A unified approach to the study of language contact:

Cross-language priming and change in adjective/noun order

Evangelia Adamou, CNRS (French National Centre for Scientific Research), France

Quentin Feltgen, INSERM and Université Sorbonne Nouvelle, France

Cristian Padure, University of Bucharest, Romania

Abstract

Aims and Objectives/Purpose/Research Questions: The connection between language contact and the bilingual speaker goes back to foundational authors in the field of contact linguistics. Yet there is very little work that combines these two levels in a single study. In this paper, we propose a unified approach to language contact by testing the role of cross-language priming on contact-induced change at the level of complex NPs.

Design/Methodology/Approach: We conducted three studies with different types of data. In Study 1, we analyse the Romani Morpho-Syntax database to identify word order preferences in Romani dialects from different countries. In Study 2, we examine a corpus of interviews in Romani from Romania. In Study 3, we conduct an experiment to test short-term priming in adjective/noun order from Romanian to Romani and within Romani.

Data and Analysis: In Study 1, we examine the word order in approximately 3,000 NPs from 119 Romani speakers. In Study 2, we analyse a speech corpus of 9,400 words from four elderly Romani-Romanian bilinguals. In Study 3, 90 Romani-Romanian bilinguals participated in a priming experiment. We used multinomial mixed-effects logistic regression, Bayesian models and Random Forests to analyse the experimental results.

Findings/Conclusions: Study 1 shows that Romani speakers from Romania stand out for their frequent use of postnominal adjectives. Study 2 confirms these uses in free speech. Study 3 reveals significant cross-language priming effects, whereby speakers favour the use of DET-N-ADJ order in Romani immediately following a N_{DET}-ADJ sentence read in Romanian.

Originality: Our study is the first to demonstrate cross-language priming effects in adjective/noun order.

Significance/Implications: We illustrate a unified approach to language contact by introducing theoretical and methodological advances from the field of bilingualism into the study of contact-induced change.

1

Introduction

The goal of this paper is to propose an integrated approach to language contact by bringing together the effects of bilingualism that bear the seeds of potential language change, with the outcomes of these effects once many speakers start using them systematically in their interactions. The connection between language contact and the bilingual speaker goes back to foundational authors in the field of contact linguistics such as Weinreich (1953) and Haugen (1953), and is widely accepted in contemporary contact linguistics (for some recent collective volumes in this direction see Kootstra & Muysken, 2017; Zenner et al., 2019; Adamou & Matras, 2021). Yet, very few researchers investigate these two levels jointly.

To achieve this overarching goal, we test the role of short-term implicit priming in ongoing contact-induced change following recent studies that single out cross-language priming as a key mechanism in language contact (see Torres Cacoullos & Travis, 2018 based on free-speech corpus data; Kootstra & Şahin, 2018 based on experimental data; also see Loebell & Bock, 2003 for the theoretical proposal). The novelty of our approach lies in the combination of methodologies from the field of bilingualism and the field of language contact, using data from a dialectological database (Study 1), a conversational oral corpus (Study 2), and a psycholinguistic experiment (Study 3). The cross-dialectal data allow us to document language change in a diachronic perspective, the corpus data to delve into naturalistic exchanges, and the experimental data to measure priming in controlled conditions.

Our case study is word order in complex NPs among bilingual speakers of Romani, an Indic language, and Romanian, a Romance language. Romani offers a rare opportunity to investigate the effects of language contact as it is a language in long-term and intense contact with a variety of languages. The reason for choosing this pair of languages in particular is that they exhibit different word order in NPs consisting of a determiner, a noun, and an adjective: the unmarked inherited Romani order for descriptive adjectives such as 'big', 'good', or 'red' is ADJ-N (Adamou & Matras, 2020) and the unmarked Romanian order is N-ADJ (Marchis & Alexiadou, 2007; Dryer, 2013). In addition to the unmarked ADJ-N order in Romani, researchers have reported that the adjectives can also follow the noun, generally in marked constructions associated to information structure either as afterthoughts (Matras, 2002: 167) or for contrastive focus (Tirard, 2017). The assumption in the Romani literature to account for

2

the intensification of N-ADJ patterns is that they reflect the order preferences of the main contact languages (see among others Igla, 1996; Tirard, 2017; Adamou & Matras, 2020). However, to our knowledge, detailed studies of the constraints of use in these Romani varieties are not available.

Before introducing our studies, let us start by providing some theoretical and linguistic background on the topic.

Priming: a mechanism in contact-induced structural change?

'Priming' is the observation that processing one stimulus (the prime) unconsciously affects the processing of a subsequent stimulus (the target). In theories that consider processing to provide indirect evidence for the nature of linguistic representations, the priming data suggest that structures of various types (syntactic, word order, semantic/pragmatic, and phonological) may share some aspects of their representation in speakers' minds.

In the present paper, we focus on structural priming where the activation of one structure affects the production or the comprehension of a subsequent structure. For example, in standard priming experiments participants are asked to repeat sentences and describe pictures (e.g., Bock, 1986). These studies overall find that when participants repeat a passive sentence like 'a compromise is being suggested by the chairman', they are 'primed' by this structure. In practice, this means that they are more likely to produce the same structure, in this case passive, when describing an unrelated upcoming picture that allows for this structure to be used, as opposed to when they have just repeated an active sentence like 'the chairman is suggesting a compromise'. Researchers have also shown that the strength of the priming effect differs depending on various factors. First, the frequency of a structure matters but it is unclear in what way. Although frequent structures may have a considerable effect overall, less frequent structures, like passives in English, appear to be more sensitive to priming, an effect known as the 'inverse-preference effect' (see among others Ferreira & Bock, 2006). In addition, it has been shown that structural priming is strengthened when there is lexical overlap in the critical content words (verbs, nouns, and adjectives) between the prime and the target; this phenomenon is dubbed 'lexical boost' (first noted in Pickering & Branigan, 1998). Moreover, even though studies have shown that the priming effect is strongest when there is no temporal delay between prime and target, it appears that structural priming can also persist over time. For example, Gries (2005) analysed large English free-speech corpora and demonstrated that syntactic priming is persistent across several utterances.

3

The structures for which priming effects have been documented vary, starting with classic studies on active and passive constructions and alternative dative constructions (Bock, 1986; Bock & Loebell, 1990; also, in the variationist frame see Weiner & Labov, 1983), verb-participle order (Hartsuiker & Westenberg, 2000), relative clause attachment (high or low) (Scheepers, 2003), and transitive and intransitive clauses (Gompel, Arai, & Pearson, 2012). Studies on noun phrases are relatively rarer, although researchers have tested NP modifier type (adjective vs. relative clause) (Cleland & Pickering, 2003), the order of two adjectives relative to one another (Goudbeek & Krahmer, 2012), and genitive constructions (Bernolet, Hartsuiker, & Pickering, 2012). To our knowledge, no study has tested adjective/noun order preferences, which is the focus of our paper.

In addition to studies on monolinguals, researchers provide compelling evidence for the existence of cross-language structural priming in these well-documented structures (Loebell & Bock, 2003 for dative constructions). A recent meta-analytic study confirms the presence of structural priming between the two languages of bilinguals although it is less robust than among monolinguals (Mahowald et al., 2016). The authors conclude that cross-language priming might benefit less from the 'lexical boost'. Indeed, while priming in a single language relies on the same lexical items, priming from one language to another involves access to different lexical items, so-called 'translation equivalents' or 'cognates' when they share the same form. This is not to say that translation equivalents are not at all related. For example, van Hell and De Groot (1998) find that translation equivalents benefit from facilitation effects even if to a lesser extent than cognates. Other studies reveal that proficiency might also modulate the impact of the priming effect, which appears to be more robust among less proficient L2 speakers than among L1 speakers, suggesting that learners of a second language might rely more strongly on item-specific representations (Kim & McDonough, 2007; Bernolet, Hartsuiker, & Pickering, 2013). This means that early bilinguals may respond differently to cross-language priming effects.

Yet, to date, studies of priming among early bilinguals are rare. Torres Cacoullos and Travis (2018) is one exception. Based on the analysis of a free-speech corpus from Spanish-English bilinguals residing in New Mexico, US, Torres Cacoullos and Travis find evidence for priming in conversations. Consistent with the literature on priming, they find cross-language priming (e.g., *I* primes *yo*) to be less robust than within-language priming (e.g., *yo* primes *yo*). Frequency of use also plays a role as the use of null Spanish pronouns favours the use of null Spanish pronouns, but the use of English null pronouns, which are very rare, does not prime Spanish null pronouns. Travis et al. (2017) also note that priming effects may

4

depend on co-occurring contextual features and constructions that can be lexically-specific or schematic. Moreover, using the experimental priming paradigm, Kootstra and Şahin (2018) test dative structures with Papiamento-Dutch bilinguals living in the Netherlands and in Aruba. They find that speakers who live in the Netherlands and are in daily contact with Dutch exhibit higher levels of priming by Dutch dative structures than those who live in Aruba and have less intense contact with Dutch. In sum, bilingual experience matters.

In this paper, we seek to add to this line of research by investigating another lesser-studied bilingual population, that of Romani-Romanian speakers residing in Romania.

Romani spoken in Romania

Romani is an Indic language of the Indo-European stock and has been spoken in Europe for more than ten centuries. The presence of Roma in Romania goes back to the late fourteenth century when Roma from the Byzantine Empire fled the wars. At present, Romania hosts the largest population of Romani speakers in Europe, with estimates ranging from 250,000 up to more than two million speakers. In the north, the Hungarian Roma and the Gabor speak Central Romani dialects. In the south, the Ursari and the Spoitori speak Balkan Romani dialects. Finally, living in various areas of Romania, the Kalderaš speak Vlax dialects (Matras, 2013). Virtually all Roma in Romania speak Romanian as they live and interact with speakers of Romanian in their everyday life. Although Romanian is the language of instruction, Romani is taught at all levels, from pre-school to university since the 1990s (Sarău, 2017). In 2009, more than 31,000 pupils attended these classes. Educational materials use a standard variety of Romani, based on Kalderaš, and a standardized alphabet, though many Roma rely on the Romanian script in informal writing.

Study 1: The Romani Morpho-Syntax Database

In a bottom up approach, we investigate the order of the adjective with respect to the noun across Romani dialects using the Romani Morpho-Syntax (RMS) database.[1] Since we lack diachronic written evidence for Romani, we compare present-day data and formulate hypotheses about what factors may account for the observed differences (see Elšík & Matras, 2006 for a detailed exposition of the methodology). If, as Adamou and Matras (2020) suggest, the preferred order of the contact language strengthens the use of (DET-)N-ADJ in Romani, this should be evidenced in a cross-dialectal sample like the RMS.

---

[1] Accessed at http://romani.humanities.manchester.ac.uk/rms/. Viktor Elšík and Yaron Matras 2001–present.

Methodology

The Romani Morpho-Syntax (RMS) database. The RMS database includes data from more than 100 speakers collected through the RMS questionnaire comprising lexicon, sentences, and verb paradigms that speakers translate into Romani from one of their main contact languages (1,061 questions in total). The recordings of their responses as well as the transcriptions are available online. The responses to the questionnaires are associated to a database where they serve to illustrate the main morpho-syntactic phenomena. A 'sample', identified by a country code and a number, corresponds to the responses of an individual speaker. The metadata available for each sample offer additional information about the current and past contact languages of the speaker but not on the language of the questionnaire.

Coding. First, we searched for information on the use of postposed adjectives in the database. As this information was available for only 37 samples, we manually searched all available samples in the database, arriving at a total of 119 samples/speakers. For each sample, we coded adjective-noun order in 25 sentences, with or without a determiner, resulting in a total of 2,975 occurrences.

Results

Analysis of the occurrences in the RMS samples shows that postnominal adjectives are frequent in some countries and are rare or absent in others (see Figure 1). The major division is between samples with less than 20% use of (DET-)N-ADJ[2] order and samples with approximately 40% to 70% of such use. It appears that adjective/noun order preferences are overall not categorical, but it is also apparent that Romani dialects exploit the variability in adjective placement to different degrees. Where do these differences come from knowing that the same adjectives in the same contexts were tested for all speakers? Do these differences result from language-internal processes of change that took place differently in different dialects throughout the centuries or do they correlate to some extent with the adjective order in the contact languages?

---

[2] We use (DET), in parentheses, because some Romani dialects show reduction or loss of the definite articles. Adamou and Matras (2021) note that these changes occur in those dialects that are in contact with languages that have no definite articles (e.g., Polska Roma in contact with Polish, Russian Lovari in contact with Russian, Finnish Romani in contact with Finnish). Indefinite articles are also being lost in some Romani dialects when the contact languages have no indefinite articles (e.g., Polish and Russian Romani).

When looking more closely at the results it appears that the Romani samples with less than 20% use of postnominal adjectives come from Finland, North Macedonia, Bulgaria, Hungary, Serbia, Russia, Slovenia, Croatia, Slovakia, Poland, Ukraine, Lithuania, Estonia, Albania, Greece, and Latvia. Based on the information in the metadata, it appears that the speakers are currently in contact with languages with a dominant ADJ-N order. This is the case of all the Slavic languages (e.g., Macedonian, Bulgarian, Serbian, Russian, Czech, Polish, Ukrainian), and the Finnic (Finnish and Estonian), Ugric (Hungarian), Baltic (Lithuanian and Latvian), and Germanic languages (German and Swedish), as well as Turkish and Greek (see Dryer, 2013). Yet, despite the absence of N-ADJ order in these languages, some speakers used the (DET-)N-ADJ order in their responses. One possibility is of course that (DET-)N-ADJ order is used for afterthoughts (Matras, 2002), independent from language contact. Another possibility is that the (DET-)N-ADJ order may stem from the influence of past contact languages as well as diffusion through contact with speakers of dialects that exhibit this feature (e.g., the speaker in the Czech Republic speaks a Vlax dialect influenced by long-term contact with Romanian).
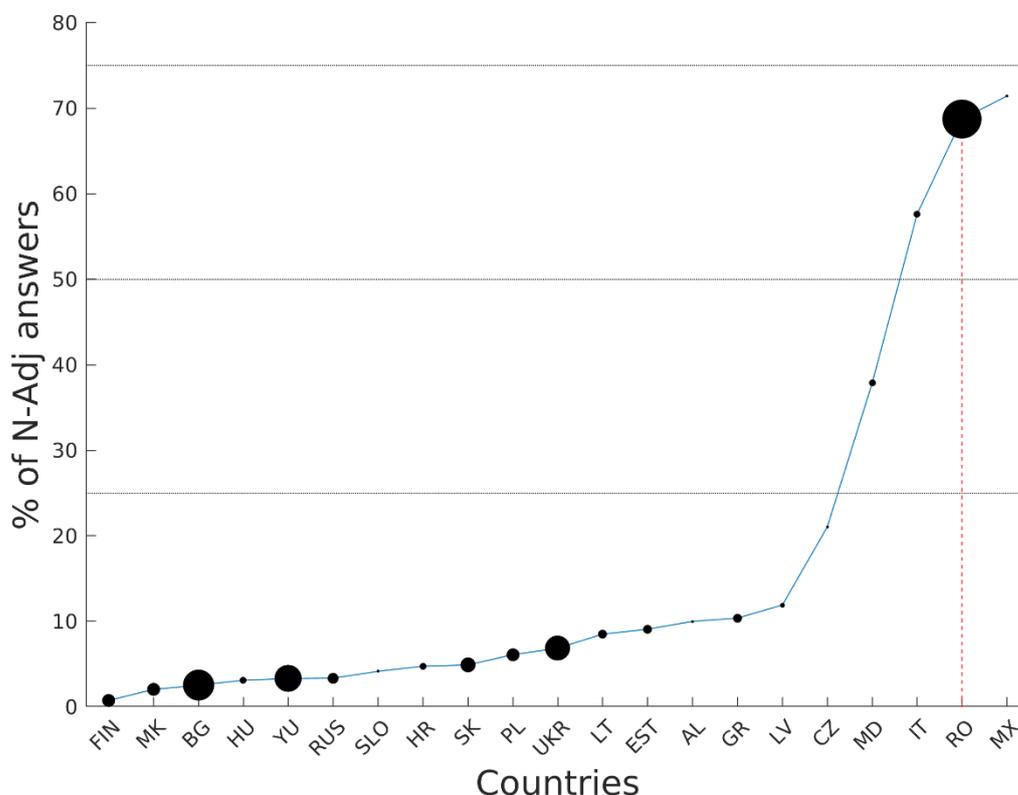

Figure 1. Percentage of all (DET-)N-ADJ responses in the RMS database by country. The size of the dots corresponds to the number of samples available for each country. Country codes: FIN (Finland); MK (North Macedonia); BG (Bulgaria); HU (Hungary); YU (former

7

Yugoslavia); RUS (Russia); SLO (Slovenia); HR (Croatia); SK (Slovakia); PL (Poland); UKR (Ukraine); LT (Lithuania); EST (Estonia); AL (Albania); GR (Greece); LV (Latvia); CZ (Czech Republic); MD (Moldova); IT (Italy); RO (Romania); MX (Mexico).

In contrast, the Romani respondents who used the (DET-)N-ADJ order more regularly, amounting to up to 70% of responses in total, are all in contact with Romance languages that have variable N-ADJ and ADJ-N order,[3] that is, Italian, Spanish (in the sample from Mexico), Romanian, and Moldovan Romanian.[4] In the samples from Romania, in particular, the postnominal order of the adjective is the preferred order for all 19 speakers, with the prenominal order found in approximately 30% of the responses among individual speakers. An example of the ADJ-N order can be seen in (1), in 'big celebration' and 'New Year' (in bold), and of the N-ADJ order in (2), in 'big dog', all from the same speaker of a Vlax (Kalderaš) variety spoken in Romania.

(1) Kalderaš, Romania (RMS, RO-008, 691)

ando	amaro	gav	sy	**jekh**	**bari**	**sərbətoarja**	sako	**nevo**	**bərš**
in	our	village	be.3SG	INDF	big	celebration	every	new	year

'There is a big celebration in our village every New Year's Eve.'

(2) Kalderaš, Romania (RMS, RO-008, 908)

jekh	**žukel**	**baro**	daravel	le	ciknoren
INDF	dog	big	frighten.3SG	DEF.OBL	children.OBL.PL

'A big dog frightens the children.'

A fine-grained analysis of this variation is still needed, but it seems that the placement of the adjective in these examples does not depend on the semantic type of the adjective ('big').

---

[3] According to Cinque (2010), in Romance languages, the position of an adjective correlates with a restrictive and a non-restrictive interpretation: non-restrictive adjectives are prenominal and restrictive ones are postnominal. Differences are also noted for Romance languages depending on the semantic type of the adjectives: intersective and ambiguous adjectives and readings are generally postnominal (e.g., in Spanish *el abogado bueno* 'the lawyer good (as a human being)' or *búscate un abogado bueno* 'get a good lawyer (good as a lawyer/good person)'), while non-intersective adjectives and readings tend to be prenominal (e.g., *el buen abogado* 'the good lawyer (good as a lawyer)') (Demonte, 2008). Moreover, in Romanian, when adjectives are ambiguous between an intersective and a non-intersective interpretation (e.g., 'important'), then the prenominal order correlates with a non-restrictive interpretation and the postnominal order with ambiguity between a restrictive and non-restrictive interpretation (Marchis & Alexiadou, 2007).

[4] Note that Romani speakers in the Republic of Moldova speak Russian in addition to Moldovan Romanian.

Discussion

Overall, our study confirms the observation in Adamou and Matras (2020) that adjective/noun order preferences in Romani closely follow the preferred order in the contact languages in addition to the variability related to information structure that has been noted for some Romani dialects. In particular, our study shows that, when translating the same sentences, Romani speakers use the inherited (DET-)ADJ-N order as the sole or dominant option when they are in contact with languages that also have ADJ-N order, like the Slavic languages. In contrast, in settings where the contact language has a preferred (DET-)N-ADJ order, as in Romance languages, we note a significant increase in the use of this order in Romani. This is most apparent for the speakers from Romania whose responses to the RMS questionnaire show up to 70% of (DET-)N-ADJ order. In addition, we note the rare use of the (DET-)N-ADJ order (less than 20% of responses) among some speakers from countries where the current contact language does not offer the possibility of using the (DET-)N-ADJ order, suggesting influence from past contact languages and diffusion. A follow-up study should be carried in each country to evaluate word order preferences with a larger number of speakers (some settings were represented in the RMS by a single speaker), by carefully examining the contact setting, and by investigating the role that information structure and other linguistic factors may be playing (e.g., intersective and non-intersective adjectives and readings, restrictive and non-restrictive interpretation).

In the remainder of this paper, we will focus on the Romanian-Romani setting as it exhibits the greatest number of speakers using the (DET-)N-ADJ order.

Study 2: A free speech corpus

The RMS sample provided useful dialectal evidence on noun and adjective ordering in Romani, reflecting diachronic changes. However, as speakers translated sentences from the contact language, their responses could be subject to cross-language priming. In this study, we investigate whether the (DET-)N-ADJ order is in use among Romani-Romanian bilinguals when they speak freely in Romani. Indeed, many researchers investigating language contact consider that naturalistic speech provides ecologically valid data that can serve as a baseline (Adamou, 2016; Torres Cacoullos & Travis, 2018). In this study, we analysed a Romani corpus of interviews with four elderly speakers recorded in Romania (Furtună, Medeleanu, & Petrilă, 2016).

9

Methodology

Participants. The corpus contains interviews of four speakers who identified their dialect as Carpathic, Kalderaš, Ursari, and Spoitori, thus providing a variety of Romani dialects spoken in Romania. Table 1 presents the characteristics of the four speakers, two male and two female, all of them aged 60 or above. The content analysis of the interviews indicates that speakers are early bilinguals, that is, they learned both Romani and Romanian in childhood. The short life stories also allow us to conclude that all four speakers had immersive experience with both Romanian and Romani throughout their lives.

Table 1. Descriptive representation of the Romani-Romanian bilinguals with respect to age, gender, dialect and group self-identification, place of residence, and profession

Name	V.MA[5]	V.ST	M.PU	V.MI
Age at time of recording	64	75	83	60
Gender	Female	Male	Male	Female
Dialect and group self-designation	Richinara (Ursari)	Kalderaš	Carpathian	Spoitori (Tinsmith)
Place of residence	Buzău, Buzău County	Piteşti, Argeş County	Glodeni, Mureş County	Giurgiu, Giurgiu County
Profession	Tailor and merchant	Writer	Basket weaver	Tinsmith and merchant
Total number of words	2820	2586	2120	1892

Corpus. The corpus of life stories was collected and published by the Centre of Culture and Social Research 'Romane Rodimata' (Furtună, Medeleanu, & Petrilă, 2016). The interviewers were all native Romani speakers. The corpus amounts to 9,418 words.

Annotation. The edition includes the transcription in Romani, and its translation into Romanian and English, accompanied by a CD with the original video and audio recordings.

---

[5] We anonymized the interviewees.

First, we coded the uses of all NPs with a noun and an adjective, with or without a determiner, in the transcribed Romani corpus. Then, we confirmed their use in the original recordings.

Results

The analysis of the Romani corpus reveals three different types of word order involving a noun, an adjective, and a determiner. The example in (3) illustrates the native DET-ADJ-N order with the colour adjective 'white'. The example in (4) illustrates the innovative DET-N-ADJ order with the size adjective 'big'. The example in (5) reveals a third construction, DET-N-DET-ADJ with the size adjective 'big'. This phenomenon is known in the literature as 'polydefinites' (Kolliakou, 2004) or 'determiner spreading' (Marchis & Alexiadou, 2007). In Romani, this is a mechanism that allows for the use of N-ADJ order by doubling the determiner. Note that in all three examples, the noun is a borrowing from Romanian, but the adjective is inherited.

(3) DET-ADJ-N

kärde	kola	parne	ròkie
had	those	white	dresses

'(women) had those white dresses…' (Speaker: M.PU)

(4) DET-N-ADJ

äk	nekàzo	baro
INDF	grief	big

'a big grief…' (Speaker: V.ST)

(5) DET-N-DET-ADJ

keldăs	i	vestea	i	bari
celebrated.3SG	DEF.NOM.F	news	DEF.NOM.F	big

's/he celebrated the great news…' (Speaker: V.MA)

Table 2 presents an overview of the various uses of adjectival phrases, including instances with or without a determiner (N=53 in total). From a quantitative perspective, we note that although examples with DET-N-ADJ order are scarce, all speakers use the N-ADJ order without a determiner. In addition, all speakers used the DET-N-DET-ADJ order at least once in their interview.

Table 2. Order of noun and adjective for individual speakers in the free-speech corpus

	V.MA	V.ST	M.PU	V.MI	*Total*
DET-ADJ-N	1	0	5	5	*11*
DET-N-ADJ	0	2	0	2	*4*
DET-N-DET-ADJ	1	1	1	1	*4*
ADJ-N	5	1	12	2	*20*
N-ADJ	4	4	2	4	*14*
Total	*11*	*8*	*20*	*14*	*53*

Discussion

The analysis of the free-speech corpus confirms the use of both the inherited (DET-)ADJ-N order and postnominal adjectives among speakers of various Romani dialects from Romania. Although the number of occurrences in this study is too low to draw any strong conclusions, it can be seen that speaker M.PU, who is an 83 year-old basket-weaver and speaker of the Carpathian dialect, relies the most on the ADJ-N order. In contrast, speaker V.ST, a 70-year-old writer and speaker of the Kalderaš dialect, used ADJ-N once. The other two speakers, two merchants in their 60s, use both orders equally.[6]

In addition, corpus analysis reveals the use of polydefinites, a phenomenon that was also reported in the literature on Romani spoken in Greece in contact with Modern Greek (Igla, 1996), and in Albania in contact with both Albanian and Greek (Tirard, 2017). These studies suggest that what drives the use of polydefinites in Romani is the use of polydefinites in the

---

[6] A reviewer notes that (DET-)ADJ-N is used more frequently in Study 2 as compared to Study 1. We believe that this difference may be due to inter-individual differences in the two samples. In addition, although we are confident in the results of the comparison of uses within Study 1, where all speakers used the same adjectives in the same sentences, we are more cautious when it comes to comparing rates of use in Study 1 with rates of use in Study 2, which is a relatively small corpus of spontaneous speech. Finally, we believe that higher rates in Study 1 can indicate that some kind of cross-language priming may be affecting the preferred order among respondents given that they are translating sentences in Romani from a contact-language with a different adjective-noun order. The role of short-term priming is tested more systematically in Study 3.

contact languages. Similarly, Romanian can serve as a model language for Romani as it offers the possibility to use two determiners in adjectival phrases, combining the noun and the definite article with the demonstrative article *cel* and the postnominal adjective (Marchis & Alexiadou, 2007). Though the Romanian construction is superficially similar to determiner spreading, it is analysed by Marchis and Alexiadou as a reduced relative clause. The exact mechanism behind this change in Romani from Romania has not yet been studied in detail and lies beyond the scope of our paper, which focuses instead on the DET-N-ADJ order.

Study 3: A priming experiment

Study 1 shows that Romani-Romanian speakers use the (DET-)N-ADJ order more frequently as compared to Romani speakers in contact with languages that do not have dominant (DET-)N-ADJ order. Study 2 confirms that the use of the (DET-)N-ADJ order is not an artefact of the translation method used in Study 1 by showing that Romani-Romanian bilinguals alternate between (DET-)ADJ-N and (DET-)N-ADJ order in natural conversations in Romani. However, information structure greatly shapes the rates of the different word orders in conversation. To test whether short-term priming in adjective/noun order from Romanian could account for the increased use of postnominal adjectives in Romani, we conducted an experiment with Romani-Romanian bilinguals. This has the advantage of offering a controlled environment, including for information structure effects.

Methodology

Design. We adapted the design from the monolingual priming experiment in Ziegler et al. (2019) where a priming trial consists of a sentence trial and a picture trial. The original experiment tests dative structures, but in our experiment, we test complex noun phrases. Knowing that cross-language priming is weaker than within-language priming, we opted for three conditions that could facilitate priming: (i) lexical boost (there is lexical overlap of the head noun between target and prime); (ii) same context (always contrastive); (iii) no temporal delay between the prime and the target; and (iv) self-priming (the participant reads aloud the prime sentence and then describes a picture designed to elicit a target sentence) (see Mahowald et al., 2016 and Travis et al., 2017).

In order to achieve sufficient statistical power, we followed sample size recommendations using 18 target items and 90 participants, corresponding to statistical power of 1 in the condition involving lexical boost (see Mahowald et al., 2016).

Participants. Ninety Romani-Romanian bilinguals participated in this study (39 female; age range 17-79, M=37, SD=16). Participants rated their proficiency in spoken Romani as 4.5 (on a scale from 1 to 5 where 1 is the lowest and 5 the highest score) and in spoken Romanian as 4.6. No participant indicated a score lower than three in any of these languages. Twelve participants had attended at most primary school, 62 had attended high school, and 16 had had a higher education. Participants rated their capacity to read Romani as 3.2 and Romanian as 4.4. For Romani, 32 participants declared a score lower than three but managed to complete the task nonetheless. Follow-up conversations suggest that the low score indicates the lower frequency with which they engage in reading Romani and their mastery of the standard alphabet rather than their capacity to read Romani in Latin script, based on Romanian orthographical conventions like the ones we used in the stimuli. 49 participants self-identified as Ursari, 32 as Argintari (speaking a variety close to Kalderaš), and the remaining 9 as belonging to various other groups.

Stimuli. Ziegler et al. (2019) conducted the priming experiment online. In our study, we opted for the implementation of the experiment using printed stimuli and running it at participants' homes. We created forty visual stimuli that appeared on a sheet of A4 paper, which was folded in two in the middle. In the upper part of the sheet, a sentence was written in Romanian or in Romani; these were the prime sentences and followed the same structure: 'Choose *NP*'. There were 18 prime trials in which the NP was always composed of a determiner (the), a colour adjective (black, blue, green, yellow, brown, purple, pink, orange) and a noun (man, woman, dog, fish, cat, ball, car, pen, chair, bag, circle, heart, triangle, square, arrow, star). Prime trials were randomly interspersed with 20 filler trials following a similar structure: 'Choose X', where X was either a noun or a noun modified by a relative clause.

Six Romanian-Romani trials tested N_{DET}-ADJ[7] prime sentences, 6 Romani-Romani trials tested DET-ADJ-N and 6 DET-N-DET-ADJ prime sentences. We did not include any DET-N-ADJ primes in Romani. At the bottom of each sheet, there was a pair of simple coloured pictures. These pictures were preceded by the Romani word *Akana…* 'Now…' that served as a reminder for the language that the participants had to use in their description.

Participants had to provide instructions to select the object that was not mentioned in the prime sentence, resulting in a target utterance with the same verb, same noun, but different adjective. As a result, we tested word order choice in the exact same information structure

---

[7] We use the subscript to note the suffixal status of the Romanian definite article.

14

condition, that is, with the colour adjective under contrastive focus. We chose the contrastive context as it allows for all three types of adjective/noun order to occur in combination with prosodic focus marking (no study is available on prosodic marking of focus in Romani from Romania but see Arvaniti & Adamou, 2011 for a Vlax variety spoken in Greece).

Figure 2 offers an example of a prime sentence followed by the key word 'now' and the set of pictures. The origin of the target colour adjective was balanced: 8 cognates were borrowed from Romanian (yellow, pink) vs. 10 native or from past-contact languages (red, black, green). This choice is motivated by a study on noun adjective order in Albanian Romani reporting that inherited adjectives are more likely to be preposed than borrowed adjectives (Tirard, 2017).

The sentences and the pictures were normed by five speakers of Romani. Following their remarks, we added a lexical variant in two of the stimuli and corrected the use of case in one sentence.
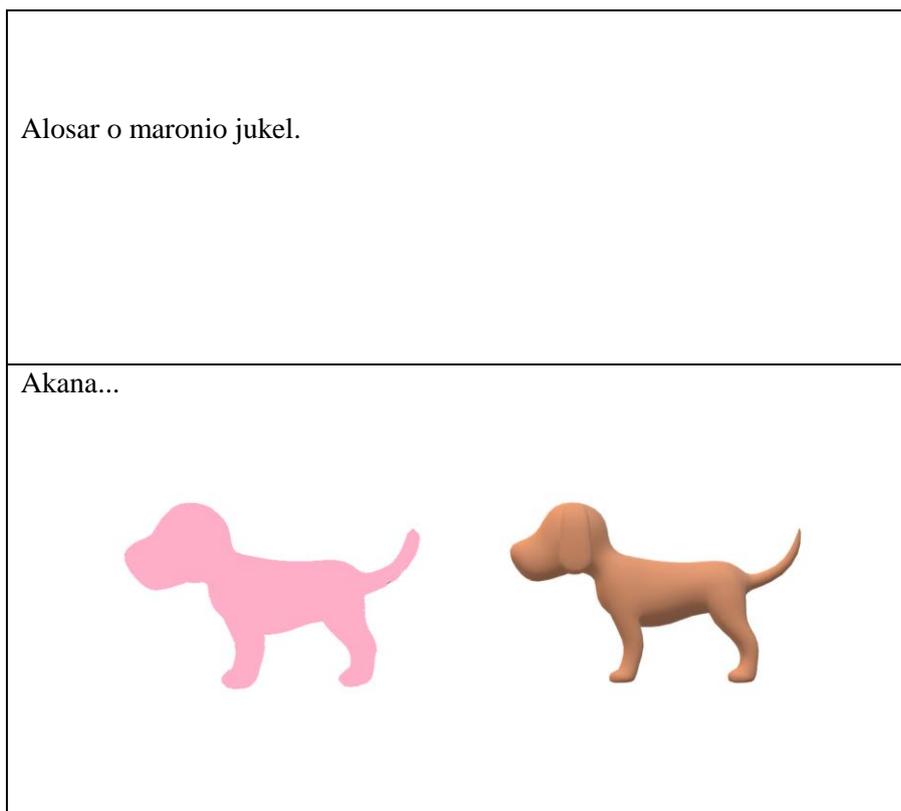

Alosar o maronio jukel.

Akana...

Figure 2. An example of a prime sentence ('Choose the brown dog', here in Romani) and a set of target pictures used in the priming experiment. The sheet was folded in the middle (where the line appears) so that participants read the prime sentence without having seen the target picture. They then turned the sheet and only saw the two pictures and the Romani word *Akana* 'now' without being able to see the prime sentence.

15

Procedure. We contacted participants through three Romani cultural associations who integrated the study as part of their activities to promote Romani language and culture. Participants did not receive payment individually, but were invited to an event that we sponsored and that was organized by the associations. Two Romani-Romanian speakers conducted the experiment, one of whom is a co-author in this paper. The experiment was conducted at participants' homes in the presence of the experimenter.

First, the experimenter went through the consent form in detail with the participants. Participants were unaware of the specific research questions of the experiment but were informed that it was part of research on Romani. Once participants signed the form, the experimenter explained the procedure of the task and asked participants to speak clearly, as if they were talking to children in a classroom. The experimenter recorded the entire session with an Olympus LS-12 Stereo Audio Digital voice recorder (LS12). The experiment started with two warm-up trials. In the sentence trial, participants read aloud a sentence that was written in Romanian or in Romani (e.g., 'Choose the brown dog'). They then turned over the page to see two pictures depicting the same object in two different colours; one of the two objects corresponded to the sentence description they had just read, the other did not (e.g., brown dog vs. pink dog). Participants were asked to provide instructions in Romani that could guide potential listeners to select the object that was *not* the one mentioned in the previously read sentence (e.g., 'Now choose the pink dog'). After describing the appropriate picture, participants placed the sheet of paper on a pile and continued with the next sheet. At the end of the session, participants responded to a language background questionnaire. The entire session lasted about 15 minutes. Every participant produced 18 target utterances, 6 in each of the priming types (DET-ADJ-N vs. N_{DET}-ADJ vs. DET-N-DET-ADJ prime).

Coding. Based on the audio recordings, the participants' target utterances were hand coded as DET-N-ADJ, DET-ADJ-N, DET-N-DET-ADJ, or null. Two researchers coded 20% of the data separately. Discrepancies between these transcriptions were due to cases of rapid speech that made it difficult to disambiguate between a DET-N-ADJ and DET-N-DET-ADJ order. Subsequently, both researchers coded all the ambiguous examples and came to an agreement.

Analysis. To analyze the data, we performed a multinomial mixed-effects logistic regression. This approach requires defining a reference category for the response (here the inherited DET-ADJ-N order) in order to perform a joint mixed-effects linear regression for the log-odds

ratio of each category compared to the reference one (here we have DET-N-ADJ vs. DET-ADJ-N and DET-N-DET-ADJ vs. DET-ADJ-N). The fixed effects are those of the PRIME and the ORIGIN OF THE VERB (inherited or other language borrowing vs. Romanian borrowing). Since the prime itself is categorical with three levels, there are 2 regressors associated with it. Here again, the DET-ADJ-N prime type was taken as the reference category and left uncoded. The results must therefore be understood with respect to this choice of reference: the weight of the PRIME N_{DET}-ADJ regressor for the DET-N-ADJ vs. DET-ADJ-N log-odds ratio is for instance interpreted as how much the N_{DET}-ADJ PRIME favors a DET-N-ADJ response over a DET-ADJ-N one, as compared to a DET-ADJ-N PRIME. The random effects are EXPERIMENTAL ITEMS and PARTICIPANTS. According to Barr et al. (2013), the best approach is to include all possible random effects: four for the PARTICIPANTS (on the intercept and on the slope of the three regressors) and only one for the EXPERIMENTAL ITEMS (on the intercept only since one item is always associated with the same origin and the same type of prime).

Furthermore, we opted for a Bayesian framework. Indeed, given the complexity of the model, the Bayesian approach is better suited for interpreting the results and assessing the significance of the parameters. Moreover, this approach has already been fruitfully applied in linguistic studies (Levshina, 2016; Grafmiller et al., 2018). We used the *MCMCglmm* package to perform the analyses. According to Raftery and Lewis' diagnostic (Raftery & Lewis, 1995), we used a total of 250,000 iterations for the Markov chains with a 10,000 burn-in period and a thinning of 60, for a total of 4,000 iterations accounted for. The prior structure followed the suggestions of Hadfield (2010) for the categorical mixed-effects models and was non-informative. We checked the convergence of the Markov chains using three different diagnoses (Ntzoufras, 2011): a) by running 10 independent chains and using Gelman and Rubin's criterion on these: the *R* factor we obtained for each of the fixed effect was always between 1.0 and 1.01, fulfilling the *R < 1.05* criterion for convergence; b) applying the Heidelberg and Welch diagnoses for run length and convergence (all fixed effects successfully passed the test); c) performing the Geweke diagnosis: the Z-score of all fixed effects were comprised between -1.47 and 1.01, therefore passing the test.

There are two ways to check the relevance of the fixed effects in this framework. The first, typically Bayesian in spirit, is to compute the 95% confidence interval of the parameter value according to the prior and check that the null value (0) lies outside of this interval. The second is more in line with standard frequentist analyses and relies on MCMC computation to derive a *p*-value of the mean posterior estimate under the null hypothesis. Both measures are provided by the MCMCglmm function and yield consistent acceptability statements.

17

The output of the model, for each trial, is not a value but a set of probabilities for each response type. The usual approach to assess agreement between the model and the data to select for the type with the highest probability (Levshina, 2016) does not work here, since the *N-ADJ* type is more probable than the other two for all primes and all adjective origins. Therefore, we sampled the fixed effects part of the model 1,000 times and compared the overall agreement of each sample with the data (47% agreement, 95% CI = [44%; 48%]). This relatively low value is due to high inter-individual variation which is not captured by the fixed effects part of the model. We compared this agreement with a sampling of a null model that keeps the probabilities constant across trials, equal to the overall proportion of each response type (38% agreement, 95% CI = [36%; 40%]), leading to a Cohen's κ coefficient equal to 0.14. The predictive improvement is therefore significant.

Since inter-individual variability is very high, we tried to identify which factors could explain this variability. To do so, we relied on a Random Forests approach (Tagliamonte & Baayen, 2012), which assesses the predictive importance of the different variables on the response variable of interest, even when these predictors are possibly correlated (as is often the case with sociolinguistic data like ours). Here, we considered the number of DET-ADJ-N type responses of each participant, and we set their gender, age, years of education, group (Argintari, Ursari, and other), oral Romani proficiency score, oral Romanian proficiency score, written Romani proficiency score and written Romanian proficiency score as predictor variables of this DET-ADJ-N type prevalence. We used the *cforest* and *varimp* functions of the party package in R for the analysis. To calculate the threshold below which the factors show no significant impact on the responses, we took the opposite value of the most negative importance score, following Tagliamonte and Baayen (2012).

Results

The statistical analyses reveal a significant effect of the Romanian prime N_{DET}-ADJ on Romani DET-N-ADJ and to a lesser extent on DET-N-DET-ADJ responses (see Table 3). In addition, we find a significant effect of Romani prime DET-N-DET-ADJ on DET-N-DET-ADJ and, to a lesser extent, on DET-N-ADJ. The origin of the adjective plays no role in the preferred word order. Figure 3 shows the descriptive statistics (see the Appendix for the raw result data). It appears that DET-N-ADJ is overall the preferred answer in the experiment. In addition, it can be seen that DET-ADJ-N responses are more frequent following another DET-ADJ-N prime than any other prime (using a two-sample t-test: p = 2.2e-16).

18

As there is high inter-speaker variability (see Figure 4), we conducted an analysis using Random Forests to identify the most relevant sociolinguistic factors. Results in Figure 5 show that group belonging and the education level of the participants are the most important variables: those who were not Argintari nor Ursari used more DET-ADJ-N in their responses, and the more schooling the participants had received, the greater the prevalence of DET-ADJ-N in their responses. To a lesser extent, the self-declared higher proficiency level in written Romanian has a negative effect on the DET-ADJ-N prevalence: the higher it is, the less participants use DET-ADJ-N. The other factors have no significant impact on the responses.

We looked at inter-speaker variation in more detail, noting that twenty-five speakers never use a DET-ADJ-N order, nine never use a DET-N-DET-ADJ order, and two never use a DET-N-ADJ order. In addition, five participants show no variation at all and use a single form consistently; in other words, their productions are not affected by priming. We then performed the same multinomial mixed-effects analysis with the subset of 59 participants who use all three response types. The results do not change the analysis as all effect sizes are similar (i.e., MCMC p-values < 0.001 for prime N DET -ADJ on DET-N-ADJ/DET-ADJ-N; prime N DET -ADJ on DET-N-DET-ADJ/DET-DET-ADJ-N; prime DET-N-DET-ADJ on DET-N-DET-ADJ/DET-ADJ-N; MCMC p-value = 0.002 for prime DET-N-DET-ADJ on DET-N-ADJ/DET-ADJ-N). Finally, we examined the correlation between all the sociolinguistic variables and the use of one, two, or three orders in the experiment. We found a significant correlation with self-declared spoken proficiency in Romani: speakers who declared having higher proficiency used more variants (r = 0.24, p = 0.02).

Table 3. Fixed effects of the multinomial mixed-effects model. Effect size, 95% confidence interval of the posterior value and MCMC *p*-values are computed through a Bayesian generalized mixed-model MCMC algorithm.

Regressor	Effect	95% CI	*p*-MCMC
Intercept on N_{DET}-ADJ/DET-ADJ-N	0.75	[-0.53; 2.07]	0.24
Intercept on DET-N-DET-ADJ/DET-ADJ-N	- 0.38	[-1.07; 0.80]	0.53
Prime N_{DET}-ADJ on DET-N-ADJ/DET-ADJ-N	**2.52**	**[1.19; 3.89]**	**0.0005 (***)**
Prime N_{DET}-ADJ on DET-N-DET-ADJ/DET-DET-ADJ-N	**2.01**	**[0.72; 3.45]**	**0.0035 (**)**
Prime DET-N-DET-ADJ on DET-N-ADJ/DET-ADJ-N	**2.54**	**[0.86; 4.47]**	**0.0060 (**)**
Prime DET-N-DET-ADJ on DET-N-DET-	**3.64**	**[1.94; 5.45]**	**0.0005**

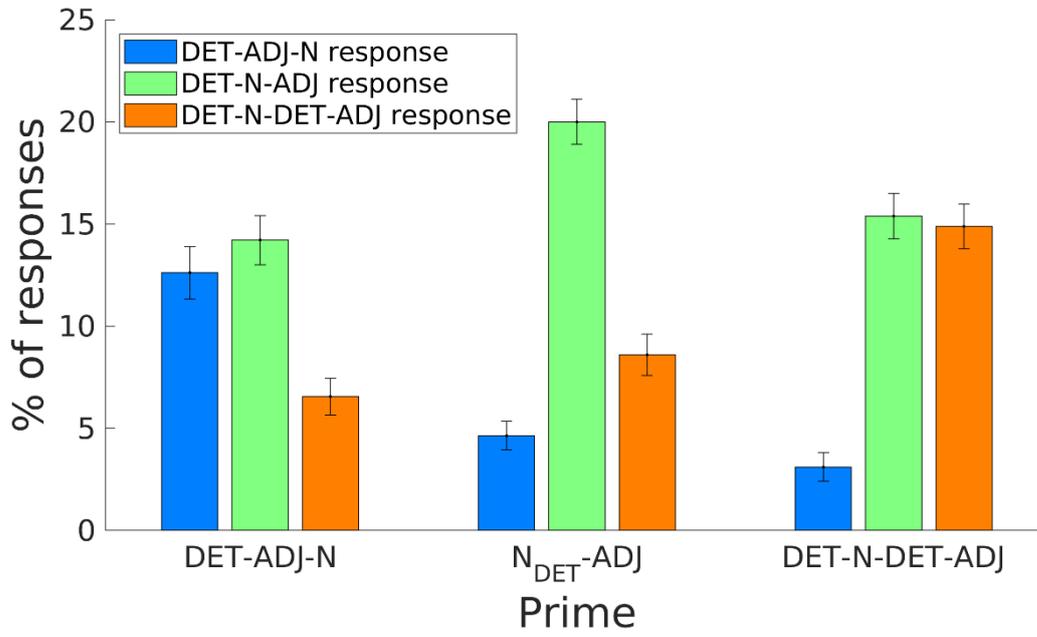|			
ADJ/DET-ADJ-N			(***)
Origin on DET-N-ADJ/DET-ADJ-N	- 0.52	[-1.85; 0.78]	0.42
Origin on DET-N-DET-ADJ/DET-ADJ-N	- 0.70	[-2.01; 0.64]	0.27

Figure 3. Percentage of responses in target utterances in Romani as a function of priming type (Romanian N_{DET}-ADJ vs. Romani DET-ADJ-N vs. Romani DET-N-DET-ADJ). Error bars are computed across participants.
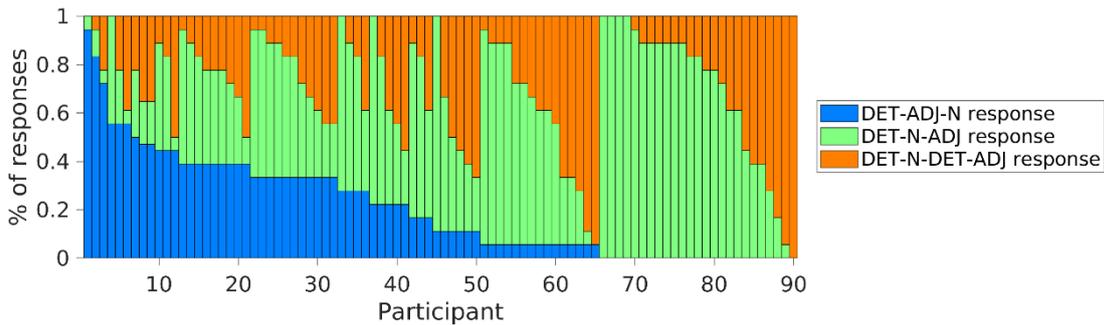

Figure 4. Percentage of responses in target utterances in Romani (DET-N-ADJ vs. DET-ADJ-N vs. DET-N-DET-ADJ)

20
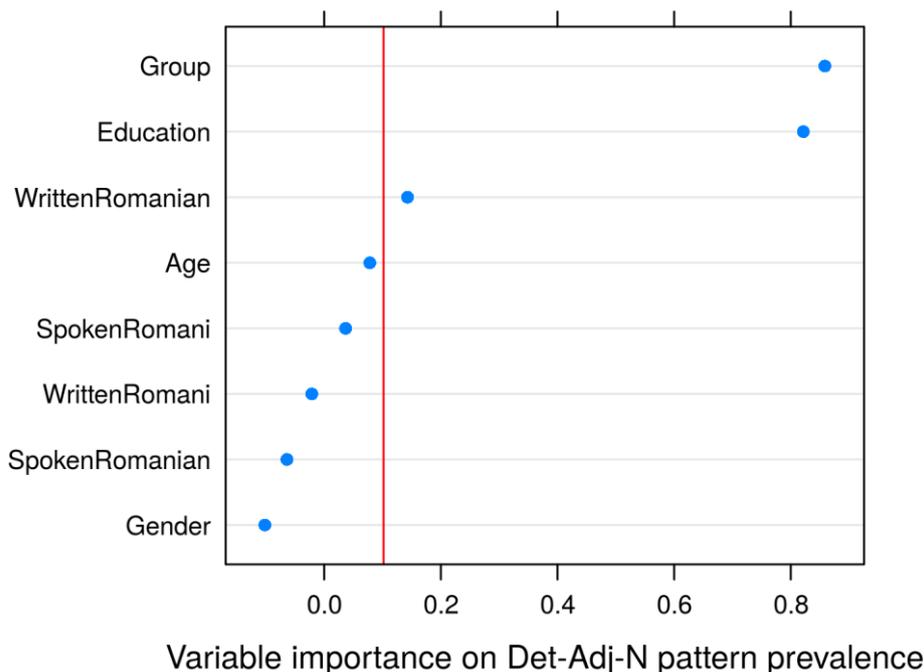
Figure 5. Importance ranking of variables (Random Forests) for DET-ADJ-N responses. The red vertical line corresponds to the threshold below which the factors show no significant impact on the responses.

Discussion

Overall, these results confirm the hypothesis that cross-language priming is a relevant mechanism in the selection of adjective/noun order in Romani at the level of the bilingual speaker, favouring the use of DET-N-ADJ order immediately following a N_{DET}-ADJ prime from Romanian. In addition, we observe that Romanian N_{DET}-ADJ also primes Romani DET-N-DET-ADJ and that, within Romani, DET-N-DET-ADJ primes DET-N-ADJ, suggesting that postnominal adjectives share part of their representation with polydefinites. This might account for the higher overall use of DET-N-ADJ order in the entire experiment as it was not only primed by the Romanian N_{DET}-ADJ primes but also by the Romani DET-N-DET-ADJ primes.

Moreover, when looking at the sociolinguistic variables, it appears that participants who had more years of schooling prefer the inherited DET-ADJ-N order. However, we note that, even if the effect is significant to a lesser extent, those who declared higher proficiency in written Romanian, use DET-ADJ-N order the least. It is therefore difficult to interpret the role of formal education, which is mainly in Romanian, and we restrain from drawing any strong

21

conclusions. It is also interesting to observe that differences in self-declared proficiency in written Romani did not correlate with the size of the priming effect, and that higher self-declared proficiency in spoken Romani correlated with the use of more variants. Overall, our results show that the tendency to use DET-N-ADJ order is modulated by formal education to some extent. The precise mechanism behind cross-language priming effects and literacy levels in the two languages needs to be further investigated by carefully controlling for the profiles of bilinguals. To our knowledge, these aspects have not yet been thoroughly discussed in the literature.

Finally, contrary to the observation in Tirard (2017) that borrowings favour an innovative order in Romani, we did not find that adjectives of Romanian origin boosted the DET-N-ADJ or DET-N-DET-ADJ order in Romani. This may be due to the fact that we have worked with different populations and different contact languages (i.e., Tirard worked with Romani-Albanian bilinguals). However, in theoretical terms, our results are coherent with the view that morphologically integrated words are also expected to be syntactically integrated (Poplack, 2018).

General discussion and conclusion

In this article, we set out to investigate the role of priming in contact-induced language change based on three different types of data.

First, we used data from a Romani dialectological database to establish word-order preferences in various dialects and contact settings. We found that an innovative (DET-)N-ADJ order is particularly salient in the responses of Romani speakers from Romania. Cross-dialectal comparison suggests that when the (DET-)N-ADJ order is available in the contact language, as is the case in Romanian and other Romance languages, then there is a boost to (DET-)N-ADJ order use in Romani. This finding agrees with observations in the literature on Romani adjective-noun use (e.g., Igla, 1996; Tirard, 2017; Adamou & Matras, 2020). Follow-up studies in these countries are needed to confirm these tendencies by examining larger samples and by taking into consideration the linguistic and sociolinguistic variables that may be at play in naturalistic speech.

We then analysed word order in adjectival phrases in a free-speech corpus from Romani speakers living in Romania to obtain ecologically valid data (Adamou, 2016; Torres Cacoullos & Travis, 2018). The analysis confirms that these speakers alternate between the (DET-)ADJ-N order and the (DET-)N-ADJ order in Romani. More research could determine

whether priming affects these uses in addition to language internal factors such as information structure and investigate other potential constraints.

Finally, to test whether exposure to a N$_{DET}$-ADJ order in Romanian would favour the use of a DET-N-ADJ order in Romani, in the same contrastive context, we conducted a priming experiment with 90 participants. Statistical analyses revealed a significant short-term priming effect from Romanian to Romani, with speakers more frequently using the DET-N-ADJ order in Romani following reading the N$_{DET}$-ADJ order in Romanian. These results are relevant to the field of psycholinguistics and bilingualism as they do not only document the existence of cross-language priming in an under-studied population of early bilinguals but also add adjective and noun ordering to the growing list of cross-language priming effects (see Kim & McDonough, 2007; Bernolet, Hartsuiker, & Pickering, 2013; Torres Cacoullos & Travis, 2018; Kootstra & Şahin, 2018). Our results are also relevant to the field of language contact as they show that priming is one mechanism that we should take into consideration when we study word order preferences. Future studies could also focus on the medium-term effects of priming on adjective-noun order; the prediction is that a priming effect will be found, albeit a weaker one.

Moreover, in our priming experiment, we found great inter-speaker variation, a result that stresses the need to work with a large number of speakers when investigating the effects of bilingualism and ongoing language change. Indeed, if we view bilingualism as a dynamic and adaptive process shaped by experience throughout the lifespan, we should expect differences in patterning to arise (see Adamou, 2021; Backus, 2021). In terms of sociolinguistic profiles, the results defy easy interpretation: we found that speakers who declared higher Romani proficiency levels tend to use more variants, that the more formally educated speakers preserve the inherited DET-ADJ-N order, and that speakers who declared higher proficiency in written Romanian tend to use this order the least. It is therefore worth exploring the precise role of formal education and literacy on cross-language priming in future studies by controlling for the participants' profiles.

In conclusion, this article offers a unified approach to language contact that unveils the connection between what bilinguals do when they speak two languages and how cognitive mechanisms in speech production can potentially impact their languages in the long run.

Abbreviations

ADJ adjective, DEF definite article, DET determiner, F feminine, INDF indefinite, N noun, NOM nominative, NP noun phrase, OBL oblique, PL plural, SG singular.

Appendix

Table A1. Raw number and ratio of responses in target utterances in Romani as a function of priming type (N_{DET}-ADJ vs. DET-ADJ-N vs. DET-N-DET-ADJ)

	DET-ADJ-N response	DET-N-ADJ response	DET-N-DET-ADJ response
DET-ADJ-N prime	204 (38%)	230 (42%)	106 (20%)
N_{DET}-ADJ prime	75 (14%)	324 (60%)	139 (26%)
DET-N-DET-ADJ prime	50 (9%)	249 (46%)	241 (45%)

References

Adamou, E. (2016). *A corpus-driven approach to language contact: Endangered languages in a comparative perspective*. Berlin: de Gruyter Mouton.

Adamou, E. (2021). *The adaptive bilingual mind: Insights from endangered languages.* Cambridge: Cambridge University Press.

Adamou, E., & Matras, Y. (2020). Romani syntactic typology. In: Y. Matras and A. Tenser, eds., *Handbook of Romani language and linguistics*, 187–227. London: Palgrave Macmillan.

Adamou, E., & Matras, Y., eds. (2021). *The Routledge handbook of language contact.* London: Routledge.

Arvaniti, A., & Adamou, E. (2011). Focus expression in Romani. In: M. Byram Washburn, K. McKinney-Bock, E. Varis, A. Sawyer and B. Tomaszewicz, eds., *Proceedings of the 28th West Coast Conference on Formal Linguistics*, 240–248. Somerville, MA: Cascadilla Proceedings Project.

Backus, A. (2021). Usage-based approaches. In: E. Adamou and Y. Matras, eds., *The Routledge handbook of language contact*, pp. 110–126. London: Routledge.

Barr, D.J., Levy, R., Scheepers, C., & Tily, H.J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68(3)*, 255–278.

Bernolet, S., Hartsuiker, R.J., & Pickering, M.J. (2007). Shared syntactic representations in bilinguals: Evidence for the role of word-order repetition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 931–949. https://doi.org/10.1037/0278-7393.33.5.931

Bernolet, S., Hartsuiker, R.J., & Pickering, M.J. (2012). Effects of phonological feedback on the selection of syntax: Evidence from between-language syntactic priming. *Bilingualism: Language and Cognition*, *15*, 503–516. https://doi.org/10.1017/s1366728911000162

Bernolet, S., Hartsuiker, R.J., & Pickering, M.J. (2013). From language-specific to shared syntactic representations: The influence of second language proficiency on syntactic sharing in bilinguals. *Cognition*, *127*, 287–306. https://doi.org/10.1016/j.cognition.2013.02.005

Bock, K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*, 355–387. https://doi.org/10.1016/0010-0285(86)90004-6

Bock, K., & Loebell, H. (1990). Framing sentences. *Cognition*, *35*, 1–39.

Cinque, G. (2010). *The syntax of adjectives: A comparative study.* Cambridge, M.A.: MIT Press.

Cleland, A. (2003). The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *Journal of Memory and Language*, *49*, 214–230. https://doi.org/10.1016/s0749-596x(03)00060-3

Demonte, V. (2008). Meaning-form correlations and adjective position in Spanish. In: L. McNally and C. Kennedy, eds., *Adjectives and adverbs: Syntax, semantics, and discourse*, 71–100. Oxford: Oxford University Press.

Dryer, M.S. (2013). Order of adjective and noun. In: M. S. Dryer and M. Haspelmath, eds., *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

(Available online at http://wals.info/chapter/87, Accessed on 2020-08-16.)

Elšík, V., & Matras, Y. (2006). *Markedness and language change*. Berlin: Mouton de Gruyter.

Ferreira, V.S., & Bock, K. (2006). The functions of structural priming. *Language and Cognitive Processes*, *21*, 1011–1029.

Furtună, A.N., Medeleanu, M.L., & Petrilă, D.S., eds. (2016). *Ŵivipnasqe rromane paramìćă; Poveşti în graiurile romilor; Life stories in different Romani dialects*. Popeşti Leordeni: Dykhta! Publishing House.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7(4)*, 457–472.

Goudbeek, M., & Krahmer, E. (2012). Alignment in interactive reference production: content planning, modifier ordering, and referential overspecification. *Topics in Cognitive Science*, *4*, 269–289.

Grafmiller, J., & Szmrecsanyi, B. (2018). Mapping out particle placement in Englishes around the world: A study in comparative sociolinguistic analysis. *Language Variation and Change*, *30(3)*, 385–412.

Gries, S.T. (2005). Syntactic priming: A corpus-based approach. *Journal of Psycholinguistic Research*, *34*, 365–399. https://doi.org/10.1007/s10936-005-6139-3

Hadfield, J.D. (2010). MCMCglmm: Markov chain Monte Carlo methods for generalised linear mixed models. *Tutorial for MCMCglmm package in R*, *125*. (Available online at: http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=A0A6F721EBD5E0E4375185F DC6EE193C?doi=10.1.1.160.5098&rep=rep1&type=pdf Accessed on 2020-08-16.)

Hartsuiker, R.J., & Westenberg, C. (2000). Word order priming in written and spoken sentence production. *Cognition*, *75*, B27–B39. https://doi.org/10.1016/s0010-0277(99)00080-3

Haugen, E. (1953 [1969]). *The Norwegian language in the Americas: A study of bilingual behaviour*. Bloomington: Indiana University Press.

Igla, B. (1996). *Das Romani von Ajia Varvara. Deskriptive und historisch-vergleichende Darstellung eines Zigeunerdialekts*. Wiesbaden: Harrassowitz.

Kim, Y., & McDonough, K. (2007). Learners' production of passives during syntactic priming activities. *Applied Linguistics*, *29*, 149–154. https://doi.org/10.1093/applin/amn004

Kolliakou, D. (2004). Monadic definites and polydefinites: their form, meaning and use. *Journal of Linguistics*, *40*, 263–323. https://doi.org/10.1017/s0022226704002531

Kootstra, G.J., & Muysken, P.C. (2017). Cross-linguistic priming in bilinguals: Multidisciplinary perspectives on language processing, acquisition, and change (introduction to special issue). *Bilingualism: Language and Cognition*, *20*, 215–218. https://doi.org/10.1017/S1366728916001127.

Kootstra, G.J., & Şahin, H. (2018). Crosslinguistic structural priming as a mechanism of contact-induced language change: Evidence from Papiamento-Dutch bilinguals in Aruba and the Netherlands. *Language*, *94*, 902–930. https://doi.org/10.1353/lan.2018.0050

Levshina, N. (2016). When variables align: A Bayesian multinomial mixed-effects model of English permissive constructions. *Cognitive Linguistics*, *27(2)*, 235–268.

Loebell, H., & Bock, K. (2003). Structural priming across languages. *Linguistics*, *41*. https://doi.org/10.1515/ling.2003.026

Mahowald, K., James, A., Futrell, R., & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. *Journal of Memory and Language*, *91*, 5–27. https://doi.org/10.1016/j.jml.2016.03.009

Marchis, M., & Alexiadou, A. (2007). On the distribution of adjectives in Romanian: the cel construction. In: E. Aboh, E. van der Linden, J. Quer and P. Sleeman, eds., *Romance Languages and Linguistic Theory*, 161–178. Amsterdam: John Benjamins.

Matras, Y. (2002). *Romani: A linguistic introduction*. Cambridge: Cambridge University Press.

Matras, Y. (2013). Mapping the Romani dialects of Romania. *Romani Studies*, *23*, 199–243.

Ntzoufras, I. (2011). *Bayesian modeling using WinBUGS* (Vol. 698). Hoboken, NJ: John Wiley & Sons.

Pickering, M.J., & Branigan, H.P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, *39*, 633–651. https://doi.org/10.1006/jmla.1998.2592

Pickering, M.J., & Ferreira, V.S. (2008). Structural priming: a critical review. *Psychological Bulletin*, *134*, 427–459.

Poplack, S. (2018). *Borrowing: Loanwords in the speech community and in the grammar*. Oxford: Oxford University Press.

Raftery, A.E., & Lewis, S.M. (1995). The number of iterations, convergence diagnostics and generic Metropolis algorithms. In: W.R. Gilks, D.J. Spiegelhalter and S. Richardson, eds., *Practical Markov Chain Monte Carlo*, 1–15. London: Chapman and Hall.

Sarău, G. (2017). *Istoricul studierii limbii rromani în învăţământul universitar din România (1992-2017)*. Bucureşti: Universitatea Bucureşti.

Scheepers, C. (2003). Syntactic priming of relative clause attachments: persistence of structural configuration in sentence production. *Cognition*, *89*, 179–205. https://doi.org/10.1016/s0010-0277(03)00119-7

Tagliamonte, S.A., & Baayen, R.H. (2012). Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change*, *24(2)*, 135–178.

Tirard, A. (2017). Syntactic doubling and variation: The case of Romani. In: I. Buchstaller and B. Siebenhaar, eds., *Language variation: European perspectives VI,* 141–156. Amsterdam and Philadelphia: John Benjamins.

Torres Cacoullos, R., & Travis, C. (2018). *Bilingualism in the community. Code-switching and grammars in contact*. Cambridge: Cambridge University Press.

Travis, C. E., Torres Cacoullos, R., & Kidd, E. (2017). Cross-language priming: A view from bilingual speech. *Bilingualism: Language and Cognition*, *20*, 283–298.

van Gompel, R.P.G., Arai, M., & Pearson, J. (2012). The representation of mono- and intransitive structures. *Journal of Memory and Language*, *66*, 384–406. https://doi.org/10.1016/j.jml.2011.11.005

van Hell, J.G., & de Groot, A.M.B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition*, *1*, 193–211. https://doi.org/10.1017/s1366728998000352

Weiner, E. J., & Labov, W. (1983). Constraints on the agentless passive. *Journal of Linguistics*, *19*, 29–58.

Weinreich, U. (1953 [1968]). *Languages in contact*. The Hague: Mouton.

Zenner, E., Backus A., & Winter-Froemel, E., eds. (2019). *Cognitive contact linguistics: Placing usage, meaning and mind at the core of contact-induced variation and change*. Berlin: Mouton de Gruyter.

Ziegler, J., Bencini, G., Goldberg, A., & Snedeker, J. (2019). How abstract is syntax? Evidence from structural priming. *Cognition*, *193*, 104045. https://doi.org/10.1016/j.cognition.2019.104045