CLS 44: 
THE PANELS

2008

Proceedings from the Panels of the Forty-fourth Annual Meeting of the Chicago Linguistic Society:

Volume 44-2

Edited by:
Max Bane
Juan José Bueno Holle
Thomas Grano
April Lynn Grotberg
Yaron McNabb

1 Introduction
Romani, which has a millennia-long history of stable multilingual contact, is problematic for theories that seek to identify structural constraints on codeswitching and, related to that, on structural borrowing. In this paper, I examine a set of such phenomena, namely the switching into and eventual copying (Johansson 2002) of Turkish conjugations into various Romani dialects in the Balkans, a phenomenon I call *code compartmentalization*. On the one hand, concepts such as Myers-Scotton’s (2002) *composite matrix language* (assuming imperfect language learning or competition for dominance) or her *matrix language turnover hypothesis* (arguing that composite matrices result from imperfect language mastery and/or language shift in progress), as well as other attempts at distinctions and constraints (Myers-Scotton 2005), fail to address problems posed by stable multilingualism as found in the Balkans in general and in some dialects of Romani in particular. On the other hand, the fact that precisely the Turkish verb is switched using both Romani and Turkish word and morpheme orders and creating new hybrid constructions poses problems for theories attempting to constrain single mixed grammars (MacSwan 2004). The borrowing of a Turkish second plural inflection into some of these Romani dialects is also worth noting here as an example of borrowed morphology that is sometimes said not to occur (most recently Labov 2007:348-349). In this paper, I bring together data from a variety of sources on Turkish verbs in Romani dialects to show that the distribution of borrowed features is hierarchical. I also argue that while the matrix/embedded language opposition, on the one hand, and the single mixed-grammar concept, on the other, are both useful in explaining some codeswitching behavior under certain sociolinguistic circumstances, in situations such as those found in the Balkans—prolonged, stable multilingualism with conditions under which social dominance is balanced by other linguistic ideologies—one must acknowledge that speakers have creative access to two grammars without recourse to puristic notions such as *imperfection* or claims that bilingualism consists only in the combination of discrete monolingual grammars.

In studying mixed languages, Bakker (1997) uses the term *interwining* to describe the integration of discrete grammatical and lexical modules from different languages to form a new code, e.g., Michif or Copper Island Aleut. Mixed languages also raise difficulties for Myers-Scotton’s ideas, cited above. In this, I agree with MacSwan (2004:29) when he writes: “Nothing constrains code switching” although I have reservations about his proviso “apart from the requirements of the mixed grammars.” At issue is whether we are forbidden to posit two individual grammars creatively in contact in the mind of the speaker.
with the potential to produce a different grammar (see also Auer 2007). Based on the data presented here, I would argue that we are not so forbidden. In the case of Turkish and Romani, we are dealing with special types of accommodation and not with incipient or partial language shift. This is demonstrated in dialects such as Ajia Varvara (Iglia 1996), where speakers no longer know Turkish but still use Turkish conjugations. At the same time, the Ajia Varvara code is no longer “mixed,” although it was at one time.

As a phenomenon, Turkish conjugation in Romani is the opposite of Bakker’s intertwining. The grammatical module is the same (verbal conjugation), but the paradigms are segregated according to source (certain Turkish verbs with Turkish grammar, all other verbs with Romani grammar). It is this adaptive strategy of integrating non-native grammar that I call code compartmentalization. Romani code compartmentalization can be attributed to a combination of sociolinguistic and language-ideological factors: Romani-speaking adults are always multilingual; traditional Romani cultural norms place a high value on both multilingualism and Romani language maintenance and ethnic differentiation.

2 Romani inflection

Virtually all Romani dialects display compartmentalization in the nominal system: all substantives that entered Romani up through the earliest contact with Byzantine Greek (presumably in Asia Minor) are declined with native stem-vowels, those entering subsequently (basically, in Europe) use other stem-vowels or attach directly to the nominative stem. The two types are illustrated in Table 1.

<table>
<thead>
<tr>
<th>Thematic, Oikoceltic or Asian</th>
<th>Athematic, Xenoceltic, or European</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘vog’ ‘soul’</td>
<td>dat. vogeskë</td>
</tr>
<tr>
<td>Chove ‘Romani boy’</td>
<td>dat. Chaveskë</td>
</tr>
</tbody>
</table>

Table 1: Basic patterns of Romani declension.

Thus, during a period of intense Romani-Greek bilingualism, speakers created a mechanism for segmenting newer loans from what was perceived as older inherited vocabulary.

In verbal categories, the Romani tense-aspect system is remarkably resilient and pan-Romani, while parts of the modal system are quite open to contact-induced change. The basic oppositions in the Romani tense-aspect system can be seen in Table 2.

For marking modality, all Romani dialects have the subjunctive clitic țe, which introduces optatives, conditionals, and verbs subordinated to ‘want’, ‘begin’, etc. These structures parallel the analytic subjunctives of the other Balkan languages that replaced older infinitives. In the Romani dialects of the Balkans—coinciding roughly with the limits of the Balkan Sprachbund and not extending beyond it—the future is normally formed by means of an invariant particle based on the root kam- ‘want’ (usually ka), plus the short present (i.e., without final -a, but see note 10). Ko plus imperfect forms an irreal conditional (also a Balkanism).

3 Romani dialects with Turkish conjugations

Turkish conjugations occur in North and South Balkan and South Vlah dialects in regions where there was close contact with Turkish throughout the Ottoman period, contact which may or may not continue into the present. The phenomenon, however, is not present in all such dialects. The specific dialects that we shall consider here are given in the Table 3. Bold faced letters are those used for the abbreviations in Table 5 and in the examples. Dialects in angled brackets are not specified in RMS (2001-2005) as having Turkish conjugations, but examples appear in the data. Unless otherwise specified, data are drawn from the RMS. Sepeći is based on Coeh & Henischink (1999), Fuduzi is based on Ivanov (2000), Ajia Varvara on Iglia (1996), and Komotini (Turkish Gümlüce) on Adamou (2006). Data from Draganova (2003) and Kaspičan, Haskovo.

1 In some cases, the line between code-copying and codeswitching is difficult to demarcate. Nonetheless, while the occasional codeswitched Turkish verb form can occur in other Balkan languages (generally with at least one other lexical item of Turkish origin), the borrowing of complete conjugations is a phenomenon specific to Romani.
Tsarevo, and Stefan Karadžov in eastern Bulgaria were also consulted but are not included in the tables.

<table>
<thead>
<tr>
<th>SOUTH BALKAN</th>
<th>NORTH BALKAN</th>
<th>SOUTH VLAX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crimean</td>
<td>Kaspičan Xoraxani</td>
<td>Aija Varvara</td>
</tr>
<tr>
<td>(Rostov, S Russia)</td>
<td>(Svilenovo, NE Bulgaria)</td>
<td>(Athens, Greece)</td>
</tr>
<tr>
<td>Sepeči</td>
<td>Sliven Nange</td>
<td>Valiž Bol</td>
</tr>
<tr>
<td>(Izmir, Turkey)</td>
<td>(=Gradski, SE Bulgaria)</td>
<td>(=Laço, NE Bulgaria)</td>
</tr>
<tr>
<td>Fiorina Arli</td>
<td>Fuładži</td>
<td>Sindel Kałbărdzi</td>
</tr>
<tr>
<td>(NW Greece)</td>
<td>(Haskovo, SE Bulgaria)</td>
<td>(NE Bulgaria)</td>
</tr>
<tr>
<td>Prizren</td>
<td>Sliven Muzikanti</td>
<td>Varna Kałajđi</td>
</tr>
<tr>
<td>(SW Kosovo)</td>
<td>(=Trakjski Kałajđi, Vlaxoria, NE Bulgaria)</td>
<td></td>
</tr>
<tr>
<td>Spotori</td>
<td>Komotini</td>
<td></td>
</tr>
<tr>
<td>(=Kalajđi, SE Romania)</td>
<td>(NE Greece)</td>
<td></td>
</tr>
<tr>
<td>Varna Găzăzăno</td>
<td>(NE Bulgaria)</td>
<td></td>
</tr>
<tr>
<td>Varna Burgăză</td>
<td>(= Parpuli, NE Bulgaria)</td>
<td></td>
</tr>
<tr>
<td>&lt;Pazardžik Malo Konare&gt;</td>
<td>(SE Bulgaria)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Romani dialects with Turkish conjugation.

Matras (2002:134-135) notes that the use of an unadapted predication has the potential to create ambiguity concerning the matrix language (in Myers-Scotton's terms). I argue that the occurrence of these paradigms and their position in Romani grammar demonstrates that predication cannot be taken as the determining of a matrix, neither is it necessarily the signal of a codeswitching. Matras also relates the paradigms to the sociolinguistic acceptance of full and prolonged bilingualism, which certainly can account for the origin but not the continuity of such paradigms, e.g., in the dialect of Aija Varvara, where knowledge of Turkish has been lost, but Turkish paradigms persist. A similar situation holds in Fiorina Arli, and these are both in marked contrast to the situation in Prizren, where Turkish has remained as a prestige language, but the use of Turkish conjugations among the younger generation is limited to quotations (RMS 2001-2005). This last may well be due to changes in language status and social relations in recent years, during which the Romani community in Prizren was one of the few in Kosovo to escape the depredations that followed the 1999 NATO bombings. We can also note that while the highest degree of Turkish conjugational integration into Romani occurs in dialects in contact with Turkish in Bulgaria, the Sepeči dialect spoken in İzmir, Turkey is among the most limited. This may well have to do with the relative position of Turkish in the two countries. In Turkey, it is the language of the state and of the dominant population, whereas in Bulgaria, Turkish is not the language of the state or the majority, but it is the prestige language among Muslims. In Bulgaria, as in Turkey, Macedonia, Albania, and Kosovo, the majority of Roma are Muslims, whereas in Greece they are Christians, except in Western Thrace, where Muslims were allowed to remain after the exchange of populations in 1923.

4 Turkish verbal inflections used in Romani
Seven of the eight Turkish tense/aspect/mood (TAM) markers (in Lewis 1967:136 terms), or of the nine position-three inflections (in Göksel & Kerslake 2005:79-83), are attested in Romani-Turkish conjugation, as is the verbal negator, the infinitive marker, and the copular preterite marker idı. Table 4 summarizes the affixes that can occur, although not all occur in all dialects.5

| (ı)yor | progressive present | (A)ı-r | general present |
| (-ı)DI | preterite | (-ı)mıJs | perfect |
| (ı)A | optative or subjunctive | AcaK | future |
| (-ı)ık | conditional | mAK | infinitive |
| m4 (2/3 aor. m4) | verbal negator | |

Table 4: Turkish tense-aspect-mood markers.6

Of the TAM markers, the simple preterite in -DI occurs in all Romani dialects with Turkish conjugation; the two presents occur in most dialects. Ccopular idı is next in order of frequency followed by the optative or subjunctive. Less frequent is the perfect, which occurs in two usages, one participial and the other finite (usually evidential). The participial use is not of interest to us here since it can be

5The necessitive or obligative in m4/I and the locitive infinite (-m4ılA) that Göksel and Kerslake's (2005:83) label imperfective, excluded by Lewis (1967) from this level of marker and limited to formal contexts in Turkish, are not attested. Copularative mıJs and ıSt are as well as the general modal particle -DIe are likewise absent. For detailed discussion of these markers in Turkish, see the aforementioned references. Terminology for Turkish TAM markers is not uniform, but the details need not concern us here.

6The (ı) indicates a marker that can occur as a finite inflection, or as a stative or free-standing copula; all other markers occur only as inflections. Capital letters indicate phonemes subject to vowel harmony or voicing alternations.
treated as a derived rather than an inflected form, but the finite perfect has important implications for the systems of the relevant dialects. Finally, the Turkish future, the verbal negator -ma-, the infinitive marker and the conditional marker occur in a relatively small number of dialects and raise issues of differentiating compartmentalization from codewitching and code mixing.7 The Turkish verbal negator occurs between the root and TAM markers and can occur in Turkish-Romani codeswitches with or without a Romani negator, as we will see below. Such occurrences are among the problems Turkish conjugation in Romani presents for constraint-oriented theories of codewitching.

5 Romani-Turkish examples8

As indicated above, all dialects with Turkish conjugation have at least the preterite, which is the TAM category (and marker, insofar as Romani also uses a dental that is sometimes realized as a stop, see Table 2) closest to the Romani simple preterite:

(1) Phenel ke rone bejen-di-k.
    say:PRS:3SG that husband:ACC like:PST-1PL
    ‘He says we liked your husband.’ (Ko)

Next in order of occurrence and frequency is some form of the present tense. Among the dialects that have only the Turkish present in addition to the Turkish preterite, some have only the general present (SP, AV), some have only the progressive (SM), and some have both (SK, perhaps PA). In general, dialects with additional TAM categories have both presents, although sometimes these dialects lack the progressive (Fu).

(2) Voj adelas paš lende taj konuš-ar lenge.
    she stand:IMPF by them:LOC and speak:PRS them:DAT
    ‘She was standing between them and talking to them.’ (VD 649)

(3) Vov konuš-uor gadbor but ta kerda
    he talk:PRS:PROG so much very and made:PST:3SG
    mam te bristavar so mangov.
    me:ACC SBJV forget:PRS:1SG what want:PRS:1SG
    ‘He talked so much, he made me forget what I wanted.’ (VD 785)

7Draganova (2005:95-96) reports the conditional in -sa for Tsarevo, Haskovo, and Kaspoin, but it does not occur in the RMS or Funzadi materials.
8Turkish words are given in bold face, words in other non-Romani languages are underlined, but only if they arguably represent invocations from a current contact language. Examples from RMS 2001-2005 are given with reference numbers to the entries in that database.

In those dialects with both presents, the general present occurs in future and subjunctive clauses in te and ka, respectively, although progressive presents can also occur in these environments.

(4) Naj man kaničk protrive te jardzmu-ar-am tuke.
    Not.is me:ACC nothing against SBJV help:PRS:1SG you:DAT
    ‘I don’t mind helping you.’ (SK 710)

(5) Odova ayla mere kevete zą da dą
    that:M come:3SG:IMPF my:OBL house:LOC for SBJV
    te konuš-ui mansa.
    SBJV speak:PRS:PROG me:INSTR
    ‘He came to my house in order to speak with me.’ (SK 440)

(6) Kan bekli-er-im tut angli carkev.
    FUT wait:PRS:1SG you:ACC before:F church
    ‘I will wait for you in front of the church.’ (SK 1009)

(7) Kan bekli-lor-am tut andi kangiri.
    FUT wait:PRS:PROG-1SG you:ACC on:F church
    ‘I will wait for you in front of the church.’ (VD 1009)

Among those dialects with both present and preterite Turkish inflection, we can distinguish those which also have cliticized ıd and those that do not. Dialects without cliticized ıd such as AV and Sp form imperfects and pluperfects using Romani agglutinative material, e.g., ček-di-sas ‘he had gone out’ (Sp 389). Unlike Turkish, which cliticizes ıd to either present, Romani only adds it to the general present (cf. Draganova 2005:94). Examples (8) and (9), however, show that both general present plus ıd and preterite alone can have imperfect meaning:

(8) But manuša sa kaj kapono, samo odova
    many person:PL were at party only that:M
    manuš anı tan-ar-do ‘meri dude
    person which knew-PRS:PST my:ACC father
    odova konuš-ar-di mansa.
    that:M talk:PRS:PST me:INS
    ‘There were many people at the party, but only the man who my father talked to me.’ (SK 389)

Note that this is an abraded form of the progressive present and not a present in -y characteristic of West Rumelian Turkish.
While the two presents of Turkish represent a relatively small deviation from the single present of Romani—all are inflected and the Romani present expresses both sets of Turkish aspectual distinctions—and the cliticization of Turkish *idi* parallels the use of Romani *as* or *sine*, the Turkish optative-subjunctive represents a major challenge to the Romani system insofar as the Turkish optative-subjunctive is synthetic whereas the Romani optative-subjunctive is constructed analytically using the particle *te* and the present tense.\(^{10}\) Example (10) shows an optative without *te*, while example (11) shows an optative with *te*.

(10) *Ov avijas me keres te konuš-sun mons.*

*He came:IMPF my:OBL house:LOC SBIV speak-OPT:3SG me:INS*

('He came to my house to talk to me.' (SN 440))

(11) *Odja ep bisterla kapa-sun o dar.*

*that:F always forget:3SG:PRS close-OPT:3SG the door*

('She always forgets to close the door.' (SK 907))

In both these examples, the Turkish form is following a Romani morphosyntax rule; Turkish itself would require an infinitive.\(^{11}\) In the case of (10), both Turkish and Romani morphological order is present, while in (11) the switch uses a purely Turkish construction in a Romani clause.

The expression of futurity with Turkish verbs in Romani has three possibilities: Romani future marker + Turkish present as seen in examples (6) and (7) above, Romani future marker + Turkish optative as in (12) and Turkish future as in (13).

(12) *Kidal kam dišir-cim e dasengo thus kam change-OPT:1PL the Bulgarian:PL:GEN*

Here again, both morpheme orders occur. Examples (15) and (16) are unusual in having infinitives where Romani (including Ks and VG) would normally require a *te*-clause. Both the dialects concerned also have *te*-optative elsewhere, and it can be argued that the preceding Turkish word in both cases has precipitated a codeswitch:

(13) *Ame naši dón-dle-s Ži kana we can't return-FUT-1PL up.to when dooral-mu-fa oda. get.well-NEG-PRS that:F*

('We cannot go back until she gets well.' (VG 384))

In (6) and (7), the Turkish is simply a translation for the Romani, but (12) begins with a Romani future marker then switches to a Turkish optative, which calques an older Romani (and Balkan) construction using an optative subjunctive marker after the future. The morpheme order is thus both Romani and Turkish.

Although most dialects choose either a Romani or Turkish negator for Turkish verbs, example (14) shows both Romani and Turkish negation.

(14) *Sja ek romni ek ta banza rom-a ani was one Rom:F:SG one and Some Rom-M:PL which me na tana-nu-dam.*

*1 NEG know-PRS-NEG-PST*

('There was a Gypsy woman and a few Gypsy men who I did not know.' (KX 560))

Here again, both morpheme orders occur. Examples (15) and (16) are unusual in having infinitives where Romani (including Ks and VG) would normally require a *te*-clause. Both the dialects concerned also have *te*-optative elsewhere, and it can be argued that the preceding Turkish word in both cases has precipitated a codeswitch:

(15) *Mi phen bašia-do baar-ma-a kana my:F sister scream-INF-DAT when tharde amaro bhor. burn:PST:3PL our:M house*

('My sister began to scream when they burned down our house.' (VG759))

(16) *Odva džanna trin dil-dle konul-ma-a.*

*that:M know-PRS three tongue-ADV speak-INF-DAT*

('He knows [how] to speak three languages.' (KX 894))

In (15), the word order is Romani, whereas in (16) the main clause and beginning of the subordinate clause begins in Romani but the subordinate clause then finishes with Turkish word order.
Example (18) shows a contrast using a Turkish preterite and a perfect of the evidential type that would occur in Turkish but not Romani.

(17) Oda artıkan git-mis ame kana git-ti-k
that:M already go-PST-PRF we when go-PST-1PL
kaj kher.
at house
‘He had already gone before we got there.’ (V G 389)

Here, Turkish normally requires the perfect in the first clause and the preterite in the second clause because, as is clear from the context, the speaker did not witness the action. Romani would require the same preterite in both clauses.12

Example (18) illustrates an effect of Turkish on Romani that I have not seen elsewhere. The perfect is used with Turkish verbs in contexts where it would be expected in Turkish, while with native Romani past tense verbs in the same contexts, the particle berim is normally placed immediately after the verb or its pronominal object if that follows the verb.13

(18) Oda kana dikjas la berim don-mal
that:M when see/PST.3SG her:ACC berim freeze-PST-PRF
pe temeste.
in place:LOC
‘When he saw her [berim], he froze in his tracks [lit. ‘place’]’(Fu)

Norbert Bořetzky (pers. com.), who called the phenomenon to my attention, considers the origin of the particle to have come from Turkish hegil ‘sure, known’, but I consider the postposition berim ‘according to this’ to be another possible candidate, as it is phonologically closer and semantically could carry a notion of secondary attribution appropriate to the evidential-like meanings of the Turkish perfect. The Futadži phenomenon is unique in that it is a native development co-occurring with an integrated foreign (Turkish) category.

6 Hierarchies of Turkish conjugations in Romani
Table 5 summarizes the types of paradigms that occur in the Romani dialects given in Table 3.

---

12 The details of the interaction of evidentiality and taxis are more complex than can be presented here (see Friedman 1978), but it suffices to say that at issue is the importation of a Turkish distinction within a Romani clause.

13 On rare occasion, berim occurs after a temporal expression in the verb phrase (5 out of 85 in a 7,000 word corpus), three times after a present tense (but always when used with historical present meaning), once each after a subject and a locative expression.
Romani dialects of the Balkans. In some cases we get both or mixed morpheme orders, in others we get the penetration of the absent category into Romani.

There is a correlation between synthetic future, negative marking, and, perhaps, infinitive marking (if this last is not strictly a code-switch). The synthetic nature of the Turkish future is alien to the structure of Romani in the Balkans, which is distinguished by a Balkan type of future using an invariant clitic particle, and once the Turkish future is integrated, the way is open for infixal negation, and perhaps even the infinitive as well. Although these last features are characteristic of dialects whose speakers are still actively bilingual in Turkish, such bilingualism is also found among speakers of dialects that do not have such a deep penetration (heavy copying) from Turkish.

7 Excursus on borrowed morphology
A recent revival of the claim that morphology is not borrowed (e.g., Labov 2007:348-349) is also contradicted by the data from Romani dialects with Turkish conjugation. In certain Romani dialects of eastern Bulgaria that have Turkish conjugation, the 2Pl marker -iz is copied (with final devoicing and a generalization in vowel harmony) to the native Romani 2Pl preterite marker, and, in some cases, it is extended to the 1Pl preterite. The evidence that this extension is a Romani innovation is the fact that the Turkish dialects with which these Romani dialects are in contact have the 1Pl ending -ê (Elik & Matras 2006:135-136; Dall 1976:119-124). Table 6 gives the person endings of standard Turkish, which, in the case of eastern Bulgaria, correspond to the endings in the local dialects, and Table 7 gives the 1Pl and 2Pl present and preterite endings of most Romani of the Balkans, as well as the 12Pl in those dialects that have borrowed Turkish morphology.

<table>
<thead>
<tr>
<th>Unmarked</th>
<th>Preterite</th>
<th>Subjunctive</th>
</tr>
</thead>
<tbody>
<tr>
<td>-lm</td>
<td>-lz</td>
<td>-m</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-k</td>
</tr>
<tr>
<td>-iz</td>
<td>-shlz</td>
<td>-n</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-nlz</td>
</tr>
<tr>
<td>-ê</td>
<td>-lar</td>
<td>-Ø</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-lar</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-slm</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-sal</td>
</tr>
</tbody>
</table>

Table 6: Turkish person markers.

<table>
<thead>
<tr>
<th>Present</th>
<th>Preterite</th>
<th>KX&amp;VG&amp;VD</th>
<th>AV=12Pl</th>
<th>SK</th>
<th>C=2Pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>-at(s)</td>
<td>-am</td>
<td>am-at(s)</td>
<td>am-us</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>-ej[s]</td>
<td>-an</td>
<td>an-at(s)</td>
<td>an-us</td>
<td>an-at(s)</td>
<td>an-us</td>
</tr>
</tbody>
</table>

Table 7: Romani 12Pl person markers (Balkan and Vlax dialects).

In this context, we can also say something about phonology. It is sometimes claimed that phonological switches do not occur word internally (e.g., MacSwan 2004:390). Aside from the fact that such switches do, in fact, occur in real codeswitching situations (as do English definite articles with Spanish nouns in the speech of Spanish dominant bilinguals living in a mixed home environment), such claims fail to take into account the fact that in centuries-long, stable multilingual situations, phonologies can converge. Such is the case in some Balkan Turkish dialects, where the high back unrounded i is pronounced like the schwa of neighboring languages and z is subject to final devoicing, while speakers of those languages can produce the i of Turkish, which then can spread into native words.

8 Excursus on Turkish-Macedonian codeswitching
As noted above, the type of code compartmentalization found in the Romani dialects of the Balkans that are in contact with Turkish does not occur in the other Indo-European Balkan languages. Nonetheless, these languages, too, show interesting codeswitches that are problematic for theories of constraints on codeswitching. One example occurs in a corpus of nineteenth century folktales collected in Prilep in west central Macedonia by Marko Cepenkov. Sentence (19) occurs as the conclusion of a story that has been narrated in Macedonian.

(19) Demek od lepeška zelnik ne bidatat! to.say from cow.turd spinach pie not is
Turc-i-te vel-at: Sej kov-ar
Turk-PL-the say-3PL: PRES lineage determines-3sg:PRES
Čovek-ot.
person-the
"That is to say you can’t make a spinach pie with a cow pie! The Turks say: Lineage determines the person." (Cepenkov 1972:109)

Here, Cepenkov explicitly announces he is about to switch into Turkish, begins in Turkish (soy ‘lineage’), and the verb is unambiguously Turkish (kov- ‘determine’) with Turkish general present inflection (-or). The word order, however, is Macedonian (Turkish would have the verb final, especially in such a proverbial expression), and the final word is Macedonian (Čovek) followed by a Macedonian grammatical morpheme (the definite article -or) that is treated in various accounts as an enclitic or an inflection (its behavior displays elements of both). Moreover, Cepenkov’s dialect of Macedonian requires a clitic pronoun (in this instance, go) before the finite verb agreeing in gender-number-case function with the definite direct object. Both standard Turkish and the local dialect of Turkish lack such a rule. What we have here, then, is a sentence that begins in Turkish, switches to Macedonian inside the VP, with Macedonian word order, but Turkish rules for a definite direct object, with a Turkish verb and a Macedonian direct object.

14 But see Kostov (1973) for a related phenomenon.
9 Conclusions

The integration of Turkish conjugation into Romani grammar—or the segregation of Turkish verbs, depending on the viewpoint—is comparable only to the affect of Greek on Romani declension, albeit the Turkish phenomenon is not as widespread and shows varying degrees of pervasiveness among Romani dialects. Nonetheless, Turkish verbs in Romani constitute a unique paradigmatic class—in some instances, long after contact with Turkish has ceased. The hierarchy of integration follows expectable structural lines modified, however, by formal, pragmatic, and sociolinguistic factors. Although the phenomenon as a whole crosses Romani dialectological lines, heavier degrees of copying are concentrated in certain regions. Nonetheless, mere geography and/ or knowledge of Turkish cannot account for the dialectal differences, and thus sociolinguistic factors must be taken into account in classifying and explaining Romani dialect typology in the context of contact with Turkish. Moreover, the data themselves—being drawn as they are from a variety of dialects in which different levels and degrees of codeswitching and code compartmentalization are present—demonstrate that at times Romani speakers create new grammars using a combination of Romani and Turkish materials. Approaches that attempt to define such codeswitching phenomena in terms of two monolingual grammars or as somehow “constrained” in such a way that this or that combination is “ungrammatical” must either fail or else ignore real data.

As Auer & Muhamedova (2005:53) have argued using Stolt’s (1964) important data on Early Modern High German/Latin and their own Kazakh/Russian corpus:

The ‘dominance’ of one language over another in intrasentential code-mixing seems rather to be a matter of gradience and can mean different things in different situations of language contact. [...] [A] neat separation between matrix and embedded language is impossible.

In both cases, the two pairs of languages operate with similarly different prestige/solidarity factors, and in both cases, the multilingualism was/is relatively stable and extending over several (or more) generations, albeit in the former the second language had no native speakers at the time of the bilingualism. Auer & Muhamedova (2005:52) also make the point that by claiming that matrix language is an abstract contract dependent on morpheme order and late system morphology—which latter two, however, cannot be cleary established—Myers-Scotton (2002) in fact leaves us with no means to define the first concept without equating it with a monolingual code.

Myers-Scotton (1999:384-385) magisterially (cf. Myers-Scotton 1999:367) tells us: “As linguists, we should all be most interested in theory building because our proper goal is to arrive at explanations of linguistic structure.” There are others, however, who would point out that since we have descriptions of only a small fraction of the world’s languages, and that since the majority of the world’s languages will be extinct by the end of this century, as linguists we should be most interested in documenting the actually occurring phenomena of human languages while we still can, since theories based on a paucity of data are often falsified when more complete data sets are examined. Obviously both documentations and explanations of linguistic structure are of value in the advancement of knowledge (or “creation of new knowledge” as cutting-edge usage has it). Moreover, it is imperative that we understand language in its social context (e.g., Heller 2007). I would argue here, however, against a Procrustean embedding of the study of codeswitching into formalist matrices. At issue is not that which “can” or “cannot” occur, but that which actually does occur and the contexts in which it occurs. And we should also remember that absence of evidence is not evidence of absence. The variety of human creativity—in language use as in other endeavors—needs description as well as analysis. As Auer & Muhamedova (2005:53) point out: “bilingual talk cannot be analysed as a mixture of two monolingual codes”; the conditions for mixing are “not defined by the linguist looking at monolingual codes, but by the bilinguals themselves” and, we can add, a speaker can do anything that a linguist can think of, and more.

References


1 Introduction

Bilingualism is difficult to fully define, but an agreement on a general definition has been reached by researchers in this field: bilingual primarily means someone with possession of two languages (Mackey 1962, Valdes & Figueroa 1994, Wei 2000). In other words, bilingual speakers are by definition able to access two languages, either sequentially or simultaneously. Since both languages can be active simultaneously, even though one of the two languages may be partially activated, the consequence may be that during speech production planning (hereafter SPP), bilingual processing is a more complex phenomenon, and bilingual speakers sometimes produce sentences that contain units from more than one language.

Code-switching (hereafter CS) is a well-known phenomenon as a pattern of bilingual speakers’ language use. This language mixing is not considered as evidence of incomplete knowledge of two languages the bilingual speaker speaks any more. Rather, it is viewed as a communicative strategy among bilingual speakers, and many researchers agree that there are structural characteristics in CS. Another phenomenon which can be observed in bilingual speakers’ language use is bilingual slips of the tongue (hereafter BSOT). This one time on-line speech error has merely been looked at in the bilingual research field because any kind of language mixing has been understood under the rubric of CS.

Consider the following example, which is a part of conversation carried between two Korean-English bilingual speakers.

1. A: ku ke a) must-see movie ya kkok pwa.
   that thing -is definitely see
   “That is a must-see movie. You gotta see it.”
B: b) What’s so good about it?
A: coach-ka ku ai-wi c) [læ:næ]-lul
   coach-NOM that kid-POSS nonword-ACC
   illayha-ca understand-when
   “As soon as the coach understood the kid’s language…”
B: d) what?

This type of language switching frequently occurs in casual conversational

---

1Korean-English data given in this paper that are not marked as cited are from author’s own corpus.