

There are no language switching costs when codeswitching is frequent

Evangelia Adamou, Rachel X. Shen

▶ **To cite this version:**

Evangelia Adamou, Rachel X. Shen. There are no language switching costs when codeswitching is frequent. International Journal of Bilingualism, 2019, 23 (1), pp.53-70. 10.1177/1367006917709094 . halshs-01522408

HAL Id: halshs-01522408
https://shs.hal.science/halshs-01522408

Submitted on 8 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

doi.org/10.1177/1367006917709094

There are no language switching costs when codeswitching is frequent

Evangelia Adamou CNRS (National Centre for Scientific Research), France
Xingjia Rachel Shen University Paris Diderot and Nanjing University, France and China

Abstract

Aims and Objectives/Purpose/Research Questions

There is ongoing discussion as to the cost of language switching, with some studies indicating high cost and others showing low or no cost. The main research question in this paper is whether there are language switching costs in communities in which codeswitching is frequent.

Design/Methodology/Approach

We conducted two on-line experiments, i.e. a picture choice with sentence auditory stimuli and a word recognition task in sentence context. We constructed 16 sentences with differing degrees of ecological validity (16 sentences x 4 versions = 64). The sentences included verbs with different language preferences in natural conversations (L1, L2, or both).

Data and Analysis

Thirty-seven simultaneous L1-Romani L2-Turkish bilinguals participated in Experiment 1 and 49 in Experiment 2. To analyze the results, linear mixed models (lmer) were constructed using the 'lme4' package in R.

Findings/Conclusions

In Experiment 1, participants responded significantly faster for the all-Turkish sentences, followed by the mixed Romani-Turkish sentences, and the two types of ecologically non-valid sentences. However, there were no processing costs for the mixed sentences when they contained Turkish verbs that are more frequently used in Turkish in the spontaneous conversations. In Experiment 2, reaction times were similar for Turkish verbs (with Turkish verb morphology) in a mixed Romani-Turkish or a unilingual Turkish sentence.

doi.org/10.1177/1367006917709094

Originality

Taken together these findings indicate that language switching costs in comprehension depend on the frequency of codeswitching in the bilingual community, as well as on exposure to specific lexical items.

Significance/Implications

In sum, the Romani-Turkish data support a usage-based approach to bilingual processing. They show that comprehenders anticipate codeswitching based on prior experience. More specifically, when the input aligns with their expectations, there are no processing costs.

Keywords: sentence processing; language switching costs; codeswitching; usage-based models; Romani; Turkish

Processing and language switching costs

When humans engage in task repetition they are faster than when they have to deal with task switching (Monsell, 2003). Some studies showed that bilinguals are more comfortable in switching between tasks than monolinguals are, a phenomenon known as the 'bilingual advantage in executive control' (see Bialystok et al. (2004), but evidence against it in Paap & Greenberg, 2013). Based on this evidence, Bialystok and colleagues further argued that the same cognitive processes are involved in both linguistic and nonlinguistic switching. However, it is still unclear whether linguistic switching has any processing costs associated to it, as some studies indicate high cost (Soares & Grosjean, 1984; Grainger & Beauvillain, 1987; Meuter & Allport, 1999; Thomas & Allport, 2000; Jackson et al., 2001; Alvarez et al., 2003; Costa & Santesteban, 2004; Proverbio et al., 2004) and others low or no cost (Moreno et al., 2002; Jackson et al., 2004; Cheng & Howard, 2008; Ibáñez et al., 2010; Gullifer et al., 2013; Mosca & Clahsen, 2016), whether in comprehension or in production.

Linguistic switching costs are generally related to the magnitude of the inhibitory control required for language selection (Green, 1998). Attention has been paid to a number of factors that can potentially influence the processing of switching such as age and type of acquisition, proficiency, socioeconomic background, as well as type of tasks and stimuli. It has been shown for example that L2 learners process switching differently from highly proficient and early bilinguals, the latter showing an unexpected processing 'advantage of L2

(L3) over L1' (Costa & Santesteban, 2004). Similarly, research based on neuroimaging shows that naming in the L1 after naming in an L2 is associated with a greater effort in reactivating the dominant L1 which had been inhibited (Abutalebi & Green, 2016). Also, although socioeconomic status has an effect on language and cognition, it appears to be independent from bilingualism (Calvo & Bialystok, 2014). Moreover, for comprehension tasks, switching costs across an L1 and an L2 are symmetrical (Thomas & Allport, 2000). However, there can be asymmetrical switching costs across an L1 and an L2 in production (Meuter & Allport, 1999; Costa & Santesteban, 2004), but these costs may be overcome with more preparation time (Mosca & Clahsen, 2016).

It has also been argued that the cost of translation is different from that of inter-sentential switching (Ibáñez et al., 2010; Gullifer et al., 2013) and that the cost of ecologically-valid codeswitching is similar to that of unilingual passages in L1 as opposed to unexpected, irregular codeswitching and L2 (Chan et al., 1983). Moreno et al. (2002) confirm that an unexpectedness effect may arise from switchings if they are not the ones usual in a given bilingual community. The authors further observe that written stimuli, which are predominantly used in the experiments, are less ecologically-valid than auditory stimuli since codeswitching generally occurs orally. Moreover, ecologically-valid stimuli should reflect the codeswitching constraints that prevail in a given community. For example, the use of an English determiner preceding a Spanish noun is possible but not frequent in naturalistic data (Herring et al., 2010). In consequence, stimuli such as "He heard a knock at the *puerta*", used in Moreno et al. (2002: 191), may be unexpected not because of the language switch to Spanish *puerta*, 'door', but because the type of switch.

The importance of the communicational habits in language processing was put forward by Green & Abutalebi (2013) through the 'adaptive control hypothesis' that stresses the capacity of the language control network to adapt to the needs of the interactional setting. In that sense, the cognitive processes involved in language switching tasks may differ from those involved in codeswitching in natural communicative settings and these in turn may differ for bilinguals who frequently codeswitch and for bilinguals who codeswitch scarcely. In particular, Green (1998, 2011) suggests that bilinguals from communities with frequent codeswitching rely more heavily on the joint activation of the two languages, as opposed to bilinguals from communities that do not frequently codeswitch and therefore rely more heavily on their language control network to avoid conflict between the two languages. In order to fully comprehend the adaptive capacities of the bilingual brain, researchers call for more research taking into consideration the conversational practices of bilinguals (see

Abutalebi & Green, 2016; Hofweber, Marinis & Treffers-Daller, in press). More generally, usage-based research argues for a strong link between cognitive representations and frequency of use (Tomasello, 2003 for language acquisition; Goldberg, 2006 for constructions; Backus, 2015 for codeswitching) and recent models demonstrate how processing is facilitated by statistical frequencies (Wells et al., 2009; Jaeger & Snider, 2013; MacDonald, 2013).

Despite the growing interest in taking into account communicational habits, there is still little research on processing among bilinguals from communities that codeswitch in a systematic manner. This paper is an attempt to help fill this gap by providing the first experimental evidence with respect to sentence processing from a community with intense, well-established codeswitching between Romani and Turkish. At a descriptive level, widening the typological database by considering these forms of intense codeswitching should allow us to better understand the cognitive processes that characterize codeswitching and relate it to the research on language switching. At a theoretical level, such a study may help disentangle the cognitive and the sociolinguistic constraints in processing.

Romani-Turkish as spoken in Greek Thrace

This paper examines sentence processing in a Romani-Turkish bilingual community that frequently codeswitches between the two languages. Romani is an Indic language of the Indo-European family. Roma had arrived in the Byzantine Empire by the 11th century, in what are now Turkey and Greece (Matras, 2002). At present, approximately 200,000 Roma live in Greece (Bakker, 2001). In this paper we focus on a Romani community settled in the Drosero neighbourhood at the outskirts of the city of Xanthi in Greek Thrace; see map in Figure 1. The community has approximately 4,000 members of low socioeconomic status.

doi.org/10.1177/1367006917709094



Figure 1. Map of the area in Thrace, Greece. The research was conducted in the town of Xanthi.

The dominant Romani dialect spoken in Xanthi is characterized by the frequent insertion of Turkish words from all parts of speech with the exception of pronouns, reaching approximately 15% of all words in natural conversation (Adamou & Granqvist, 2015:531). The resulting Romani-Turkish variety is called *Xoraxane Romane* 'Turkish Romani' by its speakers. It is the home language and the language of in-group communication in this tightly-knit community.

The use of Romani-Turkish among the Muslim Roma of Greek Thrace has attracted some attention in the literature because of the typologically rare use of Turkish verb morphology with Turkish verbs inserted in Romani dominant speech (Adamou, 2010; Adamou & Granqvist, 2015; Adamou, 2016). An example illustrating Romani with Turkish codeswitching insertions is given in (1).

Greek Thrace Romani < Romani (in regular font), Turkish (in bold**)**

(1) **ep** me ka dikh-av kale ?
always 1SG.NOM will look-1SG DEM.OBL

me **da** **səndəm** //
1SG.NOM FOC tired.PRET.1SG

me **da** mang-av **dineneəm** //
1SG.NOM FOC want-1SG rest.OPT.1SG

'Am I **always** the one to look after them? **I'm tired** of it! Me **too**, I want **to rest**.'

An analysis of a free speech corpus with Romani as the dominant language indicates that speakers use 12% of Turkish verbs and that these verbs are *always* inserted with the Turkish person and tense-aspect-modality morphology (Adamou & Granqvist, 2015).

Although the use of non-integrated contact verbs is rare cross-linguistically (Wohlgemuth, 2009), it is reported for several Romani dialects in Europe, in different language contact settings and involving typologically diverse languages, i.e. North Russian Romani in contact with Russian (Rusakov, 2001), Finnish Romani in contact with Finnish (Adamou & Granqvist, 2015), Crimean Romani in contact with Turkish (Elšík & Matras, 2006: 135), Lithuanian Romani in contact with Russian (Elšík & Matras, 2006: 135), and several dialects spoken in the Balkans in contact with Turkish (Friedman, 2013). This suggests that sociolinguistic factors are at play, such as extensive and intensive bilingualism without strong prescriptive attitudes (Adamou, 2010).

The Roma of Xanthi also speak a non-native variety of Turkish that they use in traditional trade, most likely since Ottoman times — 15th to early 20th century (Adamou, 2010). In the Greek Thrace context, Turkish is the minority language of the Greek Thrace Muslim minority. As a consequence all members of the Muslim community, independent of their L1, have the right to Greek-Turkish bilingual education in primary school. Secondary education is strictly provided in Greek, which is also the language of administration and services. In practice, the effects of formal education on Roma should be viewed with some caution as access to schooling is strongly affected by the social exclusion that characterizes the relations between the Greek State and the Romani communities in Greece. Indeed, Roma over the age of 30 from Xanthi have at best attended Greek primary school with high rates of illiteracy. Most Roma under 30 have low school attendance rates and often drop out before completing middle school.

Roma generally acquire both Romani-Turkish and Turkish before the age of 3 at home and in the community with differing degrees of exposure. Despite simultaneous acquisition, we will refer to the traditional community language, Romani-Turkish, as L1 and Turkish as L2. We note a shift to Turkish in the past decade in several families and it is not uncommon for a family to have raised a first-born child as bilingual and then shift to an all-Turkish upbringing for the second child. Despite this shift, we note that children are exposed to Romani in the community since they often grow up in households not only with their parent(s) and siblings, but also with their paternal grandparents. They also have daily contact with their parents' siblings, their spouses and children who may sometimes live in the same

house for economic reasons. It is also important to keep in mind that families receive almost daily visits from extended family, friends and neighbours and that children have a certain freedom of movement within the community. In the absence of access to day care, children of all ages often follow their primary caregivers to work and therefore are exposed to Turkish and Greek spoken by community outsiders. Table 1 summarizes the trilingual setting for most adult Roma in Xanthi.

Table 1. The trilingual setting for most adult Roma in Xanthi

Acquisition	L1 Romani-Turkish	L2 Turkish	L3 Greek
Function	Primary	Secondary	Secondary
Sociopolitical status	Minority	Minority	Majority

Experiment 1: Picture choice with auditory sentence stimuli

Experiment 1 is an on-line, bimodal picture-sentence matching task with auditory stimuli. This experiment is inspired by the visual world paradigm with eye-tracking used in Dahan & Tanenhouse (2004) in that it also uses visual and aurally presented stimuli. The research question that this experiment addresses is whether mixed Romani-Turkish sentences have higher processing costs than unilingual L2-Turkish sentences. We therefore wanted to test the reaction times (RTs) of the Romani participants with respect to Romani-Turkish mixing within sentences (local costs) in a demanding task with frequent language switches between sentences (global costs). In line with usage-based models, Romani-Turkish mixing should be treated similarly to unilingual speech as it is very frequent in the community (Adamou & Granqvist, 2015). However, in line with a number of studies on language switching costs, the mixed Romani-Turkish sentences should be associated to higher costs than unilingual speech.

Methodology

Participants

Thirty seven Roma participated in the task. They were contacted through a local non profit organization based on their proficiency in both Romani-Turkish and Turkish. All gave oral informed consent and, in agreement with the organization's representatives, they received

no compensation for their participation in the study.[1] Participants were all residents of the community of Drosero in Xanthi and were of similar socio-economic status, i.e. low income levels. All but one had low education levels, i.e. 10 attended at most primary school, 26 attended at most secondary school, and 1 had a university degree. 23 participants were female and 14 male. Age of the participants ranges from 13-51 (M = 22.59, SD = 11.13). All the participants declared understanding and speaking Romani and Turkish. They have received no formal education in Romani, a language with no written tradition, nor Turkish that they acquire in oral interactions within the community and with outsiders. All declared that they acquired Romani-Turkish and Turkish before the age of 3, with varying degrees of exposure to each language that it was not possible to measure. 27 participants declared Romani to be their primary language of communication (age M = 25.6, SD = 11.77), ten declared it was Turkish (age M = 17.87, SD = 1.34).

Materials

Sixteen two-sentence stimuli were created, consisting of a first sentence with a Turkish noun at its beginning and a second sentence with a Turkish verb at its beginning. The first and second sentences were related in meaning and were culturally adapted; see Appendix A for the list of sentences translated into English. Four versions of each sentence were recorded with a fluent female, native Romani speaker from the locality, providing a total of 64 auditory stimuli. The sentences were carefully constructed in collaboration with this speaker and were repeated orally until reaching a fluent and naturalistic version. The number of words for the four versions was kept as similar as possible, i.e. between 5-7 words per sentence. The recordings were done with a Tascam DR-100 solidstate recorder via a supercardioid head-worn microphone. Using Audacity, long hesitations, typically occurring between the first and the second sentence, were reduced to achieve similarity in the total duration of the audio stimuli. The resulting sentences for the four versions were similar in length with a mean duration of 6800ms. The length of the sound input in milliseconds was also integrated in the statistical models.

In Version a, the sentences were recorded using the Romani-Turkish mixing patterns that are frequent in the community. In this version, exemplified in (4a), the Turkish noun is

[1] This decision is in accordance with local cultural practices and takes into consideration the sensitive political context in which financial retribution may become problematic. This task was therefore integrated into the NGO's general activities.

integrated into Romani morphology while the Turkish verb maintains its original Turkish morphology.

Version b contains strictly Turkish material and was produced in the local Turkish variety. This resulted in verb initial sentences, parallel to the Romani sentences, even though Turkish is a verb final language. See (4b).

Version c, shown in (4c), is close to so-called classic codeswitching and was constructed by enunciating one part of the sentence in Romani, the other in Turkish with morphologically non-integrated nouns and verbs and with words from word classes which are otherwise used in Romani, such as pronouns. This option is possible in other Romani varieties, but is not attested in the free conversational corpus of the community (Adamou & Granqvist, 2015).

Finally, Version d included sentences in Romani with Turkish verbs adapted in the Romani morphology, like borrowings; see example (4d). This possibility is attested in other Romani varieties of the Balkans but is not used in the local Romani variety.

Romani (in regular font), Turkish (in bold), Greek (underscored)

(4) 'The neighbours were having parties very often. They drank and danced until late.'

Mixed Romani-Turkish

a.	e	**komʃ**-je	but	**seki**	ker-en-as	<u>ɣlend-ja</u>
	DEF.PL	neighbour-PL	very	often	make-3PL-IMPF	party-PL

itʃ-er-di-ler	xa-n-as	but	**getʃ**i	sao	gie
drink-AOR-PST-PL	eat-3PL-IMPF	very	late	every	day

Turkish

b.	**Komşu-lar**	**ör**	**gün**	**eylence**	**yap-ıyor-lar-dı**
	neighbour-PL	every	day	party	make-PROG-3PL-PST

doi.org/10.1177/1367006917709094

İç-er-di-ler	**ye-r-di-ler**	**ör**	**gün**	**çok**	**geç**	**vakt-a**	**kadar**
drink-AOR-PST-PL	eat-AOR-PST-3PL	every	day	very	late	time-DAT	until

Romani with Turkish codeswitching

c.	**komʃu-lar**	**er**	**gyn**	ker-en-as	**eɣlendʒe**
	neighbour-PL	every	day	make-3PL-IMPF	party

onlar	**itʃ-er-di-ler**	xa-n-as	kel-en-as	dʒi	but	**getʃ**i
3PL	drink-AOR-PST-PL	eat-3PL-IMPF	dance-3PL-IMPF	until	very	late

Romani with Turkish borrowings

d.	e	**komʃ**-je	**er**	**gyn**	ker-en-as	**eɣlendʒe**-a
	DEF.PL	neighbour-PL	every	day	make-3PL-IMPF	party-PL

itʃki-al-en-as	xa-n-as	kel-en-as	but	**getʃ**i
drink-LVM-3PL-IMPF	eat-3PL-IMPF	dance-3PL-IMPF	very	late

Sentence (4a) is the most ecologically valid. We note that the Turkish noun ‘neighbour’ is adapted to the Romani morphology through the plural suffix *–je* and the use of the definite article, but the Turkish verb ‘drink’ has Turkish morphology. In the rest of the sentence, the two other verbs are Romani, with Romani verb morphology. The word ‘party’, [ɣlendja], can be identified as a switch to Greek (in natural conversations approximately 4% of Greek words are used, mainly nouns). We also note the use of the Turkish adverb ‘late’.

In sentence (4b), which is all Turkish, we note that the word 'every' [ör] is pronounced differently from the usual Turkish phonetics, [her], but it is also different from the word as used in Romani, [er]. We also note that the verb is not in clause final position as it is in native varieties of Turkish.

Sentence (4c), which is constructed with codeswitching alternations between Turkish and Romani, results in the use of the Turkish noun 'neighbours' with the Turkish plural which is never used as such in the Romani-Turkish mixed manner illustrated in (4a). We also note the use of the free pronoun for the Turkish third person plural, *onlar*, which is also distinct from the Romani-Turkish mixed form, *onar*.

Finally, sentence (4d) provides a version where the Turkish verb is integrated in the Romani verb morphology. Nouns are also integrated into Romani morphology, e.g. the noun 'party' keeps its Turkish phonetics but receives the plural marker *-a*.

Moreover, in order to test the difference in effect between a highly predictable switch and a less expected one, we took into consideration the lexical frequency of the verbs as observed in natural speech (Adamou, 2016). Five sentences included a Turkish verb that is more frequently used in Turkish than in Romani ('Turkish'), six sentences included a verb that is attested in the corpus once and for which the Romani variant is more frequently encountered ('Romani'), and five sentences included a Turkish verb that is used with equal frequency in both Turkish and in Romani ('Variable'). We predict that the verbs that are preferably used in Turkish will be associated to short RTs, that the verbs preferably used in Romani will be associated to longer RTs, and that the verbs that are used in either Romani or Turkish will be associated to long RTs.

We further conducted a norming study for the four versions of the sentences. Four Roma from the community (ages 18-25) who did not participate in the experiment were asked to rate the sentences with respect to the likelihood of hearing them in the community. They were asked to provide their ratings on a 5-point scale, from 1 (not likely) to 5 (very likely). The ratings were noted by the researcher. Each participant listened to 16 sentences in total and versions were counterbalanced across participants. Version a sentences received a mean rating $M = 4.6$, Version b $M = 4.1$, Version c $M = 2.3$ and Version d $M = 0.4$. In combination with the naturalistic evidence, this result confirms that Versions a and b are the most natural in the community whereas Versions c and d are the most unnatural.

Sixteen pairs of pictures served as material for the target sentences and two pairs of pictures were used for the trials. See Appendix B for an example of visual stimuli. We did not

conduct a norming study for agreement between the pictures and the sentences since we did not predict any pattern with respect to accuracy in the picture choice.

Procedure

The participants were tested in the office of a local non-profit organization, in a calm environment. They were seated in front of a computer screen and wore headphones. The experiment was conducted on a computer using Open Sesame (Mathôt et al., 2012). The stimuli were fully randomized. The instructions were provided in Greek by the researcher, i.e. in the language of school, and in Romani-Turkish and Turkish by a local assistant, thus creating a multilingual environment (Soares & Grosjean, 1984). The Romani participants were invited to watch several pairs of pictures and listen to the synchronized auditory stimuli which appeared simultaneously. They were told that they would listen to various languages, Romani-Turkish, Turkish, and Greek, so that they could anticipate the frequent inter-language switchings (see Ibáñez et al., 2010 who note the capacity of translators to anticipate language switches). Participants were instructed to select the picture that appeared to be more closely related to the meaning of the audio stimuli by pressing a button on the computer: a left arrow button if they wanted to select the picture on the left of the screen, and a right arrow button for the picture showing on the right. They were told that they should press the button as soon as possible and not to wait for the completion of the sentences. Inter-trial intervals were not controlled for and were left up to the participant. The task started with two warm-up trials. Each participant responded to 16 trials consisting of four sentences for each version (a, b, c, and d). For 36 participants, this experiment took place after the Experiment 2 and in this case the participants were familiarized with the sentences; 13 participants first completed Experiment 1 before conducting the Experiment 2.[2] Changes in the order of the tasks were made to avoid any confounding effects. There were short breaks (2-5 minutes) between the two tasks. A pilot study with three participants included 5 fillers in the experiments. However, we noted that participants were distracted before the end of each task or did not wish to complete it. For this reason, we decided not to include the fillers in the final version. Once the experiment was completed, participants declared that they did not find the task difficult and appeared amused by the pictures. They were also happy to listen to their native language via a computer.

[2] Due to an error in Open Sesame configuration, the results of these participants for Experiment 2 were not recorded.

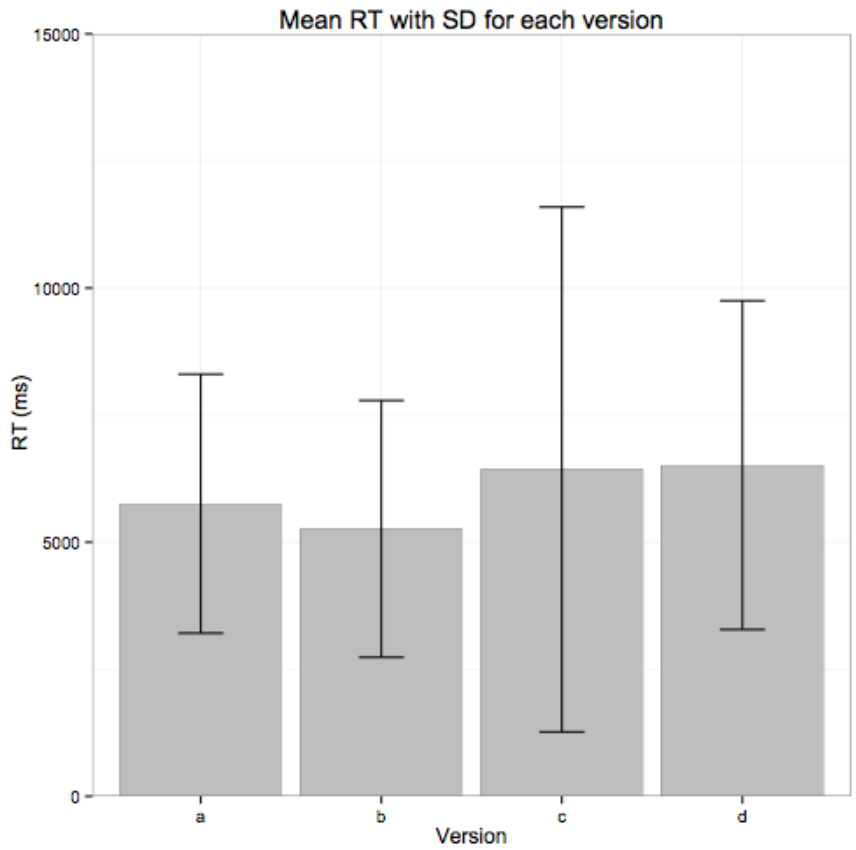
doi.org/10.1177/1367006917709094

Analysis

Linear mixed models (lmer) were constructed using the 'lme4' package (Bates et al., 2014) in R (R Core Team, 2013) to analyze the results. The dependent variable is the reaction time (RT), and the independent variables are 'Language preference' for the verb in the free-speech corpus (Variable, Romani, and Turkish), and 'Versions' of the sentences in the experiment (a, b, c, and d). 'Subjects', 'Sentences' and 'Duration' of audio files in terms of the acoustic input time in milliseconds are coded as random factors. We eliminated the outliers of the RTs that were faster than 500ms, and slower than 20000ms, which was about 0.7% of the total data.

Results

The participants responded the fastest for the unilingual Turkish sentences (Version b), followed by the Romani-Turkish mixed sentences (Version a) and the codeswitches (Version c). The mean RTs are: Version a = 5756.4ms, Version b = 5178.4ms, Version c = 5914.1ms and Version d = 6409.9ms; see Figure 2.



doi.org/10.1177/1367006917709094

Figure 2. Mean reaction times (RT) in milliseconds (ms) of different versions of the sentences (a = mixed Romani-Turkish, b = all Turkish, c = Romani with Turkish codeswitching, d = Romani with Turkish borrowings) in Experiment 1 (picture choice)

ANOVA shows that there are significant "Version" differences (logLikelihood is -6900.9, $\chi 2_3$= 19.0, $p < .001$)). The all-Turkish sentences, in Version b, are significantly shorter than all the other versions, i.e., Version a (t = 2.7), Version c (t = 2.3) and Version d (t = 4.3). The codeswitching Version c is also significantly shorter than Version d (t = 2.0), which is based on atypical Turkish verbal forms with Romani morphology. The mixed Romani-Turkish Version a is marginally shorter than Version d (t = 1.5). However, there is no significant differences between the mixed Romani-Turkish version (Version a) and Version c that involves codeswitching (t<1).

We then conducted a more fine-grained analysis by considering the RTs with respect to the language preference of the verbs in natural speech. Indeed, although this experiment does not target the verbs in particular, the RTs indicate that participants reached the beginning of the second sentence, including the Turkish verb, before pressing the button. Verbs were classified in three types depending on preferred language when they speak in the Romani-Turkish variety. They are tagged 'Variable' when speakers in the community use verbs in Romani or in Turkish, 'Romani' when they mainly use the verbs in Romani although some peripheral occurrences in Turkish can be observed, and 'Turkish' when the verbs are mainly used in Turkish. The mean RTs of the experiment are presented in Table 2.

Table 2. Experiment 1: Mean RTs of different versions of the sentences (a = mixed Romani-Turkish, b = all Turkish, c = Romani with Turkish codeswitching, d = Romani with Turkish borrowings) with different language preferences for the verbs as observed in a free-speech corpus (variable, Romani, Turkish)

	Variable	**Romani**	**Turkish**
a	7653.6	5018.8	5145.0
b	6525.8	4053.5	5647.2
c	6973.0	4110.4	6437.4
d	7111.5	5374.8	6727.0

doi.org/10.1177/1367006917709094

The ANOVA shows significant 'Version' differences (logLikelihood is -6900.9, $\chi 2_3$= 19.0, p < .001) and 'Language preference of verb in the corpus' differences (logLikelihood is -6898.3, $\chi 2_2$= 11.5, p < . \s 003).

For 'Language preference of verb in the corpus': 'Variable' have the longest RTs (ts>2.5). For 'Variable' and 'Romani' verbs, the all-Turkish Version b and the codeswitching Version c have shorter RTs than the mixed Romani-Turkish Version a and the borrowing Version d (ts>1.9). For the verbs most frequent in 'Turkish' in natural speech, both the mixed Romani-Turkish and the all-Turkish versions, i.e. Version a and Version b, are shorter than the two ecologically non-valid versions, i.e. Version c and Version d (ts>1.9). See Figure 3.

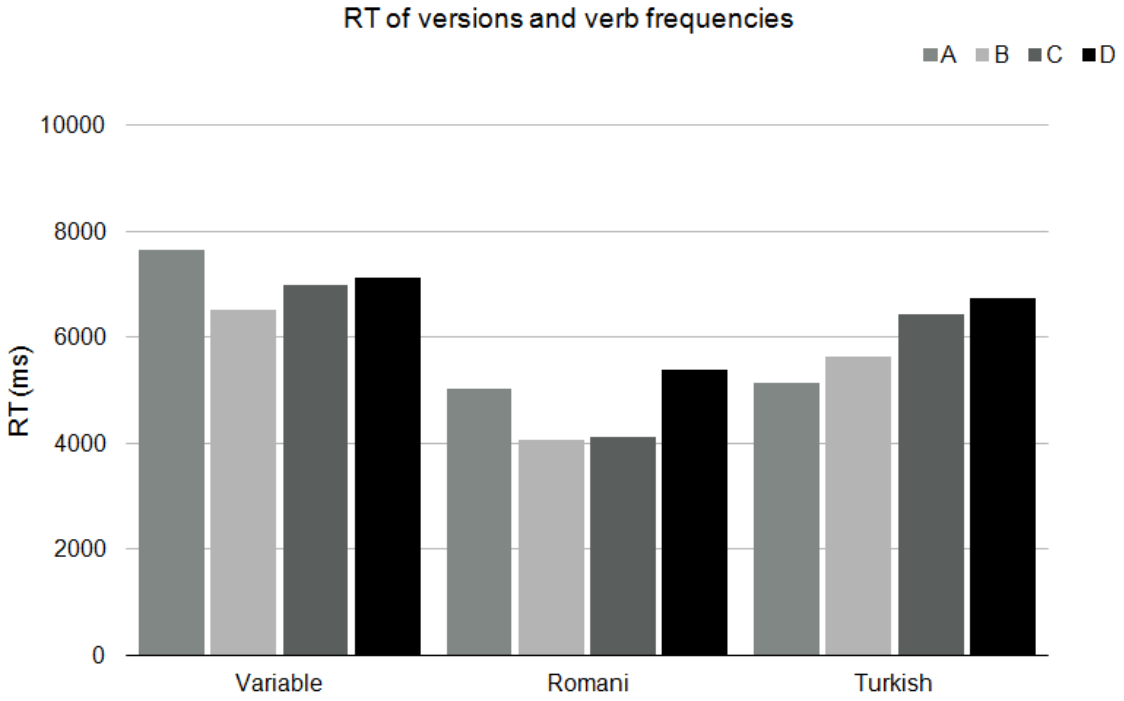


Figure 3. Mean reaction times (RT) in milliseconds (ms) of different versions of the sentences (a = mixed Romani-Turkish, b = all Turkish, c = Romani with Turkish codeswitching, d = Romani with Turkish borrowings) with different language preferences for the verbs as observed in a free-speech corpus (variable, Romani, Turkish) in Experiment 1 (picture choice)

Discussion

In this experiment, participants responded faster for the unilingual Turkish sentences (Version b) than for the mixed Romani-Turkish sentences (Version a). The mixed Romani-

Turkish sentences (Version a) were similar to codeswitching (Version c). This result can be understood in relation to other studies on bilingual processing which indicate that understanding unilingual speech is the least costly. However, when language preference of the verbs in the natural speech of the community is taken into consideration, it appears that for the verbs that were more frequently Turkish in the corpus, mixed Romani-Turkish sentences were as fast as Turkish unilingual sentences. It is interesting to note that the slowest RTs arise when both Romani and Turkish verbs are used in the community. This might indicate that variation in language choice entails processing costs in comprehension.

A limitation of this experiment is that RTs do not directly provide information on the processing costs of the Turkish verb but of the sentences as a whole since participants can press the button at any point of the sentence. Also, possible confusion may have arisen from difficulties with picture association rather than with processing. Experiment 2 was designed to overcome these limitations.

Experiment 2: Word recognition in sentence context

Experiment 2 is an on-line task with auditory sentence stimuli using the word monitoring paradigm (Kilborn & Moss, 1996). The experiment is designed to determine whether the morphologically non-integrated Turkish verbs have higher processing costs when they occur in a mixed Romani-Turkish environment than in a unilingual Turkish environment. In line with studies showing language switching costs, the prediction would be that the morphologically non-integrated verbs will have high processing costs in Romani-Turkish sentences. However, Adamou & Granqvist (2015) show that the Turkish verbs are frequent switches in the Romani community of Xanthi. In line with usage-based models, Turkish insertional switches should be processed with low costs, similar to unilingual speech.

Methodology

Participants

Forty nine trilingual Romani-Turkish-Greek speakers participated in this experiment. As in Experiment 1, participants were contacted through a local non-profit organization. All gave oral informed consent and received no compensation for their participation in the study following agreement with the organization's representatives. Participants were of similar

socio-economic status, i.e. low income levels, and low education levels, i.e. 18 attended at most primary school, 30 middle school or secondary school, and 1 had a university degree. 34 participants were female and 15 male. Ages range from 13-50 (*M* = 24.10, *SD* = 11.5). As in Experiment 1, all the participants declared that they acquired Romani-Turkish and Turkish simultaneously before the age of 3, although they had varying degrees of exposure to each language at the time. 39 declared Romani to be their primary language of communication (age *M* = 26.5, *SD* = 11.72), and ten declared that Turkish is their primary language (age *M* = 17.87, *SD* = 1.34).

Materials

We used the same auditory stimuli as in Experiment 1. The same four versions that were used in Experiment 1 were also used for this experiment; i.e. (a) mixed Romani-Turkish, (b) all Turkish, (c) Romani-Turkish codeswitching, and (d) Romani with Turkish borrowings. Similar to Experiment 1, language preference in a free-speech corpus for the Turkish verbs that were used in the experiment was also taken into consideration: variable, Romani, Turkish (see Adamou, 2016).

The target word in this experiment was the word that immediately followed the Turkish verb, the latter being the prime. As can be seen in Table 3, the prime was practically the same in Versions a, b, and c, and differed slightly in Version d. The target was the same in Versions a, c, d but different in Version b.

Table 3. Stimulus examples from Experiment 2 (Romani in regular font, Turkish in bold, Greek underscored)

Version	Auditory stimuli for the sentences: 'The neighbours were having parties very often. They drank and danced until late.'	Target	Prime
a. Mixed Romani-Turkish	E **komʃ**je but **seki** kerenas <u>ɣlendja.</u> **Itʃerdiler** xanas but **getʃ**i sao gie.	Xanas	**Itʃerdiler**
b. All Turkish	**Komşular ör gün eğlence yapıyorlardı.** **İçerdiler yerdiler ör gün çok geç vakt a kadar.**	Yerdiler	**İçerdiler**
c. Romani with Turkish	**Komʃular er gyn** kerenas **eɣlendʒe.** **Onlar itʃerdiler** xanas kelenas dʒi but **getʃi.**	Xanas	**Itʃerdiler**

codeswitching			
d. Romani	E **komʃ**je **er gyn** kerenas **eɣlendʒe**a.	Xanas	**Itʃki**alenas
with Turkish borrowings	**Itʃki**alenas xanas kelenas but **getʃi.**		

Procedure

As in Experiment 1, the participants were tested in the NGO's offices in a calm environment. They were seated in front of a computer screen and wore head-phones. The instructions were provided in Greek by the researcher as well as in Romani-Turkish and Turkish by a local assistant to create a multilingual environment. The Roma participants were told that this was a task in which both speed and accuracy were important. They were also told that they would be hearing several languages, Romani-Turkish, Turkish, and Greek, to prepare them for the frequent switches between languages. The experiment was conducted using Open Sesame (Mathôt et al., 2012). It started with two warm-up trials. The participants were invited to listen to pre-recorded instructions in Greek: 'Press the big yellow button as soon as you hear the word…'. They would then listen to the target word that was cut from the original recording and pasted in the instruction sentence. When assured that the participant was ready, the researcher pressed the button in order to start the auditory stimuli. The participant pressed the button on his/her own when appropriate. A recorded follow-up comprehension question in Greek was played by the researcher to make sure that the participants paid attention to the meaning of the sentences. The participants had to answer the question with 'yes' or 'no'. For this part of the task, the researcher pressed the appropriate buttons on the computer to record the response to simplify the experiment as a whole. Each participant completed an experimental session consisting of 16 trials, composed of four sentences from each of the four versions (a, b, c, and d).

All the participants were very excited to participate in the task. They were very motivated and expressed their amazement at listening to the stimuli. They declared no comprehension problems.

Analysis

Analysis as in Experiment 1. We eliminated the outliers of the RTs, i.e. about 0.3% of the total data.

Results

doi.org/10.1177/1367006917709094

The mean RTs in this experiment are as follows: Version a = 5360.1ms, Version b = 5195.7ms, Version c = 6376.8ms and Version d = 6355.7ms; see Figure 4. This shows that the Turkish verbs are processed similarly when followed by Romani or by Turkish material.

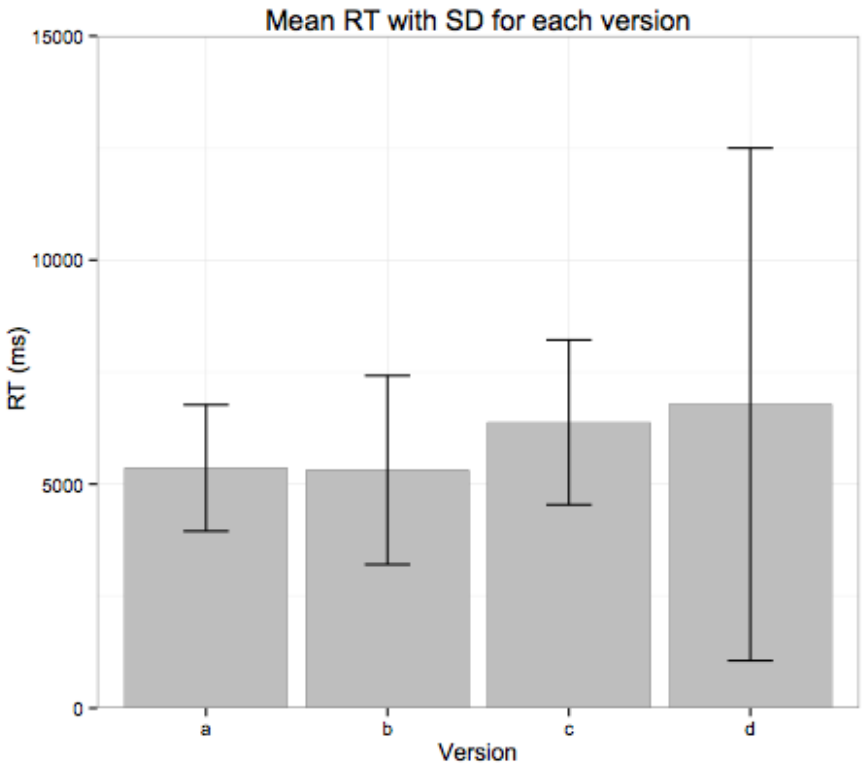


Figure 4. Mean reaction times (RT) in milliseconds (ms) to different versions of the sentences (a = mixed Romani-Turkish, b = all Turkish, c = Romani with Turkish codeswitching, d = Romani with Turkish borrowings) in Experiment 2 (word recognition)

The ANOVA shows that there are significant 'Version' differences (logLikelihood is -5314.3, $\chi 2_3 = 79.9$, $p < .001$). Version a, the mixed Romani-Turkish version, is significantly shorter than the two ecologically non-valid versions, i.e. Version c (t = 5.9) and Version d (t = 6.1). The all-Turkish version, Version b, is also significantly shorter than the two ecologically non-valid ones, i.e. Version c (t = 6.9) and Version d (t=7.1). However, there are no significant differences between the mixed Romani-Turkish version (Versions a) and the all-Turkish version (Version b) (t<1) nor between the two versions that are the least natural in the community, Versions c and d (t<1).

doi.org/10.1177/1367006917709094

As in Experiment 1, we conducted an analysis of the RTs by taking into consideration the preferred language in which the verbs are used in natural speech, i.e. both Romani and Turkish ('variable'), Romani, or Turkish. The results are shown in Table 4.

Table 4. Experiment 2: Mean RTs of different versions of the sentences (a = mixed Romani-Turkish, b = all Turkish, c = Romani with Turkish codeswitching, d = Romani with Turkish borrowings) with different language preferences for the verbs as observed in a free-speech corpus (variable, Romani, Turkish)

	Variable	**Romani**	**Turkish**
a	5615.0	5397.8	5055.8
b	5439.2	4691.4	5494.3
c	6289.8	6229.5	6653.3
d	5300.4	7838.2	5589.8

The ANOVA shows that there are significant 'Version' * 'Language preference of verb in the corpus' interactions (logLikelihood is 5224.3, $\chi 2_6 = 13.1$, $p < .04$). For verbs that are either 'Variable' (Romani or Turkish) or 'Turkish' in the spontaneous corpus, Version c has the longest RT (ts>2), but for verbs that are more frequently used in 'Romani' in natural speech, version d has the longest RT (ts>1.9). See Figure 5.

doi.org/10.1177/1367006917709094

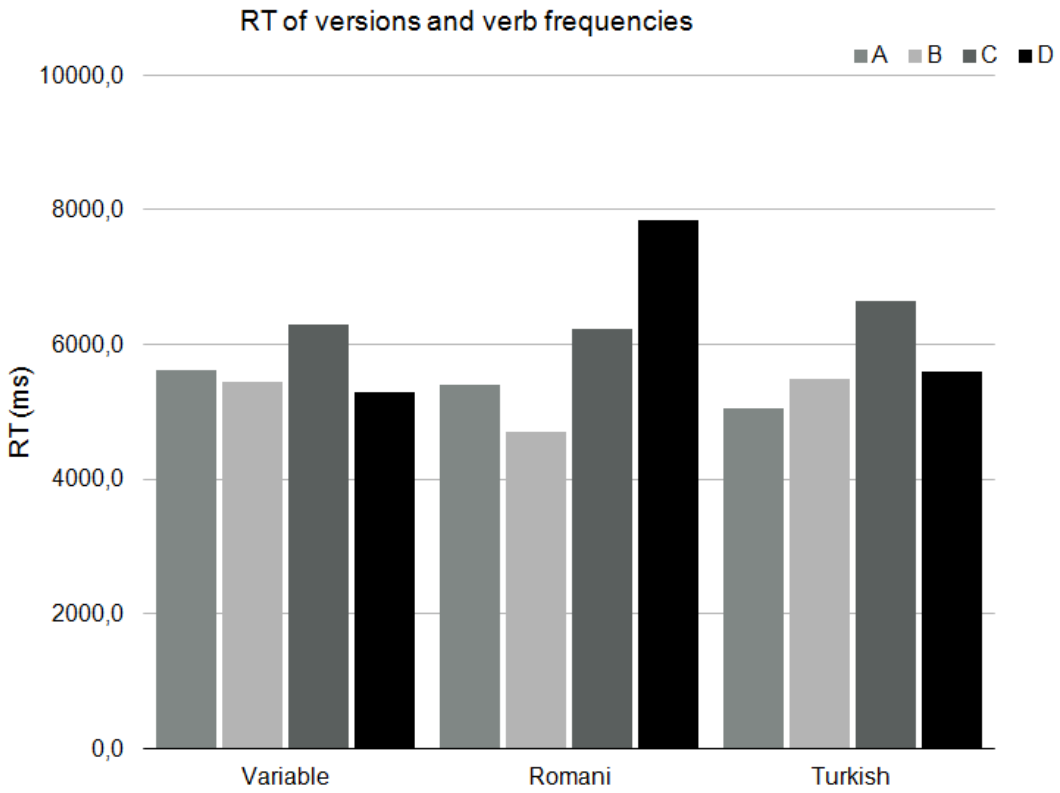


Figure 5. Mean reaction times (RT) in milliseconds (ms) of different versions of the sentences (a = mixed Romani-Turkish, b = all Turkish, c = Romani with Turkish codeswitching, d = Romani with Turkish borrowings) with different language preferences for the verbs as observed in a free-speech corpus (variable, Romani, Turkish) in Experiment 2 (word recognition)

Discussion

Results of Experiment 2 show that participants responded as fast for the Turkish verbs in a unilingual Turkish environment (Version b) as in mixed Romani-Turkish sentences (Version a). This indicates that the morphologically non-integrated Turkish verbs are processed similarly whether they are surrounded by Romani or by Turkish material.

Summary and general discussion

This paper offers experimental evidence from an under-investigated type of language contact setting. Indeed, the Romani population under study is atypical when compared to the populations generally examined in studies of bilingual processing for using the two languages in contact, Romani and Turkish, for over a century in a manner that involves systematic

language mixing. Moreover, the present study involves participants who have simultaneously acquired the two languages in contact, Romani and Turkish, without any exposure to formal learning for either language, contrasting with most studies on language processing that involve languages with written traditions and exposure to formal teaching at least in one of these languages.

In this paper, two research questions were addressed. The first was whether codeswitching is associated to high processing costs when comprehenders are frequently exposed to it. The second research question was whether verbs inserted with their native morphology are processed more slowly when they appear in a unilingual sentence environment or in a mixed sentence environment.

Taken together, the findings from the two experiments show that ecological validity and frequency of use in natural conversations play a major role. Results from Experiment 1 showed that mixed Romani-Turkish sentences are more costly than sentences in unilingual Turkish speech. In accordance with other psycholinguistic studies on language switching, these results indicate that the comprehension of unilingual speech is less costly than the comprehension of speech involving language switching. However, a closer look at the results revealed that the mixed Romani-Turkish sentences are processed faster than all-Turkish sentences when they contain lexical items that respect the language preference found in the spontaneous speech.

Experiment 2 showed that Turkish verbs with Turkish verb morphology are processed similarly whether in a mixed Romani-Turkish sentence or in a Turkish unilingual sentence. It therefore appears that highly proficient, simultaneous bilinguals do not experience any difficulties in inhibiting the L1 in order to process a codeswitching insertion from L2, even when this takes place in the context of a demanding comprehension task involving frequent language switches between sentences.

Our study therefore underlines the role of ecological validity in understanding processing of language switching and lends support to the usage-based approaches of codeswitching and bilingual processing. More broadly, this study strengthens the models that relate processing costs to unexepectedness effects and prediction errors. Finally, our study calls for more psycholinguistic research to be conducted with under-studied populations, involving communities with different communicational habits than those found in well-described populations in order to better comprehend the bilingual brain and cognition.

doi.org/10.1177/1367006917709094

Acknowledgments

xxx

Abbreviations

1, 2, 3 first, second, third person; AOR aorist; COM comitative; DEM demonstrative; FOC focus; Q question marker; IMPF imperfective; LVM loan verb marker; NFUT non-future; NOM nominative; OBL oblique; OPT optative; PL plural; PRET preterit; PST past; SG singular.

Appendix A

This Appendix contains the sentences that were used as auditory stimuli in both experiments. They are presented here in English but four different versions of them were presented to the participants involving Romani and Turkish. The nouns and the prime verbs that were consistently Turkish in the original are presented in italics.

Prime verbs that are either Romani or Turkish in the free-speech corpus:

1. The *strangers* came today. They *waited* for Sabiha in the village.
2. Their *daily wage* is good. But they *work* 10 hours per day.
3. The *monkeys* were very smart. They *understood* very well whatever I told them.
4. The *soldiers* were very brave. They *waited* for days hiding in the woods.
5. The *neighbours* were having parties very often. They *drank* and danced until late.

Prime verbs that are more frequently Turkish in the free-speech corpus:

6. The *couple* was very young. They *married* as soon as they turned fourteen.
7. The *enemies* were not so strong. They *returned* home after a few days.
8. The *tales* were quite popular among the children. They *read* new ones every day.
9. Their *hearts* beat waiting for the news. They *thought* that he might be dead.
10. The *kings* have to be wise. They *write* all the laws.

Prime verbs that are more frequently Romani in the free-speech corpus:

11. The *door* finally opened. They all *came* home for dinner.
12. The *doctors* didn't say a word. They *went* straight at home to look at my brother.
13. The *anger* made them red. They *left* the house in silence.
14. The *country* was small but wealthy. They *got* whatever they wanted.

15. The *cars* were old. They *did/fixed* them by themselves when they broke.
16. The *hammers* were rusty for some time. They *put* them at the storage room.

Appendix B

This Appendix contains the visual stimuli associated with the sentence 'The neighbours were having parties very often. They drank and danced until late.'





References

Abutalebi, J., & Green, D. (2016). Neuroimaging of language control in bilinguals: neural adaptation and reserve. *Bilingualism: Language and Cognition*, *19*, 689–698.

Adamou, E. (2010). Bilingual Speech and Language Ecology in Greek Thrace: Romani and Pomak in Contact with Turkish. *Language in Society*, *39*, 147–171.

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

Adamou, E., & Granqvist, K. (2015). Unevenly mixed Romani languages. *International Journal of Bilingualism*, *19*, 525–547.

Alvarez, R. P., Holcomb, P. J., & Grainger, J. (2003). Accessing word meaning in two languages: an event-related brain potential study of beginning bilinguals. *Brain and Language*, *87*, 290–304.

Backus, A. M. (2015). A usage-based approach to codeswitching: The need for reconciling structure and function. *Code-switching Between Structural and Sociolinguistic Perspectives*, 19–37. Berlin: Mouton de Gruyter.

Bakker, P. (2001). Romani in Europe. In Extra G. & D. Gorder (Eds.), *The other languages of Europe: Demographic, sociolinguistic and educational perspectives* (pp. 293–313). Cleveland: Multilingual Matters.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67,* 1–48.

Bialystok, E., Craik, F., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control. *Psychology and Aging, 19,* 290–303.

Calvo, A., & Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition, 130,* 278–288.

Chan, M-C., Chau, H. L., & Hoosain, R. (1983). Input/output switch in bilingual code switching. *Journal of Psycholinguistic Research, 12,* 407–416.

Cheng, Y.-L., & Howard, D. (2008). The time cost of a mixed-language processing: an investigation. *International Journal of Bilingualism, 12*, 209–222.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language, 50,* 491-511.

Dahan, D., & Tanenhouse, M. (2004). Continuous Mapping from Sound to Meaning in Spoken-Language Comprehension: Immediate Effects of Verb-Based Thematic Constraints. *Journal of Experimental Psychology, 30*, 498–513.

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

Friedman, V. (2013). Compartmentalized Grammar: The Variable (Non)-Integration of Turkish Verbal Conjugation in Romani Dialects. *Romani Studies 23*, 107–120.

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language.* Oxford: Oxford University Press.

Grainger, J., & Beauvillain, C. (1987). Language blocking and lexical access in bilinguals. *Quarterly Journal of Experimental Psychology, A(39),* 295–319.

Green, D. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*, 67–81.

Green, D. (2011). Language control in different contexts: the behavioural ecology of bilingual speakers. *Frontiers in Psychology, 2*. Retrieved from http://journal.frontiersin.org/article/10.3389/fpsyg.2011.00103/full

Green, D., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology, 25,* 515–530.

Gullifer, J., Kroll, J. F., & Dussias, P. E. (2013). When language switching has no apparent cost: Lexical access in sentence context. *Frontiers in Psychology, 4,* 1–13.

Herring, J. R., Deuchar, M., Parafita Couto, M. C., & Moro Quintanilla, M. (2010). 'I saw the madre': evaluating predictions about codeswitched determiner-noun sequences using Spanish-English and Welsh-English data. *International Journal of Bilingual Education and Bilingualism, 13,* 553–573.

Hofweber, J., Marinis, T., & Treffers-Daller, J. (in press) Effects of dense code-switching on executive control. *Linguistic Approaches to Bilingualism*.

Ibáñez, A., Macizo, P., & Bajo, M. T. (2010). Language access and language selection in professional translators. *Acta Psychologica, 135,* 257–266.

Jackson, G. M., Swainson, R., Cunnington, R. & Jackson, S. R. (2001). ERP correlates of executive control during reported language switching. *Bilingualism: Language and Cognition, 4*, 169–178.

Jackson, G. M., Swainson, R., Mullin, A., Cunnington, R., & Jackson, S. R. (2004). ERP correlates of a receptive language-switching task. *The Quarterly Journal of Experimental Psychology*, *57A(2)*, 223–240.

Jaeger, F., & Snider, N. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition, 127,* 57–83.

Kilborn, K., & Moss, H. (1996). Word monitoring. *Language and Cognitive Processes, 11*, 689–694.

Lipski, J. (2016a). Palenquero and Spanish: A first psycholinguistic exploration. *Journal of Pidgin and Creole Languages 31,* 42–81.

Lipski, J. (2016b). Language switching constraints: more than syntax? Data from Media Lengua. *Bilingualism: Language and Cognition.* Retrieved from CJO2016. doi:10.1017/S1366728916000468.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*, 314–324.

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

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology, 4*. http://dx.doi.org/10.3389/fpsyg.2013.00226

Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language, 40,* 25–40.

Monsell, S. (2003). Task switching. *Trends in Cognitive Science, 7*, 134–140.

Moreno, E. M., Federmeier, K. D., & Kutas, M. (2002). Switching languages, switching palabras (words): an electropysiological study of code switching. *Brain and Language, 80*, 188–207.

Mosca, M., & Clahsen, H. (2016). Examining language switching in bilinguals: The role of preparation time. *Bilingualism: Language and Cognition, 19,* 415–424.

O'Shannessy, C., & Meakins, F. (2012). Comprehension of competing argument marking systems in two Australian mixed languages. *Bilingualism: Language and Cognition, 15*, 378–396.

Paap, K., & Greenberg, Z. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology, 66,* 232–258.

Poplack, S., & Dion, N. (2012). Myths and facts about loanword development. *Language Variation and Change, 24*, 279–315.

Proverbio, A. M., Leoni, G., & Zani, A. (2004). Language switching mechanisms in simultaneous interpreters: An ERP study. *Neuropsychologia*, *42*, 1636–1656.

R Core Team. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Rusakov, A. (2001). The North Russian Romani dialect: interference and code switching, In Dahl Ö. & M. Koptjevskaja-Tamm (Eds.), *Circum-Baltic languages* (pp. 313–338). Amsterdam & Philadelphia: Benjamins.

Soares, C., & Grosjean, F. (1984). Bilinguals in a monolingual and a bilingual speech mode: The effect on lexical access. *Memory and Cognition, 12*, 380–386.

Thomas, M. S. C., & Allport, A. (2000). Language switching costs in bilingual visual word recognition. *Journal of Memory and Language, 43*, 44–66.

Tomasselo, M. (2003). *Constructing a language: A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.

Wells, J., Christiansen, M., Race, D., Acheson, D., & MacDonald, M. C. (2009). Experience and sentence comprehension: Statistical learning and relative clause comprehension. *Cognitive Psychology, 58*, 250–271.

Wohlgemuth, J. (2009). *A typology of verbal borrowings*. Berlin & New York: Mouton de Gruyter.