Chapter 2

Word order in the speech of Kurmanji-Turkish bilinguals

Kateryna Iefremenko
Leibniz-Centre General Linguistics & University of Potsdam

The paper investigates word order, particularly the domain of post-predicate position, in Turkish and Kurmanji as two languages located in the Western Asian Transition Zone that are in an intense and long-term contact with each other. Both languages are OV; however, each of them allows placement of constituents in post-predicate position. The results of the analysis show that there is an effect of flagging and semantic role in Kurmanji, which is in line with previous research on word order in Kurmanji, and an effect of weight in Turkish on the employment of post-predicate elements. At the same time, the qualitative analysis showed that there are instances of noncanonical placement of case-flagged goals in Kurmanji that occur in one particular construction *erdê ketin* 'to fall on the ground.'

1 Introduction

In this paper, I investigate a possible structural convergence in the word order, namely the post-predicate domain, of Turkish and Kurmanji as two languages located in the Western Asian Transition Zone (for a definition, see Haig et al. 2024 [this volume]). Both languages are OV languages, but each of them employs the post-predicate position to some extent. In Turkish, word order is determined by information structural requirements, and thus the post-predicate position is reserved for background information. In Kurmanji, the post-predicate position is driven by verb semantics, i.e., Goal arguments are placed in the post-predicate position.

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

1.1 Factors of language change

In situations of language contact, it is typical for languages to influence each other. According to Thomason (2001), linguistic factors of contact-induced language change determine what will change once social factors have created the situation where something will change. However, linguistic factors can be overridden by social factors, and therefore Thomason & Kaufman (1988) define the intensity of contact as the most important social factor for the prediction of a contact-induced language change. Thomason (2001) provides a four-stage borrowing scale depending on the intensity of contact. In casual contact, when borrowers are not fluent in the source language, only content words are borrowed. In a slightly more intense contact setting when borrowers are fluent bilinguals but form a minority in the community, function words as well as content words can be borrowed at the lexical level, and some minor structural borrowing is possible. In a moderately intense contact setting when there are more borrowers than in the previous stage and social factors favor borrowing, more change can be expected on the structural level; word order features, clause-combining strategies, or inflectional affixes might be borrowed. And finally, in an intense contact setting when there is extensive bilingualism among the speakers and social factors strongly favor borrowing, anything can be borrowed, resulting even in major typological change of the borrowing language.

The intensity of contact is determined by the duration of the contact and the number of speakers in the community. The contact between Turkish and Kurmanji has begun in the Ottoman Empire and intensified with the establishment of the Turkish Republic where Turkish became the country's sole official language (Yağmur 2001). Nowadays, Turkish is dominant and even the only language in official spheres as education, government, business, while Kurmanji is used mostly in the families. As for the number of speakers in the community, the size of the Kurmanji-speaking communities varies depending on the region: in the eastern and south-eastern regions of Turkey, Kurmanji is spoken by the majority of the population, while in the western regions, Turkish prevails even in informal contexts. Thus, on the borrowability scale, the contact between Kurmanji and Turkish in Turkey can be referred to as stage 3 (a setting of moderately intense contact) where structural changes, including word order alternations, can be expected.

Another social factor, which is not provided in Thomason & Kaufman (1988) as one of the most important factors for predicting the outcome of the contact, but which I find relevant for my particular research is the societal status of a language: minority and majority language. It is believed that typically a more prestigious donor language influences a less prestigious recipient language (Johanson

2002). As has been noted above, Turkish in Turkey is the dominant language in such spheres of life as education, business, media, and in the western regions of Turkey it prevails also in the informal settings. Thus, Turkish is clearly the majority language of the society. As for Kurmanji in Turkey, its societal status heavily depends on the region: while in the eastern and south-eastern regions Kurmanji is the dominant language of the majority of speakers (though only in informal settings), in the western regions, it is the minority language. For my study, the data from Kurmanji-Turkish bilinguals in Turkey were collected in Ankara, the capital of Turkey, where Turkish is the dominant language, and Kurmanji is a minority language in this context.

1.2 Word order in language contact situations

A number of studies (Thomason & Kaufman 1988; Thomason 2001; Heine 2008) pointed that word order is prone to change in language contact scenarios. Following this, there are studies that exemplify this claim for Turkic languages. For example, the word order in Karaim changed from OV to VO due to the contact with Slavic and Baltic languages (Csató 2002). Similar to Karaim, Gagauz, which has stayed in a long term contact with Slavic languages, underwent change that resulted in a VO order becoming dominant (Menz 1999). Another study by Keskin (2023) investigated word order across numerous Turkic varieties in the Balkans where Turkish has been in contact mostly with Indo-European languages for centuries. Analysis of the post-predicate domain in these Turkic varieties showed that the further the Turkic variety is located from the borders of Turkey, the higher is the frequency of the VX order in this variety. On the other hand, studies that investigated comparatively recent contact of Turkish with Indo-European languages, i.e., Turkish as a heritage language in the Netherlands, Germany and the U.S., did not find a shift from OV to VO order in these varieties (Doğruöz & Backus 2007; Schroeder et al. In press). Finally, as for research on word order in Turkish in contact with another language in a context where Turkish is the majority language of the society, there are no studies, at least to my knowledge, apart from those done by our research group based on the same data set as the one presented in this paper (Iefremenko et al. 2023; Iefremenko Submitted).

With respect to word order in Kurmanji in Turkey, there are studies that investigate the post-predicate position across different dialects of Kurmanji (Haig 2015; Haig 2019; Gündoğdu 2019), but the focus is more on word order typology and its diachronic change due to contact with other languages. Besides, Asadpour (2022) investigated the placement of goals in Mukri Kurdish spoken in Iran and in the contact languages (such as Armenian, Azeri Turkic and Northeastern

Neo-Aramaic) and found that in fact the combination of several factors such as information structure, semantics and morphosyntax explains the placement of constituents in the post-predicate position in the analyzed languages.

Thus, this study will try to fill a gap first by investigating possible changes in the word order in Turkish in contact with another language in a context where Turkish is the majority language of the society; second, by investigating word order, namely the post-predicate position, in Kurmanji, that is in intense contact with Turkish and is the minority language of the society. The analysis is based on the data that come from 30 Kurmanji-Turkish adult bilingual speakers residing in the Turkish-dominant region in Turkey (Iefremenko 2021a,b).

2 Word order in Kurmanji

Kurdish is a macro-language that consists of a continuum of closely related languages spoken by Kurds over a large geographic area spanned across several countries, such as Turkey, Iran, Iraq, Syria, among others (Sheyholislami 2015). Northern Kurmanji is one of the Kurdish languages, predominantly spoken in southeast Turkey, northwest and northeast Iran, northern Iraq, and northern Syria. It is classified as a member of the northwest Iranian branch of the west Iranian languages, within the Iranian branch of Indo-European language family (Haig & Matras 2002).

Like other West Iranian languages, Kurmanji is an OV language, although it is not always verb-final (Haig 2015). Kurmanji indeed systematically places certain elements after the verb. According to Haig & Thiele (2014), the post-predicate position is reserved for "goals", where it is a cover term for:

a) locational goals of verbs of motion (e.g., *go, run, fall*) and caused motion (e.g., *put, place, take*)

(1) Northern Kurdish Yavuzeli (Matras et al. 2016: K002)
ez çû-m-e mal-ê
1SG go.PST-1SG-DRCT house-OBL.F
'I went home.'

(2) Karakoçan (Matras et al. 2016: K075)
jinik qutîk-ek anî mal-ê
woman box-INDEF bring.PST.3SG house-OBL.F
'The woman brought a box into the house.'

b) recipients of verbs of transfer (e.g., *give*)

(3) Northern Kurdish Elbistan (Matras et al. 2016: K022)
we ew ne-dê mi
2PL this NEG-give.PST.3SG 1SG.OBL
'You didn't give it to me.'

c) addressees of verbs of speech (e.g., *say, speak, promise*)

(4) Northern Kurdish Siirt (Matras et al. 2016: K008)
min got-e wî
1SG.OBL say.PST.3SG-DRCT 3SG.OBL
'I said it to him.'

In other words, the word order of Kurmanji is not a pure OV, but rather OVX. Goals are systematically yet not consistently placed in the post-predicate position in Kurmanji. Haig (2022) states that the position of post-predicate elements in Kurmanji is syntactically fixed and is not the result of pragmatically driven scrambling or stylistic variation, i.e., factors that account for example for the post-predicate position in Turkish. Factors that influence the position of goals are flagging and regional variation (Haig 2015; Gündoğdu 2019). On the other hand, a recent study by Asadpour (2022) on Mukri Kurdish spoken in Iran showed that information structure also plays a role in the placement of Goal arguments. Namely, accessible inferable information occurs in the post-predicate position, while topicalized goals are placed in the preverbal position. Nonetheless, considering that Kurdish varieties are spoken across several countries and as a result come into contact with different languages, the variations observed in the studies could potentially stem from distinctions between these varieties.

I will now elaborate on the relation between the word order in Kurmanji and the type of flagging of Goal arguments. Goals in Kurmanji can be flagged by case (as demonstrated in 5) or by an adposition (see examples 7–9).

(5) Northern Kurdish Tatvan (Matras et al. 2016: K024)
min xarin anî od-ê
1SG.OBL food bring.PST.3SG room-OBL.F
'I brought the food to the room.'

Case-flagged Goal arguments are always placed right after the predicate and cannot be separated by an adverb or any other argument. Importantly, the Goal

argument cannot be placed immediately before the verb. As Gündoğdu (2019: 110) emphasizes, "in a Kurmanji clause, at most two case-flagged NPs (subject and direct object) are licensed in the preverbal position". Thus, example 6 below where the argument *odê* (to the room) is placed in the immediate pre-predicate position would be considered noncanonical.

(6) Northern Kurdish Tatvan (constructed example)
min xarin od-ê anî
1SG.OBL food room-OBL bring.PST.3SG
'I brought the food to the room.'

The second means of marking goals is by the help of an adposition. In Kurmanji, there are several types of adpositions, namely basic prepositions, locational nouns which can be used together with a preposition (see example 7), postpositions, and circumpositions.

(7) Northern Kurdish Pertek (Matras et al. 2016: K028)
lawik-ê qiçik di-bez-e ber_bi dî-ya xwe
boy-EZ.M little PRS-run.3SG-DRCT towards mother-EZ.F own
'The little boy is running to his mother.'

In general, the position of Goal arguments flagged by adpositions is more flexible compared to those flagged by case. The only exception are locational nouns that are not preceded by a preposition: they are always placed in post-predicate position. The reason for such position is that this type of adpositions historically evolved from nouns. Hence, similar to case-flagged goals, locational nouns are placed in the post-predicate position (Haig 2015; Haig & Thiele 2014). As for the position of the other types of adpositions, it is largely dependent on dialect. As it is noted in Haig (2015) and Haig & Thiele (2014), initially the OVX word order of Kurmanji emerged due to the contact with early Aramaic/Iranian languages. As a result, there are different preferences in placing goals across the modern dialects of Kurmanji: namely, goals are predominantly post-predicative in the southernmost dialects where the language stayed in a long-lasting contact with Neo-Aramaic, while in the northern and western dialects goals are overwhelmingly pre-predicative due to their extensive contact with Armenian and Turkish. Besides, in the south, there is a tendency for post-predicative goals to be accompanied by a preposition or a circumposition. At the same time, in the dialects of Central Anatolia, the combination of a post-predicative Goal accompanied by a preposition is very restricted.

3 Word order in Turkish

Turkish is considered to have a relatively free word order, with basic word order being (S)OV, which means that even though word order variation is possible, in some instances word order must stay fixed. In Turkish, variation in word order serves pragmatic purposes such as signaling topics and distinguishing between old and new information (Erguvanlı 1984). Hence, word order in Turkish is strongly motivated by information structure: a link to the previous context or topicalized information appears sentence-initially, new information occurs immediately before the verb, and backgrounded information can be placed post-predicatively (Erguvanlı 1984; Kornfilt 1997). Thus, (8a) has pragmatically neutral unmarked order, whereas (8b–8f) are pragmatically marked.

(8) a. Turkish (self-constructed examples)
SOV
Murat araba-yı sat-tı.
Murat car-ACC sell-PST.3SG
'Murat sold the car.'

b. OSV
Araba-yı Murat sat-tı.
car-ACC Murat sell-PST.3SG
'It is Murat who sold the car.'

c. SVO
Murat sat-tı araba-yı.
Murat sell-PST.3SG car-ACC
'Murat sold the car.' (emphasis on the verb)

d. OVS
Araba-yı sat-tı Murat.
car-ACC sell-PST.3SG Murat
'It is the car that Murat sold.'

e. VSO
Sat-tı Murat araba-yı.
sell-PST.3SG Murat car-ACC
'It (the car) was sold by Murat.'

f. VOS
Sat-tı araba-yı Murat.
sell-PST.3SG car-ACC Murat
'The car was sold (by Murat).'

Furthermore, it is important to note that there are differences in the employment of post-predicate structures in spoken and written modes as well as formal and informal language. Particularly, there seem to be restrictions on the use of post-predicate structures in formal written situations. For example, in legal documents as well as news items (both written in newspapers and journals as well as read on radio or TV), post-predicate structures are extremely rare (Erguvanlı 1984: 67). Exceptions are columns in papers when authors deliberately choose more informal style, or literature pieces where an author uses their own style and may employ backgrounding techniques for specific pragmatic reasons (Erguvanlı 1984: 67). On the other hand, elements placed in post-predicate position frequently occur in informal spoken mode, which is usually characterized as spontaneous and unplanned (Schroeder 1995). Thus, spoken language is full of (self-)corrections and afterthoughts, while the same process in written language can be employed with the help of editing and corrections. Another reason of a higher number of post-predicate structures in spoken language compared to written one is that utterances are limited in size due to speaker's awareness of the listener's capacity limitations (Chafe 1985). Thus, as Schroeder (1995: 206) emphasizes, the employment of the post-predicate position helps the listener to keep track of the topical development and the deictic framework in which the predication holds.

4 Research questions

To summarize the preceding discussion, both languages in contact that are under research here have OV word order, but both employ the post-predicate position in a different way. While Kurmanji systematically places Goal arguments — particularly those flagged with case — in the post-predicate position, word order in Turkish is determined by information structural requirements, and the post-predicate position is reserved for background information, regardless of the semantic role of the elements.

As discussed in Section 1, the contact between the two languages has lasted for centuries and is fairly intense, and according to Thomason & Kaufman's (1988) borrowability scale, we can expect structural changes, including the ones concerning word order. In the paper, I will investigate potential convergences in word order, in particular in the post-predicate domain, in Turkish and Kurmanji in Turkey, basing the analysis on the variables encoded in WOWA (see Haig et al. 2024 [this volume]). For the reason that WOWA categories do not incorporate information structure as one of the variables, I do not explore its effect on the word order changes in the investigated languages. This is a limitation of the current paper.

5 Methodology

5.1 Participants

The data for this study come from 30 Kurmanji-Turkish bilinguals (9 females and 21 males). The participants were exposed to Kurmanji from birth or an early age in their family and started acquiring Turkish mainly when they entered school (though some participants were already exposed to Turkish in their families). The place of birth of the participants varies: the majority of the speakers were born and raised in the east and the south-east of Turkey, but some were born and grew up in the western cities of Turkey. At the time of data collection, all participants were living in Ankara, where the dominant language of the society is Turkish. In an urban city, such as Ankara, speakers of different dialects interact between each other on a daily basis; hence, such contact may lead to dialect levelling. Therefore, my data were not controlled for the dialect of Kurmanji. The age of the participants varies between 23 and 37 years, with the mean age being 28.1; thus, all the participants are young adults. All 30 participants had a high school degree, and most of them completed bachelor's or master's degree. So, all the speakers are highly proficient in Turkish since it is the language of education in Turkey. As for the education in Kurmanji, 14 participants stated that they had taken at least a several-month language course in Kurmanji, one participant had earned a master's degree from a Kurmanji language and literature department, and several participants indicated that they actively used Kurmanji for business purposes. However, the results of the section on the self-rated proficiency in the questionnaire showed that on average the speakers rate their proficiency in Kurmanji lower ($\bar{x}$ = 17.53 out of 20 possible) compared to their proficiency in Turkish ($\bar{x}$ = 19.32 out of 20 possible).

5.2 Data collection

The data used for this study were collected with the help of the "Language Situations" method (Wiese 2020). This method combines controlled elicitation with spontaneous data, and thus is suitable for systematic comparisons across contact-linguistic constellations as well as different languages. It captures quasi-naturalistic productions across different communicative situations, including informal versus formal, and written versus spoken communicative situations. The elicitation comprised two sessions (one in Kurmanji and one in Turkish), with at least three days between the sessions. Participants were shown a short video of a car accident and were asked to imagine themselves having witnessed it. After that they were asked to recount the incident in four different imagined situations:

to a friend via a WhatsApp voice message (informal spoken), to a friend via a WhatsApp text message (informal written), to the police via a voice mail (formal spoken), to the police in a form of a written witness report (formal written). To exclude a possible effect of priming, the order of the communicative situations and the order of the languages were balanced.

At the end of the second session, each participant was asked to fill out an extensive questionnaire. The questionnaire comprised nine sections: participants' general information, educational and professional background of the participants, family information, information about the languages, self-assessment of their language skills (in Kurmanji, Turkish, as well as in foreign languages, on a five-point scale), participants' language use with family members and peers, a section concerning media use and free time (texting WhatsApp messages or writing emails in Kurmanji, Turkish, as well as in foreign languages, three scores of frequency), questions concerning personal character traits, and feedback on the participation in the study. The questionnaire was available in Turkish as it is the language for official documents in Turkey. It was always filled out by the participants themselves in the presence of an elicitor.

5.3 Annotation and statistical analysis

For the current study on the post-predicate elements in Iranian and neighboring languages, only the informal spoken productions of the speakers were analyzed.[1] The Kurmanji data comprised 745 communication units[2] , out of which 507 were analyzed because they contained pre- or post-predicate arguments (Iefremenko 2021a). The Turkish data consisted of 799 communication units, out of which 587 were analyzed (Iefremenko 2021b). Data coding was done in line with the WOWA coding method (Haig et al. 2024 [this volume]) and annotated for animacy, weight, semantic role, flagging, and position. What is important for the current analysis is that afterthoughts and self-repairings, which usually occur in the post-predicate position in OV languages, were not excluded from the analysis. Besides, there are several aspects of coding that are important for the analysis of Turkish and Kurmanji data. First, adpositions are not counted as part of the weight because they are considered to be part of the flagging. Thus, for instance, in the example (9) the token *rê da* (on the street) is considered to have weight 1 because the postposition *da* is not counted.

[1](In-)formality and mode as factors that might influence the word order in Turkish and Kurmanji are analyzed and discussed in another study by Iefremenko (Submitted).

[2]A communication unit is an independent clause with its modifiers and dependent (subordinate) clauses.

(9) Northern Kurdish Ankara (Iefremenko 2021a: A, 3)
iii ez rê da di-çû-m
hmm 1SG street on PROG-go.PST-1SG
'I was going on the street.'

Second, subjects were excluded from the analysis across all the analyzed languages in WOWA. This has to be taken into account when interpreting the results for Turkish because as indicated in Section 3, unlike Kurmanji, the word order in Turkish is determined by information structure and as a result, subjects can also be placed in the post-predicate position as long as they are not new information.

The statistical analyses were run in R (R Core Team 2023). In addition to the base package of R, I used tidyverse for data manipulation and visualization (Wickham et al. 2019). I ran binomial generalized linear regression models using the lme4 package (Bates et al. 2015).

6 Results

6.1 Kurmanji

First, I will start with the analysis of post-predicate constituents in Kurmanji.

Figure 1 shows the percentage of tokens in the pre- and post-predicate position in the Kurmanji data. The line inside each box indicates the median, while a dot represents the percentage of tokens used in pre- or post-predicate position by one speaker, calculated in relation to the total number of communication units for each speaker. Figure 1 demonstrates that the majority of the communication units have pre-predicate arguments, with around 20% taking post-predicate arguments. Also, we can see that individual variability in the group is high: some participants place tokens post-predicatively in up to 70% of their utterances, while in some speakers all utterances are pre-predicate.

As stated in Section 2, word order in Kurmanji is determined by verb semantics where goals are placed in post-predicate position. Therefore, high individual variability in the group might be caused by some speakers' frequent use of goals. In order to establish whether there is an effect of one of the variables from the WOWA coding scheme, namely animacy, weight, role or flagging (as described in Section 5.3), I ran four binomial generalized linear regression models with the dependent variable Position (preverbal, coded as 0 vs. postverbal, coded as 1) and the independent variables Animacy, Weight, Flagging, and Role.

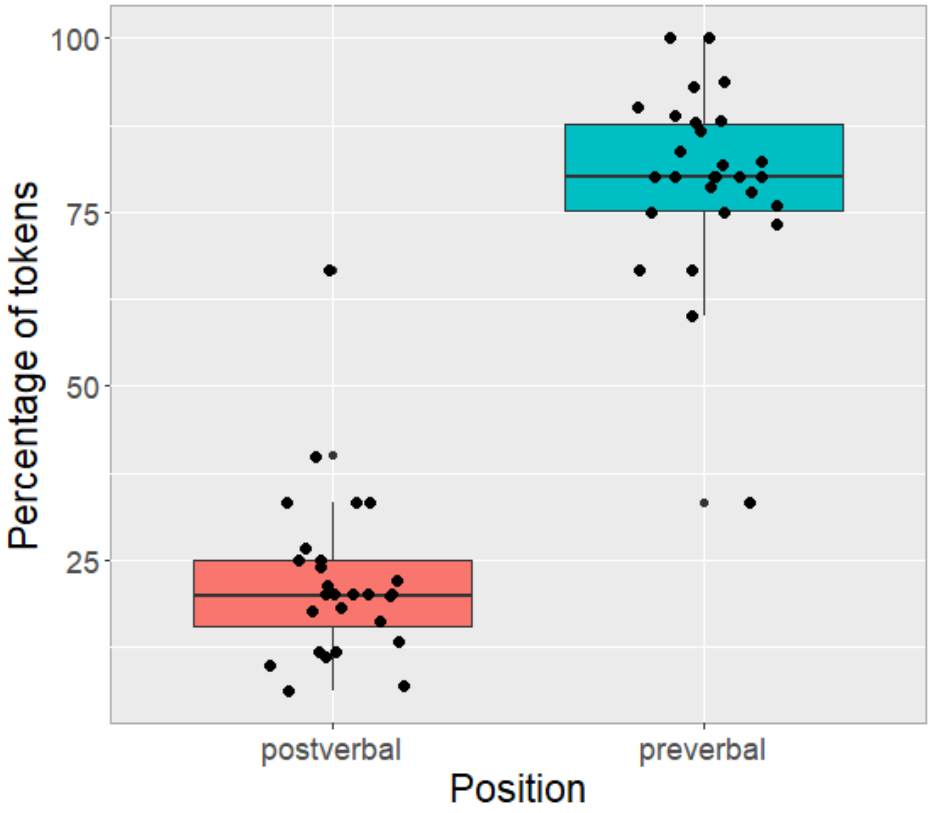

Figure 1: Percentage of tokens in pre- or post-predicate position in Kurmanji

6.1.1 Animacy

Animacy is a categorical independent variable with seven levels in the Kurmanji data. The results of the regression given in Table 1 show that there is no correlation between animacy and the position of the tokens in the Kurmanji data.

6.1.2 Weight

Weight is a discrete numeric variable, based on number of phonological words. Four values are distinguished: one to three phonological words, and more than three (which is coded as "four"). The results of the regression provided in Table 2 show that there is no correlation between the weight of the constituents and the position of the tokens in relation to the verb.

Table 1: Regression table for binomial GLM with the dependent variable Position and the independent variable Animacy in Kurmanji.

fixed effect	β	σ	p-value
(intercept)	-1.65e+01	5.656e+02	0.97 (ns)
Anim-adv	1.28e+01	5.656e+02	0.98 (ns)
Anim-anim	1.36e+01	5.656e+02	0.98 (ns)
Anim-bp	1.35e+01	5.656e+02	0.98 (ns)
Anim-hum	1.45e+01	5.656e+02	0.97 (ns)
Anim-inan	1.56e+01	5.656e+02	0.97 (ns)
Anim-other	-9.53e-09	2.465e+03	1.00 (ns)

Table 2: Regression table for binomial GLM with the dependent variable Position and the independent variable Weight in Kurmanji

fixed effect	β	σ	p-value
(intercept)	-1.26	0.30	< .0001
Weight	-0.15	0.19	0.43 (ns)

6.1.3 Flagging

Flagging is a categorical independent variable with eight levels in the Kurmanji data. The results of the regression analysis presented in Table 3 show that there is an effect of Flagging on the placement of arguments in the pre- or post-predicate position in Kurmanji.

The model showed that arguments flagged with case (see example 10), locational nouns (see example 11), or a preposition together with a locational noun (see example 12) are more likely to be placed in the post-predicate position than in the pre-predicate one.

(10) Northern Kurdish Ankara (Iefremenko 2021a: Y, 629)
kûçik xe *avêt-e* *top-e*
dog oneself throw.PST.3SG-DRCT ball-OBL.F
'The dog threw itself towards the ball.'

Table 3: Regression table for binomial GLM with the dependent variable Position and the independent variable Flagging in Kurmanji

fixed effect	β	σ	p-value	
(intercept)	-2.67	0.36	< .0001	
Flag-case	2.23	0.43	< .0001	***
Flag-circ	-0.09	0.69	0.88	(ns)
Flag-lvc-poss	-11.89	882.74	0.98	(ns)
Flag-postp	0.84	0.57	0.11	(ns)
Flag-prep	-0.06	0.51	0.89	(ns)
Flag-prep-relnoun	1.91	0.58	0.001	**
Flag-relnoun	3.84	0.52	< .0001	***

(11) Northern Kurdish Ankara (Iefremenko 2021a: Y, 626)
çû iii ser rê
go.PST.3SG hm to street
'He went to the street.'

(12) Northern Kurdish Ankara (Iefremenko 2021a: C, 57)
iii di rê da iii jin-ekî kûçik-ê wî ew ê derbas
hm in street CIRC hm woman-INDEF dog-EZ.M her they FUT passing
bû-n-a li_hemberê rê
become.PST-3PL-DRCT across street
'Hm, on the street, (there was) a woman (and) her dog, they were about to cross the street to the opposite side.'

6.1.4 Role

Role is a categorical independent variable with 15 levels in the Kurmanji data. The results of the regression provided in Table 4 show that there is an effect of role on the placement of arguments in the post-predicate position.

Figure 2 shows the percentage of tokens with a particular role in the pre- and post-predicate position in the Kurmanji data, which was calculated in relation to the total number of tokens in the data. The line inside each box indicates the median, while dots represent the percentage of tokens used in pre- or post-predicate position by one speaker.

The model as well as Figure 2 show that goals of motion (see example 13) and caused motion (see example 14) are more likely to be placed in the post-predicate position than in the pre-predicate one.

Table 4: Regression table for binomial GLM with the dependent variable Position and the independent variable Role in Kurmanji

fixed effect	β	σ	p-value	
(intercept)	-2.85	0.59	< .0001	
Role-addr	0.28	1.19	0.8	(ns)
Role-ben	0.9	1.22	0.45	(ns)
Role-com	1.06	0.96	0.27	(ns)
Role-cop	-14.71	907.61	0.98	(ns)
Role-cop-loc	-14.71	3956.1	0.99	(ns)
Role-do	-1.61	1.16	0.16	(ns)
Role-do-def	-14.71	1978.09	0.99	(ns)
Role-goal	3.65	0.63	< .0001	***
Role-goal-c	2.39	0.66	< .0001	***
Role-instr	-14.71	1318.72	0.99	(ns)
Role-loc	-0.15	0.72	0.83	(ns)
Role-other	-0.51	1.17	0.66	(ns)
Role-rec	20.41	3956.10	0.99	(ns)
Role-stim	2.85	1.53	0.06	(ns)

(13) Northern Kurdish Ankara (Iefremenko 2021a: C, 69)
çû ber_bi ereba
go.PST.3SG towards car
'He went towards the car.'

(14) Northern Kurdish Ankara (Iefremenko 2021a: E, 108)
û kûçik jî xe awêt-e top-ê
and dog also oneself throw.PST.3SG-DRCT ball-OBL.F
'And the dog threw itself to the ball.'

Thus, the analysis demonstrated that in Kurmanji goals of motion and caused motion, which are flagged with case, locational noun, or sometimes a preposition and a locational noun, are likely to be placed in post-predicate position. This is in fact what is known from previous literature on the post-predicate position in Kurmanji (Haig 2015; Haig 2019; Gündoğdu 2019).

Regarding the other types of goals, as shown in Figure 2, there are no examples of addressees placed in post-predicate position in my data, and all examples containing an addressee argument are placed in pre-predicate position. The reason

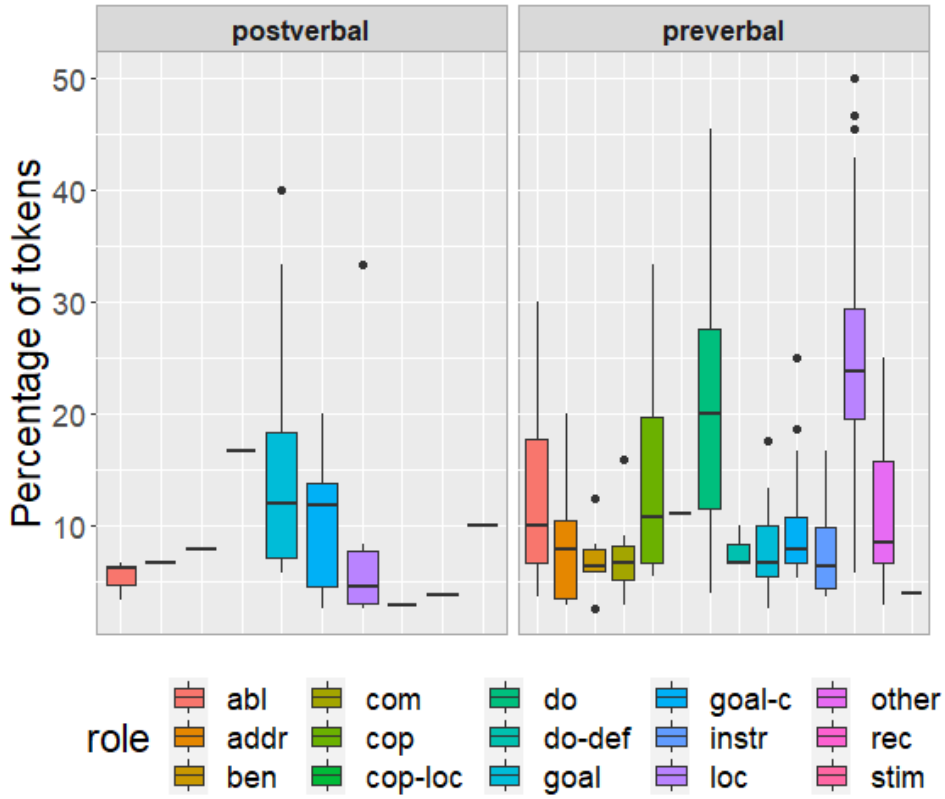

Figure 2: Percentage of tokens with a particular role in the pre- and post-predicate position in Kurmanji.

for this might be flagging: all pre-predicate examples are flagged with a circumposition (example 15) or a postposition (example 16).

(15) Northern Kurdish Ankara (Iefremenko 2021a: F, 134)
mi go e ji te ra bi-bêj-im
1SG.OBL say.PST hm to 2SG.OBL CIRC SUBJ-say.PRS-1SG
'I said, I am telling (this) to you.'

(16) Northern Kurdish Ankara (Iefremenko 2021a: F, 114)
mi go ez te re bêj-im
1SG.OBL say.PST 1SG 2SG.OBL POSTP SUBJ.say.PRS-1SG
'I said, I am telling (this) to you.'

With regard to recipients, there are only two examples in the Kurmanji data, and they are placed post-predicatively (example 17). Presumably, the reason for a small number of recipients in the data is the method: the video shown to the participants did not trigger the use of verbs of transfer.

(17) Northern Kurdish Ankara (Iefremenko 2021a: T, 498)
xeber-ê bi-di-m te
news-OBL SUBJ-give.PRS-1SG 2SG.OBL
'Let me tell (lit. give) you the news.'

Apart from goals of (caused) motion, addressees, and recipients, Figure 2 shows that locations (abbreviated in Figure 2 as "loc" – see example 18) and sources of motion (abbreviated in Figure 2 as "abl" – see example 19) can also be placed in post-predicate position in Kurmanji.

(18) Northern Kurdish Ankara (Iefremenko 2021a: F, 126)
du erebe piştî hevdu di-çû him iii li_ser rê da
two car after each.other PROG-go.PST.3SG both hmm on street CIRC
'Two cars were coming one after another on the street.'

(19) Northern Kurdish Ankara (Iefremenko 2021a: A, 28)
e du erebe di-hat-in ji wî alî
hm two car PROG-come.PST-3PL from his side
'Hm, two cars were coming from his side.'

Another interesting observation is placement of case-flagged goals of motion. As the previous literature (Haig 2015; Haig 2019; Gündoğdu 2019) shows, goals of motion flagged with case are always placed in immediate post-predicate position. Yet there are examples in my data where case-flagged goals of motion are placed in pre-predicate position (see example 20 and 21).

(20) Northern Kurdish Ankara (Iefremenko 2021a: N, 373)
nan û av erd-î ket
bread and water ground-OBL.M fall.PST.3SG
'Bread and water fell on the ground.'

(21) Northern Kurdish Ankara (Iefremenko 2021a: O, 386)
pişte zilam iii gok-ê xwe imm (–) ji dest-ê xwa imm (–)
after man hmm ball-EZ.F own hmm from hand-EZ.M own hmm
erdêk/_erd-ê ket
ground-OBL.M fall.PST.3SG
'Afterwards, the man, hm, his ball, hm, fell from his hands on the ground.'

In total in the data there are four examples with a case-flagged Goal of motion placed in pre-predicate position, and all four examples are actually the same construction[3] *erdê ketin* 'to fall on the ground'. In the contact language Turkish, *erdê ketin* is rendered as *yere düşmek*. In Turkish, *yere düşmek* is a fixed construction. This is also supported by the Turkish data from Kurmanji-Turkish bilinguals in Ankara: there are 10 instances of *yere düşmek*, and in all of them the argument *yere* 'on the ground' is placed in the immediate pre-predicate position. Besides, I have also checked other data collected with the help of the same method, namely Turkish monolingual speakers from Izmir and Eskişehir (Wiese et al. 2021) as well as Kurmanji-Turkish heritage speakers in Berlin, and found that in all the instances of *yere düşmek*, the argument *yere* was placed in the immediate pre-predicate position. Thus, I argue that the stable position of the argument *yere* in Turkish leads to transfer of the whole construction into Kurmanji and as a result, the argument *erdê* is placed pre-predicatively. Besides, in Kurmanji there are other constructions with the verb *ketin*, which are set phrases and where the Goal is always placed in pre-predicate position, e.g., *bi rê ketin* 'to set off (on a journey)'. This might explain the fact why I do not find examples of transfer of constructions with other verbs.

Thus, the analysis of the post-predicate position in Kurmanji of Kurmanji-Turkish bilingual speakers in Ankara showed that Kurmanji is an OV language where goals of motion and caused motion flagged with case, a locational noun, or a preposition used together with a locational noun, are systematically placed in post-predicate position.[4] However, the analysis showed that other elements, such as location and source of motion, can also be placed in post-predicate position in Kurmanji. Apart from this, the data demonstrated that some speakers employ the pre-predicate position for case-flagged goals of motion, which is a non-canonical position for such arguments.

6.2 Turkish

In this section, I will present the results of the analysis of the majority language of the society the speakers live in – which is Turkish.

Figure 3 shows frequency distribution of the coded tokens used in pre- or post-predicate position. The percentage was calculated in relation to the total number of utterances that contain tokens placed pre- or post-predicatively. Each dot represents the mean use of tokens by a speaker. In Figure 3, we can see that in

[3]For a definition of a construction see Croft & Cruse (2004: 257–262).

[4]As for addressees and recipients, there is not enough data in the data set to make conclusions about their position in a sentence in relation to the verb.

general the pre-predicate position of arguments is preferred: only 10% of utterances contain post-predicate elements. However, the figure also shows that there is individual variability: some speakers use up to 25% of post-predicate structures in their narrations.

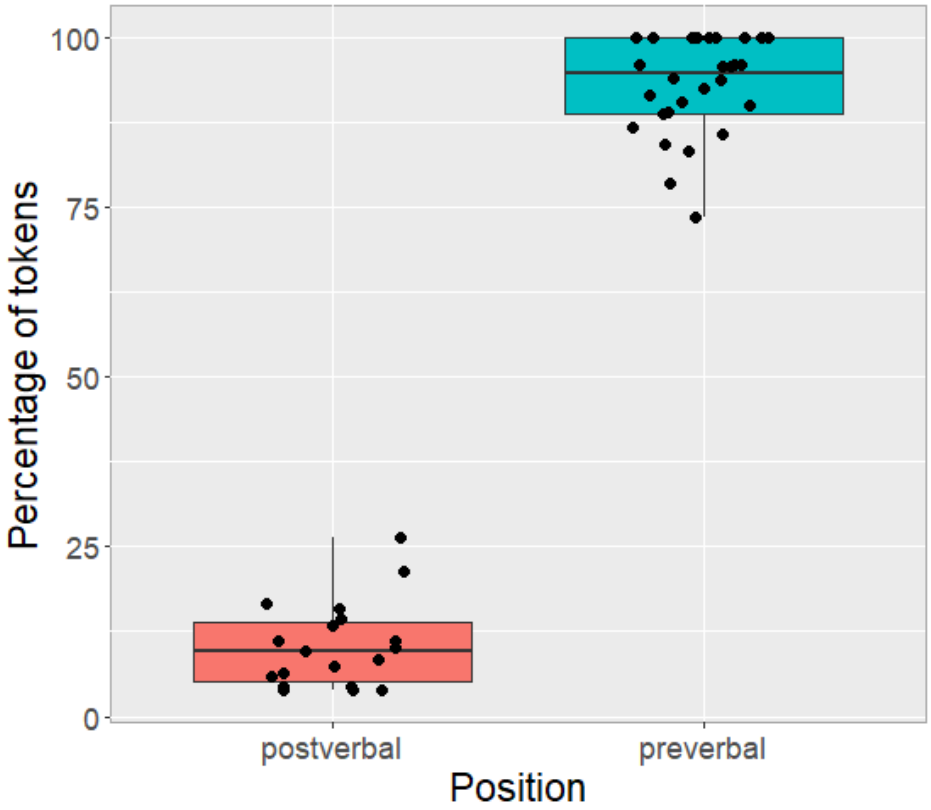

Figure 3: Percentage of tokens in pre- or post-predicate position in Turkish.

Similar to Kurmanji, in Turkish, each token was annotated for animacy, weight, role and flagging (see Haig et al. 2024 [this volume]). To find out whether there is an effect of one of these variables, I ran four binomial generalized linear regression models with the dependent variable Position (preverbal, coded as 0 vs. postverbal, coded as 1) and the independent variables Animacy, Weight, Flagging, and Role.

6.2.1 Animacy

Animacy is a categorical independent variable with six levels in the Turkish data. The results of the regression given in Table 5 show that there is no correlation between Animacy and the Position of the tokens in the Turkish data.

Table 5: Regression table for binomial GLM with the dependent variable Position and the independent variable Animacy in Turkish

fixed effect	β	σ	p-value
(intercept)	-1.757e+01	2.284e+03	0.99 (ns)
Anim-adv	1.512e+01	2.284e+03	0.99 (ns)
Anim-anim	-5.966e-08	2.412e+03	1.00 (ns)
Anim-bp	1.498e+01	2.284e+03	0.99 (ns)
Anim-hum	1.612e+01	2.284e+03	0.99 (ns)
Anim-inan	1.446e+01	2.284e+03	0.99 (ns)

6.2.2 Weight

Weight is a discrete numeric variable, ranging from one to four and more phonological words. The results of the regression provided in Table 6 show that there is a positive effect of weight on the placement of the elements in post-predicate position, meaning that heavier elements (those that consist of three or four phonological words) are more likely to be placed in post-predicate position than those with the weight of one or two phonological words.

Table 6: Regression table for binomial GLM with the dependent variable Position and the independent variable Weight in Turkish

fixed effect	β	σ	p-value
(intercept)	-3.32	0.35	< .0001
Weight	**0.5**	**0.19**	**0.008** **

Thus, for instance, in example (22) the structure *yolun diğer tarafından* 'from the other side of the street' consists of three phonological words and is placed in post-predicate position.

(22) Turkish (Iefremenko 2021b: U, 622)
bi tane araba gel-iyor-du yol-un diğer taraf-ın-dan
one piece car come-PROG-PST.3SG road-GEN other side-POSS-ABL
'There was a car coming from the other side of the street.'

6.2.3 Flagging

Flagging is a categorical independent variable with four levels in the Turkish data. The result of the regression analysis presented in Table 7 shows that there is no effect of flagging on the placement of arguments in pre- or post-predicate position in Turkish.

Table 7: Regression table for binomial GLM with the dependent variable Position and the independent variable Flagging in Turkish

fixed effect	β	σ	p-value
(intercept)	-2.25	0.37	< .0001
Flag-case	-0.38	0.41	0.35 (ns)
Flag-postp	0.05	0.71	0.94 (ns)
Flag-postp-relnoun	-0.5	0.81	0.47 (ns)

6.2.4 Role

Role is a categorical independent variable with 16 levels in the Turkish data. The results of the regression provided in Table 8 show that there is no correlation between the Role of a token and its Position in relation to the verb.

Thus, the analysis of word order in Turkish of Kurmanji-Turkish bilingual speakers showed that the only significant predictor is the weight of the elements: those constituents consisting of three and four phonological words are more likely to be placed in post-predicate position. Unlike Kurmanji, in Turkish such variables as Role or Flagging do not have an effect on the placement of arguments in post-predicate position in my data.

Before proceeding to the discussion of results, I would like to briefly discuss the findings from the Turkish data from Erzurum (and Erzurum Province) (Dogan 2021) that were analysed using the same scheme of the WOWA project. From the first look at the normalized numbers of different roles of the tokens, it is evident that the data sets from Ankara and Erzurum differ in the distribution of post-predicate arguments. For this reason, I have conducted similar analyses for the

Table 8: Regression table for binomial GLM with the dependent variable Position and the independent variable Role in Turkish

fixed effect	β	σ	p-value
(intercept)	-2.68	0.46	< .0001
Role-addr	1.38	0.79	0.08 (ns)
Role-ben	0.93	0.71	0.19 (ns)
Role-com	1.29	0.79	0.10 (ns)
Role-cop	-15.88	3261.3	0.99 (ns)
Role-cop-loc	-15.88	4612.2	0.99 (ns)
Role-do	-0.67	0.74	0.36 (ns)
Role-do-def	-15.88	1171.5	0.98 (ns)
Role-goal	0.16	0.56	0.77 (ns)
Role-goal-c	0.19	1.13	0.86 (ns)
Role-instr	-15.88	1966.64	0.99 (ns)
Role-loc	-0.3	0.65	0.63 (ns)
Role-other	0.65	0.61	0.28 (ns)
Role-poss	-15.88	6522.6	0.99 (ns)
Role-rec	-15.88	4621.2	0.99 (ns)
Role-stim	21.24	6522.6	0.99 (ns)

Erzurum data as I did for the data from Ankara (see Appendix for the results of the regression analyses).

First of all, the Turkish data from Erzurum is substantially pre-predicate (with around 12% of the utterances containing post-predicate positions), that is very similar to the Turkish data collected in Ankara. However, the regression analyses run on the Turkish data from Erzurum showed that there is no effect of weight, meaning that regardless of the number of phonological words that tokens consist of, they can be placed in the pre- or post-predicate position with no preference. But the two variables that have a positive effect on the employment of the post-predicate position are Role and Flagging. Namely, the model with Role as an independent variable showed that goals of motion are often placed in post-predicate position in Turkish. At the same time, the regression model with Flagging as an independent variable demonstrated that tokens flagged with case or a postpositional relational noun are likely to be placed in post-predicate position.

Looking back at the results of the Kurmanji data from Ankara, we see that these two variables, Role and Flagging, were also found to be significant. Namely,

the analysis of the Kurmanji data showed that goals of motion, and tokens flagged with case or a locational noun (sometimes in a combination with a preposition) are likely to be placed in post-predicate position in Kurmanji. Hence, similar factors influence the employment of the post-predicate position in the data sets from Kurmanji in Ankara and Turkish in Erzurum, while there are no common factors in the Turkish data sets collected in Ankara and Erzurum.

Nevertheless, it must be acknowledged that the two Turkish data sets (from Ankara and Erzurum) are not exactly comparable. First, the Turkish data from Erzurum come from three speakers, whereby one speaker is a young adult and two speakers are in their sixties, and their educational status is not known, unlike the speakers in Ankara who are all educated young adults. Second, the data were collected in the nineties, nearly 30 years prior to the data collection in Ankara. But what is most important is that it is not known whether the speakers from Erzurum are in fact bilingual in Turkish and Kurmanji, or if they are Turkish monolingual speakers. But what definitely differentiates the speakers in Erzurum from the speakers in Ankara is the number of Kurmanji speakers in the community in general and the societal status of Turkish. Even though Turkish certainly remains the language of formal contexts (such as education, business, etc.), Kurmanji is used more extensively in informal contexts, compared to Ankara (though we cannot compute the index of language use for the Erzurum data due to the absence of metadata for the speakers).

7 Discussion

In the study, I investigated word order, namely the post-predicate domain, in Turkish and Kurmanji that have been in a long-lasting contact with each other. Both languages are OV, but each of them employs the post-predicate position in a different way. While in Turkish the word order is determined by information structural requirements and the post-predicate position is reserved for background information, in Kurmanji the post-predicate position is the position for Goal arguments, particularly those flagged with case. Taking into account the different conditions encoded in WOWA, I investigated whether there is a structural convergence in word order in Turkish and in Kurmanji in Turkey.

First of all, the quantitative analysis showed that both languages predominantly place arguments in the pre-predicate position: with around 20% of the utterances in Kurmanji and 10% of the utterances in Turkish being post-predicate. Thus, both languages retain OV word order.

For Kurmanji, the regression analyses showed that the employment of the post-predicate position depends on the semantic role of the elements in the respective

clause and their flagging. Namely, goals of motion and caused motion are likely to be placed in post-predicate position, and this is in line with what has been described in previous research on Kurmanji word order (Haig 2015; Haig 2019; Gündoğdu 2019). As for addressees and recipients, the regression analysis did not show that there is a tendency to place them post-predicatively in my data. The main reason for this is different types of flagging: while the post-predicatively placed goals are flagged with case, all the pre-predicate examples of addressees and recipients are flagged with a circumposition or a postposition. Besides, in general I have not found many examples of addressees and recipients in my data possibly because the video shown to the participants did not trigger the use of verbs of speech and verbs of transfer. In fact, another variable that showed an effect on the placement of tokens in post-predicate position in Kurmanji is flagging. That is, tokens that are flagged with case, locational noun, or a preposition together with a locational noun are likely to be placed in post-predicate position. Thus, so far the analysis showed that the Kurmanji data is in line with most of the previous studies on word order in Kurmanji: Goal arguments flagged with case or a locational noun are placed post-predicatively.

At the same time, the qualitative analysis of the Kurmanji data showed that, albeit infrequently, such elements as location and source of motion are placed in the post-predicate position. Nevertheless, it is not clear whether the use of these elements in post-predicate position is language contact-induced, for two reasons: First, I did not find enough examples of this kind, and second, the analysed data is spoken and therefore is full with afterthoughts and self-repairings, and it is not fully clear whether certain post-predicative elements are in fact afterthoughts or self-corrections.

Another interesting observation from the Kurmanji data is that there are examples where case-flagged goals of motion are placed in pre-predicate position. Remember that in Section 2, I emphasized that case-flagged goals are licensed in post-predicate position. However, my data demonstrate that such constructions in Kurmanji can also be pre-predicate, but this concerns only one construction *erdê ketin* 'to fall on the ground', whereby *erdê* 'on the ground' as a case-flagged Goal argument is placed before the verb *ketin* 'to fall'. Interestingly, the same construction in the contact language Turkish is rendered as *yere düşmek*, whereby the Goal argument *yere* 'on the ground' is always placed in pre-predicate position. I argue that the pre-predicate position of the argument *erdê* occurs due to transfer of the whole construction from the dominant language Turkish. Besides, the reason why transfer occurs only with this particular construction might lie in the verb itself: in Kurmanji there are other set phrases with the verb *ketin* where

the Goal is always placed in pre-predicate position, e.g., *bi rê ketin* 'to set off (on a journey)'.

As for the post-predicate position in Turkish, the analysis showed that the only significant predictor of the employment of the post-predicate position, among those encoded in WOWA, is the weight of tokens, that is, tokens that consist of three and four phonological words are likely to be placed in post-predicate position. The placement of heavier constituents in post-predicate position is generally a phenomenon typical of spoken language. As Schroeder (1995) emphasizes, in spoken discourse information is conveyed in smaller chunks to make it more accessible to the hearer and the post-predicate position allows the hearer to keep track of the topical development and the deictic framework in which a predication takes place. On the other hand, the analysis of the Turkish data from Erzurum, where the contact with Kurmanji is more intense and Kurmanji is present in more spheres of life compared to Ankara, showed that Turkish speakers in Erzurum tend to place goals of motion, particularly those flagged with case and relational noun, in post-predicate position. At the same time, weight of constituents did not have an effect on the placement of constituents in the data from Erzurum.

In sum, the results of the analysis based on the WOWA encoding scheme show a minimal degree of convergence between the two languages of Kurmanji-Turkish speakers in Ankara. In Kurmanji semantic role of constituents and their flagging are determining factors in the placement of the constituents in relation to the verb, which is in line with the previous research (Haig 2015; Haig 2019; Gündoğdu 2019). In Turkish weight was proven to be a significant factor: longer constituents are more likely to be placed in the post-predicate position. The result that points to possible signs of an ongoing contact-induced language change is the changes in the word order of particular constructions. However, such changes are observed only in the minority language Kurmanji and not the majority language Turkish. Thus, such results point to the effect of societal status of the languages on the direction of the language change: a more prestigious language Turkish influences a less prestigious Kurmanji in Ankara.

At the same time, the analysis of the Erzurum Turkish data suggests that another social factor – the intensity of contact – has an impact on the occurrence of changes in word order of the languages in contact. Unlike Turkish in Ankara, in Erzurum Turkish semantic role and flagging have an effect on the employment of the post-predicate position. Hence, when Kurmanji is present in more spheres of life and the community size is bigger, changes also happen in the majority language Turkish. Even though it is not clear whether the speakers in the Erzurum data set are in fact bilinguals, the predictors for the placement of tokens post-predicatively are the same as in Kurmanji.

Finally, it is important to emphasize that information structure was not encoded in WOWA and, consequently, has not been analyzed in this chapter. Therefore, it remains unclear whether information structural constraints are loosened in the Turkish of bilinguals and whether information structure plays a role in the word order of Kurmanji as a result of contact with Turkish. This presents a limitation of the current study; however, the role of information structure on word order in Turkish and Kurmanji of the same speakers is discussed in another study by Iefremenko (Submitted).

Abbreviations

ABL	ablative	NEG	negation
ACC	accusative	OBL	oblique
CIRC	circumposition	PL	plural
DRCT	directional	POSS	possessive
EZ	ezafe	POSTP	postposition
F	feminine	PROG	progressive
FUT	future	PRS	present
GEN	genitive	PST	past
INDEF	indefinite	SG	singular
M	masculine	SUBJ	subjunctive

Acknowledgements

I would like to thank all members of the project "Post-predicate Elements in Iranian: Inheritance, Contact, and Information Structure" and the participants of the project's workshops for their invaluable feedback and comments. I also extend my gratitude to Prof. Dr. Christoph Schroeder, the anonymous reviewer and the editors of this volume for their helpful and highly constructive suggestions. Special thanks to Dr. Shuan Karim for the assistance with LaTeX. The research was supported through funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) for the Research Unit "Emerging Grammars in Language Contact Situations", project P9 (313607803) and the Alexander-von-Humboldt-Stiftung.

References

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

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67. 1–48.

Chafe, Wallace. 1985. Linguistic differences produced by differences between speaking and writing. In David Richard Olson, Nancy Torrance & Angela Hildyard (eds.), *Literacy, language and learning: The nature and consequences of reading and writing*, 105–123. Cambridge: Cambridge University Press.

Croft, William & D. Alan Cruse. 2004. *Cognitive linguistics*. Cambridge: Cambridge University Press.

Csató, Éva Ágnes. 2002. Karaim: A high-copying language. In Mari C. Jones & Edith Esch (eds.), *Language change. The interplay of internal, external and extralinguistic factors*, 315–328. Berlin: De Gruyter Mouton.

Dogan, Mahîr. 2021. Oghuz (Erzurum). In Geoffrey Haig, Donald Stilo, Nils N. Schiborr & Mahîr Dogan (eds.), *WOWA — Word Order in Western Asia: A spoken-language-based corpus for investigating areal effects in word order variation*. (accessed on October 20, 2022). Bamberg: University of Bamberg. multicast.aspra.uni-bamberg.de/resources/wowa.

Doğruöz, A. Seza & Ad Backus. 2007. Postverbal elements in immigrant Turkish: Evidence of change? *International Journal of Bilingualism* 11(2). 185–220. DOI: 10.1177/13670069070110020301.

Erguvanlı, Eser. 1984. *The function of word order in Turkish grammar*. Berkeley: University of California Press.

Gündoğdu, Songül. 2019. Asymmetries in Kurmanji morphosyntax. In Songül Gündoğdu, Ergin Öpengin, Geoffrey Haig & Erik Anonby (eds.), *Current issues in Kurdish linguistics*, 93–115. Bamberg: University of Bamberg Press.

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

Haig, Geoffrey. 2019. Northern Kurdish (Kurmanjî). In Geoffrey Haig & Geoffrey Khan (eds.), *The languages and linguistics of Western Asia: An areal perspective*, 106–158. Berlin: De Gruyter Mouton.

Haig, Geoffrey. 2022. Post-predicate constituents in Kurdish. In Yaron Matras, Geoffrey Haig & Ergin Öpengin (eds.), *Structural and typological variation in the dialects of Kurdish*, 335–377. Cham: Springer International Publishing. DOI: 10.1007/978-3-030-78837-7_8.

Haig, Geoffrey & Yaron Matras. 2002. Kurdish linguistics: A brief overview. *STUF-Language Typology and Universals* 55(1). 3–14.

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

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

Heine, Bernd. 2008. Contact-induced word order change without word order change. In Peter Siemund & Noemi Kintana (eds.), *Language contact and contact languages*, 33–60. Amsterdam: Benjamins. DOI: 10.1075/hsm.7.04hei.

Iefremenko, Kateryna. 2021a. Kurdish (Northern, Ankara). 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.* (accessed on October 20, 2022). Bamberg: University of Bamberg. multicast.aspra.uni-bamberg.de/resources/wowa.

Iefremenko, Kateryna. 2021b. Oghuz (Ankara). 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. multicast.aspra.uni-bamberg.de/resources/wowa/#turkic.

Iefremenko, Kateryna. Submitted. *Word order in Turkish and Kurmanji Kurdish in language contacts.* Potsdam: University of Potsdam. (Doctoral dissertation).

Iefremenko, Kateryna, Cem Keskin & Christoph Schroeder. 2023. The post-predicate position in Turkish in language contact situations: How resistant to change? Paper presented at the *45th Annual Meeting of the German Linguistic Society* (DGfS) in Cologne, Germany 7-10 March 2023.

Johanson, Lars. 2002. *Structural factors in Turkic language contacts.* Richmond/Surrey: Curzon Press.

Keskin, Cem. 2023. On the directionality of the Balkan Turkic verb phrase: Variationist and theoretical perspectives. *Languages* 8(1). 2. DOI: 10.3390/languages8010002.

Kornfilt, Jaklin. 1997. *Turkish.* London/New York: Routledge.

Matras, Yaron et al. 2016. *The dialects of Kurdish.* University of Manchester. https://www.kratylos.org/~raphael/kurdish/database/.

Menz, Astrid. 1999. *Gagausische Syntax: Eine Studie zum kontaktinduzierten Sprachwandel*. Wiesbaden: Harrassowitz.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna. https://www.R-project.org/.

Schroeder, Christoph. 1995. Post-predicate arguments in Turkish. *Sprachtypologie und Universalienforschung* 48. 204–218.

Schroeder, Christoph, Kateryna Iefremenko & Mehmet Öncü. In press. The postverbal position in heritage Turkish. A comparative perspective with a focus on non-clausal elements. In Zeynep Kalkavan-Aydın & Yazgül Şimşek (eds.), *Zweisprachigkeit Deutsch-Türkisch*. Münster: Waxmann.

Sheyholislami, Jaffer. 2015. Language varieties of the Kurds. In Wolfgang Taucher, Mathias Vogl & Peter Webinger (eds.), *The Kurds: History–religion–language-politics*, 30–51. Vienna: Austrian Federal Ministry of the Interior.

Thomason, Sarah G. 2001. *Language contact*. Edinburgh: Edinburgh University Press.

Thomason, Sarah G. & Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy McGowan D'Agostino McGowan, Roman François, Garrett Grolemund, Alex Hayes, Lionel Henry, Jim Hester, Max Kuhn, Thomas Lin Pedersen, Evan Miller, Stephan Milton Bache, Kirill Müller, Jeroen Ooms, David Robinson, Dana Paige Seidel, Vitalie Spinu, Kohske Takahashi, Davis Vaughan, Claus Wilke, Kara Woo & Hiroaki Yutani. 2019. Welcome to the Tidyverse. *Journal of Open Source Software* 4(43). 1686. DOI: 10.21105/joss.01686.

Wiese, Heike. 2020. Language situations: A method for capturing variation within speakers' repertoires. In Yoshiyuki Asahi (ed.), *Methods in dialectology XVI*, 105–117. Frankfurt a. M.: Lang.

Wiese, Heike, Artemis Alexiadou, Shanley Allen, Oliver Bunk, Natalia Gagarina, Kateryna Iefremenko, Esther Jahns, Martin Klotz, Thomas Krause, Annika Labrenz, et al. 2021. *RUEG Corpus*. Available online. DOI: 10.5281/ZENODO.3236068.

Yağmur, Kutlay. 2001. Turkish and other languages in Turkey. In Guus Extra & Durk Gorter (eds.), *The other languages of Europe. Demographic, sociolinguistic and educational perspectives*, 407–426. Clevedon: Multilingual Matters.

Appendix A Erzurum Turkish (Turkic, Dogan 2021)

Table 9: **Animacy:** Regression table for binomial GLM with the dependent variable Position and the independent variable Animacy in Turkish spoken in Erzurum

fixed effect	β	σ	p-value	
(intercept)	-1.557e+01	1.029e+03	0.98	(ns)
Anim-adv	1.439e+01	1.029e+03	0.98	(ns)
Anim-anim	1.279e+01	1.029e+03	0.99	(ns)
Anim-bp	1.339e+01	1.029e+03	0.99	(ns)
Anim-hum	1.331e+01	1.029e+03	0.99	(ns)
Anim-inan	1.383e+01	1.029e+03	0.98	(ns)
Anim-other	2.205e-08	1.188e+03	1.00	(ns)

Table 10: **Weight:** Regression table for binomial GLM with the dependent variable Position and the independent variable Weight in Turkish spoken in Erzurum

fixed effect	β	σ	p-value	
(intercept)	-2.10	0.29	< .0001	
Weight	0.16	0.19	0.39	(ns)

Table 11: **Flagging:** Regression table for binomial GLM with the dependent variable Position and the independent variable Flagging in Turkish spoken in Erzurum

fixed effect	β	σ	p-value	
(intercept)	-3.20	0.45	< .0001	***
Flag-case	1.46	0.47	0.02	**
Flag-postp-relnoun	2.25	0.63	0.0003	***

Table 12: **Role**: Regression table for binomial GLM with the dependent variable Position and the independent variable Role in Turkish spoken in Erzurum

fixed effect	β	σ	p-value	
(intercept)	-2.30	6.05	< .0001	***
Role-addr	5.129e-02	9.588e-01	0.95	(ns)
Role-becm	-9.531e-02	1.207e+00	0.93	
Role-ben	1.386e+00	8.466e-01	0.1	(ns)
Role-cop	-1.526e+01	9.890e+02	0.98	(ns)
Role-do	-7.793e-01	7.588e-01	0.3	(ns)
Role-do-def	-2.632e-02	6.632e-01	0.96	(ns)
Role-goal	2.042e+00	6.438e-01	0.001	**
Role-goal-c	1.022e+00	7.032e-01	0.14	(ns)
Role-instr	-1.526e+01	1.251e+03	0.99	(ns)
Role-loc	-6.592e-01	1.251e+03	0.43	(ns)
Role-other	1.372e-14	9.574e-01	1.00	(ns)
Role-poss	-1.526e+01	2.797e+03	0.99	(ns)
Role-rec	-3.365e-01	9.499e-01	0.72	(ns)
Role-stim	-1.526e+01	3.956e+03	0.99	(ns)