contact patterns, is relevant, as Stilo (2018) presents data from various dialects spoken in different parts of the Armenian-speaking world. However, even his data from Yerevan show a far lower percentage of post-predicate arguments than those of Samvelian et al. (2023).

In this context, it is important to note that formal literary MEA, as used in writing and formal speech, shows significant differences in morphology, phonology, and syntax[1] from any form of colloquial spoken Armenian, including the colloquial MEA spoken in Yerevan from which Stilo's data are taken. As stated in the introduction, VO order is promoted as the unmarked word order in grammars of the formal literary language.[2] It is likely that the difference between the results of Stilo (2018), even for speakers from Yerevan, and those of Samvelian et al. (2023) reflects the fact that the former deals with colloquial language, and the latter includes data with characteristics of formal literary language,[3] indicating another potential syntactic difference between literary and colloquial MEA. The present study provides further evidence in support of the proposal that register is a key factor in the difference between the results of Stilo (2018) and Samvelian et al. (2023), since the "outlier" speaker with the highest percentage of VO (speaker 3, with 44% VO, as opposed to an average of 17.4% VO for all the other speakers) is

[1]Morphological differences between formal literary EA and colloquial EA as spoken in Yerevan include the form of the 3sg. present auxiliary (formal *e* vs. colloquial *a*) and certain forms of the "emphatic" pronoun (formal nom.pl. *irenkʿ* vs colloquial *irankʿ*, formal gen.sg. *ir*, dat.sg. *iren* vs. colloquial gen.sg. *ira*, dat.sg. *iran*) among many others. Phonological differences include the change of the diphthong /ay/ > [e], as in the demonstratives (distal), which have the forms *es*, *ed*, *en* in the colloquial language; see also the formal/colloquial correspondences *hayr/her* 'father,' *mayr/mer* 'mother,' *dzayn/dzen* 'voice,' *layn/len* 'wide,' etc. Syntactic differences include the use of different cases with certain adpositions, for example the use of the dative of 1st and 2nd person pronouns with postpositions such as *het* 'with,' *mot* 'close to, by,' *hamar* 'for' etc. in the formal literary language, as opposed to the genitive in the colloquial language: formal *indz het* 'with me,' *indz mot* 'close to me,' *indz hamar* 'for me' vs. colloquial *im mot, im het, im hamar*, as well as differences in relativization strategies (notably use of indeclinable complementizer instead of declined relative pronoun) described in Hodgson (2019).

[2]Donabédian (p.c.) even reports hearing a teacher of formal literary Western Armenian state that OV order is "Turkish," and therefore incorrect.

[3]For example, examples (18–21) in Samvelian et al. (2023), taken from their corpus, show features characteristic of formal literary language (3sg. aux. *e*, genitive of "emphatic" pronoun *ir* etc.), and while the words used in the experiments could be described as register-neutral, the instructions that appear on the screen as shown in Samvelian et al. (2023) Fig. 1 are in formal literary EA, which is likely to have prompted the use of this form of language in the responses.

also the only one who uses certain word forms associated with the formal literary language (see Table 3 and associated discussion). These findings echo similar findings from Persian, where different grades of formality (and the difference between spoken and written modalities) have a significant impact on certain aspects of word order, see Rasekh-Mahand et al. (2024 [this volume])) for details.

The discussion of the placement of DOs in Stilo (2018) takes place in context of the wider issue of post-predicate constituents in many OV languages of Western Asia. These languages differ from "rigid" OV languages in that they regularly allow at least certain constituents to appear in post-predicate position, giving (O)VX word order. Stilo (2018) and subsequent studies show that the type of constituent that is most common and widespread in post-predicate position is Goal of verbs of motion or caused motion (see Haig et al. 2024 [this volume], for an overview). In some languages, similar behaviour is shown by other constituents which, like goals, can be considered to have the semantic property of "endpoint," namely recipient, benefactive, adddressee, and object of change of state predicates such as 'become'. Arguments without this semantic property, such as ablative, instrumental, locative, and comitative, do not typically appear in post-predicate position in these languages (see Haig & Khan 2019, Stilo 2018). The data in Stilo (2018), which include four different dialects of Armenian (Erzurum, Lori, Stepanakert, and colloquial Yerevan), indicate that in all of these, goals show a preference for post-predicate position. In some dialects, recipient and benefactive also show a tendency to appear in post-predicate position, though not as frequently as goals. Addressee arguments do not seem to be affected by this tendency in Armenian, showing 90–100% pre-predicate position in all the dialects investigated. In fact, most of the historically OV languages in this volume exhibit the same split (see Haig et al. 2024 [this volume], Section **??**), distinguishing addressees from spatial goals. Instrumental, ablative, comitative, and locative are predominantly pre-predicate across the whole area.

This pattern is associated with a particular geographic location, with the epicentre in the Mesopotamian region (modern Northern Iraq, Western Iran, and southeastern Turkey) (Haig & Khan 2019, Haig et al. 2024 [this volume]). As for the typological profile of languages showing OVX word order, these have been characterised as OV languages showing some properties typical of VO languages, such as prepositions and/or initial complementizers. The phenomenon reflects the area's status as a "transition zone" from Turkic-type head-final, through Iranian mixed typologies, to Semitic head-initial (Stilo 2006, Haig & Khan 2019, Haig et al. 2024 [this volume]). Armenian, too, could be described as a language with mixed typology, having undergone a change from mainly head-initial Classical Armenian (prepositions, probable VO preference) to mainly head-final modern

Armenian, which does, however, retain some properties associated with head-initial languages, such as initial complementizers. Various proposals have been made regarding the causes of the OVX phenomenon, including iconicity (arguments with endpoint semantics appearing in final position), and contact with languages such as Aramaic (or Russian) that typically have post-predicate arguments (Haig & Khan 2019, Haig et al. 2024 [this volume]). However, in the case of Armenian, it is possible that post-predicate arguments represent a conservative feature retained from Classical Armenian. This could apply to the presence of VO orders as well as OVX, as suggested by both Samvelian et al. (2023) and Stilo (2018), although the fact that post-predicate Goals, but not objects, seem to appear less frequently in subordinate clauses, which have been claimed to show a general tendency for more conservative syntax, perhaps suggests the possibility that OVX may be a more recent phenomenon.

2 The Methodology and the corpus

2.1 Overview

The corpus is composed of 7 oral narratives by 7 participants (6 women and a man), who narrate their favorite movies. It was compiled within the framework of the Eastern Armenian National Corpus[4] (henceforth EANC ArmFilmNarr). The recording was done in Yerevan in 2007–2008, and the total recording time is 1.35h. The corpus contains 11,832 tokens, divided into 2,241 clauses. The corpus was originally created for the study of topic accessibility and continuity of subjects and DO in MEA (in a distinct research project) based on a framework inspired by Givón (1975), Givón (1983) and Du Bois (1987) (for more information on this study, see Hodgson et al. In press). The purpose of the present study is to investigate factors affecting the position (pre- vs. post-predicate) of DOs, as well as certain other types of arguments that may occur variably in pre- or post-predicate position. Thus, we note the position (pre-or post-predicate) of monotransitive DO, Goal, addressee, recipient, benefactive, endpoint of 'become,' ablative, instrumental, locative, and comitative arguments. For each of these, we note other factors that could potentially affect the position of the argument: weight (number of words/intonation words), type of anaphoric element (e.g pronoun vs.

[4]Eastern Armenian National Corpus is a comprehensive corpus of Modern Eastern Armenian comprising approximately 110 million tokens. It encompasses written and oral data starting from the mid 19th century to the present. The corpus is provided with full morphological annotation, offering robust search functionalities, and is openly accessible at www.eanc.net. For more details on Eastern Armenian National Corpus, see Khurshudyan et al. (2022).

definite NP vs. indefinite NP), animacy, and givenness. The presence and position of an overt subject argument is also noted, in order to test for crowding effects, i.e. the tendency to avoid having more than one overt argument on the same side of the predicate. The type of verb form is also noted, as the distinction between periphrastic and simple verbs could potentially affect the position of arguments, as could clause type (main vs. subordinate clause). For DO, figures are also given for topic persistence (number of mentions of the referent in the following 10 clauses) and referential distance (distance in clauses to previous mention of the referent, noted as 1, 2, or 3, with 3 indicating 3 or more). Figure 1 provides an illustrative example.

#	text	Predicate type	Voice	Order	Arg.O	Arg.A	Pers.O	Pers.A	H.O	H.A	NG.O	NG.A	Pre/PostV O	Pre/PostV A	Weight O	Weight A	Other role	Weight	Type	Animacy	Given/New	Pre/post	Others	Verb form	Clause type	Force, Negation	RD.O	RD.A	TP.O	TP.A
59	2_F	TR		LOCOVC	NP.DEF	V.AGR	3	3	NH	H	N	G	Pre		1		Loc	1	PRO.LOC	I	N	Pre		IPT-C			3	1	0	9

/ ənd̪eɣ / d̪uɾ̪ə bacʿum a
'there she opens the door'

Figure 1: Example of annotation of a sentence

In the first column (clause-arm), we find the clause (in Figure 1, the clause and its translation are shown below the columns). The second column shows the type of predicate (here "tr," i.e. monotransitive). There is a column for voice, which is left blank when active. The word order of the clause as a whole is shown (LocOVC, i.e. Locative, Object, (lexical) Verb, Copula/auxiliary). Then, the properties of O are noted in the grey columns, and A in the green columns. These include type of argument (O is NP.DEF, i.e. definite NP, and A is V.agr, i.e. verb agreement), person (3sg for both), animacy (H = human, NH = non-human), givenness (G = given, N = new), the position of overt arguments (Pre, i.e. pre-predicate, for O, A is left blank, because it is expressed by verb agreement alone), and weight of overt arguments (blank for A, 1, i.e. one word, for O). The next column shows that the clause also contains a locative (Loc). In the following columns, its values for weight (1), type of argument, i.e. type of anaphoric element (PRO, i.e. pro-form), animacy, givenness, and position are given. Other columns show the verb type (IPT-C, i.e. imperfective participle + copula/auxiliary), which is of interest particularly in terms of simple (monolectic) vs. complex (participle + auxiliary) verb forms, and the clause type (left blank for main clauses). Referential distance (RD) and topic persistence (TP) are noted in separate columns for subjects and objects of main clauses in which both are 3rd person. Examples of the main categories of anaphoric elements (examples 1-10) and verb types (examples 11-17) are given below:

2.2 Anaphoric elements

(1) **Verb agreement** (for subject only)
Eastern Armenian
*čʿ-git-**em***
NEG-know-1SG.PRS
'**I** don't know.'

(2) **Zero anaphora** (especially for DO, but also some other object-like elements: example (2) shows zero anaphora for both DO and recipient)
Eastern Armenian
dra hamar a nvir-um
DEM2.GEN for be.3SG.PRS give-IPFV
'He gives **it to her** because of that.'

(3) **Agreement marker** (effectively genitive clitic, can be used for objects of some adpositions)
Eastern Armenian
*ayjik-ə het-**ə** xos-um a*
girl-DEF with-AGR3 speak-IPFV be.3SG.PRS
'The girl talks with **him**.'

(4) **Personal pronoun**
Eastern Armenian
***kʿez** šat em sir-um*
2SG.DAT much be.1SG.PRS love-IPFV
'I love **you** very much.'

(5) **Demonstrative**
Eastern Armenian
***ed** bacʿ-um en*
DEM2 open-IPFV be.3PL.PRS
'They open **that**.'

(6) **"Emphatic"**[5].
Eastern Armenian
*menkʿ **iran** k-gtn-enkʿ*
1PL.NOM EMP.DAT FUT-find-1PL.PRS
'We will find **him**.'

[5]For a discussion of this element and its behaviour in this corpus, see Hodgson et al. (In press). See also Donabedian-Demopoulos (2007), Sigler (2001)

(7) **Pro-adverb**
Eastern Armenian
əndeɣ *duŕ-ə* *bacʿ-um* *a*
there door-DEF open-IPFV be.3SG.PRS
'**There** she opens the door.'

(8) **Bare NP**
Eastern Armenian
takʿsi *a* *gal-is*
taxi be.3SG.PRS come-IPFV
'**A taxi** comes.'

(9) **NP with indefinite article**
Eastern Armenian
tencʿ ***mi takʿsi*** *gal-is* *a*
thus IA taxi come-IPFV be.3SG.PRS
'So **a taxi** comes.'

(10) **Definite NP**
Eastern Armenian
takʿsi-n *kangn-acʿn-um* *en*
taxi-DEF stand-CAUS-IPFV be.PL.PRS
'They stop **the taxi**.'

2.3 Verb types

2.3.1 Simple

(11) **Monolectic present of some verbs**
Eastern Armenian
un-i *ir* *ŕistaran-ə*
have-3SG.PRS EMP.GEN restaurant-DEF
'He has his restaurant.'

(12) **Aorist**
Eastern Armenian
dukʿ *mard* ***span-ecʿikʿ***
2PL.NOM person kill-2PL.AOR
'You killed a person.'

(13) **Subjunctive**
Eastern Armenian
vor iran dur ***ga***
COMP EMP.DAT like come.3SG.SUB
'so that she likes it'

(14) **Future/conditional forms with the prefix k-** are also classed as "simple" here
Eastern Armenian
k-ogn-enkʿ
FUT-help-1PL
'We will help.'

2.3.2 Complex (participle + auxiliary)

(15) **Present**
Eastern Armenian
heto taŕ-er-ə ***gr-um a***
after letter-PL-DEF write-IPFV be.3SG.PRS
'After (that) it writes the letters.'

(16) **Perfect**
Eastern Armenian
heto arden bolševik-ner-ə ***mt-el en*** *Hayastan*
after already bolshevik-PL-DEF enter-PFV be.3PL.PRS Armenia
'Then the Bolsheviks have already entered Armenia.'

(17) **Future**
Eastern Armenian
Yes ***gn-alu em***
1SG.NOM go-FPT be.1SG.PRS
'I will go.'

3 Results and analyses

3.1 Word order variation

Our corpus of oral narratives of favorite films comprising 2,241 clauses shows, overall, twelve possible word order combinations (see Table 1).

Table 1: The distribution of word order configurations in EANC Arm-FilmNarr corpus

#	WO	#	%
1.	SV	415	41.8%
2.	OV	275	27.7%
3.	AOV	94	9.5%
4.	VO	69	6.9%
5.	VS	52	5.2%
6.	AVO	32	3.2%
7.	AV	30	3.0%
8.	OVA	10	1.0%
9.	OAV	10	1.0%
10.	VA	3	0.3%
11.	VAO	2	0.2%
12.	VOA	1	0.1%
	Total	993[a]	100%

[a]The difference between the total in this table (993) and the total number of clauses in the corpus (2,241) is due to the fact that the table only counts clauses that contain overt subjects and/or objects, and a verb. Since Armenian makes frequent use of zero anaphora for both subjects and objects, many clauses have no overt subject or object. There are also some incomplete or syntactically anomalous clauses that have not been counted here.

According to the corpus results, the most frequent word order is SV, with around 42% (415) of all occurrences, which could be accounted for by the abundant number of intransitive verb constructions typical of the narrative genre.

3.2 Direct object

3.2.1 General overview

The second most frequent word order is OV. The absence of the agent (A) can be accounted for by MEA's pro-drop character, which is particularly frequent in oral discourse. OV word order also indicates the prevalence of preverbal position for objects in general. Since almost all clausal objects in MEA are postverbal, it is more informative to limit ourselves to non-clausal objects. Therefore, we concentrate on non-clausal objects, and every mention of objects will be understood as referring to monotransitive, non-clausal objects, unless otherwise stated. The overall distribution of non-clausal objects is 78.7% preverbal vs. 21.3% postverbal

(see Table 2). The data do not show a significant difference between the behaviour of lexical NP and pronominal objects, with the former showing 21.9% VO, the latter 18.9% VO (see Table 2). Armenian does not possess clitic pronouns, so the pronouns in question are "strong," independent pronouns, with the equivalent of "weak" unstressed pronouns being zero anaphora.

Table 2: The distribution of all overt monotransitive non-clausal DOs in EANC ArmFilmNarr corpus

DO type	Total	VO	% VO
Lexical NP	374	82	21.9%
Pronominal	106	20	18.9%
Total	480	102	21.3%

The net preference for preverbal position (21.3% VO, i.e. 78.7% OV) shown in these data is contrary to the study by Samvelian et al. (2023), but roughly consistent with the data in Stilo (2018). Potential factors that could account for this difference are discussed above in Section 2: discourse mode (oral vs. written), genre (oral narratives vs. out-of-the-blue sentences), and register (colloquial vs. formal). The current oral corpus of favorite film narratives potentially includes all sentence modalities which can be part of a structured narrative, and not simply declarative, out-of-the-blue sentences, as was the case for the data used in the study by Samvelian et al. (2023).

Another important difference is register, since Samvelian et al. (2023) include language with characteristics of formal register, whereas the current corpus covers generally colloquial register, although with some semi-formal elements (the context of the recordings was generally informal, though the speakers were aware that they were being recorded, which may have prompted them to use some elements of formal/literary EA). In Samvelian et al. (2023), it has already been proposed that EA word order variation could also be correlated with register variation, as also discussed for Persian in Rasekh-Mahand et al. 2024 [this volume]; the significant morphological, phonological, and syntactic differences between colloquial and formal/literary EA are discussed in Section 2.[1]

Evidence that register is indeed a significant factor in this context is provided by the fact that Speaker 3[6] (see Table 3) is the only speaker to use formal/literary

[6]Speaker 3 is the "outlier" with a significantly higher percentage of postverbal DOs than the other speakers (44%, compared to an average of 17.4% for all other speakers, and 21.3% in the corpus as a whole, see Table 3)

forms of demonstratives. For example, her speech contains two examples of medial demonstrative *ayd* as opposed to colloquial *ed/et*) and of the 3rd person "emphatic" pronoun nominative plural *irenkʿ* as opposed to colloquial *irankʿ* (gen.sg. *ir* is also used once by Speaker 1 as well as by Speaker 3).

Table 3: The distribution of postverbal DOs according to speaker in EANC ArmFilmNarr corpus

Speaker	Total O	VO	% VO
1	23	5	21.7%
2	79	13	16.5%
3	45	20	44.4%
4	51	12	23.5%
5	34	6	17.6%
6	43	7	16.3%
7	63	8	12.7%

3.2.2 Impact factors

The following set of impact factors that could potentially be relevant for EA word order variation have been included in our analysis:

1. definiteness
2. givenness
3. animacy
4. topic persistence
5. referential distance
6. weight (heavy NP shift)
7. lexical vs. pronominal Os
8. crowding / null subject (pro-dropping) effects /overt A
9. main vs. subordinate clauses
10. simple vs. complex verb forms

3.2.2.1 Definiteness

better use automatic sectioning and counters

"Definiteness" here refers to the presence of the definite article (which is enclitic on the noun). According to Samvelian et al. (2023), EA word order variation is directly correlated with definiteness, with definite DOs being mainly postverbal, and indefinite/bare DOs preverbal. We checked the distribution of preverbal and postverbal bare, indefinite and definite DOs in the current study, and the results confirm the correlation between the definiteness of DOs and postverbal position (see Table 4).

Table 4: The distribution of preverbal and postverbal bare, indefinite (with article) and definite DOs in EANC ArmFilmNarr corpus

Object type	Total	VO	% VO
Definite NP	232	76	33%
Indefinite NP (with indefinite article)	21	2	10%
Bare NP (indefinite without article)	103	2	2%
Total	356[7]	275	mean 15%

As Table 4 shows, the distribution of definite DOs is 67% preverbal vs. 33% postverbal, whereas that of bare DOs is 98% preverbal vs. 2% postverbal. The indefinite DOs have an intermediate position with 90% preverbal vs. 10% postverbal. Therefore, there is a certain hierarchy of postverbal word order possibility depending on definiteness, in which definite DOs are the most postverbal, followed by indefinite (with article) DOs, and bare DOs are the least postverbal of all. This hierarchy corresponds to the grammatical semantics of definite ⟹ specific ⟹ non-specific. The peculiarity of EA indefiniteness is that it is bipartite with indefinite (specific) and bare (non-specific) semantics, and the "classical" indefinite noun would tend to be bare, i.e., non-specific, in EA, rather than with an indefinite article, i.e., specific. As the indefinite article is mainly used for marking (specific) indefiniteness, its usage is fairly rare. Hence, the results of the current study show general tendencies that are consistent with Samvelian et al. (2023) in that EA postverbal word order is associated with definiteness and that the more definite/specific the DO is, the more likely it is to be postverbal.

[7] This total is smaller that of Table 4, because it only includes lexical NPs, while Table 4 also includes pronominal objects.

3.2.2.2 Givenness

"Given," as opposed to "new," is used here to indicate referents that have been previously mentioned in the discourse. Given Os are considerably more likely to appear in post-predicate position than Os which represent new information. However, this effect is weaker than the effect of grammatical definiteness (definite NP O = 32.9% post-predicate, indefinite (including bare) NP O = 3.3% post-predicate), although the net preference of post-predicate Os for givenness is still considerable (over 80% of post-predicate Os are given) (see Table 5). Thus, it is likely that the givenness effect is merely a reflection of the effect of definiteness (note that in Armenian, certain categories of NPs may be grammatically definite even if not given, e.g. possessives, partitives, and some nominalized non-nominal constituents):

Table 5: The distribution of given and new DOs in EANC ArmFilmNarr corpus

	Total	VO	% VO
O given	289	86	29.8%
O new	190	17	8.9%

3.2.2.3 Animacy

We observe that animate Os are more likely to appear in post-predicate position than inanimate ones. Note, however, that animate referents are considerably more likely to be definite than inanimate ones (animate DO = 79.7% definite, inanimate DO = 59.5% definite), so in order to understand whether the effect of animacy is significant, we need to investigate it in combination with definiteness. The results are shown in Table 6.

We find that among definite Os, animate referents appear slightly more frequently in post-predicate position than inanimate ones, although a Chi-square 2x2 contingency table shows that the difference between definite animate and definite inanimate DOs is not statistically significant. This is similar to the results of Samvelian et al. (2023: 481), who find that the main effect of animacy is not significant, but "there is a marginally significant interaction … indicating that animacy has an effect on the order of definite DOs". The number of indefinite DOs in post-predicate position is, as we have seen, very small, both for animate and inanimate referents.

Table 6: The distribution of animate and inanimate DOs in EANC ArmFilmNarr corpus

	Total	VO	% VO
Animate O definite NP	63	27	42.8%
Inanimate O definite NP	170	51	30%
Animate O indefinite NP	16	0	0%
Inanimate O indefinite NP	113	4	3.5%

3.2.2.4 Topic persistence

Based on the methodology of Givón (1983), topic persistence was measured as the number of occurrences of that referent in the following 10 clauses (for more details on the methodology and on MEA data see Hodgson et al. In press). The average topic persistence of preverbal Os is equal to 1.6, whereas that of postverbal ones is 1.9. Therefore, postverbal Os have higher topic persistence than preverbal Os. Taking into account the previous impact factors, it could be argued that prototypically postverbal Os are mostly definite, given, and animate, hence have higher topic persistence. The average topic persistence of postverbal Os is higher than that of preverbal definite Os, which show an average topic persistence of 1.5 (the same as that of definite Os as a whole). Human Os have on average a considerably higher topic persistence (3.7) than non-human Os (1), so the fact that postverbal position is possibly associated with animacy/humanity could be a factor in the higher average topic persistence of postverbal as opposed to preverbal Os. Indeed, if we look at objects that are both definite and human, there is no significant difference in average topic persistence between preverbal and postverbal examples (preverbal: 3.08, postverbal: 3.13).

Table 7: Average topic persistence of pre- and post-predicate DOs in EANC ArmFilmNarr corpus

	Average topic persistence
Pre-predicate O	1.6
Post-predicate O	1.9

3.2.2.5 Referential distance

There is only a very small difference between the average referential distance (distance in clauses to previous mention of the referent) of pre-predicate and post-predicate Os, with post-predicate Os showing a slightly smaller average referential distance (2.3) than pre-predicate Os (2.5) (see Table 8).

Table 8: Average referential distance of pre- and post-predicate DOs in EANC ArmFilmNarr corpus

	Average referential distance
Pre-predicate O	2.5
Post-predicate O	2.3

3.2.2.6 Weight (heavy NP shift)

In order to check the correlation between the weight of DOs and their word order variation, the length of DOs in words was measured, more particularly for 1-word, 2-word and 3+ word Os (see Table 9). The results showed a higher frequency of postverbal position for 2-word Os (35% postverbal) vs. 1-word Os (15.4% postverbal). The fact that non-specific indefinites, lacking the indefinite article, which, as we have seen, overwhelmingly appear in preverbal position, are often one-word phrases is a probable factor in the higher frequency of preverbal position for 1-word Os. When O is composed of more than three words, the frequency of postverbal position is actually lower than that for 2-word DOs (23% postverbal), though still higher than that for 1-word DOs.

Subsequently, the measurements were refined in order to explain the distribution of 3+word Os as well as the difference in behaviour between 2-word and 3+word Os. To do this, supplementary factors of definiteness and the pronominal character of Os were added to the analysis of 2-word and 3+ word Os (see Table 9).

According to the results, 2-word definite Os are 43.4% postverbal (cf. 21.3% postverbal for all Os[8]) and 3+ word definite Os are 31.6% postverbal. We then zoomed in on indefinite 3+word Os, distinguishing those with the simple indefinite article *mi* 'a' and the indefinite expression composed of the indefinite article *mi* 'a' and the quantifier *hat* 'unit' which is often in reality the equivalent

[8]This figure excludes those Os that appear in both pre- and post-predicate position.

Table 9: The distribution of 1-word, 2-word, and 3+ word Os correlated with O type in EANC ArmFilmNarr corpus

	1-word O			2-word O			3+-word O		
	Total	VO	%VO	Total	VO	%VO	Total	VO	%VO
Bare	73	0	0%	22	2	9.1%	5	0	0%
Indefinite	0	0	0%	5	1	20%	14	1	7.1%
Definite	105	27	25.7%	83	36	43.4%	38	12	31.6%
Pronominal	88	14	16.7%	10	3	30%	4	1	25%
Total	266	41	15.4%	120	42	35%	61	14	23%

of the indefinite article in colloquial register (see Table 11). 21% of all 3+ word Os include the simple indefinite article *mi* 'a,' whereas 79% include the indefinite article with a quantifier (see Table 10). The indefinite article in EA being principally unstressed, the expression of the indefinite article with a quantifier *mi hat* 'a unit' could be considered as one prosodic word, hence it could be included in our measurements as one word rather than two due to its "real weight." However, even when *mi hat* is counted as one word rather than two, 2-word objects still show a higher percentage of OV order (36.2%) than 3+-word objects (26.9%), something which is unexpected from the point of view of heavy NP shift. The generally inconclusive evidence for an effect of weight are consistent with the findings from other corpora of spontaneous spoken language investigated in this volume, which report only marginal effects of weight on object placement (see Haig et al. 2024 [this volume], Section ??).

Table 10: The distribution of 3+ word indefinite Os in EANC ArmFilmNarr corpus

3+-word indefinite O	Total	VO	% VO
mi	3	1	33.3%
mi hat	11	0	0%

3.2.2.7 Lexical vs. pronominal O

The linguistic character of DOs showed little impact on the word order distribution, with lexical DOs being 21.9% postverbal, and pronominal DOs being 18.9%

postverbal (see Table 11). Overall, preverbal DOs are largely dominant in both cases and the distribution proportion in line with that of all DOs in the corpus (21.3% postverbal, Table 2).

Table 11: The distribution of preverbal and postverbal lexical and pronominal DOs in EANC ArmFilmNarr corpus

Type of O	Total	VO	% VO
Lexical NP	374	82	21.9%
Pronoun	106	20	18.9%

3.2.2.8 Crowding / Null Subject Effect

One of the hypotheses concerning the impact factors was that argument crowding or its opposite, null subject, could affect word order variation so that the presence of an overt subject could induce postverbal DOs to avoid crowding, or the presence of a preverbal DOs could be correlated to null subject effect.

Table 12: The distribution of DOs with and without other overt arguments in EANC ArmFilmNarr corpus

	Total overt DO	VO	% VO
Overt S	148	35	23.6%
No overt S	339	69	20.4%
No overt S or other role[a]	267	51	19.1%
Overt preverbal S or other role	175	39	22.3%

[a]"Other role" refers to those elements discussed in section 3.3.

To check this, we first observed the distribution of overt/no overt subjects (S) which, however, manifested insignificant impact on the position of DO (see Table 12). Overall, the distribution is almost identical, therefore, no correlation is observed between the crowding/null subject effect and the word order variation. The same is true if we take into account the presence of elements with other roles as well as subject; the presence of a preverbal argument does not seem to be associated with any increase in postverbal position for DO.

3.2.2.9 Main vs. subordinate clauses

In some languages, object position differs depending on whether the object is in a main clause or a subordinate clause. Our MEA data do not show any significant difference in the position of objects between main clauses and subordinate clauses:

Table 13: The distribution of DOs in main and subordinate clauses in EANC ArmFilmNarr corpus

Clause type	Total O	VO	% VO
Main clause	391	80	20.5%
Subordinate clause	87	20	22.9%

3.2.2.10 Simple vs. complex verb forms

In some languages, different types of verb forms are associated with different positions of arguments. For example, some verb forms of nominal origin may show argument positions analogous to those of noun modifiers, which in modern Armenian almost invariably precede the element they modify. MEA has complex (periphrastic) verb forms involving participles, which have some nominal characteristics, as well as simple verb forms inherited from Classical Armenian.

Thus, we might expect a stronger preference for pre-predicate arguments with periphrastic verb forms. However, direct objects show no evidence for such a pattern, with little difference between different types of verb forms (see Table 14). In fact, periphrastic verb forms show a slightly higher percentage of post-predicate Os than simple verb forms (21.3% vs. 18.3%).

Table 14: The distribution of DOs with complex and simple verbs in EANC ArmFilmNarr corpus

Verb type	Total O[9]	VO	% VO
Complex verb	314	67	21.3%
Simple verb	115	21	18.3%

[9]Here, only clauses with a single participle + auxiliary or a single simple verb are counted.

3.3 Other post-predicate arguments

3.3.1 General overview

In addition to objects, other types of arguments also show variable position in Armenian and neighbouring languages of Western Asia. As discussed in Haig & Khan (2019), previous studies find that across the area, there is a tendency for Goal arguments of verbs of motion and caused motion to appear in post-predicate position. In some languages, this tendency is extended to other elements that could be considered to share "endpoint" semantics, namely recipient, benefactive, endpoint of change of state verbs such as 'become,' and addressee; see Haig et al. (2024: Section ?? [this volume]) for an updated overview.

Stilo (2018) finds that in his Armenian data, there is indeed a tendency for goals to appear in post-predicate position. He finds a weaker tendency for benefactive and recipient, but no such tendency for addressee, which strongly prefers pre-predicate position. He finds that instrumental, ablative, locative, and comitative arguments show a preference for pre-predicate position across the area, being generally unaffected by the tendency for post-predicate position associated with goals and goal-like elements. We investigate all these types of arguments, and find that goals do indeed show a preference for post-predicate position (approximately 70% post-predicate). This tendency also seems to apply to benefactive, though the numbers involved are small. The numbers for recipient of 'give,' and of verbs with similar meanings, are also small, but unlike Stilo's (2018) data, show few examples in post-predicate position. Like Stilo (2018), we find that addressee arguments, which are rarely expressed overtly, show no tendency to appear in post-predicate position in our Armenian data. In contrast, the name in sentences such as 'they call him/her/it X,' which are also infrequent, shows a marked preference for post-predicate position (see Table 15).

If we consider goals separately from the other roles in Table 15, we see that they have a much higher instance of post-predicate position (for the other categories with very high post-predicate figures, i.e. name and benefactive, the numbers are too small to draw firm conclusions):

(18) Goal: 68.9% post-predicate
Other roles: 29.1% post-predicate

The relatively high percentage of post-predicate position (40%) for ablative is unexpected in the light of proposals that interpret the tendency for post-predicate position as an iconic expression of "endpoint" semantics. It is possible that a different type of analogy is at work here, with ablative equated with Goal

Table 15: Frequency of post-predicate placement, other roles

Other roles	Total	Post-predicate	% Post-predicate
Name ('they call him/her/it X')	5	4	80%
Benefactive	11	8	72.7%
Goal	122	84	68.9%
Ablative	50	20	40%
Comitative	44	13	29.5%
Location	116	30	25.8%
Become	9	2	22.2%
Instrumental	71	15	21.1%
Recipient (transfer of possession in general)	11	2	18%
Recipient (of verb *tal* 'give')	6	1	16.7%
Addressee	6	0	0%
Total other roles[a]	445	178	40%

[a]This figure represents the total of 11 roles, minus Recipient (of verb *tal* 'give'), as this is already included in Recipient (transfer of possession in general).

as both are typical arguments of verbs of motion (ablative as starting point, Goal as endpoint). However, it is more likely that the relatively high post-predicate percentage of ablative in these texts simply reflects the fact that it is more likely than average to be given (76% vs. 65% total other roles) and animate (32% vs. 27% total other roles), factors which seem to favour post-predicate position (grammatical definiteness is not relevant here, as the ablative case ending in MEA cannot co-occur with the definite article).

3.3.2 Impact factors

3.3.2.1 Definiteness

As in the case of DOs, "definiteness" is used here to indicate the presence of the definite article. Note that many of the roles investigated here (those that take ablative, instrumental, locative, or genitive case, and most nominative goals and locations) grammatically exclude the definite article. The fact that goals, which show a marked preference for postverbal position, are often grammatically indefinite for reasons of morphology rather than semantics gives a large number

of postverbal indefinites, masking the effect of definiteness in itself. Therefore, goals are shown separately from other roles in Table 16 below. However, even when goals are excluded, we still find that definites are more likely than indefinites to appear in post-predicate position (38.1% vs. 26.9%). Indeed, goals themselves also show a higher percentage of post-predicate position when grammatically definite (75% vs. 64.9% for indefinite). Nonetheless, the effect of definiteness on roles other than DO is not statistically significant.

Table 16: The distribution of definite and indefinite NP other roles in EANC ArmFilmNarr corpus

	Total	Post-predicate	% Post-predicate
Indefinite NP Goal	57	37	64.9%
Definite NP Goal	52	39	75%
Indefinite NP other	119	32	26.9%
Definite NP other	105	40	38.1%

3.3.2.2 Givenness

In the light of the fact that many other roles cannot take the definite article for grammatical reasons, it might be expected that givenness could show a stronger correlation than definiteness with post-predicate position. However, for roles other than Goal, we find no apparent effect of givenness at all. Note that the correlation between givenness and post-predicate position is much weaker for other roles than for direct object, with new other roles showing 30.1% post-predicate position, compared to 8.9% for new direct objects (see Table 17).

Table 17: The distribution of given and new other roles in EANC ArmFilmNarr corpus

	Total	Post-predicate	% Post-predicate
Given Goal	77	56	72.7%
New Goal	39	24	61.5%
Given other	221	64	29.0%
New other	103	31	30.1%

3.3.2.3 Animacy

As with direct objects, we see that animate referents appear more frequently in post-predicate position than inanimate ones, although, once again, the effect does not reach statistical significance. However, the difference is smaller here than that found with direct objects, where animates are 28% post-predicate and inanimates 18%. A possible reason for this is that many of the inanimate other role referents are goals, which differ from all other roles in showing a preference for post-predicate position (goals are almost exclusively inanimate). Thus, in Table 18, we present the data for other roles excluding goals. It can be seen that if we discount goals, the effect of animacy becomes more apparent.

Table 18: The distribution of animate and inanimate other roles in EANC ArmFilmNarr corpus

	Total	Post-predicate	% Post-predicate
Total other roles animate	122	45	36.9%
Total other roles inanimate	331	135	40.7%
Other roles animate -Goal	113	38	33.6%
Other roles inanimate -Goal	218	58	26.6%

As we have seen, animate referents are more likely to be definite than inanimate ones, so in order to accurately gauge the effect of animacy, definiteness must also be taken into account, as shown in Table 19. As with DOs, we find that for definite NPs animate referents do show a higher proportion of post-predicate position than inanimates, although again, this does not reach statistical significance, while for indefinite NPs, which show lower frequency of post-predicate position, there does not seem to be any effect of animacy. For the reasons discussed above, these figures do not include goals.

3.3.2.4 Weight

As in the case of DOs, we find that other roles with weight 1 (composed of one word) show a lower frequency of post-predicate position than those which are longer, indicating that weight may be a factor promoting post-predicate position. This is the case even if we exclude pronouns, which are typically composed of one word and show a stronger preference for pre-predicate position than lexical NPs (see following Section e) “Lexical vs. pronominal”). Very heavy elements (weight 4+) show the highest percentage of post-predicate position, indicating

Table 19: The distribution of animate and inanimate other roles in EANC ArmFilmNarr corpus divided by definiteness

Definite NP						Indefinite NP					
Animate			Inanimate			Animate			Inanimate		
N	VX	%VX	N	VX	%VX	N	VX	%VX	N	VX	%VX
41	19	46.3%	64	21	32.8%	25	7	28%	94	25	26.6%
Total definite NP: 105 Total definite VX: 40 % definite VX: 38.1%						Total indefinite NP: 119 Total indefinite VX: 32 % indefinite VX: 26.9%					

the possibility of heavy NP shift. The effect of weight is not particularly strong, but other roles present somewhat stronger evidence than DOs for its relevance as a factor favouring post-predicate position (see Table 20).

Table 20: The distribution of other roles according to weight in EANC ArmFilmNarr corpus

Weight	Total	Post-predicate	% Post-predicate
1	239	87	36.4%
1 (NP only)	19	7	36.8%
2	139	59	42.4%
3	51	21	41.4%
3+	74	32	43.2%
4+	23	11	47.8%

3.3.2.5 Lexical vs. pronominal

A comparison of lexical NPs and pronouns shows that the former appear more frequently in post-predicate position in the EANC ArmFilmNarr corpus. However, note that certain types of pronouns, such as interrogative and relative pronouns, show particular syntactic behaviour that places them in pre-predicate position, so it will be more informative to investigate those types of pronouns which can appear in either pre- or post-predicate position. For this reason, we look at demonstrative pronouns, personal pronouns, and the "emphatic" pronoun *inkʿə*

(for a discussion of this element and its behaviour in this corpus, see Hodgson et al. In press; see also Donabedian-Demopoulos 2007). Since all these elements are inherently definite, it is also informative to compare them with definite NPs. We still find that lexical NPs show a clearly higher percentage of post-predicate occurrences than any of these pronouns, and if we compare these pronouns with definite NPs, the difference is greater still. It is interesting that other roles seem to show a clearer difference between the behaviour of lexical and pronominal elements than DOs, which show only a small difference in the frequency of post-predicate position (see Table 21).

Table 21: The distribution of lexical and pronominal other roles in EANC ArmFilmNarr corpus

Type of element	Total	Post-predicate	% Post-predicate
Total pronouns	116	31	26.7%
Total lexical NP	337	149	44.2%
Demonstratives	12	4	33.3%
“Emphatic” pronoun	10	2	20%
Personal pronouns	17	6	35.3%
Definite NP	157	79	50.3%

3.3.2.6 Crowding effect

As discussed in the corresponding section on DOs, it has been proposed that the presence of another pre-predicate argument could favour post-predicate position, in order to avoid “crowding” of more than one argument on the same side of the predicate. However, our data do not provide evidence for this, as in fact the percentage of post-predicate arguments is higher (51.3%) when there is no other overt argument than when there is an overt pre-predicate subject and/or object (36.8%). When both subject and object are pre-predicate, other arguments are less frequent still in post-predicate position (33.3%). In the presence of a post-predicate subject, other roles appear more infrequently in post-predicate position (20%), implying that there may be a tendency to avoid more than one post-predicate argument. However, when there is a postverbal object, the figures are very close to the average for other roles as a whole (37.9%, as compared to ≈ 40% for other roles in general). Thus, these data do not provide conclusive evidence for any type of crowding effect (see Table 22).

Table 22: The distribution of other roles according to the presence of other overt arguments in EANC ArmFilmNarr corpus

	Total	Post-predicate	% Post-predicate
No overt S, A, or O	146	75	51.3%
Overt preverbal S	149	51	34.2%
Overt preverbal A	7	3	42.9%
Overt preverbal O	46	21	45.6%
Overt preverbal A and O	21	7	33.3%
Total overt preverbal only	223	82	36.8%
Overt preverbal A, postverbal O	8	2	25%
Overt preverbal O, postverbal A	0	0	0%
Overt postverbal S	25	5	20%
Overt postverbal A	0	0	0%
Overt postverbal O	29	11	37.9%
Overt postverbal A and O	0	0	0%
Total overt postverbal only	54	16	29.6%

3.3.2.7 Main vs. subordinate clause

In the section on direct objects, we saw that in these data, there is no significant difference between the position of direct objects in main vs. subordinate clauses. However, other roles, and especially goals, show a higher frequency of post-predicate position in main clauses as compared to subordinate clauses. Subordinate clauses have been observed to show more conservative word order patterns than main clauses, for example the persistence of OV in subordinate clauses in Germanic languages such as German. Thus a possible explanation for these findings is that the tendency for post-predicate Goals is a relatively recent phenomenon that has spread by contact from other languages of the area, such as Iranian languages and Neo-Aramaic, and has not spread fully to subordinate clauses. In this context, it is interesting that direct objects do not show such a difference, suggesting that their variable position could indeed be a conservative feature inherited from Classical Armenian, rather than a more recent contact-induced phenomenon. However, note that the percentage of postverbal goals (48%, see Table 23) is still much higher than that of postverbal DOs (22.9%, see Table 13) in subordinate clauses, so we can say that the tendency for post-predicate Goals is still present. In addition, we only have a small number of subordinate clauses with Goal arguments, so it is not possible to draw firm conclusions on this issue.

Table 23: The distribution of other roles in main and subordinate clauses in EANC ArmFilmNarr corpus

	Total	Post-predicate	% Post-predicate
MC total other roles	341	144	42.2%
SC total other roles	84	27	32.1%
MC goals only	94	70	74.5%
SC goals only	25	12	48%

3.3.2.8 Simple vs. complex verb form

We observe that complex (participle + auxiliary) verb forms show a lower percentage of post-verbal arguments, including goals, than simple verb forms. A possible explanation for this phenomenon is that arguments of participial (nominalized) verb forms tend to show positional characteristics of noun modifiers, i.e. preceding the element they modify. It is also possible that the phenomenon is linked to properties of focus marking in Eastern Armenian, as the auxiliary in complex verb forms can mark focus when the focused element precedes the lexical verb, but not when it follows. However, neither of these explanations is particularly convincing given the fact that direct objects do not seem to show this pattern (as seen in Section 3.2.2, in this corpus, complex verb forms in fact show a slightly higher percentage of post-predicate DOs than simple verb forms). Further research is clearly needed to clarify the interaction between verb type, word order, and information structure in Armenian.

Table 24: The distribution of other roles with complex and simple verbs in EANC ArmFilmNarr corpus

	Total	Post-predicate	% Post-predicate
Complex verb total other roles	288	110	38.2%
Complex verb goals only	77	53	68.8%
Simple verb total other roles	53	27	50.9%
Simple verb goals only	18	14	77.8%

3.3.3 Summary of other roles

The data from our MEA corpus confirms that in MEA, as in other languages of the wider area, goals of verbs of motion and caused motion show a preference for post-predicate position (approximately 70%). However, in contrast to some languages of Western Asia (see Haig et al. 2024 [this volume]), there is no such tendency observed for other constituents sharing the semantics of "endpoint," such as recipient or addressee. An apparent exception is benefactive, which shows an even stronger preference for post-predicate position (72.7%), although the small number of examples makes this less reliable. It is also worth noting that the number of overt recipient and addressee referents in this corpus is very low. Apart from Goal, virtually all the other roles investigated which have more than 20 examples (comitative, location, instrumental) show similar figures, of 20–30% post-predicate position, similar to those for direct objects overall. Ablative has a slightly higher figure (40% post-predicate), but this may be because the ablative referents in this corpus show higher than average figures for givenness and animacy, which may have some effect favouring post-predicate position.

Definiteness, too, is shown to favour post-predicate position, but indefinite other roles do not show the extreme preference for pre-predicate position that is characteristic of indefinite direct objects. Even bare indefinite other roles show 38.7% post-predicate position, while indefinite other roles in general (including those with the indefinite article or an indefinite proform) show 39.2%, virtually identical to the average of other roles as a whole (40%). Weight appears to have an effect on position, with heavier elements appearing somewhat more frequently in post-predicate position. Other roles show more evidence of an effect of weight than direct objects, although this is still not particularly strong. Other roles also show more clearly than direct objects the differences in behaviour between pronouns and lexical NPs, with the latter being more likely to appear in post-predicate position (44.2%, vs. 26.7% for pronouns). As with direct objects, there is no evidence for a crowding effect, whereby the presence of other pre-predicate arguments could promote post-predicate position in order to avoid "crowding" of arguments on one side of the predicate. Unlike direct objects, other roles present possible evidence that post-predicate constituents may be more common in main than subordinate clauses, and with simple rather than complex verb forms. A possible explanation for the former could be that the post-predicate Goal phenomenon is a recent contact-induced development that has spread more slowly to subordinate clauses, although the numbers are too small to draw any firm conclusions. The link between argument position and verb form is a topic for further research.

4 Conclusions

As regards direct objects, definiteness proved to be a key impact factor for the postverbal position: 33% (def) vs. 10% (indef) vs. 2% (bare). This is consistent with the previous study by Samvelian et al. (2023), with the difference that in the present study, the percentage of post-predicate definite Os is considerably lower than those reported by Samvelian et al. (2023), who find 82.7% of definite Os in post-predicate position in their first experiment. The percentage of post-predicate definite DOs in the present study is intermediate between the very high figures found by Samvelian et al. (2023), and the very low figures (around 14% for colloquial Yerevan) reported by Stilo (2018). One probable factor behind the difference is that the data in the experimental studies of Samvelian et al. (2023) include only out-of-the-blue sentences; note that a similar pattern has been observed in Romeyka, where those studies based on elicitation of out-of-the-blue sentences yield predominantly VO structures, while the data from connected spontaneous spoken discourse show a much higher rate of OV (Schreiber & Janse 2024 [this volume]). As Samvelian et al. (2023) note in their conclusion, it is very likely that the rate of SOV is higher in spontaneous oral discourse.

Another relevant factor, also proposed by Samvelian et al. (2023), is register. Samvelian et al. (2023) includes data with characteristics of formal literary language, which show significant morphological, phonological, and syntactic differences from colloquial Yerevan EA, such that the two should be considered different forms of language. The association of post-predicate Os with formal registers is supported by the fact that the speaker who uses by far the highest percentage of postverbal Os in the present study (44%, as opposed to an average of 17.4% for all the other speakers, and 12.7% for the speaker with the lowest percentage) is also the only one who uses certain word forms associated with the formal literary language. Thus, if we discount this one speaker, who uses a more formal register, the percentage of postverbal Os in this study is not so different from that recorded by Stilo (2018) for colloquial Yerevan. In addition, a similar figure (approx. 90% OV) is obtained for nominal direct objects in the Agulis corpus of spoken vernacular Armenian (Hodgson In press).

The effect of grammatical definiteness is stronger than that of the pragmatic property of givenness (29.8% of given Os appear in post-predicate position, vs. 8.9% of new information Os). Among definite Os, animate referents show a slightly higher percentage of post-predicate position (42.8% for animates vs. 30% for inanimates), although this does not reach statistical significance. This, too, is consistent with the findings of Samvelian et al. (2023). Given that definiteness and animacy are characteristics associated with topical referents, it is un-

surprising that topic persistence is higher for postverbal Os (1.9 (postverbal) vs. 1.6 (preverbal)). Post-predicate Os are also associated with slightly lower average referential distance than pre-predicate ones (2.3 vs. 2.5), which is also to be expected given the association of post-predicate position with topicality in general. Some effect of heavy NP shift effect was observed, with longer NPs being more frequent in post-predicate position. (13% (1-word Os) vs. 31% (2-word Os) vs. 19% (3+word Os)). This is also broadly consistent with the findings of Samvelian et al. (2023). Neither the object type (lexical or pronominal) nor the crowding / null subject effect had any evident impact on the position of DOs. Overall, the present corpus study of oral narratives showed that OV word order is more frequent in MEA than VO (79% preverbal vs. 21% postverbal). A prototypical postverbal O in this corpus is definite, given (91%), human (41% vs. 27% of preverbal objects), with higher topic persistence.

As regards other post-predicate constituents, the same factors that have been found to be associated with post-predicate position for DOs (definiteness, animacy, and weight) seem to have a slight effect in the case of other roles too, although the numbers involved cannot be considered statistically significant. The effect of definiteness is less pronounced than for DOs, with 50.3% of definite NP other roles appearing in post-predicate position, vs. 39.2% of indefinites. As we can see, indefinite other roles do not show such a strong tendency to avoid post-predicate position as indefinite DOs. As with DOs, animate definite other roles (excluding goals) are more likely to appear in post-predicate position than inanimate definite ones (46.3% vs. 32.8%). As with DOs, animacy does not seem to affect the position of indefinites. Other roles present slightly stronger evidence of heavy NP shift to post-predicate position than DOs, with those comprised of 4+ words showing the highest percentage of post-predicate position (weight 1 = 36.4% post-predicate, weight 2 = 42.4%, weight 3 = 41.4%, weight 4+ = 47.8%), although the observed effect is still fairly weak. Other roles present evidence that pronominal arguments are less likely to appear in post-predicate position than lexical NPs (26.7% post-predicate, vs. 44.2% for lexical NPs), while for DOs there is no apparent difference. As for direct objects, no evidence is found of a crowding effect.

The preference for goals to appear in post-predicate position is a separate issue, that has been shown to have an areal dimension. This preference is confirmed by the data in this study, where 68.9% of goals appear in post-predicate position. This study presents no evidence that the preference for post-predicate position is extended to other constituents with "endpoint" semantics, such as recipients or addressees, with the possible exception of benefactive. However, the numbers

of overt examples of all these types of argument (recipient, addressee, benefactive) are very small, so we cannot draw a firm conclusion here. The fact that the tendency for goals to appear in post-predicate position is more pronounced in main than subordinate clauses is possible evidence that it is a relatively recent, contact-induced phenomenon. The position of direct objects shows no significant difference between main and subordinate clauses, and it is possible that the existence of postverbal objects, and perhaps other arguments, too, is a conservative characteristic inherited from Classical Armenian (see Samvelian et al. 2023, Stilo 2018), which may perhaps also explain its apparent association with formal register. This is a topic for future research. In any case, Modern Eastern Armenian can be said to fit the typological profile of an "OVX" language, in that despite showing mainly head-final characteristics, it also has some characteristics associated with typically head-initial languages, such as initial complementizers.

Abbreviations

1	first person	FUT	future
2	second person	GEN	genitive
3	third person	IA	indefinite article
AGR3	third-person agreement marker	IPFV	imperfective
AOR	aorist	NEG	negative
CAUS	causative	NOM	nominative
DAT	dative	PFV	perfective
DEM2	medial demonstrative	PRS	present
DEF	definite	PL	plural
EMP	emphatic	SG	singular
FPT	future participle	SUB	subjunctive

References

Badikyan, Khachik. 1976. *Žamanakakicʿ hayereni parz naxadasutʿyan šaradasutʿyunə*. Yerevan: Yerevani hamalsarani hratarakchutyun.

Der-Houssikian, Haig. 1978. The semantic significance of S, O, V sequences in Armenian. In Mohammad A. Jazayery, Edgar C. Polomé & Werner Winter (eds.), *Linguistic and literary studies, vol. 2: Descriptive linguistics*, 227–236. Lisse: Peter de Ridder Press.

Donabedian-Demopoulos, Anaïd. 2007. A la recherche de la logophoricité en arménien. In Jocelyne Fernandez-Vest (ed.), *Combat pour les langues du monde/Fighting for the world's languages, hommage à Claude Hagège*, 165–176. Paris: Editions L'Harmattan.

Dryer, Matthew S. 1998. *Aspects of word order in the languages of Europe*. Berlin: De Gruyter Mouton.

Dryer, Matthew S. 2013. Order of object and verb. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/83.

Du Bois, John. 1987. The discourse basis of ergativity. *Language* 63(4). 805–855. DOI: 10.2307/415719.

Dum-Tragut, Jasmine. 2002. *Word order correlations and word order change: An "applied-typological" study on literary Armenian variants*. Munich: Lincom.

Dum-Tragut, Jasmine. 2009. *Armenian: Modern Eastern Armenian*. Amsterdam: Benjamins. DOI: 10.1075/loall.14.26.

Faghiri, Pegah & Pollet Samvelian. 2020. Word order preferences and the effect of phrasal length in SOV languages: Evidence from sentence production in Persian. *Glossa: A journal of general linguistics* 5(1). DOI: 10.5334/GJGL.1078.

Giorgi, Alessandra & Sona Haroutyunian. 2016. Word order and information structure in Modern Eastern Armenian. *Journal of the Society for Armenian Studies* 25. 185–201.

Givón, Talmy. 1975. Focus and the scope of assertion: Some Bantu evidence. *Studies in African Linguistics* 6. 185–205.

Givón, Talmy. 1983. Topic continuity in discourse: An introduction. In Talmy Givón (ed.), *Topic continuity in discourse: A quantitative cross-language study*, 1–42. Amsterdam: Benjamins Publishing Company.

Haig, Geoffrey & Geoffrey Khan. 2019. Introduction. In Geoffrey Haig & Geoffrey Khan (eds.), *The languages and linguistics of Western Asia: An areal perspective*, 1–29. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110421682-001.

Haig, Rasekh-Mahend, Stilo, Schreiber & 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: ??.

Hawkins, John A. 1979. Implicational universals as predictors of word order change. *Language* 55. 618–648.

Hawkins, John A. 1983. *Word order universals*. New York: Academic Press.

Hayrapetyan, Sergo A. 1981. Uɣiɣ xndri šaradasutʿunə dasakan grabarum. *Herald of Social Sciences* 9. 71–84.

Hodgson, Katherine. 2013. *Discourse configurationality in Eastern Armenian: The nominal projection*. Cambridge: University of Cambridge. (MA).

Hodgson, Katherine. 2019. *Relative clauses in colloquial Armenian: Syntax and typology*. Paris: Institut national des langues et civilisations orientales. (Doctoral dissertation).

Hodgson, Katherine. In press. Armenian, dialect of Agulis. In Geoffrey Haig, Donald Stilo, Mahîr C. Doğan & Nils N. Schiborr (eds.), *WOWA — Word Order in Western Asia: A spoken-language-based corpus for investigating areal effects in word order variation*. Bamberg: University of Bamberg.

Hodgson, Katherine, Victoria Khurshudyan & Anaïd Donabedian. In press. Word order, anaphora, and topicality in Eastern Armenian. In Claudine Chamoreau, Enrique Palancar & Anaïd Donabedian (eds.), *Topicality and the shaping of grammar: New perspectives from lesser-known languages*. Amsterdam: Benjamins.

Kahnemuyipour, Arsalan & Karine Megerdoomian. 2011. Second position clitics and the vP phase: The case of the Armenian auxiliary. *Linguistic Inquiry* 42(1). 152–162. DOI: 10.1162/ling_a_00033.

Kahnemuyipour, Arsalan & Karine Megerdoomian. 2017. On the positional distribution of an Armenian auxiliary: Second-position clisis, focus, and phases. *Syntax* 20(1). 77–97.

Khurshudyan, Victoria. 2006. *Средства выражения хезитации в устном армянском дискурсе в типологической перспективе: Hesitation means in oral Armenian discourse in typological perspective*. Moscow: Russian State University for the Humanities. (Doctoral dissertation).

Khurshudyan, Victoria, Timofey Arkhangelskiy, Misha Daniel, Vladimir Plungian, Dmitri Levonian, Alex Polyakov & Sergei Rubakov. 2022. Eastern Armenian national corpus: State of the art and perspectives. In Victoria Khurshudyan, Nadi Tomeh, Damien Nouvel, Anaid Donabedian & Chahan Vidal-Gorene (eds.), *Proceedings of the workshop on processing language variation: Digital Armenian (DigitAm) within the 13th Language Resources and Evaluation Conference*, 28–37. Marseille: European Language Resources Association. https://aclanthology.org/2022.digitam-1.5.

Kozintseva, Natalia. 1995. *Modern Eastern Armenian*. Munich: Lincom Europa.

Minassian, Martiros. 1980. *Grammaire d'Arménien Oriental*. New York: Delmar.

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: ??.

Samvelian, Pollet, Pegah Faghiri & Victoria Khurshudyan. 2023. On the persistence of SVO: The case of Modern Eastern Armenian. *Linguistics* 61(3). 459–497. DOI: 10.1515/ling-2021-0013.

Schreiber, Laurentia & Mark Janse. 2024. Word order variation in Romeyka. 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*, 337–373. Berlin: Language Science Press. DOI: ??.

Sigler, Michele. 2001. A logophoric pronoun in Western Armenian. *Annual of Armenian Linguistics* 21. 13–30.

Stilo, Donald. 2006. Circumpositions as an areal response: The case study of the Iranian zone. In Lars Johanson & Christiane Bulut (eds.), *Turkic-Iranian contact areas: Historical and linguistic aspects*, 310–333. Wiesbaden: Harrassowitz.

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.

Tamrazian, Armine. 1991. Focus and WH-movement in Armenian. *UCL Working Papers in Linguistics* 3. 102–121.

Tamrazian, Armine. 1994. *The syntax of Armenian: Chains and the auxiliary*. London: University College London. (Doctoral dissertation).