Modality in Romani

Viktor Elšík (Charles University, Prague)
Yaron Matras (University of Manchester)

1. Introduction

Romani[1] (Indo-Aryan) is an inflectional language with a 'thin' layer of agglutinating morphology. Grammatical material is generally suffixed. Lexical roots that entered the language in the Byzantine period or later typically draw to some extent at least on Greek-derived morphology, which has remained productive since Byzantine contact for all subsequent loans from European languages. In its syntactic typology, Romani closely resembles typical Balkan formations, having undergone intense convergence with Greek, and later with other languages of the area. Romani is the only New Indo-Aryan language that relies exclusively on prepositions as analytic markers of semantic case roles, and the only Indo-Aryan language that possesses a definite article (pre-posed, as in Greek). Word order is flexible, generally alternating between verb-middle and verb-initial. Attributes are pre-posed. Subordinations are based on conjunctions, usually deriving from interrogatives. In complement clauses a distinction is made between factual and non-factual complements. Relative clauses are post-posed, usually introduced by an uninflected relativiser, and they usually require resumptive pronouns in the main clause.

The impact of diverse contact languages is a major feature differentiating the individual Romani dialects, and it is therefore difficult to generalise when discussing the structures of Romani. In the present contribution we try to exploit this structural diversity within Romani in order to draw some generalisations about the language based on a sample of dialects.[2] Our subsequent discussion is therefore devoted to general tendencies in Romani. Alongside internal pathways of grammaticalisation, we shall also make use of the opportunity offered by the Romani sample and pay special attention to grammatical borrowing and language convergence and their role in shaping the inventory of modality-related structures in the language. Our data examples draw on both published sources

---

[1] Information on the history of Romani and dialect differentiation within the language, accompanied by maps, examples, and sound samples, can be found on the website of the Manchester Romani Project: http://www.llc.manchester.ac.uk/Research/Projects/romani/. See also Matras (2002), Elšík & Matras (2006).

[2] See also Elšík & Matras (2006), Matras (2002), and the Romani Morpho-Syntax (RMS) database on the Manchester Romani Project website. The sample contains over 350 recordings and questionnaire elicitation of different Romani dialects from across Europe, collected between 2000-2007. We acknowledge support for our work during this period from the Arts and Humanities Research Council, the Economic and Social Research Council, and the Open Society Institute.

wherver indicated; where a source is not identified, data derive from our own ongoing survey of European Romani dialects, as part of the RMS (Romani Morpho-Syntax) Database project (see footnote 2, and see Manchester Romani Project). Much of these data were collected by questionnaire elicitation, which in some cases puts certain limitations on the discussion and the scope for interpretation.

Our present work may be understood as a kind of follow-up to Boretzky's (1996b) earlier work devoted to modals in Romani. In the present contribution, a vast amount of data is taken into consideration on dialects that had not been described and were in fact unknown in the research context at the time of Boretzky's publication. Our discussion is also anchored in a wider morphosyntactic context, one which takes into account not just the etymology of individual expressions of modality, but attempts a more typologically informed evaluation of modal constructions, inspired by models in grammaticalization theory.

2. Modals in Romani

Early Romani, the ancestor of all Romani dialects, can be reconstructed to have possessed three or four dedicated modals: a volition modal and two or three possibility modals. In all likelihood, there were no dedicated necessity modals, although periphrastic constructions that could express necessity must have been available. Today we find a bewildering variety of modals and modal constructions within Romani, especially in the domain of necessity and, to a lesser extent, possibility. Volition is clearly the most stable form of modality. There are several sources of dialect divergence with regard to modals. The majority of necessity modals, numerous possibility modals and a couple of volition modals are dialect-specific loanwords from Romani's European L2s. In addition to lexical borrowing, there are also several instances of contact-induced or 'replica' grammaticalization (cf. Heine & Kuteva 2005) of modals based (mostly) on indigenous lexical material. Finally, several modals may result from autonomous, contact-independent, grammaticalization within Romani. In order to abbreviate the subsequent discussion we first provide a morphosyntactic classification of Romani modal constructions (Section 2.1), which will then be referred to in the actual data subsections on volition, possibility and necessity modals (Sections 2.2–4). The last part of Section 2 is devoted to the interaction between modality and negation (Section 2.5).

2.1. Modal constructions

Romani modal constructions consist of two major parts: the modal and the main verb of the clause that encodes the modalized proposition. Since the clause need not be a complement of

the modal in syntactic terms (though it frequently is) and since the main verb in this clause need not be lexical, we use the term 'modalized verb'. Romani modal constructions vary on several parameters, both across and within dialects. Most importantly, there are differences with regard to the location of **subject marking**. The modal subject (the subject of the modalized proposition) is expressed through an NP, which is mostly optional if the subject is pronominal. In addition, person, number and, rarely, gender of the modal subject are, in most constructions, also cross-referenced on the modal and/or on the modalized verb. We distinguish between *personal* vs *impersonal* constructions, according to whether the modal inflects for the subject categories or not; and between *finite* vs *nonfinite* constructions, according to whether the modalized verb inflects for the subject categories or not. The criterion of subject cross-referencing thus renders four types of modal constructions.

The following examples from Hameln Sinti illustrate three of these types: the impersonal–finite construction in (1a) may express different kinds of possibility, while the personal–finite (1b) and the personal–nonfinite (1c) constructions are specialized for participant-internal possibility.

(1)   Hameln Sinti (Northwestern, Germany; Holzinger 1993)

a.   *Me   givau   našte.*
      I.NOM   sing.PFUT.1SG   can
      'I can sing.' (p. 94)

b.   *Me   hajevau   te   givap.*
      I.NOM   understand.PFUT.1SG   COMP sing.SUBJ.1SG

c.   *Me   hajevau   te   givel.*
      I.NOM   understand.PFUT.1SG   COMP sing.SUBJ.3SG[=INF]
      'I am able to sing.' (p. 94)

The structural difference between the two synonymous personal constructions is noteworthy. In (1b) both the modal and the (subjunctive) modalized verb cross-reference the first-person singular subject, whereas in (1c) only the modal does so. The modalized verb instead assumes a default, third-person singular, subjunctive form. While the finite construction has been inherited from Early Romani, the nonfinite subjunctive, or the 'new infinitive' (Boretzky 1996b, cf. also Matras 2002: 161–162), is an imperfect copy of the infinitive used in corresponding constructions in German, the current L2 of Hameln Sinti. Numerous further Romani dialects outside of the Balkans have developed the 'new infinitive' in same-subject non-factual complement clauses (and some other tightly-integrated subordinate clauses), due to pattern borrowing from L2s that, like German, possess an infinitive verb

form.[3] Nonfinite modal constructions are restricted to those Romani dialects that have developed the 'new infinitive'. Personal–finite modal constructions, on the other hand, are restricted to those dialects that retain finite non-factual complements. Impersonal–finite modal constructions can occur in dialects of both types.

The one type of modal construction that has not been illustrated so far is the impersonal–nonfinite construction. In this type there is no cross-referencing of the modal subject, and so it must be overtly encoded through an NP even if it is pronominal. No subject marking at all may only occur if generic modal subject is intended. In the example from modern Finnish Romani (2) the impersonal modal takes a nonfinite (default third-person singular subjunctive) modalized verb and the first-person singular modal subject is marked solely through an accusative pronoun. In the example from Selice Romani (3) the impersonal modal takes nonfinite (default third-person plural subjunctive) modalized verbs and the generic subject remains unexpressed.

(2)   Finnish Romani (Northwestern, Finland)
       Mān   mote   lel   tauva   tram.
       I.ACC   must   take.SUBJ.3SG[=INF]   this   medicine.NOM
       'I have to take this medicine.'

(3)   Selice Rumungro (South Central, Slovakia)
       Te   te   dživen   kampe,na   čak   te
       also   COMP live.SUBJ.3PL[=INF]   need   NEG   only   COMP
       dōgozinen.
       work.SUBJ.3PL[=INF]
       'One also needs to live, not only to work.'

Romani modal constructions also differ with regard to the location of **TAM marking**. While personal modals always do, impersonal modals may but need not inflect for TAM categories. In most Romani dialects modalized verbs in personal constructions and in TAM-inflected impersonal constructions assume a (finite or nonfinite) subjunctive form that does not encode tense or aspect. If however impersonal modals do not inflect for TAM, then TAM categories must be marked on the modalized verb. In other words, while the location of TAM marking is usually predictable from the location of subject marking in most constructions, the second criterion does differentiate between two types of impersonal–finite

---

[3] Dialects differ in what the default form of the nonfinite subjunctive is: it may be, like in Hameln Sinti, the third-person singular, but also the third-person plural or the second-person singular. Some varieties of Ukrainian Romani have reduced the second-person singular nonfinite subjunctive into a form that is now distinct from any subjunctive form, e.g. 2SG subjunctive *t'ir-ex* '[that] you do' vs infinitive *t'ir-e* 'to do'.

constructions: those where TAM categories are marked on the modal and those where they are marked on the modalized verb. In the dialect of Sliven Muzikanta, for instance, past reference can be marked on the impersonal modal *trjabvizela*, in which case the modalized verb assumes the tenseless subjunctive (4a), or it can be marked on the modalized verb, in which case *trjabvizela* assumes the default present form (4b). For the functional difference between the two examples see Section 2.4.

(4)   Sliven Muzikanta (Balkan, Bulgaria)

a.   *Trjabvizelas*   *te*   *užaras.*
        must.IMPF.3SG   COMP   wait.SUBJ.1PL
        'We had to wait.'

b.   *Trjabvizela*   *te*   *dičhan*   *les.*
        must.PRES.3SG   COMP   see.PRET.2SG   he.ACC
        'You must have seen him.'

All of the above patterns show TAM inflection either on the modal, or on the modalized verb. Nevertheless, instances of double TAM marking are also attested, though restricted to a few dialects. In the dialect of Parakalamos, as elsewhere in Romani, modalized verbs are in the subjunctive when the modal is in the present form (5a–b). However, unlike most dialects, Parakalamos Romani shows imperfect (rather than tenseless subjunctive) marking on the modalized verb when the modal is in the imperfect form. This kind of tense 'agreement' occurs irrespective of whether the modal is personal (5c) or impersonal (5d).

(5)   Parakalamos Romani (Balkan, Greece; Matras 2004)

a.   *Kamama*   *te*   *avav*   *demosiɣráfos.*
        want.PRES.1SG   COMP come.SUBJ.1SG   journalist
        'I want to become a journalist.' (p. 89)

b.   *Prepi*   *te*   *džas*   *othe.*
        must.PRES   COMP go.SUBJ.1PL there
        'We have to go to town,' (p. 72)

c.   *Kamamas*   *te*   *džavas*   *ti*   *poli.*
        want.IMPF.1SG   COMP go.IMPF.1SG to   town
        'I wanted to go to town.' (p. 75)

d.   *Eprepe*   *te*   *džakerasas.*
        must.IMPF   COMP wait.IMPF.1PL
        'We had to wait.' (p. 88)

The NP that expresses the modal subject may be in the nominative case or in an oblique case (the dative, the locative, or the markerless oblique, whose main function is to mark animate direct objects and which is usually labelled the 'accusative'). The two types of **case marking** of the modal subject are close to complementary. Nominative (canonical) marking always occurs in personal constructions (cf. 1b–c) and in those impersonal–finite constructions where TAM categories are marked on the modalized verb (cf. 1a). Oblique (noncanonical) marking, on the other hand, is obligatory in impersonal–nonfinite constructions (cf. 2). Nevertheless, we find variation between nominative and oblique marking in those impersonal–finite constructions where the modalized verb does not inflect for TAM categories. Typically the differentiation is lexically determined, though instances of (apparently) free variation are also attested. For example, in the dialect of Servy Ghympeny there are several impersonal modals, all of which take finite but tenseless complements: while *našty* always has a nominative subject (6a) and *javel-pe* always has a dative subject (6b), the subjects of *treb'i* alternate between nominative (6c) and dative marking (6d).

(6)   Servy Ghympeny (Northeastern, Ukraine)

a.   *Me   našty   isys   te   otčhak'iraw   udara.*
     I.NOM   cannot COP.3.PRET COMP uncover.1SG.SUBJ   door.PL
     'I could not open the door.'

b.   *Mange   javja   -pe   te   užakiraw.*
     I.DAT   come.PRET.3SG   REFL   COMP wait.SUBJ.1SG
     'I had to wait.'

c.   *Tu   treb'i vark'edys'   te   javes   ke   me.*
     thou.NOM   need   sometimes   COMP come.SUBJ.2SG   at/to I.NOM
     'You must come to me sometimes.'

d.   *Tuke   treb'i dor'ik te   džas.*
     thou.DAT   need   there   COMP go.SUBJ.2SG
     'You must go there.'

The role of finiteness in the case marking of the modal subject can be illustrated from a dialect that shows a different type of variation. In Kubanskie Servy the subject of impersonal modals *nasči* and *trebun'i* is nominative when the modalized verb is finite (7a–b) but dative when the verb is nonfinite (7c–d).

(7)   Kubanskie Servy (Ukrainian, Ukraine and Russia)

a.   *Me   nasči   st'irav   e   ural'i.*
     I.NOM cannot repair.SUBJ.1SG   DEF   car

'I can't repair the car.'

b. *Tu   trebun'i   var'ikala   mande te   avex.*
   thou.NOM   must   sometimes   I.LOC COMP come.SUBJ.2SG
   'You should visit me sometimes.'

c. *Mand'i nasči te   makhe   o   cer.*
   I.DAT   cannot COMP smear.INF   DEF   house
   'I can't paint the house.'

d. *Tut'i   trebun'i   korde te   dža.*
   thou.DAT   must   there COMP go.INF
   'You have to go there.'

Next, there are differences with regard to the internal **constituency** and structural **complexity** of modal expressions. While modals that inflect neither for TAM nor for subject categories are the least complex, one can further distinguish between *synthetic* vs *analytic* inflected modals. In synthetic modals TAM and/or subject categories are marked within the same word as modality. In analytic modals, on the other hand, the actual modality word is uninflected and inflectional categories are marked separately, on an inflectional auxiliary, which is structurally identical to the copula (and the verb of existence). The following examples from Podhradie Romani illustrate the three types of modals: an uninflected modal (8a), a synthetic inflected modal (8b) and an analytic inflected modal (8c). In most instances of analytic modals the inflectional auxiliary assumes the default, third-person singular, subject category, thus only inflecting for TAM categories, and it is omitted in the present. The personal and obligatory auxiliary in (8c) is thus rather untypical in both respects. Nominal expressions of modality tend to require the copula, and so they resemble the analytic inflected modals. Adjectival modals, which are rather infrequent in Romani, show distributed marking of subject categories: the adjectival cross-references the number and (sometimes) gender of the modal subject, while the copula encodes the TAM categories and mostly cross-references the subject's person and number, as in Crimean Romani (9).

(8)   Podhradie Romani (North Central, Slovakia)

a. *Šaj   džav   khēre.*
   can   go.PRES.1SG home
   'I can go home.'

b. *Kamav   khēre te   džan.*
   want.PRES.1SG   home COMP go.SUBJ.3PL[=INF]
   'I want to go home.'

c. *Musaj   som   khēre te   džan.*
   must   COP.PRES.1SG   home COMP go.SUBJ.3PL[=INF]

'I have to go home.'

(9) Crimean Romani (Balkan; Ukraine and Russia)

Tume	*sanusas*	*dolžn'a*	*te*	*raskeld'ijen*	*les.*
you	COP.IMPF.2PL	obliged.PL	COMP	meet.SUBJ.2PL	he.ACC

'You guys must have met him.'

Although modalized verbs are typically introduced by the non-factual **complementizer** *te*,[4] there are several kinds of exceptions. Some modals simply do not allow a complementizer. These include, but are not restricted to, expressions where an original complementizer has become an integral part of the modal. In Čáry Romani, for instance, some modals require the complementizer (10a), whereas others do not allow it (10c). Despite an etymological presence of *te in the modal *moste*, the construction (10b) is of the latter type: while the complementizer *te* is separable from the modal and inseparable from the modalized verb (cf. 10a), the opposite holds for *te in (10b). In many dialects some modals that usually take the complementizer may allow its occasional dropping. Importantly, the obligatory lack of a complementizer is restricted to uninflected modals, i.e. to constructions where TAM categories are marked solely on the modalized verb. There is no such limitation with the optional dropping of *te*.

(10) Čáry Romani (North Central, Slovakia)

a.
Kamlom	*ole*	*čhavenca*	*te*	*vakerel.*
want.PRET.1SG	that.OBL.PL	boy.INSTR.PL	COMP	
speak.SUBJ.3SG[=INF]				

'I wanted to talk to those boys.'

b.
Moste	*prekal*	*odā*	*prelezinďom.*
must	over	that	crawl.over.PRET.1SG

'I had to climb over it.'

c.
Šaj	*tut*	*vareso*	*dās.*
can	you.ACC	something	give.IMPF.1SG

'I could give you something.'

Even if licenced by the modal, the complementizer is mostly omitted when the modalized verb is borrowed and retains its L2 inflection. For example, the modal *može* in Kaspičan Xoraxane requires *te* with indigenous modalized verbs (11a) but the

---

[4] Apart from introducing non-factual complements, this polysemous connector may also introduce optative predications, conditional clauses, purpose clauses and further types of adverbial subordinations.

complementizer must be omitted with Turkish-inflected verbs (11b). Similarly, some Russian Romani modals require *te* with indigenous modalized verbs but there is no complementizer when the verb is borrowed from Russian, in which case it retains its Russian infinitive form. There is one dialect, Dolenjski Romani, where the complementizer *te* has been lost altogether, in modal constructions and elsewhere (see Matras 2002: 210 for details), having been replaced by the Slovene complementizer *da* (12a–b) or dropped altogether, especially in nonfinite constructions (12c).

(11) Kaspičan Xoraxane (Balkan, Bulgaria)

a. *Može   te   džas   kaj   gav.*
   can   COMP go.SUBJ.1PL to   village
   'We can go to town.'

b. *Može   inanasınıs   leske.*
   can   believe.SUBJ.2PL   he.DAT
   'You can believe him.'

(12) Dolenjski Romani (Slovene/Istrian, Slovenia; Cech & Heinschink 2001)

a. *Triba   da   leske   dav   love.*
   need   COMP he.DAT   give.SUBJ.1SG   money.PL
   'I need to give him money.' (p. 357)

b. *Hočemo   da   lam   duj   phabaja.*
   [want.PRES.1PL]$^{\text{SLOVENE}}$ COMP take.SUBJ.1PL   two   apple.PL
   'We want to take two apples.' (p. 357)

c. *Morinave   lake   del   love.*
   must.IMPF.1SG she.DAT   give.SUBJ.3SG[=INF]   money.PL
   'I had to give her money.' (p. 356)

Several generalizations regarding the **linear order** within modal constructions can be made. The inflectional auxiliary, if there is one, immediately follows the modality word proper, thus conforming to the position of synthetic verb inflections. This order, however, is not required in nominal modal constructions with a copula (cf. 9). The complementizer, if any, precedes the modalized verb, usually immediately, though it may be separated from the verb by negators or pronominal clitics in some dialects. In pragmatically neutral contexts modals precede modalized verbs; they may be separated by other constituents, including the subject NP. However, most dialects probably allow pragmatically motivated fronting of modalized verbs, which results in postposed modals, cf. focus fronting in (3). German Sinti stands out in allowing postposition of uninflected modals even in neutral contexts, cf. (1a).

Table 1 summarizes the structure of five major types of modal constructions in Romani (COMP+ means that the complementizer is at least optionally present with Romani-inflected modalized verbs):[5]

Table 1: *Major types of modal constructions in Romani*

constr.	SUBJECT INFL.		TAM INFL.		NOM NP	AUX	COMP
	modal	verb	modal	verb			
PF	+	+	+	±	+	±	+
PnF	+	−	+	−	+	±	+
nPF1	−	+	+	±	±	±	+
nPF2	−	+	−	+	+	−	±
nPnF	−	−	+	−	−	(−)	+

Personal–finite (PF) constructions mark TAM categories on the modal, use the complementizer and have a nominative subject; they vary with respect to the marking of TAM categories on the modalized verb and with respect to the presence of an inflectional auxiliary. Personal–nonfinite (PnF) constructions mark TAM categories solely on the modal, use the complementizer, and have a nominative subject; there may but need not be an auxiliary. There are two major types of impersonal–finite constructions. Those that mark TAM categories on the modal (nPF1) always allow the complementizer but vary on a number of parameters: TAM marking on the modalized verb, case marking of the subject NP and the presence of an inflectional auxiliary. Those impersonal–finite constructions that do not mark TAM categories on the modal (nPF2) are rather uniform: they mark TAM on the modalized verb, have a nominative subject, and cannot have an inflectional auxiliary since the modal is uninflected; the complementizer may be obligatorily lacking. Finally, impersonal–nonfinite (nPnF) constructions mark TAM categories solely on the modal, have an oblique subject, and allow the complementizer.[6]

2.2. Volition

There are two widespread expressions of volition in Romani. The verb *kam-* is the exclusive expression of volition in most Romani dialects, and can hence be reconstructed as the Early

---

[5] As it is not restricted to modal constructions, the obligatory lack of a complementizer in Dolenjski Romani nonfinite constructions is disregarded in Table 1.

[6] Inflectional auxiliaries are unattested in Romani impersonal–nonfinite modal constructions. This does not appear to be due to any structural reason, however.

Romani form. This verb can still be used with nominal objects, and in most dialects it retains its lexical meanings 'to love, to like' and 'to desire'. The lexical *kam-* 'to love' is certainly pre-European, though it might be a loanword from West Iranian or Armenian (cf. Boretzky 1995b: 141) rather than continuing an Old Indo-Aryan etymon (as suggested by Sampson 1926, II: 131). In the Romani dialects of the Balkans *kam-* competes with a more recent volition expression, the verb *mang-*. The latter has its origin in the common Romani verb *mang-* 'to ask for, to demand', also used for 'to beg' (and from the latter meaning sometimes generalised for 'to make a living by hawking, fortune-telling or begging'). The two volition verbs appear alongside each other in several Romani dialects of the Balkans (see also Section 2.5), whereas in other dialects of this area *mang-* has replaced *kam-*, taking over all of its functions, both lexical and modal. While the grammaticalization of *kam-* into a volition modal must pre-date the split of Romani dialects, the grammaticalisation of *mang-* appears to be a fairly recent phenomenon that emerged after the outwards migration from the Balkans of some of the groups, and one that is still spreading within the area. Its origin can be traced in all likelihood to replica grammaticalisation of the South Slavic construction, e.g. Bulgarian *iskam* 'I want' as well as 'I demand'. Lexical borrowings of volition modals are infrequent: we only find the verb *wånt-(a)s-* from English in Welsh Romani (Sampson 1926, II: 401) and the verb *hoč-* from Slovene in Dolenjski Romani of Slovenia, both of which alternate with indigenous volition verbs.

All volition verbs are personal (Type PF or PnF) and mostly take the complementizer *te* (cf. 8b, 10a, 13a, 53), with the usual exceptions (see Section 2.1, ex. 12c). Occasional dropping of the complementizer is attested, for example, in Piemontese Sinti (Franzese 1985: 126). In some Romani dialects of Slovakia, including Klenovec Rumungro, the personal construction (13a) alternates with a less agentive volitional construction, which is impersonal (Type nPnF). The impersonal volition verb contains middle marking: either synthetic (13b), or analytic, reflexive-like (cf. 37c), as in Slovak, the source of this construction. In several dialects of northeastern Europe and the Balkans the verb *kam-* preserves an archaic first-person singular non-perfective suffix *-am*, rather than the regular *-(a)v*. In the past tense, the modal verb may, in some dialects, take the simple preterite (14), but the imperfect (a non-perfective remote tense) is often preferred (15), which is typical of the portrayal of mental states. The Dolenjski volition loanword from Slovene retains its L2 inflection (cf. 12b).

(13) Klenovec Rumungro (South Central, Slovakia)

a. *Na*   *kames*   *te*   *džan*   *ano*   *gaw.*
    NEG   want.PRES.2SG   COMP go.SUBJ.3PL[=INF] in.DEF town
    'You don't want to go to town.'

b. *Musaj*   *odoj*   *te*   *džas*

	must	there	COMP	go.SUBJ.2SG	

te   *iš*   *tuke*   *na*   *kamisaľol*   *te*

if   too   thou.DAT   NEG   want.MIDDLE.PRES.3SG   COMP

džan.

go.SUBJ.3PL[=INF]

'You have to go there, even if you don't want to go.'

(14) Kohila Romani (Northeastern, Estonia)

Me   *kamjom*   *khere*   *te*   *džal.*

I.NOM   want.PRET.1SG   home   COMP go.SUBJ.3SG[=INF]

'I wanted to go home.'

(15) Kaspičan Xoraxane (Balkan, Bulgaria)

Mangavas   *te*   *džav*   *mange khere.*

want.IMPF.1SG COMP go.SUBJ.1SG I.DAT home

'I wanted to go home.'

2.3. Possibility

Two or three possibility modals can be reconstructed for Early Romani: certainly the affirmative *šaj* and the negative *našti*, and probably also the affirmative *ašti*. Though *šaj* is clearly of pre-European origin in Romani, its precise etymology remains open: it may either continue some form of the Old Indo-Aryan possibility verb *śakno-* (Boretzky 1996a: 3), or it may be a borrowing of the Persian possibility modal *šāje* into Proto-Romani (Matras 2002: 162–163, also mentioned in Boretzky & Igla 1994: 268). The origin of *ašti* is likewise disputed. Matras (2002: 162–163, following Sampson 1926: 216) hints at the Old Indo-Aryan third-person singular present copula form *asti* as its source,[7] thus suggesting that Early Romani possessed both the original *ašti* and the pre-European borrowing *šaj*. Boretzky (1996a: 5–6), on the other hand, considers *ašti* to result from analogical decomposition of the negative *našti*, whereby *n-* of *našti* was identified with the indicative negator *na*. He also claims that this decomposition took place in individual dialects rather than before the dialect split of Romani, thus implying that *šaj* was the only possibility modal in Early Romani.[8]

---

[7] As Boretzky (1995a: 5) points out correctly, the OIA copula form *asti* is usually considered to be the source of the Romani third person present copula form *isi*. Nonetheless, a split development might also be considered, accounting for parallel forms.

[8] The current dialect distribution of *ašti* suggests inheritance from Early Romani rather than independent dialect-specific innovations. While *šaj* has been retained in a geographically contiguous area (stretching from

Boretzky & Igla (1994: 196) derive *našti* from a construction consisting of the indicative negator, the affirmative modal *šaj* and the non-factual complementizer *te*, i.e. from **na šaj te*. However, *našti* may also continue a regularly negated form of the possibility modal *ašti* (Matras 2002: 162–163). The disputed modal *ašti* has two etymologically unclear variants: *šašty* or *sašty* in Lithuanian and Latvian Romani and *dašti-* in Kalderaš and related North Vlax varieties (also diffused into some Ursari varieties). Assuming that both *šaj* and *ašti* were present in Early Romani, the question arises of what the functional difference between them was. This issue will be taken up in Section 3.1.

In most Romani dialects the inherited possibility modals are uninflected and occur in impersonal–finite constructions with TAM marking on the modalized verb (Type nPF2). They do not allow a complementizer in some dialects (those spoken in central-eastern Europe and occasionally in the Balkans and elsewhere), whereas in other dialects the complementizer is required or at least allowed. Several dialects spoken in, or originating from, Romania show synthetic personal inflection (Type PF) of the inherited possibility modals, at least variantly, e.g. Pikulešti–Kurturare *šaj(-in)-* and *nasči(-n)-*, Kalderaš-type *dašti(-sar)-* and *našti(-sar)-*, and Ursari *(d)ašti(-z)-*. In fact, uninflected *dašti* is attested in a single dialect (Boretzky & Igla 2004: 183). Though indigenous, these subject-inflected modals display morphological adaptation typical of loanverbs. Impersonal inflection (Type nPF1) of the inherited possibility modals is rare: we find synthetic marking of TAM categories in Piedmontese Sinti, e.g. *stik-o-l-a* [-middle-3SG-PFUT], alternating with uninflected *stik* < **ašti*; and analytical marking of TAM categories in some dialects of Ukraine (cf. 6a). The inflected possibility modals, as a rule, require the complementizer.

In addition to the possibility modals inherited from Early Romani, many dialects employ possibility expressions that result from dialect-specific grammaticalizations of pre-modal constructions or lexical verbs. An impersonal copula construction, which has been grammaticalized into a dedicated necessity modal in several dialects (see Section 2.4), is reported to have developed into a possibility expression in Crimean Romani (cf. Boretzky 1999a: 113). A negative impersonal copula construction probably also underlies the negative possibility modals *nahi* or *naj* < **na hi* [NEG COP.PRES.3] in a few dialects of the Balkans and in Core Sinti (Boretzky & Igla 2004: 184). Note that if we are to assume that *ašti* developed from an Old Indo-Aryan copula, then the dialect-specific development of possibility modals from the Romani copula represents a 'second round' of the same type of grammaticalization.

Quite commonly, participant-internal possibility (ability or capability) is expressed by personal verbs meaning 'to know', 'to understand' or 'to manage', which as a rule retain

---

Czechia to northwestern Bulgaria plus recent out-migrant dialects), reflexes of *ašti* are attested within most dialect groups of Romani.

their pre-modal meanings as well. The ability function of the indigenous verb *(a)xaljov-* 'to understand' is only attested in some dialects of Ukraine, Germany, and Finland, e.g. in Kotka Romani (16). On the other hand, the (cap)ability function of the indigenous verb *džan-* 'to know' is widespread within Romani, to the extent that Boretzky (1999b: 178) considers it to be reconstructable for Early Romani. At least in some Romani dialects the extension of these lexical verbs to (cap)ability functions is likely to have resulted from contact-induced grammaticalization, as this kind of polyfunctionality is common Romani's L2s. In several dialects spoken in and around Hungary, possibly due to pattern borrowing from Hungarian (cf. Boretzky & Igla 2004: 184), the (cap)ability verbs (17a) have been extended to express participant-external possibility functions as well (17b).

(16) Kotka Romani (Nortwestern, Finland)
Hajuveha   -ko   tu   cērel   butti touveria?
understand.PFUT.2SG -Q   thou do.SUBJ.3SG[=INF]work   ax.INSTR
'Are you able to work with an ax?'

(17) Gabor Romani (North Vlax, Romania)
a. *Čokanosa   žane   te   kere buči?*
hammer.INSTR know.PRES.2SG   COMP do.SUBJ.2SG work
'Are you able to work with a hammer?'

b. *Či   žanel   te   žal   -tar   pînë   na*
NEG   know.PRES.3SG   COMP go.SUBJ.3SG away   until   NEG
arakël   e   čie.
find.SUBJ.3SG DEF   key
'He can't leave until he finds the key.'

In some Romani dialects of central Europe the Early Romani verb *troma-* 'to dare' of Greek origin has developed deontic possibility (permission) functions (cf. Boretzky & Igla 2004: 187). An identical semantic development is attested in Slavic languages (cf. Hansen 2003: 11) and German, and so contact-induced grammaticalization may be the source of this extension (Boretzky & Igla 1994: 183). While the permission modal retains its personal verb inflection in the Central dialects, it has fused with the original complementizer in Sinti, losing all inflection on the way, e.g. *troma- te > Hungarian Sinti *trunti* (18; Type nPF2). In some Rumungro dialects the verb *troma-* is now a dedicated permission modal (19a; Type PnF), while the lexical meaning 'to dare' is expressed by a construction involving a secondary nominal back-formation from the verb (19b).

(18) Hungarian Sinti (Northwestern, Hungary; Mészáros 1980: 10)

    Trunti   *džal*   *dren.*

    be_allowed   go.SUBJ.3SG in

    'S/he is allowed to come in.'

(19) Klenovec Rumungro (South Central, Slovakia)

a.   *Me*   *tromaw*   *odā*   *te*   *keren.*

    I.NOM   be_allowed.PRES.1SG   that   COMP do.SUBJ.3PL[=INF]

    'I am allowed to do it.'

b.   *Man*   *hi*   *troma odā*   *te*   *keren.*

    I.ACC   COP.PRES.3 courage that   COMP do.SUBJ.3PL[=INF]

    'I dare [lit. I have courage] to do it.'

Several possibility modals have quite interesting lexical sources. For example, the verb *vydža-* 'to go out' (consisting of a Slavic aktionsart prefix and the indigenous verb *dža-* 'to go') is used impersonally, with locative or dative marking of the subject NP, as the basic possibility modal in the Xandžary dialect of Ukraine. The construction is probably based on a Ukrainian idiomatic expression ('it goes out to someone' meaning 'it works out for someone'), though the extension to possibility ('it is possible [can work out] for someone') is an autonomous innovation of the Romani dialect. The modal is attested in a wide range of possibility uses: ability (20a; Type nPnF), participant-external possibility (20b; Type nPF1) and permission (20c; Type nPnF). Further curious examples of lexical resources for possibility include the personal verbs *dol-* 'to get, become; to get, receive; to get [somewhere]' (21a; Type PnF), which consists of a Slavic aktionsart prefix and the indigenous verb *l-* 'take', and *pēr-* 'to fall; to get [somewhere]' (21b) in Kohila Romani. Both expressions share the meaning 'to get [somewhere]', which might be the immediate pre-modal meaning. [9]

(20) Xandžary (Ukrainian, Ukraine)

a.   *Mande*   *na*   *vydžal*   *te*   *t'ire*   *ural'i.*

    I.LOC   NEG   go_out.PRES.3SG   COMP do.INF car

    'I am not able to repair the car.'

b.   *Lest'i*   *na*   *vydžal*   *te*   *udžal*

    he.DAT NEG   go_out.PRES.3SG   COMP leave.SUBJ.3SG

    poka   *vov*   *na*   *rakhel*   *t'ii.*

---

[9] The grammaticalization of motion verbs meaning 'to arrive at, to reach' into ability modals is well attested (cf. Heine & Kuteva 2002: 45–46).

	until	he.NOM	NEG	find.PRES.3SG	key

'He can't leave until he finds the key.'

c. *Tute*   *vydžala*   *te*   *dža*   *mansa.*

thou.LOC   go_out.fut.3SG   COMP go.INFI.INSTR

'You may come with me.'

(21) Kohila Romani (Northeastern, Estonia)

a. *Me*   *dolā*   *tuke*   *lōve*   *tašša*   *te*

I.NOM   get.PFUT.1SG   thou.DAT   money.PL   tomorrow   COMP

plajskyrel.

pay.SUBJ.3SG[=INF]

'I can pay you [back] tomorrow.'

b. *Jov*   *kindža*   *peske*   *neve*   *idža*

he.NOM buy.PRET.3SG   REFL.DAT   new.PL   clothes.PL

sob   *te*   *pērel*   *džal*   *ande*   *foros.*

so_that   COMP fall.SUBJ.3SG   go.SUBJ.3SG in   town

'He bought himself new clothes so that he could go into town.'

Lexical **borrowing** of possibility modals into Romani is well attested, though many dialects do without a possibility loanword. Loans of the following possibility modals are attested within Romani: Greek *boro*, Macedonian and Bulgarian *može*, Slovene *lahko*, Polish *móc*, East Slavic *moč'* and *možno*, German *dürfen*, Italian *potere*, and Finnish *voida*. Possibility loanwords tend to be borrowed from the current L2s of the relevant Romani dialects, although Slovene-derived *lako* and *lax* are also attested in some Italian Romani varieties, e.g. Venetian Sinti (cf. Boretzky & Igla 2004: 183). Most loanwords have a wide range of possibility functions, with the exception of the loan of German *dürfen* into German Sinti, which is, like the source form, specialized for deontic possibility (permission).

    The Greek, Polish and Finnish possibility modals plus East Slavic *moč'* are personal verbs and they are always borrowed as such into Romani. Most commonly the loans are morphologically integrated into Romani verb inflection in the same way as other borrowed verbs are, through adaptation of a frequent inflectional stem of the L2 modal by means of a dialect-specific loanverb suffix, e.g. *bor-in-* or *bor-iz-* (< Greek) in some varieties of Greek Romani; *mog-in-*, *moγ-in-* or *mož-in-* (< Slavic) in numerous Northeastern dialects of Poland, Lithuania and Ukraine; and *voj-uv-* or *vojp-uv-* (< Finnish) in Finnish Romani. Nevertheless, in some dialects in current contact with Greek and East Slavic, the possibility loanverbs are borrowed together with their L2 inflection, e.g. Parakalamos Romani (22), Russian Romani, some Servy dialects of Ukraine, and Crimean Romani (23). In all of these latter dialects this borrowing strategy is also common with lexical verbs.

(22) Parakalamos Romani (Balkan, Greece; Matras 2004)

a.   *Boro*   *te*   *džav.*
     can.PRES.1SG COMP go.SUBJ.1SG
     'I can go.' (p. 87)

b.   *Na*   *borusa*   *te*   *phiravavas*   *i*   *vudar.*
     NEG   can.PRET.1SG   COMP open.IMPF.1SG   DEF   door
     'I could not open the door.' (p. 94)

(23) Crimean Romani (Balkan, Ukraine and Russia)

     Tu   *možeš*   *te*   *pak'as*   *leske.*
     thou.NOM   can.PRES.2SG   COMP believe.SUBJ.2SG   he.DAT
     'You can believe him.'

Although the East South Slavic possibility verb *može* may also be subject-inflected, it is frequently used as an impersonal modal, assuming the default, third-person singular, subject category. While some Romani dialects that borrow this modal allow both constructions, others have generalized the impersonal construction. Personal loanwords of *može* are always morphologically adapted through dialect-specific loanverb suffixes, e.g. *mož-in-*, *mož-ij-*, *mož-iz-*, *mož-is-*, *mož-i(sar)-*, i.e. there is no retention of L2 subject inflection. Impersonal loanwords of the verb, on the other hand, may but need not be adapted. Most dialects borrow the third-person singular present L2 form *može*, although they may also allow its adaptation and indigenous third-person singular inflection, especially in more complex TAM categories. Obligatory adaptation of the impersonal verb even in the present is exceptional, attested for example in Rakitovo Yerli, cf. *mož-ij-əl-a/as* [-LOAN-3SG-PRES/REM] (24).

(24) Rakitovo Yerli (Balkan, Bulgaria)

a.   *Možijəla li*   *tə*   *kərəs*   *buti*   *čukosa?*
     can.PRES.3SG Q   COMP do.SUBJ.2SG work   hammer.INSTR
     'Can you work with a hammer?'

b.   *Of*   *na*   *daralas*   *če*   *možijəlas*   *tə*
     he.NOM NEG   fear.IMPF.3SG   COMP can.IMPF.3SG   COMP
     pərəl.
     fall.SUBJ.3SG
     'He was not afraid that he could fall.'

The East Slavic construction of *možno* is retained in Romani: though the modality word itself is uninflected, it takes an impersonal inflectional auxiliary that must be omitted in the present indicative (Type nPF). The auxiliary is either borrowed in its East Slavic form or rendered by the Romani copula. The uninflected possibility modal *lako* in Dolenjski Romani is, like its Slovene source, used in impersonal–finite constructions with TAM marking on the modalized verb (Type nPF2) and without a complementizer (25). The German Sinti permission modal *darfte* occurs in the same type of construction, as the complementizer *te* has become an integral part of the modal; *darf-* is a petrified singular present stem of the German modal *dürfen*.

(25) Dolenjski Romani (Slovene/Istrian; Slovenia)

a. *I   brzo   lende   lako   živinamo.*
    and   without   they.LOC   can   live.PRES.1PL
    'And we can live without them.' (p. 357)

b. *Pe   domislindža   da   ruv   lako   haja   porana*
    REFL.ACC   realize.PRET.3SG   COMP   wolf   can   eat.PRET.3SG   old.ACC
    daja
    mother.ACC
    'He realized that the wolf may have eaten the grandmother.' (p. 359)

2.4. Necessity

In all likelihood, there were no dedicated necessity modals in Early Romani. The necessity modals we find in present-day Romani are loanwords or result from recent, dialect-specific, grammaticalizations. Nevertheless, we also find in various Romani dialects several weakly grammaticalized modal constructions that indicate how necessity could have been expressed in Early Romani. These constructions do not contain a dedicated modal word (or affix) and quite often they are not even specialized for expressing a certain kind of modality. For example, in Sepečides the **independent subjunctive**, introduced by the non-factual connector *ti* < *te*, can be used to express not only necessity (26a), but also permission (26b), volition (26c) and other modal meanings (Cech & Heinschink 1999b: 119–120). And in Welsh Romani a personal construction involving the **copula** and a finite (subjunctive or present indicative) complement, also introduced by *te*, can be used to express not only necessity (27a), but also possibility (27b) or general engagement[10] in an event (27c). A similar range of modal functions is also attested for impersonal constructions involving the

---

[10] As Sampson's (1926, II: 119) lexicographic definition goes: 'to be in the act of, to be occupied or engaged in, to be the fact that'.

copula in some dialects of the Balkans (cf. Boretzky 1999a: 113). Necessity is not explicitly encoded in examples such as (26a) and (27a), and so their necessity reading (and translation) must involve some sort of pragmatic inference.

(26) Sepečides (Balkan, Turkey; Cech & Heinschink 1999b: 120)

a. *Okulestar   ti   daras!*
   that.ABL   COMP fear.SUBJ.1PL
   'We must be afraid of that!'

b. *Ti   bešav   akate?*
   COMP   sit.SUBJ.1SG here
   'May I sit here?'

b. *Akana   me   tuke   ti   vakerav…*
   now   I.NOM thou.DAT   COMP speak.SUBJ.1SG
   'Now I want to tell you…'

(27) Welsh Romani (British, Wales; Sampson 1926, II: 119)

a. *Šomaš   te   lā   phabā   kjathakja*
   COP.PRET.1SG   COMP take.SUBJ.1SG   apple.PL   such_and_such
   thaneste.
   place.ABL
   'I had to fetch apples from such and such a place.'

b. *Na   šom   mē   te   reperava   kek.*
   NEG   COP.PRES.1SG   I.NOM COMP remember.PRES.1SG NEG
   'I cannot remember.'

c. *Šom   mē   te   dživava   lesa.*
   COP.PRES.1SG   I.NOM COMP live.PRES.1SG   he.INSTR
   'I am living with him.'

Nevertheless, the above modal constructions may become specialized for necessity, especially deontic necessity (but also for possibility, see Section 2.3). In Hameln Sinti, for instance, the independent subjunctive without *te* always conveys obligation (28; Holzinger 1993: 92), as does the impersonal copula construction in Welsh Romani (29a–b). In the latter construction the modal subject is not only cross-referenced on the subjunctive verb but also encoded through an obligatory accusative-marked NP. Necessity here is modelled on predicative possession, whereby the modal subject is construed as the possessor of the modalized event: in most Romani dialects, including Welsh Romani, the predicative possessive construction consists of a nominative-marked possessee NP, the copula agreeing with it in number (and person, if the possessee is pronominal, which is rare) and an oblique-

marked possessor NP (29c). The possession model for necessity is exactly what we find in English, one of the L2's of Welsh Romani. Boretzky & Igla (2004: 185) suggest that Romani possessive-like necessity constructions, which are well attested in several dialects, may result from contact-induced grammaticalization on the model of possessive-like necessity modals in various European languages. In some North Central dialects of Slovakia, including Čáry Romani, a personal (rather than impersonal–possessive) copula construction is specialized to express weak necessity, especially weak obligation (30). Even though Slovak, the current L2 of the dialect, expresses weak necessity by the verb of possession, here too we might be dealing with an instance of pattern borrowing, though a less straightforward one.

(28) Hameln Sinti (Northwestern, Germany; Holzinger 1993: 325)

Štrap		*i*	*grajes,*	*krap*	*leske*
brush.SUBJ.1SG		DEF.OBL	horse.ACC	do.SUBJ.1SG	he.DAT

futera.
animal_food
'I must brush the horse, I must prepare food for him.'

(29) Welsh Romani (British, Wales; Sampson 1926)

a.

Šī	*man*	*te*	*dā*	*les*	*būt*	*lōvō.*
COP.PRES.3	I.ACC	COMP	give.SUBJ.1SG	he.ACC	many	money

'I have to give him a lot of money.' (II, p. 119)

b.

Sas	*ī*	*Džakes*	*te*	*del*	*les*	*vavēr*	*stādī*
COP.PRET.3	DEF.OBL	Jack.ACC	COMP	give.SUBJ.3SG		other	hat

sunakaj.
gold
'Jack had to give him another hatful of gold.' (I, p. 213)

c.

Sas	*ī*	*phurīā*	*trin*	*gurunīā.*
COP.PRET.3	DEF.OBL	old.F.ACC	three	cow.PL

'The old woman had three cows.' (I, p. 213)

(30) Čáry Romani (North Central, Slovakia)

Sal	*tuke*	*te*	*pametinel*	*so*
COP.PRES.2SG	thou.DAT	COMP	remember.SUBJ.3SG[=INF]	what

tuke *phenā.*
thou.DAT say.fut.1SG
'You ought to remember what I'm going to tell you.'

In several dialects the weakly grammaticalized copula constructions have undergone considerable grammaticalization. In most South Central the impersonal copula construction developed into a strongly grammaticalized necessity modal *site* or *iste*, which continues the third-person present indicative copula form *(i)si* plus the complementizer *te*. And in the Core Sinti dialects the personal copula construction developed into the strongly grammaticalized necessity modal *homte*, *hunte*, *unti* etc.,[11] which continues the generalized first-person singular[12] present indicative copula form *hom < *(i)som* plus the complementizer (cf. already Pott 1844). In either case the two etymological components have fused together and the resulting univerbal modal is separable from the modalized verb. In Selice Rumungro *site* may constitute a separate utterance (31b) and in Hameln Sinti *hunte* may be postposed to the verb (32b). Although both modals are historically based on a present indicative copula form, neither inflects for TAM categories: their marking has shifted from the copula to the modalized verb (31b, 32c; Type nPF2). The Rumungro modal takes a regular negator rather than involve the irregular negative third-person present copula form, i.e. *na site* rather than *nane te (31b). Both modal constructions now show nominative marking of the modal subject (31a, 32a–b), although in older Rumungro oblique marking was still available in the impersonal *si te* construction (cf. Müller 1869: 161). Unlike their less grammaticalized source constructions, *site* and *homte* serve as the basic necessity modals in the respective dialects and have a wide range of necessity functions.

(31) Selice Rumungro (South Central, Slovakia)

a. *Amen   site   džas   onďa.*
    we.NOM   must   go.SUBJ.1PL there.
    'We have to go there.'

b. *Na   site   mange papaleg   palikerďal.   De,   site!*
    NEG   must   I.DAT again   thank.PRET.2SG   but   must
    '[A:] You need not have thanked me again. [B:] I did have to!'

(32) Hameln Sinti (Northwestern, Germany; Holzinger 1993: 92–93)

a. *Tu   hunte kres   kova.*
    thou.NOM   must   do.SUBJ.2SG that

b. *Tu   kres   kova   hunte.*

---

[11] The modal *humte* has also diffused from Sinti into Bohemian (North Central) Romani. It appears to have been present also in some Para-Romani varieties (cf. Boretzky & Igla 2004: 184–185).

[12] Boretzky & Igla (2004: 185) find the generalization of this person–number category suprising. It should be noted, however, that the first person singular is, cross-linguistically, the second most frequent person–number category in verbal paradigms (Bybee 1985).

thou.NOM   do.SUBJ.2SG that   must

'You have to do that.'

c.   *Hunte   džajom   khere.*

     must   go.PRET.1SG home

     'I had to go home.'

Like the South Central dialects, the Lovari-type dialects spoken in and around Hungary, including Austrian Lovari, possess the necessity modal *site* or *iste*, which is used in constructions with TAM marking on the modalized verb (33a; Type nPF2). Unlike its South Central counterpart, however, the Lovari modal is not separable from the modalized verb. In addition, the dialect has also retained a less grammaticalized copula construction in the past, where tense is marked on the copula rather than on the modalized verb and where the complementizer clearly remains a separate syntactic unit (33b; Type nPF1). The latter construction is only attested as an expression of participant-external necessity, while *site* appears to have a wider range of necessity functions.

(33) Austrian Lovari (North Vlax, Austria)

a.   *Site   ašasas   khere.*

     must   stay.IMPF.1PL   home

     'We had to stay at home.' (Cech *et al.* 1999: 28)

b.   *Me   sas   te   dikhav   pel   grast.*

     I.NOM   COP.PRET.3 COMP see.SUBJ.1SG   on.DEF.PL   horse

     Sas   te   pijavav,   te   xaxavav

     COP.PRET.3   COMP make_drink.SUBJ.1SG   COMP make_eat.SUBJ.1SG

     le.

     they.ACC

     'I had to take care of the horses. I had to give them to drink, to feed them.' (Fennesz-Juhasz & Heinschink 1999: 66)

There are several further sources of dialect-specific grammaticalization of necessity modals in Romani. A few Balkan and Appenine dialects of Romani have grammaticalized the verb *ther-* 'to hold' into a personal verb of **possession** ('to have') and in one of these dialects, Rumelian Romani, this verb may also express necessity (34; Type PF). Both developments are contact-induced (cf. Boretzky & Igla 2004: 185), the latter once again, as with the impersonal copula construction, exemplifying the possessive model for necessity. Another common source of necessity expressions in Romani are **volition** verbs (cf. Boretzky 1996a: 14–17). Personal volition verbs may occasionally be used to express necessity in some Romani dialects of Greece, Slovakia, Ukraine (35; Type PF) and perhaps

elsewhere. Pattern borrowing from Greek can be invoked to explain these uses at least in Greek Romani (cf. Boretzky 1996a: 16). In quite a few dialects the polysemous verb *kam-* (see Section 2.2) may also mean 'to owe' and this meaning, rather than the verb's volition meaning, might be the immediate source of its necessity uses (cf. English *ought*, the historical past of *owe*).

(34) Rumelian Romani (Balkan, Turkey; Paspati 1973 [1870], Boretzky 1999a: 113)

Therava	*te*	*phenav*	*tuke.*
have.PRES.1SG	COMP	say.SUBJ.1SG	thou.DAT

'I have to tell you.'

(35) Servy Ghympeny (Northeastern, Ukraine)

Me	*na*	*kamam*		*dala*	*draba*
I.NOM	NEG	like.PRES.1SG		this.PL	drug.PL

save		*me*	*kamam*		*te*	*prelaw.*
which.PL		I.NOM	want.PRES.1SG		COMP	take.SUBJ.1SG

'I do not like these pills that I have to take.'

Impersonal de-volitional construction show a greater degree of grammaticalization. In these constructions the volition verb takes the default, third-person singular form and is accompanied by a reflexive pronoun in the accusative case, which is a way of decreasing the construction's valency; the modal subject NP is in an oblique case. In Prizren Arli the verb *mang-* encodes volition in a personal construction, but necessity in an impersonal–reflexive construction; this may have been modelled on Albanian (Boretzky 1999a: 114). In Kumanovo Kovački and Crimean Romani, on the other hand, the two verbs are now distinct: the original volition verb *kam-* had acquired a wide range of necessity functions within the impersonal–reflexive construction, but its volition meaning was later taken over by the verb *mang-* (see Section 2.2). Though the reflexive necessity construction (Type nPF1) is still attested in Crimean Romani (36a), the dialect now prefers dropping of the reflexive pronoun (36b), apparently without any semantic effect.

(36) Crimean Romani (Balkan, Ukraine and Russia)

a.
Tumenge	*kamela*	*pes*	*te*	*dž'an*	*othe.*
you.DAT	necessary.3SG.PRES	REFL	COMP	go.2PL.SUBJ	there

'You have to go there.'

b.
Amenge	*kamelas*	*te*	*udž'aras*	*e*	*brešındestar.*
we.DAT	necessary.3SG.IMPF	COMP	wait.1PL.SUBJ	DEF.OBL	rain.ABL

'We had to wait because of the rain.'

The Central dialects of Romani exhibit further developments of the impersonal–reflexive de-volitional construction (cf. Boretzky 1996a: 14–15). The accusative form of the singular reflexive pronoun *pe(s)* has been agglutinated to the volition verb *kam-*. In the Vendic subgroup of the Central dialects the reflexive was prefixed and some of the forms then underwent further phonological erosion, e.g. present **pe + kam-el* [REFL + want-3SG] > *peka-l*. In the Rumungro and North Central subgroups, on the other hand, the reflexive was suffixed to the volition verb and the forms underwent not only phonological erosion but also externalization of verb inflections, e.g. present **kam-el + pe(s)* [want-3SG + REFL] > *ka(m)p-e(l)*. Typically, these impersonal necessity modals are used as expressions of participant-internal necessity, though in several Central dialects they have been extended to other necessity functions as well. Some of these latter dialects (viz. the outliers Gurvari, Cerhari and Plaščuny) retain the impersonal construction in all necessity functions, while others reserve the impersonal construction (Types nPnF or nPF1) for expressing participant-internal necessity (37a, 38a) but develop a functionally distinct personal construction (Type PnF) for other types of necessity. For example, in numerous Central dialects of Slovakia, e.g. Lučivná Romani, subject-inflected *kamp-* expresses weak obligation (37b) and in the dialect of Dobšiná it expresses participant-external necessity in general (38b).[13] Note that the strongly grammaticalized de-volitional necessity modal (e.g. 37a) is functionally distinct from the less grammaticalized impersonal–reflexive construction involving the verb *kam-* (37c; Type nPnF), which still expresses volition (see Section 2.2).

(37) Lučivná Romani (North Central, Slovakia)

a. *Kampel   mange te džal   andre   skľepa.*
    be_needed.PRES.3SG I.DAT COMP go.SUBJ.3SG[=INF] in.DEF   shop
    'I need to go to the shop.'

b. *Furt   kampes   avka te   kerel.*
    always be_needed.PRES.2SG so   COMP   do.SUBJ.2SG
    'You should always act like this.'

c. *Kamľas   pes   mange odoj te   džal.*
    want.PRET.3SG   REFL.ACC I.DAT there COMP go.SUBJ.3SG[=INF]
    'I wanted to go there.'

(38) Dobšiná Romani (North Central, Slovakia)

a. *Kampel   mange te zav   andro foro.*
    be_needed.PRES.3SG I.DAT COMP go.SUBJ.1SG in.DEF town

---

[13] See Section 2.5, ex. (53) for a more complex example from Burgenland Roman.

'I need to go to town.'

b. *Kampav*   *te*   *zal*   *andro foro.*
    must.PRES.1SG   COMP go.SUBJ.3SG[=INF]in.DEFtown
    'I have to go to town.'

The majority of Romani necessity modals are dialect-specific **lexical borrowings**. Only few Romani dialects do without a necessity loanword, namely some of those that possess strongly grammaticalized de-copular and/or de-volitional necessity modals. Loans of the following necessity expressions are attested within Romani: Turkish *lâzım*, Greek *prepi*, South Slavic *trjabva treba(ti), mora(ti)* and *valjati*, Romanian *trebui* and *musai*, Hungarian *muszáj*, Slovak *musieť, maťʼ* and *potrebovaťʼ,* Polish *musieć, trzeba* and *powinien*, East Slavic *treba, dolžen, nado* and *prixoditsja*, German *müssen* and *brauchen*, English *ought*, Swedish *måste* and *må*, and Finnish *pitää*.[14] The dialectal distribution of these loanwords is too complex to be discussed here (see Boretzky & Igla 2004: 184–187 for a partial overview). Some loanwords are restricted to dialects whose speakers are currently bilingual in the source language, while others have been retained even after the Romani speakers shifted to a different L2. For example, the loan of Romanian *trebui* is found not only in most Romani dialects of Romanian bilinguals but also in most non-Romanian Vlax dialects, some of which lost contact with Romanian centuries ago; it has also diffused from Vlax into some non-Vlax dialects of Macedonia and Bulgaria.

Many, perhaps most, borrowed necessity modals retain the range of functions they have in their source language, especially when the Romani speakers are currently bilingual in it. For example, the loan of English *ought* in Welsh Romani expresses weak obligation or probability and the loan of Slovak *potrebovaťʼ*in some varieties of Slovak Romani expresses participant-internal need. However, the function of Romani necessity loanwords may also differ from that of the source modal, especially if the source language is not the dialect's current L2 and/or if the loan acquires a different morphosyntactic status than it has in the source language. For example, the loan of the Polish uninflected and generic modal *trzeba* 'one must, one should' (cf. Hansen, in press) need not have a generic subject in Kohila Romani (39a) and may even show personal inflection (39b). On the other hand, loanwords from an older L2 may also show functional specialization, often due to competition with loanwords from the current L2. For example, Ub Gurbet *trubu-* appears to be specialized for weak obligation (40a), although its source modal, Romanian *trebui,* is much more general;

---

[14] Several of these necessity modals are themselves borrowings into the languages that serve as the immediate source for Romani: Turkish *lâzım* from Arabic, Romanian *trebui* from Slavic, West Slavic *musieť musieć,* Hungarian *muszáj* and Romanian *musai* from German (the latter two cases from *muss sein* 'must be').

the current Serbian loan *mora-* has taken over most necessity functions, including strong obligation (40b).

(39) Kohila Romani (Northeastern, Estonia)
a. *Jöj   čebi   lōve   te   janel.*
   she   must   money.PL   COMP bring.SUBJ.3SG
   'She must bring the money.'
b. *Tu   čebisos   te   jēs   an   gošti.*
   thou   must.PRES.2SG   COMP come.SUBJ.2SG   in   guest.PL
   'You must come for a visit.'

(40) Ub Gurbet (South Vlax, Serbia)
a. *Tu   trubus   te   posetis   ma   jekar.*
   thou   ought.PRES.2SG   COMP visit.SUBJ.2SG   I.ACC once
   'You should visit me once.'
b. *Uvek   moraš   t-   aes   pažljivo.*
   always   must.PRES.2SG   COMP come.SUBJ.2SG   careful
   'You must always be careful.'

Necessity loanwords inflecting for subject categories are mostly morphologically adapted through loanverb suffixes, e.g. Dolenjski Romani *mor-in-* < Slovene *mor-*, and show Romani verb inflection. Retention of L2 subject inflection is rare, being only attested in some of those Romani dialects that borrow the Serbian or Macedonian personal verb *mora-*. For example, in Kosovo Bugurdži (Boretzky 1993: 74) singular present forms of the modal retain their Serbian inflection, i.e. 1SG *mora-m*, 2SG *mora-š*, and 3SG *mora.* However, the third-person singular form *mora* is also used with plural subjects of any person, like in colloquial Serbian. In further dialects of Serbian and Macedonian Romani the generalization of the third-person singular form has been completed and *mora* is now used irrespective of person or number of the modal subject (Type nPF1). The impersonal *mora* mostly encodes TAM categories, either analytically, through the inflectional auxiliary, e.g. Serbian Kalderaš past *mora sas* [must COP.PRET.3], or synthetically, through indigenous TAM morphology, e.g. Šutka Arli imperfect *mora-ine* [must-REM]. On the other hand, impersonal modals may also acquire subject inflection in Romani. For example, the loan of *trebui*, which only inflects for TAM categories in Romanian, has developed subject inflection in several Vlax dialects.[15] In some of them the borrowed modal shows defective

---

[15] However, *trebui* has personal inflections in some non-standard varieties of Romanian, and so some instances of the loanword's personal inflection in Romani may in fact simply retain this property.

subject inflection and personal constructions alternate with impersonal ones. For example, the Austrian Lovari loan *trubu-* only has third-person and second-person forms (Cech & Heinschink 1999a: 63), which are used personally (41a–b; Type PF). First-person subjects, on the other hand, require an impersonal construction (41c–d; Type nPF1). Though there are distinct second-person forms, the impersonal construction may also be employed with second-person subjects (41e).

(41) Austrian Lovari (North Vlax, Austria)

a. *Trubun*   *te*   *den*   *e*   *love.*
    need.PRES.3PL COMP give.SUBJ.3PL   DEF.PL   money.PL
    'They need to give money.' (Cech & Heinschink 1999a: 73)

b. *Trubundanas*   *sama*   *te*   *les*   *tu.*
    need.PLPF.2SG attention   COMP take.SUBJ.2SG   thou
    'You should have paid attention.' (Cech *et al.* 1999: 64)

c. *Me*   *trubujas*   *lašo*   *šavo*   *t-*   *avav.*
    I.NOM   need.IMPF.3SG   good   son   COMP come.SUBJ.1SG
    'I should be a good son.' (Cech *et al.* 1999: 48)

d. *Trobundas*   *i*   *hera*   *opre*   *te*   *las.*
    need.PRET.3SG   DEF   clover   up   COMP take.SUBJ.1PL
    'We needed to lift the clover.' (Cech *et al.* 1999: 82)

e. *Trubujas*   *tu*   *t-*   *aves*   *o*   *kraj*   *ande*   *gado*
    need.IMPF.3SG thou   COMP come.SUBJ.2SG   DEF   king   in   this
    them.
    country
    'You should become the king of this country.' (Cech *et al.* 1999: 50)

Those loans of impersonal necessity verbs that remain impersonal in Romani are often morphologically integrated as Romani-inflected verb forms of the default, third-person singular, subject category, e.g. Kumanovo Arli present *treb-el* [-3SG], *treb-el-aine* [-3SG-REM], preterite *treb-in-g'-a* [-LOAN-PFV-3SG.PFV] < Macedonian *treba*. Only rarely do we find internally derived TAM forms that do not contain the default subject morphology. For example, the Sliven Muzikanta imperfect *trjabv-as* is derived directly from the Bulgarian present form *trjabva* by means of a Romani tense/mood suffix. Quite commonly, on the other hand, impersonal necessity verbs are borrowed without any morphological integration and retain their L2 inflection for TAM categories, e.g. present *prepi* and imperfect *eprepe* from Greek, present *trjabva* and past *trjabvaše* from Bulgarian, present *trebuje* < *trebuie* and imperfect *trebuja* < *trebuea* from Romanian, or present *pittää* and conditional *pitäs* < *pittäis* from (colloquial) Finnish. Several Romani modals inflect for

TAM categories even though their source forms do not. For example, the uninflected Swedish modal *måste* has been borrowed as an impersonal verb (Type nPnF) into Finnish Romani, showing synthetic TAM marking, e.g. present–future *most-u-l-a* [-LOAN-3SG-PFUT] and preterite *most-u-dīl-o* [-LOAN-PFV-3SG.PFV.M].

A common development is a complete loss of inflection of borrowed modal verbs. For example, in Srem Gurbet the loan of the Serbian personal verb *mora-* lost not only subject inflection but also TAM inflection: the original third-person singular present form *mora* is now used in all TAM categories, which are thus marked solely on the modalized verb (Type nPF2). Similarly, Welsh Romani and some varieties of Slovak and Lithuanian Romani have generalized forms such as *mus muš mos musi*, which go back to inflectional stems or the third-person singular present forms of the German personal modal *müssen* and its loans into Western Slavic, e.g. *musi* < Slovak 3SG *mus-í*. These generalized forms may then fuse with the complementizer *te*, resulting in uninflected univerbal modals such as *mos-te* in Čáry Romani (see Section 2.1, 10b), which parallel the strongly grammaticalized indigenous modals *hun-te* and *si-te* in their morphosyntax (Type nPF2). A similar development also gave rise to the uninflected modal *braux-te* in German Sinti < German *brauchen*.

Sometimes we find, within a single Romani variety, competition or complementary distribution between different constructions involving necessity loanwords. Several dialects reveal a TAM asymmetry here: the present indicative is more susceptible to borrowing, more likely to remain unintegrated, and less likely to show subject inflection than other TAM categories. In Vidin Kalajdži, for instance, the integrated impersonal loanverb *trubu-z-* (from Romanian, an old L2) is used in all TAM categories but in the present indicative it is currently being replaced by the unintegrated *trebva* (from Bulgarian, the current L2). In Slavošovce Romani the Hungarian-origin modal *mušaj* must show analytical subject and TAM inflection (Type PnF) in most TAM categories (42a), while in the present the personal construction (42b) is rare and the modal is usually uninflected (42c; Type nPF2). Similar patterns are also attested in Finnish Romani, where the uninflected *mote* (probably from Swedish *må* plus the complementizer *te*) is restricted to the present, while the impersonal *most-u-* (from Swedish *måste*) is used in all TAM categories; in some varieties of Macedonian Arli, where the loan of Macedonian *mora-* inflects for subject categories in the past, e.g. *mora-n-dil-jum* [must-LOAN-PFV-1SG.PFV] 'I had to', even though the generalized impersonal *mora* is used in the present; and elsewhere.

(42) Slavošovce Romani (North Central, Slovakia)

a. *Mušaj*   *ščamaš*   *te*   *užarel.*
   must   COP.PRET.1PL   COMP wait.SUBJ.3SG[=INF]
   'We had to wait.'

b. *Mušaj*   *ščom*   *odoj*   *te*   *dzal.*

	must	COP.PRES.1SG	there	COMP go.SUBJ.3SG[=INF]

'I have to go there.'

c. *Adádžive mušaj megin te dzav odoj.*

today	must	again	COMP go.SUBJ.1SG	there

'Today I have to go there again.'

Several dialects of Bulgarian and Greek Romani, including the dialect of Vălči Dol, show competition in the past between impersonal constructions with TAM marking on the modal (43b; Type nPF1) and impersonal constructions with TAM marking on the modalized verb (43c; Type nPF2).[16] In some of these dialects, e.g. Cerovo Cocomanja, Sliven Muzikanta (cf. Section 2.1, ex. 4) or Kalamata Romani, the former construction expresses situational necessity (44a), while the latter construction appears to be specialized for epistemic necessity (44b).

(43) Vălči Dol Romani (South Vlax, Bulgaria)

a. *Lazımi te džas ote.*

necessary	COMP go.SUBJ.2SG

'You have to go there.'

b. *Lazımi sasdı te nakav latar.*

necessary	COP.PRET.3	COMP walk.SUBJ.1SG	she.ABL

'I had to go around her.'

c. *Lazımi te ačasas sostar delas brıšind.*

necessary	COMP stay.IMPF.1PL	what.ABL	give.IMPF.3SG	rain

'We had to stay because it was raining.'

(44) Kalamata Romani (Balkan, Greece)

a. *Eprepe te cikaras jati delas brisim.*

must.IMPF	COMP wait.SUBJ.1PL	because	give.IMPF.3SG	rain

'We had to wait because it was raining.'

b. *Siγura prepi te bidzardan -es ta apo prin.*

surely	must.PRES	COMP get_to_know.PRET.2SG	OBJ.3SG.M	also	of

before

'Surely you must have got to know him before.'

---

[16] Note that the Turkish modal *lâzım* must be morphologically adapted by the nominal suffix *-i* in those dialects that allow its analytic inflection through the inflectional auxiliary, including in Vălči Dol Romani. The loan remains unadapted in those dialects that do not allow its analytic inflection, i.e. in those where TAM categories are marked solely on the modalized verb (Type nPF2).

In addition to borrowed necessity modals we also find their loan translations. While Russian and Ukrainian dialects of Romani usually borrow the East Slavic impersonal verb *prixoditsja* (lit. reflexive 'come') in its L2 form, also retaining its L2 inflection, some varieties prefer to use its 'semicalqued' or 'calqued' equivalent, e.g. present *pri-dža-l-pe* [AKTIONSART-go-3SG-REFL] or *jav-el-pe* [come-3SG-REFL]. Similarly, the East Slavic adjectival modal *dolžen* 'obliged', which still also retains its lexical meaning 'indebted' in the source languages, is mostly borrowed (and, in some varieties, morphologically integrated in the way borrowed adjectives are) but can also be translated by the adjective *bango* 'crooked; indebted'. Interestingly, since the Russian copula does not encode person in the past (45a), the default, third-person, form of the Romani copula is selected in the 'calqued' construction (45b; Type nPF1) in Russian Romani. However, personal copula occurs in Crimean Romani (cf. 9).

(45) Russian Romani (Northeastern, Russia)
a.   *Tu   dolžen byl   te   dykxes les.*
     thou   [obliged.SG.M   COP.PTC.SG.M]Russian   COMP see.SUBJ.2SG
     he.ACC
b.   *Tu   bango sys   te   dykxes les.*
     thou   obliged.SG.M COP.PRET.3 COMP see.SUBJ.2SG he.ACC
     'You must have seen him.'

2.5. Modality and negation

Negation in modal constructions may have scope over the modalized proposition or over the modality itself (cf. Palmer 2001: 90–98). In Romani the differences in the scope of negation are reflected especially in the linear order of the negator, as exemplified for the Rumungro (South Central) variety of Selice, Slovakia. In (46b) and (47b) the negator immediately precedes the main verb, having scope over the modalized proposition: the constructions express, respectively, possibility and necessity of a negative proposition. In (46c) and (47c), on the other hand, the negator immediately precedes the modal, being cumulated with it in (46c), and has scope over the modality: the constructions express, respectively, negative possibility and negative necessity of an affirmative proposition. Finally, in (46d) and (47d) both negators are used: the constructions express, respectively, negative possibility and negative necessity of a negative proposition.[17]

(46) Possibility   (47) Necessity

---

[17] The example (47d) has been elicited but might not occur in natural discourse.

a. *šaj le keras*
can he.ACC do.SUBJ.1PL
'we can/may do it'

b. *šaj le na keras*
can he.ACC NEG do.SUBJ.1PL
'it is possible that we don't do it'
i.e. 'we need not do it'

c. *naštig le keras*
cannot he.ACC do.SUBJ.1PL
'it isn't possible that we do it'
i.e. 'we cannot/may not do it'

d. *naštig le na keras*
cannot he.ACC NEG do.1PL.SUBJ
'it isn't possible that we don't do it'
i.e. 'we must do it'

a. *site le keras*
must he.ACC do.SUBJ.1PL
'we must do it'

b. *site le na keras*
must he.ACC NEG do.SUBJ.1PL
'it is necessary that we don't do it'
i.e. 'we must not do it'

c. *na site le keras*
NEG must he.ACC do.SUBJ.1PL
'it isn't necessary that we do it'
i.e. 'we need not do it'

d. *na site le na keras*
NEG must he.ACC NEG do.SUBJ.1PL
'it isn't necessary that we don't do it'
i.e. 'we can/may do it'

In most Romani dialects, as in Selice Rumungro, negators preceding (or cumulated with) the modal have scope over modality. A single exception is attested: in Welsh Romani the negator preceding the necessity modal *mus* (48a) has scope over the proposition, so that *na mus* (48b) encodes necessity of a negative proposition ('must not') rather than negative necessity ('need not'). This is clearly due to pattern borrowing from English, although *mus* itself is more likely to be of German origin (Sampson 1926: I, 216). In the remainder of this section we will be concerned solely with negative modality, focusing on the formal relationship between affirmative and negative modality.

(48) Welsh Romani (British, Wales; Sampson 1926, II: 236)
a. *Mus   te   ves   mansa.*
must   COMP come.SUBJ.2SG   I.INSTR
'You must come with me.'

b. *Na   mus   te   čas   odoj   čə́ŕlā.*
NEG   must   COMP stay.SUBJ.1PL   there   long
'We must not stop there long.'

Although negative **volition** is expressed by regular negation of the affirmative volition modal in most Romani dialects, a few dialects of Bulgaria employ a different modal for negative volition than for affirmative volition. For example, in Pazardžik Malokonare the original volition modal *kam-* is retained in negative contexts (49b), while the newly grammaticalized volition modal *mang-* (see Section 2.2) is used in affirmative contexts

(49a). The Rešitare dialect of Velingrad represents a further stage in the replacement of *kam-* by *mang-*: while the former is restricted to negative contexts, the latter may now be used in both polarities. Another rare type of irregularity concerns the shape of the negator: in Cerovo Cocomanja most verbs take the indicative negator *na*, while the volition verb *kam-* takes *nan*, which probably derives from the third-person present negative copula form *nanaj*.

(49) Pazardžik Malokonare (Balkan, Bulgaria)

a. *Mangava   te   žav   andi   zi.*
   want/demand.PRES.1SG   COMP go.SUBJ.1SG in.DEF town
   'I want to go to town.'

b. *Na   kamamte   žav   andi   zi.*
   NEG   want/love.PRES.1SG COMP go.SUBJ.1SG in.DEF town
   'I don't want to go to town.'

Negative **possibility** is frequently irregular in that it is often not expressed by regular negation of the affirmative possibility modal. Although the Early Romani negative possibility modal *našti* may continue a regularly negated form of the affirmative possibility modal *ašti* or derive from a construction involving the negator and the affirmative possibility modal *šaj* (see Section 2.3), in neither case is *našti* completely regular in synchronic terms. Indeed, pairs such as *šaj* vs *našti* can be even considered to be suppletive. Nevertheless, many Romani dialects have discontinued this Early Romani heritage as a result of hyperanalysis (cf. Croft 2000: 121–126), analogical regularization, grammaticalization and borrowing.

The first source of regular negation of possibility is the semantic shift of negative possibility modals into modals of affirmative possibility. This development, which has occurred in most modern Sinti dialects, consists in an addition of a more transparent negator (the German-origin *nit/nix* or *gar*) to a former negative possibility modal (*našti* or *naj*) and a subsequent hyperanalysis of the inherent negative value of the modal as a property of the negator alone (cf. Boretzky & Igla 2004: 183–184, Elšík & Matras 2006: 159). Examples from Austrian Sinti, a dialect in a transitional stage of the development, are illustrative: the modal *naj* may still express the negative possibility function by itself (50a), though it is more likely to be negated by a borrowed negator in this function (50b); the prevailing function of plain *naj* is affirmative possibility (50c). Those Sinti varieties that have undergone the complete shift now possess regularly negated possibility modals, e.g. Auvergne Manuš *našti* 'can' vs *našti gar* 'cannot'. Although the addition of a more transparent negator to the negative possibility modal has also occurred in several Balkan dialects of Romani, giving rise to forms such as *na-našti*, there is no hyperanalysis of

negation and so plain *našti*, if there is one, continues to function as a negative possibility modal in these dialects.[18]

(50) Austrian Sinti (Northwestern, Austria)

a. *Naj   bešē   imər   kheri.*
   cannot sit.PRES.2SG always home
   'You can't sit at home all the time.'

b. *Naj   fārena   butər nit   kheri.*
   can[not] drive.PRES.3PL   more NEG   home
   'They can't drive home anymore.'

c. *Kola   nicərens   man   vri   kaj   naj   nicərens   lə   man   vri.*
   that.PL use.IMPF.3PL I.ACC out   where can   use.IMPF.3PL they   I.NOM out
   'They used me as much as they were able to use me.'

Next, the negative possibility modal *našti* may be re-analyzed as containing the regular verb negator and a novel affirmative possibility modal may be created by analogy. Boretzky's (1996a: 5–6) claim that this kind of analogical decomposition gave rise to the affirmative *ašti* in general (see Section 2.3) is somewhat problematic in that it does not account well for the form's initial vowel. Nevertheless, analogical decomposition is likely to be the source of *šti ← na-šti* in Moravian and Abbruzian Romani and of *sti(k) ← na-sti(k)* in Lombardian and Piedmontese Sinti (these dialects retain the Early Romani negator *na*).[19] Regular negation of inherited possibility modals may also result from an analogy in the opposite direction. Thus several Romani dialects of, or originating, in Romania have supplemented or even replaced the original negative possibility modal *našti* by a construction consisting of a regular verb negator and affirmative possibility modal, e.g. Kurturare *či šaj-*, Šanxajcy *či dasči(-sar)-*, Laješa–Kišinjevcy–Rakarenge *ni dašti(-sar)-*, Ursari *ni (d)ašti(-z)-*, and probably also Kərəmidarja *na aštik*.

Finally, regular negation is also found with those possibility modals that have been grammaticalized from lexical verbs (see Section 2.3), e.g. *džan-* 'to be able' vs. *na džan-* 'not to be able', and in those dialects that borrow a possibility modal and negate it by a regular negator. The negator may be indigenous, in which case Romani 'semicalques' the regular negative possibility expression of the source language of the possibility modal, e.g.

---

[18] In Taikon Kalderaš of Sweden inflected *našti-* 'not to be capable' requires an additional analytical negator, while uninflected *našti* 'cannot' does not allow one (cf. Section 2.3). Again, there is no hyperanalysis of negation in the former case.

[19] Diffusion from another Italian Romani dialect is likely to be the source of *šti* in Italian Xoraxane, unless one wants to assume that the analogical decomposition of *našti* took place before the development of the innovative (South Vlax) indicative negator *ni*.

Parakalamos Romani *na bor-*, Finnish Romani *na vojuv-*, Lithuanian Romani *na možyn-*. Or, less commonly, the possibility modal may be borrowed together with its L2 negator, as in Russian Romani and in some Ukrainian dialects, e.g. in Kiev Servy (51).

(51) Kiev Servy (Ukrainian, Ukraine)

Ame	*n'it'ev*	*n'e*	*maɣl'i*		*kada*	*te*	*arakhe.*
we.NOM	nowhere	[NEG	can.PRET.PL][Russian]		this	COMP	find.INF

'We couldn't find this anywhere.'

However, semicalquing and matter borrowing of negative possibility modals only represents a minor pattern within Romani: most dialects that borrow an affirmative possibility modal retain the indigenous negative *našti*, and so they re-iterate or introduce suppletion between their affirmative and negative possibility modals. Also, the use of a borrowed possibility modal in negative possibility does not automatically result in a regularly negated negative possibility expression. Both points may be illustrated by examples from Karditsa Romani. This dialect borrows its possibility modal from Greek (52a) but retains the indigenous negative possibility modal (52b). There is nevertheless another way to express negative possibility: the borrowed affirmative possibility modal may combine with the indigenous negative possibility modal, which in effect functions as an irregular negator in the domain of possibility (52c).

(52) Karditsa Romani (Balkan, Greece)

a.
Borinea	*te*	*keres*	*buti*	*me*	*sfiri?*
can.PRES.2SG	COMP	do.SUBJ.2SG	work	with	hammer

'Can you work with a hammer?'

b.
Ame	*nasti*	*džaa*	*pala.*
we.NOM	cannot	go.PRES.1PL	back

'We can't go back.'

c.
Nasti	*borinava*	*ti*	*dav*	*boja*	*o*	*kher.*
cannot	can.PRES.1SG	COMP	give.SUBJ.1SG	dye	DEF	house

'I can't paint the house.'

Negative **necessity** is mostly expressed by regular negation of affirmative necessity modals. Exceptions to this general pattern are contact induced. For example, in Hameln Sinti the negative counterpart of the necessity particle *hunte* is a regular negation of the modal *brauxte*, which cannot occur without a negator (Holzinger 1993: 92).[20] This restriction

---

[20] The source (Holzinger 1993) is not explicit about whether *hunte* can be regularly negated, or not.

derives from an identical restriction on the source element, the German verb *brauchen* 'to need', when used with a clausal complement. A similar, though somewhat more complex, pattern applies to the necessity verb *peka(m)-* (see Section 2.4) in Burgenland Roman. When used impersonally (Type nPnF) the verb expresses participant-internal necessity (53a) and can be regularly negated (53b). However, when the verb is inflected for subject categories within a personal construction (Type PnF) it must be accompanied by a negator and the construction expresses negative participant-external necessity (53c). Regular negation of the participan-external necessity modal *iste* appears to be extremely rare (53d).[21] Halwachs (1998: 158) explains this asymmetry as a result of pattern borrowing from German.

(53) Burgenland Roman (South Central, Austria; Halwachs *et al.* 1999)

a. *Le   fačunge   eklik   te   hal*
   DEF.OBL   child.PL.DAT a_little COMP eat.SUBJ.3SG[=INF]
   pekalahi.
   be_needed.IMPF.3SG
   'The children would need to eat a bit.' (p. 48)

b. *Ni   te   hal,   ništa   lake   na*
   not_even COMP eat.SUBJ.3SG[=INF] nothing   she.DAT   NEG
   pekal.
   be_needed.PRES.3SG
   'She does not even need to eat or anything else.' (p. 66)

c. *Na   pekajs   mandar   te   daral.*
   NEG   must.PRES.2SG   I.ABL   COMP   fear.SUBJ.3SG[=INF]
   'You need not be afraid of me.' (p. 96)

d. *Na   iste   uso   sikajipe   džan.*
   NEG   must   to.DEF teaching   go.PRES.3PL
   'They need not go to the course.' (Romani Patrin 1999: 8)

3. Grammaticalization and borrowing asymmetries

3.1. Functional aspects of grammaticalization

Most of those Romani modal expressions that are not material borrowings derive from Romani lexical verbs or other **pre-modal** constructions. The sources for modals were discussed in the data sections (2.2–4) and are summarized in (54–56). Most of the

---

[21] It is unattested in a large collection of authentic Burgenland Roman narratives (Halwachs *et al.* 1999) and not discussed in the grammar of the dialect (Halwachs 1998).

developments are dialect-specific, with the exception of the pre-split extension of the verb 'to love' to volition functions (see Section 2.1). Modal uses of copula constructions are also likely to have existed in Early Romani, though grammaticalization of these constructions into dedicated modals is certainly dialect-specific (see Section 2.4). The majority of the developments in (54–56) involved contact-induced grammaticalization on the model of Romani's L2s and some grammaticalized modals even result from straightforward calquing or semicalquing (e.g. those based on Russian *prixoditsja*, see Section 2.3).

(54) volition   < 'love, like', 'ask for, demand'
      less agentive volition   < 'want' reflexive/middle

(55) possibility   < 'be/have', 'work out, go out', 'get [somewhere]'
      capability   < 'know', 'understand', 'manage'
      permission   < 'dare'

(56) necessity   < 'be/have', 'have, hold', 'want, owe', 'come' refl., 'indebted'
      need, weak obligation < 'want' reflexive

Some modals are less grammaticalized than others in that they still retain their lexical meanings as well, thus possessing greater semantic weight or integrity (cf. Lehmann 2002: 112). This holds especially for the volition modals, including the verb *kam-* (e.g. Selice Romani 'to love, to like; to wish, to desire; to owe; to want'), whose modal function must have existed since Early Romani and which thus shows remarkable semantic stability. Lexical meanings are also retained in several more recently grammaticalized modals, such as those expressing capability (e.g. Crimean Romani *hal'ov-* 'to understand; to feel; to guess; to manage; to be able') and those that calque polysemous source expressions (e.g. Russian Romani *bango* 'crooked; indebted; obliged', see Section 2.4). Modals expressing participant-internal necessity mostly also possess the lexical meaning 'to need [sth.]', though several dialects differentiate between expressions of event-oriented/modal need and object-oriented/lexical need, e.g. de-volitional *kamp-* vs. borrowed *potrebin-* in some dialects of Slovakia.

Semantic developments between different **kinds of modality** include the extension of modals specialized for participant-internal possibility and necessity (capability and need) into other possibility and necessity functions as well (cf. the discussion of *džan-* and *kamp-*, respectively, in Sections 2.3 and 2.4). This is in line with van der Auwera & Plungian's (1998) modality's semantic map. We have suggested in Section 2.3 that Early Romani possessed two possibility modals, *šaj* and *ašti*. The evidence from Kalderaš, a dialect cluster that preserves both *šaj* and *dašti-* (most likely a variant of *ašti*), shows that the latter tends to

be less grammaticalized semantically (according to van der Auwera & Plungian's map).[22] In Piteşti Kalderaš *šaj* is only attested as an expression of epistemic possibility, the most grammaticalized semantic function, whereas *dašti-* dominates all possibility functions. In Taikon Kalderaš of Sweden, on the other hand, the modal *šaj* expresses participant-external, deontic or epistemic possibility (57a–c; Type nPF2), while *dašti-* is specialised for participant-internal possibility (cf. the glossary item 'be able, be capable, can, manage' in Gjerdman & Ljungberg 1963: 222; Type PF). This distinction is parallelled by a similar, if not identical, functional distinction in the negative domain, probably due to morphological analogy: uninflected *našti* can express most impossibility functions (57d–e; Type nPF2), but inflected *našti-* is specialized for participant-internal impossibility (57f; Type PF). The Kalderaš data suggest that the Early Romani *ašti* was specialized for participant-internal possibility (perhaps due to a relatively late grammaticalization from a copula construction) and that its semantic extension to participant-external possibility is a recent, dialect-specific, development.[23]

(57) Taikon Kalderaš (North Vlax, Sweden; Gjerdman & Ljungberg 1963: 127)

a. *Šaj*   *dav*   *tu*   *deš*   *kronurjä.*
    can   give.PRES.1SG   thou.ACC   ten   crown.PL
    'I can give you ten crowns.'

b. *Šaj*   *keres*   *sar*   *kames.*
    can   do.PRES.2SG   how   want.PRES.2SG
    'You may do as you like.'

c. *Šaj*   *avel*   *voj*   *tehara.*
    can   come.PRES.3SG   she   tomorrow
    'She may/might come tomorrow.'

d. *Našti*   *ankərdine*   *lako*   *trajo.*
    cannot   hold.PRET.3PL   she.GEN   life
    'They couldn't save her life.'

e. *Našti*   *vazdel*   *o*   *gono*   *korkoro.*
    cannot   lift.PRES.3SGDEF   bag   alone.M
    'He can't lift the bag himself.'

---

[22] In some Kalderaš varieties both possibility modals appear to have identical functions (e.g. in a Šanxajcy variety of Ukraine and in the Bunkuleš and Markuleš varieties of Banat and northern Serbia, cf. Boretzky 1994: 137). Unfortunately, the sources on other dialects that possess reflexes of both modals, Moravian Romani and Italian Xoraxane with *šaj* and *šti*, contain no detailed information on modal functions.

[23] The evidence from a variety of Šanxajcy Kitajake is puzzling: in this Kalderaš-type dialect *dašti-* only has the lexical meaning 'to dare', while all possibility functions are expressed by *šaj*. This either assumes de-grammaticalization of *dašti-* or questions the etymological connection between *ašti* and *dašti-*.

f. *Či   naštin   te   keren   amenge   khanči.*
    NEG   unable.PRES.3PL COMP do.SUBJ.3PL we.DAT   nothing
    'They aren't able to cause any harm to us.'

Of all modalities, volition is most prone to further grammaticalisation processes in Romani. The development of volition modals into necessity modals has been discussed in Section 2.4. In addition, in some North Vlax dialects we find the epistemic possibility adverb *kam* 'perhaps', which might derive from the verb *kam-*, although a Romanian etymology is also possible (Boretzky 2003: 67, Boretzky & Igla 2004: 173). Volition modals can also grammaticalize into **post-modal** meanings. Some varieties of Slovak Romani have developed the de-volitional free-choice postfix *-kam* in de-interrogative indefinites, e.g. *so-kam* 'anything whatsoever' (< *'what one wants'). More importantly, in all Romani dialects of the Balkans the volition verbs have given rise to the future proclitics *kam kan ka* < *kam-* or *ma mə* < *mang-*, which are modelled on the future markers deriving from volitional verbs in other Balkan languages (Greek *θa* and so on; Boretzky 1999b: 170–172).

Future meanings may also be encoded by the impersonal copula construction, which othewise tends to grammaticalize into possibility and, more often, necessity modals (see Sections 2.3–4). In several Romani dialects of Macedonia and Bulgaria this development has occurred especially in negative polarity due to pattern borrowing from East South Slavic (Boretzky 1999b: 172–173). For example, in Montana Kalajdži affirmative future is marked by the de-volitional proclitic *ka(m)* (57a), while the negative future is encoded through a construction containing the negative third-person present copula form and the non-factual complementizer (57b), calquing Bulgarian *šte* and *njama da*, respectively. In several dialects of Romania and Kosovo impersonal copula constructions encode future in both polarities, due to an imperfect replication of the personal possessive-like future construction in Romanian and Albanian. For example, we find future constructions with present copula forms in Ursari (cf. Boretzky 1999a: 113) and constructions with future copula forms in Kərəmidarja (58a–b). Although the development of future from necessity is well attested cross-linguistically, it is not obvious that this is what happened in the above Romani dialects. It is quite possible that, at least in some of the dialects, the semantically vague copula construction acquired the future function directly, due to pattern borrowing from an L2, without ever passing through the necessity stage. In any case, post-modal developments of strongly grammaticalized possibility and necessity modals are unattested in Romani.

(57) Montana Kalajdži (Balkan, Bulgaria)
a. *Tase   kam- ovav   amende.*
    tomorrow   FUT   become.SUBJ.1SG   we.LOC
    'Tomorrow I will be at home [at our place].'

b. | *Tase* | *nanaj* | *te* | *ove* | *tumende.* |
|---|---|---|---|---|
| tomorrow | COP.NEG.PRES.3 | COMP | become.SUBJ.2SG | you.LOC |

'Tomorrow you won't be at home [at your place].'

(58) Kərəmidarja (North Central, Romania)[24]

a. | *Tese* | *ala* | *te* | *avaw* | *khere* |
|---|---|---|---|---|
| tomorrow | come.FUT.3SG | COMP | come.SUBJ.1SG | home |

'Tomorrow I will be at home.'

b. | *Tese* | *na* | *ala* | *te* | *aves* | *khere.* |
|---|---|---|---|---|---|
| tomorrow | NEG | come.FUT.3SG | COMP | come.SUBJ.2SG | home |

'Tomorrow you won't be at home.'

3.2. Formal aspects of grammaticalization

The least grammatical Romani modals are fully fledged morphological verbs that, like most lexical verbs, inflect for TAM and subject categories. Syntactically, these modal verbs have scope over the whole clause that expresses the modalized proposition: the modalized verb occurs in a subordinate form and is, disregarding some exceptions that are not specific to modal constructions (see Section 2.1), introduced by the non-factual complementizer. Also, the nominative marking of the modal subject NP, which is the only option with personal modals, can be interpreted as marking of the grammatical subject of the modal verb. There are three major diachronic classes of personal modal verbs in Romani. First, many are lexical borrowings of L2 personal modal verbs, thus simply retaining the morphosyntactic properties of their source forms. Second, several modal verbs such as volition *kam-* and *mang-*, (cap)ability *džan-* and *axaljov-*, permission *troma-* or necessity *ther-* derive from lexical verbs. Significantly, this class of modals overlaps almost perfectly with the class of modals that still retain their pre-modal meanings as well, and so the low degree of morphosyntactic grammaticalization of these modal verbs parallels their low degree of semantic grammaticalization.

Finally, and perhaps most interestingly, dedicated modals, too, may be personal verbs. This is the case of the possibility modals *šaj(-in)-*, *(d)ašti(-z/sar)-* and *našti(-n/sar)-* in a few (formerly contiguous) dialects (see Section 2.2). These possibility verbs do not represent an earlier, less grammaticalized, stage of the uninflected *šaj*, *ašti* and *našti*, which are found in the rest of Romani. Rather, they must have acquired their personal inflection secondarily,

---

[24] Note that the Kərəmidarja future construction is not an instance of (directly) de-allative future (cf. Heine & Kuteva 2005: 103). Although *ala* is a form of the verb 'to come', this verb generally provides some TAM (viz. non-indicative and future) forms of the copula and the verb of existence ('to be, to become') in this dialect.

through analogy, one kind of evidence being their morphology otherwise typical of loanverbs. Similarly, the personal inflection of *kamp-* and *pekam-* in some dialects must be an analogical innovation based on the impersonal *kamp-* and *pekam-* (see Section 2.3), since the forms contain a reflex of the reflexive pronoun, which is only used in the third person. Despite their analogical rather than grammaticalization origin, the above possibility verbs tend to have less grammatical functions than their uninflected counterparts, provided they co-exist within a variety (cf. 57). On the other hand, if both personal and impersonal *kamp-* or *pekam-* co-exist within a variety, the personal (participant-external) modal is more grammatical semantically than the impersonal (participant-internal) modal, which appears to contradict the degree of their morphosyntactic grammaticality. The fact that the impersonal (participant-internal) necessity modals *kamp-* or *pekam-* also function as lexical verbs of need seems to be relevant here.

Impersonal modal verbs, i.e. modals inflecting for TAM but not subject categories, are obviously less differentiated morphologically than personal modal verbs. The lack of subject inflection may result from obligatorification of an impersonal modal construction (as in the case of the loans of Bulgarian *može* in some dialects, see Section 2.2) or from fossilisation of a frequent person–number form of a personal modal (as in the case of the borrowed *mora* in some dialects, see Section 2.3), and so, given that obligatorification and fossilization are recognized paradigmatic aspects of grammaticalization (cf. Croft 2000: 157), impersonal modals can be considered to be more grammatical than personal modals. If there is alternation between personal and impersonal modal forms, the latter tend to occur in the more frequent TAM categories. Like personal modal verbs, impersonal modal verbs have scope over the whole clause that expresses the modalized proposition: the modalized verb occurs in a subordinate form and is introduced by the complementizer. However, some impersonal modal verbs with finite complements (Type nPF1) have nominative marking of the modal subject NP, which is to be interpreted as the grammatical subject of the modalized verb rather than of the modal. This indicates a certain degree of syntactic degradation of the modal. Impersonal verb inflection is common especially with possibility and necessity modals, often due to the impersonality of the source or model expressions in Romani's L2s.

Clearly the most grammatical are those modals that do not inflect either for TAM or subject categories (Type nPF2), thus functioning as invariant markers of modality that modify a fully inflected verb. Syntactically, such modals do not influence the argument structure of the modalized verb, and so the modal subject is invariably marked as the grammatical subject of the modalized verb (i.e. nominative-marked with most verbs). Also, they are the only type of modals that allow an obligatory absence of the non-factual complementizer, which is the last remnant of the embedded status of the modalized verb. The development of uninflected modals thus involves both paradigmatic and syntagmatic morphosyntactic processes of grammaticalization: fossilization and/or morphological loss of

verbal inflection on the one hand and structural condensation, scope decrease and loss of independent syntactic status on the other (cf. Croft 2000: 157). In addition to modals that were inherited as uninflected from Early Romani (*šaj*, *ašti* and *našti*) and borrowed modals that are uninflected in the source language as well, there are quite a few instances of recently grammaticalized uninflected modals in the domain of possibility (*naj*, *trunti*, *darfte*) and especially necessity (*hunte*, *site*, *mus*, *moste*, *brauxte*, *mote* and more). A complete loss of inflection and independent syntactic status in expressions of volition, on the other hand, is associated with post-modal functions. If there is alternation between constructions with inflected vs uninflected modals, the latter tend to occur in the more frequent TAM categories. More importantly, they also tend to occur in the more grammaticalized functions within a certain modality, as the distinction between the the tensed modals expressing situational necessity and the default-tense (and so in a sense uninflected) modals expressing epistemic necessity in (4) and (44) illustrates.

All Romani modals are free forms, i.e. none is bound to the modalized verb as a clitic or even as an affix. This also holds for those Romani dialects that have been in contact with languages that possess bound modals (cf. possibility suffixes in Turkish and Hungarian): unlike free modals, bound modals are never borrowed in their material form; and if there is pattern replication it does not concern the degree of coalescence. Coalescence with the modalized verb has only affected the post-volition future markers (see Section 3.1), which are proclitics. Instances of coalescence of a modal with the non-factual complementizer *te* are restricted to uninflected modals, and so to possibility and necessity. Coalescence with the reflexive pronoun is typical of de-volitonal necessity.

3.3. Borrowing asymmetries

Modality is a domain that is conspicuously susceptible to structural borrowing in Romani. However, the borrowing behaviour is asymmetrical: Some modality categories are more likely to be borrowed than others. The overall likelihood modality categories to be affected by structural borrowing is expressed by the following hierarchy:

(59) necessity > possibility > volition

Necessity appears at the top of the borrowability scale, meaning that it is both the most frequently borrowed modality category in the cross-dialect sample, and that lower positions on the hierarchy – volition and possibility – are not likely to be borrowed unless necessity is borrowed too. The hierarchy seems compatible to some degree with the hierarchy of grammaticalization: Both possibility and necessity are targets in the grammaticalization process, rather than points of departure; in other words, they occupy higher positions on the

grammaticalization hierarchy. This appears to suggest that the absence of a highly grammaticalized, semantically dedicated expression for modality attracts borrowing as an enrichment to the system. Such an argument follows a traditional "gap-hypothesis" approach to borrowing: items that are "missing" in the recipient system are more likely to be borrowed from a donor system. It does not, however, explain the relative low susceptibility of volition to borrowing – which in Romani is the least grammaticalized modality category. It appears, rather, that the trigger for borrowing is to be found in the semantic-pragmatic functions involved. Prone to borrowing are those modality categories that are participant-external, rather than participant-internal. Arguably, necessity involves greater intensity of external pressure than possibility, allowing us to postulate the following hierarchy:

(60) external pressure (necessity) > external facilitation (possibility) > internal (volition)

Borrowing, in effect the speaker accommodating 'intrusions' from a (dominant) external language, is thus in a sense a metaphorical reflection of the actor succumbing to external pressure. In relation to language processing in communicative interaction, it is yet another expression of the susceptibility to borrowing (in an oral, minority language) of structures expressing contrast, discontinuity, uncertainty, or exemption. As discussed already in Elšík & Matras (2006: 385-386) and in Matras (1998), we may conclude that borrowing targets firstly those grammatical devices that are employed to process instances of potential tension between the message conveyed by the speaker, and the hearer's expectations. It is, therefore, the 'stress-effect' in processing discourse that leads the bilingual speaker to lose control over the 'correct' selection of language form, a production error which, unaffected by normative attitudes may lead to convergence or 'fusion' of the structural representation of the category in both languages – i.e. to structural 'borrowing'.

A rather different picture emerges from the distribution of borrowed inflection accompanying borrowed modality structures. While such borrowed inflection is common with modals expressing possibility, it is rather rare with necessity – in a way reversing the hierarchy observed in (59). (Borrowed inflection with volition is attested in only one dialects, Dolenjski). The explanation is likely to be related to the availability of impersonal modality markers in the contact languages, as well as the overall semantic-pragmatic tendency to express necessity as an external, impersonal force (cf. even the language-internal grammaticalisation from volition in *kam-* to an impersonal, reflexive necessity expression *kampel*).

4. Conclusions

Figure (1) presents an integrated map of the grammaticalization paths of Romani modals:

Figure 1: Romani modality: An integrated map of grammaticalization paths
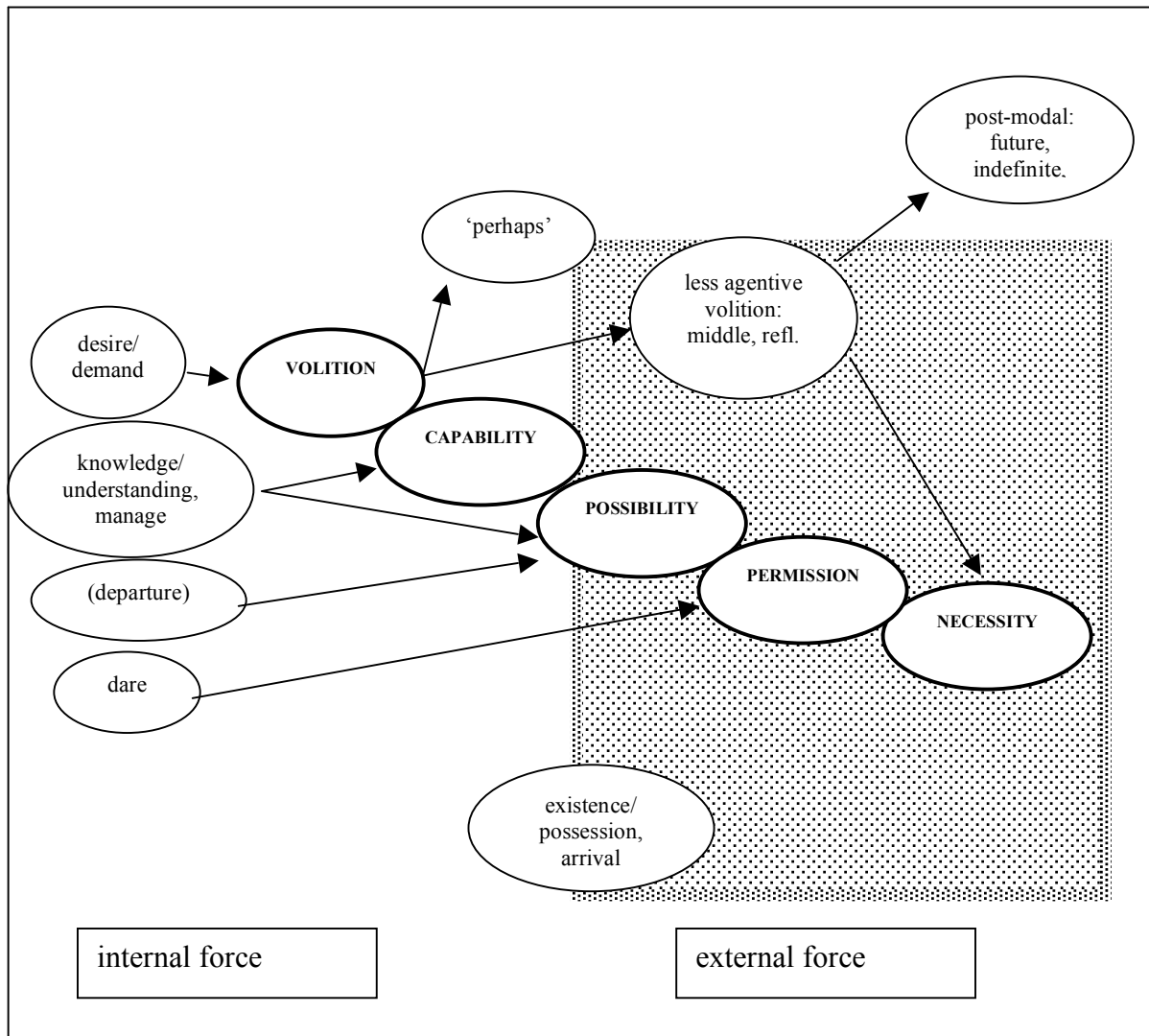
Two principal dimensions appear as relevant to the evaluation of modality and its grammaticalization paths. Modality categories can firstly be arranged on a cline reflecting the degree of `intensity` or `determination` – a conscious force driving the modal evaluation of a target action This cline interacts with a split between the anchoring of modality through `internal` and `external` force. On the internal-force side, determination is expressed as `agentivity`: It is strongest in volition, weaker in capability, and ambiguous or irrelevant in possibility. On the external-focre side, determination is expressed as intensive outside `pressure`: This is strong with necessity, weaker with permission, and ambiguous or irrelevant in possibility. Internal-force modality is of course present in volition and capability, but ambiguous in possibility, while external-force modality is present in necessity and permission, but also ambiguous in possibility. Possibility thus takes up a rather neutral position on the chart, free of explicit determination, and undefined for either internal-agentive or external-causative involvement.

The grammaticalization paths of Romani modals generally lead from predications depicting a mental-emotional state, a material state, or, marginally, movement. There is a strong tendency for internal-force predications to feed into internal-force modality, and for predications that are situated at least on the borderline between internal and external semantic conceptualizations to feed into external-force modality. Thus, existence is arguably externally-driven, as it is taken at face value without reference to the speaker's own emotions or intentions. Possession is somewhat ambiguous, as it entails a claim for possession on the part of the participant, but arguably also recognition of the possession relationship from an external source. In terms of its own grammaticalization path, Romani possession tends to draw on existence, supporting this ambiguity, at the very least, and quite possibly further biasing its external orientation. While the (marginal) grammaticalization of 'to dare' for permission is an exception to this tendency, the split between internal and grammaticalization paths receives support from the fact that possibility, occupying an ambiguous position, shows mixed source deriving both from expressions of internal and external states-of-affairs.

The two principal dimensions depicted in Figure (1) also serve as a useful map to interpret the role of structural borrowing in the Romani modality system. Borrowing is most frequently encountered when there is close correlation of external-force modality with intensity (in this case, intensive outside pressure), as discussed in 3.3.

On the entire chart, volition clearly stands out as the category with the greatest potential for further grammaticalization. Its grammatical derivations may cross the agentivity line to serve as impersonal, external-force modals, which in turn may serve as sources for the grammaticalization of necessity expressions. Volition expressions may also

end up in functions beyond strict modality, such as the epistemic future or the marking of indefiiteness.

References

Boretzky, Norbert. 1993. *Bugurdži. Deskriptiver und historischer Abriss eines Romani-Dialekts.* (= Balkanologische Veröffentlichungen, 21.) Wiesbaden: Harrasowitz.

Boretzky, Norbert. 1994. *Romani: Grammatik des Kalderaš-Dialekts mit Texten und Glossar.* Berlin: Harrassowitz Verlag.

Boretzky, Norbert. 1995a. Die Entwicklung der Copula im Romani. *Grazer Linguistische Studien* 43, 1–50.

Boretzky, Norbert. 1995b. Armenisches im Zigeunerischen (Romani und Lomavren). *Indogermanische Forschungen* 100, 137–155.

Boretzky, Norbert. 1996a. Zu den Modalia in den Romani-Dialekten. *Zeitschrift für Balkanologie* 32, 1, 1–27.

Boretzky, Norbert. 1996b. The "new" infinitive in Romani. *Journal of the Gypsy Lore Society*, V, 6, 1–51.

Boretzky, Norbert. 1999a. *Die Verwandschaftbeziehungen zwischen den Südbalkanischen Romani-Dialekten. Mit einem Kartenanhang.* Frankfurt am Main: Peter Lang.

Boretzky, Norbert. 1999b. Grammatical interference in Romani: Loan formations for foreign categories. *Acta Linguistica Hungarica* 46, 3–4, 169–200.

Boretzky, Norbert, and Birgit Igla. 1994. *Wörterbuch Romani-Deutsch-Englisch für den südosteuropäischen Raum: mit einer Grammatik der Dialektvarianten.* Wiesbaden: Harrassowitz.

Boretzky, Norbert, and Birgit Igla. 2004. *Kommentierter Dialektatlas des Romani.* Wiesbaden: Harrassowitz.

Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form.* Amsterdam: John Benjamins.

Cech, Petra, Christiane Fennesz-Juhasz, and Mozes F. Heinschink (eds.) 1999. *Lovarenge paramiči taj textura anda Österreich – Texte Österreichischer Lovara.* (= Arbeitsbericht 2 des Projekts *Kodifizierung der Romanes-Variante der Österreichischen Lovara*). Wien: Verein Romano Centro. Second, revised edition.

Cech, Petra, and Mozes F. Heinschink. 1999a. *Basisgrammatik.* (= Arbeitsbericht 1a des Projekts *Kodifizierung der Romanes-Variante der Österreichischen Lovara*). Wien: Verein Romano Centro.

Cech, Petra, and Mozes F. Heinschink. 1999b. *Sepečides-Romani: Grammatik, Texte und Glossar eines türkischen Romani-Dialekts.* (= Balkanologische Veröffentlichungen, 34.) Wiesbaden: Harrassowitz.

Cech, Petra, and Mozes F. Heinschink. 2001. Sinti Istriani, Lički Šijaci, Gopti und Cigani Brajdiči. Die Dolenjski Roma in Slowenien. In: Birgit Igla & Thomas Stolz (eds.) *"Was ich noch sagen wollte …": A multilingual Festschrift for Norbert Boretzky on occasion of his 65th birthday.* Berlin: Akademie Verlag. 341–369.

Croft, William. 2000. *Explaining language change: An evolutionary approach.* Harlow: Longman.

Elšík, Viktor, and Yaron Matras. 2006. *Markedness and language change: The Romani sample.* Berlin: Mouton de Gruyter.

Fennesz-Juhasz, Christiane, and Mozes F. Heinschink (eds.) 1999. *Lovarenge paramiči taj textura anda Österreich, kotor II – Texte Österreichischer Lovara II.* (= Arbeitsbericht 4 des Projekts *Kodifizierung der Romanes-Variante der Österreichischen Lovara*). Wien: Verein Romano Centro.

Franzese, Sergio. 1985. *Il dialetto dei Sinti Piemontesi. Note grammaticali. Glossario.* Torino: centro studi Zingari di Torino.

Gjerdman, Olof, and Erik Lundberg. 1963. *The language of the Swedish Coppersmith Gipsy Johan Dimitri Taikon: Grammar, texts, vocabulary and English word index.* Uppsala: Lundequist.

Halwachs, Dieter W. 1998. *Amaro vakeripe Roman hi — Unsere Sprache ist Roman. Texte, Glossar und Grammatik der burgenländischen Romani-variante.* Klagenfurt: Drava.

Halwachs, Dieter W. in co-operation with Gerd Ambrosch, Ursula Glaeser, Katharina Martens, and Michael Wogg. 1998. *Morphologie des Roman.* (= Arbeitsbericht 3a des Projekts *Kodifizierung und Didaktisierung des Roman*). Oberwart: Verein Roma.

Halwachs Dieter W., Gerd Ambrosch, Michael Wogg in co-operation with Ursula Glaeser, and Katharina Martens. 1999. *Märchen und Erzählungen der Burgenland-Roma.* (= Arbeitsbericht 8a des Projekts *Kodifizierung und Didaktisierung des Roman*). Oberwart: Verein Roma.

Hansen, Björn. 2000. The German modal verb *müssen* and the Slavonic languages: The reconstruction of a success story. *Scando-Slavica* 46, 77–92.

Hansen, Björn. 2003. The Grammaticalization (and Degrammaticalization?) of Modals in Slavonic. In: Berger T. / Gutschmidt K. (Hrsg.) Funktionale Beschreibung slavischer Sprachen. Beiträge zum 13. Internationalen Slavistenkongress Ljubljana 2003. München, 97-115.

Hansen, Björn. In Press. Modals. In: Kosta, P. / Berger, T. / Gutschmidt, K. / Kempgen, S. (eds.) Slavic languages: An international handbook of their history, their structure and

their investigation (Handbücher zur Sprach- und Kommunikationswissenschaft). Berlin: De Gruyter.

Heine, Bernd, and Tania Kuteva. 2005. *Language contact and grammatical change.* Cambridge: Cambridge University Press.

Holzinger, Daniel. 1993. *Das Romanes: Grammatik und Diskursanalyse der Sprache der Sinte.* (= Innsbrucker Beiträge zur Kulturwissenschaft, 85.) Innsbruck: Verlag des Instituts für Sprachwissenschaft der Universität Innsbruck.

Lehmann, Christian. 2002. *Thoughts on grammaticalization.* Erfurt: Seminar für Sprachwissenschaft der Universität (ASSidUE, 9). Second, revised edition.

Manchester Romani Project website: www.llc.manchester.ac.uk/Research/Projects/romani

Matras, Yaron. 1998. Utterance modifies and universals of grammatical borrowing. *Linguistics* 36, 281-331.

Matras, Yaron. 2002. *Romani: a linguistic introduction.* Cambridge: Cambridge University Press.

Matras, Yaron. 2004. Romacilikanes – The Romani dialect of Parakalamos. *Romani Studies* 5, 14:1, 59–109.

Mészáros, György. 1980. *A Magyarországi szinto cigányok (történetük és nyelvük).* Budapest: Magyar nyelvtudományi társaság.

Müller, Friedrich. 1869. Beiträge zur Kenntniss der Rom-Sprache. *Sitzungsberichte der philosophisch-historischen Classe der Kaiserlichen Akademie der Wissenschaften zu Wien* 61, 1, 149–206.

Palmer, F. R. 2001. *Mood and modality.* Second edition. Cambridge: Cambridge University Press.

Paspati, Alexandre G. 1973 [1870]. *Études sur les Tchinghianés ou Bohèmiens de l'empire ottoman.* Osnabrück: Biblio Verlag.

Pott, August. 1844–1845. *Die Zigeuner in Europa und Asien. Ethnographisch-linguistische Untersuchung vornehmlich ihrer Herkunft und Sprache.* Halle: Heynemann.

RMS = Romani Morpho-Syntactic Database. Viktor Elšík & Yaron Matras 2001–present; current version supported by Christopher White, Charlotte Jones, and Christa Schubert. See: Manchester Romani Project.

Romani Patrin, December 1999, No. 4. Oberwart: Verein Roma.

Sampson, John. 1926. *The dialect of Gypsies of Wales, being an older form of British Romani preserved in the speech of the clan of Abram Wood.* Oxford: Clarendon Press.

van der Auwera, Johann, and Plungian V. A. 1998. Modality's semantic map. *Linguistic Typology* 1:2, 79–124.